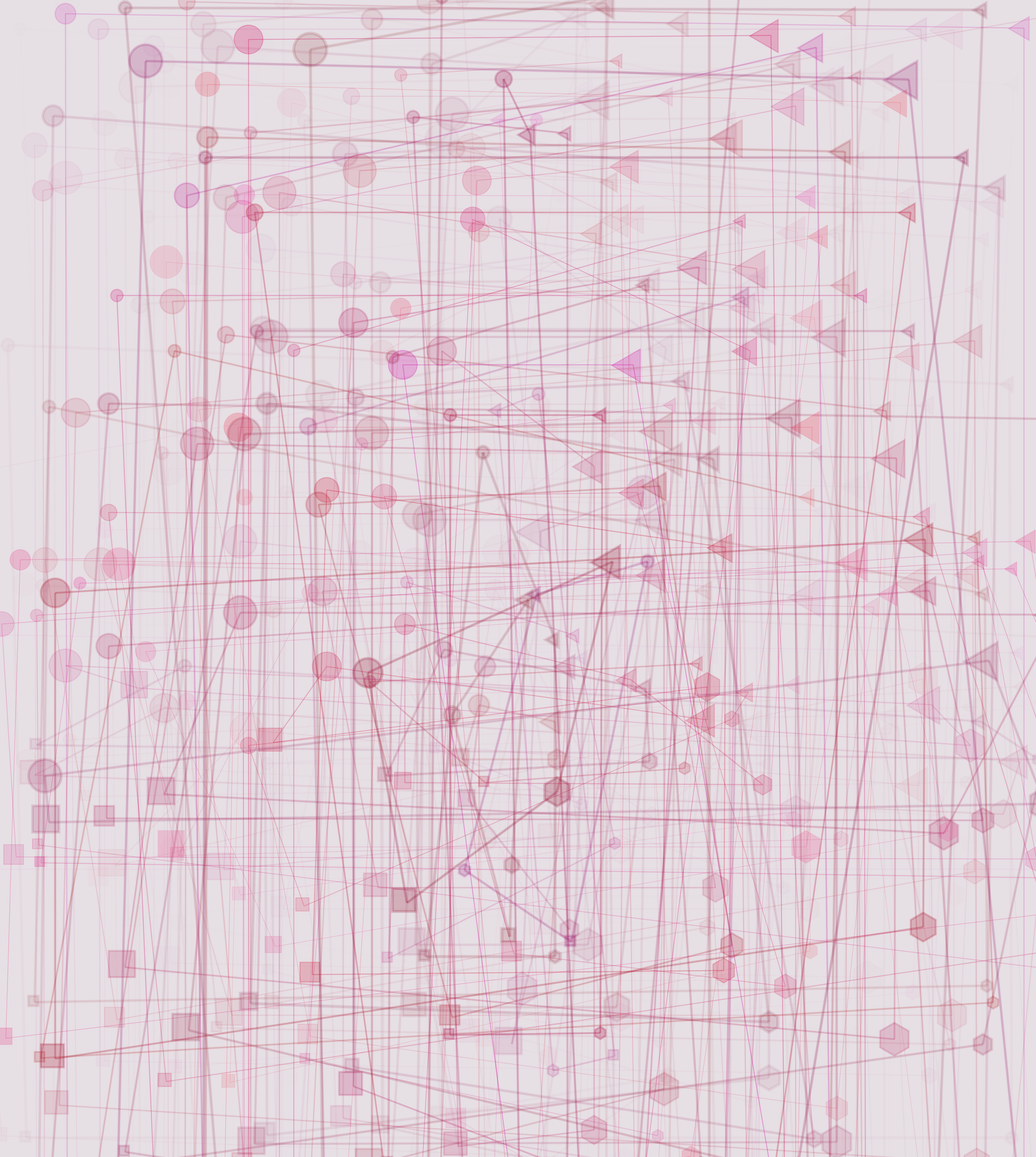


PROOF THEORY, RULES & MEANING

Greg Restall

PROOF THEORY, RULES & MEANING

Greg Restall

VERSION OF OCTOBER 26, 2018

Comments on this draft are most welcome.
Email me at restall@unimelb.edu.au.

CONTENTS

ACKNOWLEDGEMENTS

This manuscript has been a very long time coming. Many people have contributed to my understanding of these topics over the last fifteen years. My debts are plentiful.

[Acknowledgements of specific people to come.]

Greg Restall

Melbourne

OCTOBER 26, 2018

WHERE TO BEGIN

INTRODUCTION

I should like to outline an image which is connected with the most profound intuitions which I always experience in the face of logistic. That image will perhaps shed more light on the true background of that discipline, at least in my case, than all discursive description could. Now, whenever I work even on the least significant logistic problem, for instance, when I search for the shortest axiom of the implicational propositional calculus I always have the impression that I am facing a powerful, most coherent and most resistant structure. I sense that structure as if it were a concrete, tangible object, made of the hardest metal, a hundred times stronger than steel and concrete. I cannot change anything in it; I do not create anything of my own will, but by strenuous work I discover in it ever new details and arrive at unshakable and eternal truths. Where is and what is that ideal structure? A believer would say that it is in God and is His thought. — Jan Łukasiewicz

This is a draft of a monograph on proof theory and philosophy. The focus will be a detailed examination of the different ways to understand *proof*, and how understanding the norms governing logical vocabulary can give us insight into questions in the philosophy of language, epistemology and metaphysics. Along the way, we will also take a few glances around to the other side of logical consequence, the kinds of *counterexamples* to be found when an deduction fails to be valid.

The book is designed to serve a number of different purposes, and it can be used in a number of different ways. In writing the book I have two distinct aims in mind.

GENTLY INTRODUCING KEY IDEAS IN PROOF THEORY FOR PHILOSOPHERS: There are a number of very good books that introduce proof theory: for example, Bostock's *Intermediate Logic* [27], Tennant's *Natural Logic* [237], Troelstra and Schwichtenberg's *Basic Proof Theory* [240], and von Plato and Negri's *Structural Proof Theory* [153] are all excellent books, with their own virtues. However, they all introduce the core ideas of proof theory in what can only be described as a rather complicated fashion. The core *technical* results of proof theory (normalisation for natural deduction and cut elimination for sequent systems) are *relatively* simple ideas at their heart, but the expositions of these ideas in the available literature are quite difficult and detailed. This is through no fault of the existing literature. It is due to a choice. In each book, a proof system for the whole of classical or intuitionistic logic is introduced, and then, formal properties are demonstrated about such a system. Each proof system has different rules for each of the connectives, and this makes the proof-theoretical results such as normalisation and cut elimination case-ridden and lengthy. (The standard techniques are complicated inductions with different cases for each connective: the more connectives and rules, the more cases.)

In this book, the exposition will be rather different. Instead of taking a proof system as given and proving results about *it*, we will first look at the core ideas (normalisation for natural deduction, and cut elimination for sequent systems) and work with them in their simplest and purest manifestation. In Section 1.3 we will see a two-page normalisation proof. In Section 2.2 we will see a two-page cut-elimination proof. In each case, the aim is to understand the key concepts behind the central results. Then, we show how these results can be generalised to a much more abstract setting, in which they can be applied to a wide range of logical systems, and once we have established these general results, we apply them to specific systems of interest, including first order predicate logic, propositional modal and temporal logics, and quantified modal logics.

EXPLORING THE CONNECTIONS BETWEEN PROOF THEORY AND PHILOSOPHY: The central part of the book (Chapters 4 to 6) answer a central question in philosophical proof theory: When do inference rules define a logical concept? The first part of the book (Chapters 1 to 3) introduces the tools and techniques needed to both *understand* and to *address* the question. The central part of the book formulates the problem and offers a distinctive solution to it. A very particular kind of inference rule (a rule we will describe as a *defining rule*) defines a concept satisfying some very natural conditions—and there are good reasons to think of concepts satisfying these conditions as properly *logical* concepts. Then the remainder of the book (from Chapter 7) explores consequences and applications of these ideas for particular issues in logic, language, epistemology and metaphysics. Alone the way, we will explore the connections between proof theories and theories of meaning. What does this account of proof tell us about how we might *apply* the formal work of logical theorising? All accounts of meaning have something to say about the role of inference. For some, it is what things *mean* that tells you what inferences are appropriate. For others, it is what inferences are appropriate that helps constitute what particular words *mean*. For everyone, there is an intimate connection between inference and semantics.

The precise defintion is spelled out, along with its consequences, in Chapter 6.

I have in mind the distinction between *representationalist* and *inferentialist* theories of meaning. For a polemical and provocative account of the distinction, see Robert Brandom's *Articulating Reasons* [31].

The book includes marginal notes that expand on and comment on the central text. Feel free to read or ignore them as you wish, and to add your own comments. Each chapter (other than this one) contains definitions, examples, theorems, lemmas, and proofs. Each of these (other than the proofs) are numbered consecutively, first with the chapter number, and then with the number of the item within the chapter. Proofs end with a little box at the right margin, like this: ∎

The manuscript is divided into three parts, each of which is divided into chapters. The first part, *Tools*, covers the basic concepts, arguments and results which we will use throughout the book. These chapters can be used as a gentle introduction to proof theory for anyone who is interested in the field, perhaps supplemented by (or supplementing) one or more of the texts mentioned earlier in this chapter. The second part, *The Core Argument*, introduces Prior's puzzle concerning inference rules and definitions, and presents and defends a distinct answer to that question. The answer takes the form of an *argument*, to the effect that a particular kind of rule—what I call a *defining rule*—can be used to introduce a logical concept into a discourse, and shows that this concept in an important sense both *free to add*, and *sharply delineated*. The third part, *Insights* then draws out the consequences of this argument to different kinds of logical concepts (the connectives, quantifiers, identity, and modal operators) and for different issues in the philosophy of language, epistemology, metaphysics and the philosophy of mathematics.

A slogan: A logical concept is one that can be introduced by means of a *defining rule*.

In addition to these three major parts, the book contains a small introduction designed to set the scene (this chapter) and a coda, which points forward to issues to be explored in the future.

Some chapters in the *Tools* section contain exercises to complete. Logic is never learned without hard work, so if you want to *learn* the material, work through the exercises: especially the basic and intermediate exercises, which should be taken as a guide to mastery of the techniques we discuss. The advanced exercises are more difficult, and should be dipped into as desired, in order to truly gain expertise in these tools and techniques. The *project* questions are examples of current research topics.

The book has an accompanying website: `http://consequently.org/writing/ptp`. From here you can look for an updated version of the book, leave comments, read the comments others have left, check for solutions to exercises and supply your own. Please visit the website and give your feedback. Visitors to the website have already helped me make this volume much better than it would have been were it written in isolation. It is a delight to work on logic within such a community, spread near and far.

MOTIVATION

Why? My first and overriding reason to be interested in proof theory is the beauty and simplicity of the subject. It is one of the central strands of the discipline of logic, along with its partner, model theory. Since the flowering of the field with the work of Gentzen, many beautiful definitions, techniques and results are to be found in this field, and they deserve a wider audience. In this book I aim to provide an introduction to proof theory that allows the reader with only a minimal background in logic to start with the flavour of the central results, and then understand techniques in their own right.

It is one thing to be interested in proof theory in its own right, or as a part of a broader interest in logic. It's another thing entirely to think that proof theory has a role in philosophy. Why would a *philosopher* be interested in the theory of proofs? Here are just three examples of concerns in philosophy where proof theory finds a place.

EXAMPLE 1: MEANING. Suppose you want to know when someone is using "or" in the same sense that you do. When does "or" in their vocabulary have the same significance as "or" in yours? One answer could be given in terms of *truth-conditions*. The significance of "or" can be given in a rule like this one:

Perhaps you've heard of the difference between 'inclusive' and 'exclusive' disjunction. And maybe you're worried that 'or' can be used in many ways, each meaning something different.

$$\ulcorner p \text{ or } q \urcorner \text{ is true if and only if } p \text{ is true or } q \text{ is true.}$$

Perhaps you have seen this information presented in a truth-table.

p	q	p or q
0	0	0
0	1	1
1	0	1
1	1	1

Clearly, this table can be used to distinguish between some uses of disjunctive vocabulary from others. We can use it to rule out *exclusive* disjunction. If we take ⌜p or q⌝ to be *false* when we take p and q to be both true, then we are using "or" in a manner that is at odds with the truth table.

However, what can we say of someone who is ignorant of the truth or falsity of p and of q? What does the truth table tell us about ⌜p or q⌝ in that case? It seems that the application of the truth table to our *practice* is less-than-straightforward.

It is for reasons like this that people have considered an alternate explanation of a logical connective such as "or." Perhaps we can say that someone is using "or" in the way that you do if you are disposed to make the following deductions to reason *to* a disjunction

$$\frac{p}{p \text{ or } q} \qquad \frac{q}{p \text{ or } q}$$

and to reason *from* a disjunction

$$\frac{p \text{ or } q \quad \begin{matrix}[p]\\ \vdots \\ r\end{matrix} \quad \begin{matrix}[q]\\ \vdots \\ r\end{matrix}}{r}$$

That is, you are prepared to infer *to* a disjunction on the basis of either disjunct; and you are prepared to reason by cases *from* a disjunction. Is there any more you need to do to fix the use of "or"? That is, if you and I both use "or" in a manner consonant with these rules, then is there any way that our usages can differ with respect to *meaning*?

Clearly, this is not the end of the story. Any proponent of a *proof*-first explanation of the meaning of a word such as "or" will need to say something about what it is to accept an inference rule, and what sorts of inference rules suffice to define a concept such as disjunction (or negation, or universal quantification, and so on). When does a definition work? What are the sorts of things that can be defined using inference rules? What are the sorts of rules that may be used to define these concepts? We will consider these issues in Chapter 6.

EXAMPLE 2: GENERALITY. It is a commonplace that it is impossible or very difficult to *prove* a nonexistence claim. After all, if there is *no* object with property F, then *every* object fails to have property F. How can we demonstrate that every object in the entire universe has some property? Surely we cannot survey each object in the universe one-by-one. Furthermore, even if we come to believe that object a has property F for each object a that happens to exist, it does not follow that we ought to believe that *every* object has that property. The universal judgement tells us more than the truth of each particular instance of that judgement, for given all of the objects $a_1, a_2, \ldots$, it certainly seems *possible* that a_1 has property F, that a_2 has property F and so on, without *everything* having property F since it seems possible that there might be some *new* object which does not *actually* exist. If you care to talk of 'facts' then we can express the matter by saying that the fact that everything is F cannot amount to just the fact that a_1 is F and the fact that a_2 is F, etc., it must also include the fact that $a_1, a_2, \ldots$ are all of the objects. There seems to be some irreducible *universality* in universal judgements.

If this was all that we could say about universality, then it would be very difficult to come to universal conclusions. However, we seem to manage to derive universal conclusions regularly. Consider mathematics: it is not difficult to prove that *every* whole number is either even or odd. We can do this without examining every number individually. Just how do we do this?

It is a fact that we *do* accomplish this, for we are able to come to universal conclusions as a matter of course. In the course of this book we will see how such a thing is possible. Our facility at reasoning with *quantifiers*, such as 'for every' and 'for some,' is intimately tied up with the structures of the claims we can make, and how the formation of judgements from *names* and *predicates* gives us a foothold which may be exploited in reasoning. When we understand the nature of proofs involving quantifiers, this will give us insight into how we can gain *general* information about our world.

It is one thing to know that $2+3=3+2$. It is quite another to conclude that for *every* pair of natural numbers n and m that $n+m=m+n$. Yet we do this sort of thing quite regularly.

EXAMPLE 3: MODALITY. A third example is similar. Philosophical discussion is full of talk of *possibility* and *necessity*. What is the significance of this talk? What is its logical structure? One way to give an account of the logical structure of possibility and necessity talk is to analyse it in terms of possible worlds. To say that it is possible that Australia win the World Cup is to say that there is some possible world in which Australia wins the World Cup. Talk of possible worlds helps clarify the logical structure of possibility and necessity. It is possible that either Australia or New Zealand win the World Cup only if there's a possible world in which either Australia or New Zealand win the World Cup. In other words, either there's a possible world in which Australia wins, or a possible world in which New Zealand wins, and hence, it is either possible that Australia wins the World Cup or that New Zealand wins. We have reasoned from the possibility of a disjunction to the disjunction of the corresponding possibilities. Such an inference seems correct. Is talk of possible worlds required to explain this kind of derivation, or is there some other account of the logical structure of possibility and necessity? If we agree with Arthur Prior that we understand possible worlds be-

"… possible worlds, in the sense of possible states of affairs are not *really* individuals (just as numbers are not *really* individuals). To say that a state of affairs obtains is just to say that something is the case; to say that something is a possible state of affairs is just to say that something could be the case; and to say that something is the case 'in' a possible state of affairs is just to say that the thing in question would necessarily be the case if that state of affairs obtained, i. e. if something else were the case … We understand 'truth in states of affairs' because we understand 'necessarily'; not *vice versa*." — Arthur Prior, *Worlds, Times and Selves* [177].

cause we understand the concepts of possibility and necessity, then it's incumbent on us to give some explanation of how we come to understand those concepts—and how they come to have the structure that makes talk of possible worlds appropriate. I will argue in this book that when we attend to the structure of proofs involving modal notions, we will see how this use helps determine the concepts of necessity and possibility, and this *thereby* gives us an understanding the notion of a possible world. We don't first understand modal concepts by invoking possible worlds—we can invoke possible worlds when we first understand modal concepts, and the logic of modal concepts can be best understood when we understand what modal reasoning is *for* and how we do it.

EXAMPLE 4: A NEW ANGLE ON OLD IDEAS Lastly, one reason for studying proof theory is the perspective it brings on familiar themes. There is a venerable and well-trodden road between truth, models and logical consequence. Truth is well-understood, models (truth tables for propositional logic, or Tarski's models for first-order predicate logic, Kripke models for modal logic, or whatever else) are taken to be models of truth, and logical consequence is understood as the preservation of truth in all models. Then, some proof system is designed as a way to give a tractable account of that logical consequence relation. Nothing in this book will count as an argument *against* taking that road from truth, through logical consequence, to proof. However, we will travel that road in the other direction. By starting with proofs we will retrace those steps in reverse, to construct *models* from a prior understanding of proof, and then with an approach to *truth* once we have a notion of a model in hand. This is a very different way to chart the connection between proof theory and model theory. At the very least, tackling this terrain from that angle will allow us to take a different perspective on some familiar ground, and will give us the facility to offer new answers to some perennial questions about meaning, metaphysics and epistemology. Perhaps, when we see matters from this new perspective, the insights will be of lasting value.

However, the notion of truth is beset by paradox, and this should at least serve as a warning sign. Using the notion of truth as a starting point to define core features of logic may not provide the most stable foundation. It is at least worth exploring different approaches.

These are four examples of the kinds of issues that we will consider in the light of proof theory in the pages ahead. To broach these topics, we need to learn some proof theory, so let's dive in.

PART I

Tools

NATURAL DEDUCTION 1

We start with modest ambitions. In this section we focus on one way of understanding proof—natural deduction, in the style of Gentzen [81]—and we will consider just one kind of judgement: *conditionals*.

Gerhard Gentzen, German Logician: Born 1909, student of David Hilbert at Göttingen, died in 1945 in World War II. http://www-groups.dcs.st-and.ac.uk/~history/Biographies/Gentzen.html For an extensive account of Gentzen's work, and life under the National Socialist regime, see Eckart Menzler-Trott's biography [141].

1.1 | CONDITIONALS

Conditional judgements have this shape

If . . . then . . .

where we can fill in both "..." with other judgements. Conditional judgements are a useful starting point for thinking about logic and proof, because conditionals play a central role in our thinking and reasoning, in reflection and in dialogue. If we move beyond judgements about what is the case to reflect on how our judgements hang together and stand with regard to one another, it is very natural to form *conditional* judgements. You may not want to claim that the Number 58 tram is about to arrive, but you may at be in a position to judge that *if the timetable is correct*, the Number 58 tram is about to arrive. This is a conditional judgement, with the antecedent "the timetable is correct," and consequent "the Number 58 tram is about to arrive."

"The aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term." — Wilfrid Sellars [220]

In the study of *formal* logic, we focus on the *form* or *structure* of judgements. One aspect of this involves being precise and attending to those structures and shapes in some detail. We will start this by defining a grammar for conditional judgements. Any grammar has to start somewhere, and we will start with labels for atomic judgements—those judgements which aren't themselves conditionals, but which can be used to build conditionals. We'll use the letters p, q and r for these atoms, and if they're not enough, we'll use numerical subscripts to make more—that way, we never run out.

$$p,\ q,\ r, \quad p_0,\ p_1,\ p_2, \ldots \quad q_0,\ q_1,\ q_2, \ldots \quad r_0, r_1,\ r_2, \ldots$$

Each of these formulas is an ATOM. Whenever we have two formulas A and B, whether A and B are ATOMs or not, we will say that $(A \to B)$ is also a formula. In other words, given two judgements, we can (at least, in theory) form the conditional judgement with the first as the antecedent and the second as consequent. Succinctly, this *grammar* can be represented as follows:

$$\text{FORMULA} ::= \text{ATOM} \mid (\text{FORMULA} \to \text{FORMULA})$$

This is BNF, or "Backus Naur Form," first used in the specification of formal computer programming languages such as ALGOL. http://cui.unige.ch/db-research/Enseignement/analyseinfo/AboutBNF.html

That is, a FORMULA is either an ATOM, or is found by placing an arrow (written like this '$\to$') between two FORMULAs, and surrounding the result with parentheses.

So, the next line contains four different formulas

$$p_3 \quad (q \to r) \quad ((p_1 \to (q_1 \to r_1)) \to (q_1 \to (p_1 \to r_1))) \quad (p \to (q \to (r \to (p_1 \to (q_1 \to r_1)))))$$

but these are not formulas:

$$t \qquad p \to q \to r \qquad p \to p$$

You can do without parentheses if you use 'prefix' notation for the conditional: 'Cpq' instead of '$p \to q$'. The conditional are then $CpCqr$ and $CCpqr$. This is *Polish notation*.

The first, t, fails to be a formula since is not in our set ATOM of atomic formulas (so it doesn't enter the collection of formulas by way of being an atom) and it does not contain an arrow (so it doesn't enter the collection through the clause for complex formulas). The second, $p \to q \to r$ does not enter the collection because it is short of a few parentheses. The only expressions that enter *our* language are those that bring a pair of parentheses along with every arrow: "$p \to q \to r$" has two arrows but no parentheses, so it does not qualify. You can see why it *should* be excluded because the expression is ambiguous. Does it express the conditional judgement to the effect that if p then if q then r, or is it the judgement that if it's true that if p then q, then it's also true that r? In other words, it is ambiguous between these two formulas:

$$(p \to (q \to r)) \quad ((p \to q) \to r)$$

We really need to distinguish these two judgements, so we make sure our formulas contain parentheses. Our last example of an offending non-formula, $p \to p$, does not offend nearly so much. It is not ambiguous. It merely offends against the letter of the law laid down, and not its spirit. I will feel free to use expressions such as "$p \to p$" or "$(p \to q) \to (q \to r)$" which are missing their outer parentheses, even though they are, strictly speaking, not FORMULAS.

If you like, you can think of them as including their outer parentheses very *faintly*, even more faintly than this: $((p \to q) \to (q \to r))$.

Given a formula containing at least one arrow, such as $(p \to q) \to (q \to r)$, it is important to be able to isolate its main connective (the last arrow introduced as it was constructed). In this case, it is the middle arrow. The formula to the left of the arrow (in this case $p \to q$) is said to be the *antecedent* of the conditional, and the formula to the right is the *consequent* (here, $q \to r$).

We can think of formulas generated in this way in at least two different ways. We can think of them as the sentences in a very simple language. This language is either something completely separate from our natural languages, or it is a fragment of a natural language, consisting only of atomic expressions and the expressions you can construct using a conditional construction like "if . . . then . . .".

On the other hand, you can think of formulas as not constituting a language in themselves, but as constructions used to display the *form* of expressions in a language. Both of these interpretations of this syntax are open to us, and everything in this chapter (and in much of the rest of the book) is written with both interpretations in mind. Formal languages can be used to describe the forms of different languages, and they can be thought to be languages in their own right.

The issue of interpreting the formal language raises another question: What is the relationship between languages (formal or informal) and the judgements expressed in those languages? This question is not unlike the question concerning the relationship between a name and the bearer of that name, or a term and the thing (if anything) denoted by that term. The numeral '2' is not to be identified with number 2, and the formula $p \to q$ (or a sentence with that shape) is not the same as the conditional judgement expressed by that formula. Talk of judgements is itself ambiguous between the act of judging (my act of judging that the Number 58 tram is coming soon is not the same act as your act of judging this), and the *content* of any such act. When it comes to interpreting and applying the formal language of logic, it is important to reflect on not only the languages that you and I might speak (or write, or use in computer programs, etc.) but also attend to the *content* expressed when we use such languages [244].

The term "$1 + 1$" to be identified with the numeral "2", though both denote the same number. One term contains the numeral "1" and the other doesn't.

» «

Often, we will want to talk quite generally about all formulas with a given shape. We do this very often, when it comes to logic, because we are interested in the forms of valid arguments. The structural or formal features of arguments apply generally, to more than just a particular argument. (If we know that an argument is valid in virtue of its possessing some particular form, then other arguments with that form are valid as well.) So, these formal or structural principles must apply *generally*. Our formal language goes some way to help us express this, but it will turn out that we will not want to talk merely about specific formulas in our language, such as $(p_3 \to q_7) \to r_{26}$. We will, instead, want to say things like

> A *modus ponens* inference is the inference from a conditional formula and the antecedent of that conditional, to its consequent.

This can get very complicated very quickly. It is not easy to understand

> *Given a conditional formula whose consequent is also a conditional, the conditional formula whose antecedent is the antecedent of the consequent of the original conditional, and whose consequent is a conditional whose antecedent is the antecedent of the original conditional and whose consequent is the consequent of the conditional inside the first conditional follows from the original conditional.*

Instead of that mouthful, we will use *variables* to talk generally about formulas in much the same way that mathematicians use variables to talk generally about numbers and other such things. We will use capital letters, such as

$$A, B, C, D, \ldots$$

as variables ranging over the FORMULAS. So, instead of the long paragraph above, it suffices to say

Number theory books don't often include lots of *numerals*. Instead, they're filled with *variables* like 'x' and 'y.' This isn't because these books aren't about numbers. They are, but they don't merely list *particular* facts about numbers. They talk about *general* features of numbers, and hence the use of variables.

From $A \to (B \to C)$ *you can infer* $B \to (A \to C)$.

which seems much more perspicuous and memorable. The letters A, B and C aren't any *particular* formulas. They each can stand in for any formula at all.

Now we have the raw formal materials to address the question of deduction using conditional judgements. How may we characterise proofs reasoning using conditionals? That is the topic of the next section.

1.2 | PROOFS FOR CONDITIONALS

Start with some of reasoning using conditional judgements. One example might be reasoning of this form:

> *Suppose* $A \to (B \to C)$. *Suppose* A. *It follows that* $B \to C$. *Suppose* B. *It follows that* C.

This kind of reasoning has two important features. We make *suppositions*. We also infer *from* these suppositions. From $A \to (B \to C)$ and A we inferred $B \to C$. From this new information, together with the supposition that B, we inferred a new conclusion, C.

One way to represent the structure of this piece of reasoning is in this *tree diagram* shown here

$$\dfrac{\dfrac{A \to (B \to C) \qquad A}{B \to C} \qquad B}{C}$$

The *leaves* of the tree are the formulas $A \to (B \to C)$, A and B. They are the assumptions upon which the deduction rests. The other formulas in the tree are *deduced* from formulas occurring above them in the tree. The formula $B \to C$ is written immediately below a line, above which are the formulas from which we deduced it. So, $B \to C$ didn't have to be supposed. It *follows from* the leaves $A \to (B \to C)$ and A. Then the *root* of the tree (the formula at the bottom), C, follows from that formula $B \to C$ and the other leaf B. The ordering of the formulas bears witness to the relationships of inference between those formulas in our process of reasoning.

The two steps in our example proof use the same kind of reasoning. The inference from a conditional, and from its antecedent to its consequent. This step is called *modus ponens*. It's easy to see that using *modus ponens* we always move from more complicated formulas to less complicated formulas. However, sometimes we wish to infer the conditional $A \to B$ on the basis of our information about A and about B. And it seems that sometimes this is legitimate. Suppose we want to know about the connection between A and C in a context in which we are happy to grant both $A \to (B \to C)$ and B. What kind of connection is there (if any) between A and C? It would seem that it would be appropriate to infer $A \to C$, since we can derive C if we are willing to grant

"*Modus ponens*" is short for "*modus ponendo ponens*," which means "the mode of affirming by affirming." You get to the affirmation of B by way of the affirmation of A (and the other premise, $A \to B$). It may be contrasted with *Modus tollendo tollens*, the mode of denying by denying: from $A \to B$ and *not* B to *not* A.

A as an assumption. In other words, we have the means to conclude C from A, using the other resources we have already granted. But what does the conditional judgement $A \to C$ say? That if A, then C. So we can make that explicit and conclude $A \to C$ from that reasoning. We can represent the structure of chain of reasoning in the following way:

$$\dfrac{\dfrac{\dfrac{A \to (B \to C) \qquad [A]^{(1)}}{B \to C} \qquad B}{C}}{A \to C}\,[1]$$

This proof can be read as follows: At the step marked with [1], we make the inference to the *conditional* conclusion, on the basis of the reasoning up until that point. Since we can conclude C *using* A as an assumption, we can make the further conclusion $A \to C$. At this stage of the reasoning, A is no longer active as an assumption: we *discharge* it. It is still a leaf of the tree (there is no node of the tree above it), but it is no longer an active assumption in our reasoning. So, at this stage we bracket it, and annotate the brackets with a label, indicating the point in the demonstration at which the assumption is discharged. Our proof now has two assumptions, $A \to (B \to C)$ and B, and one conclusion, $A \to C$.

$$\dfrac{A \to B \qquad A}{B}\,{\scriptstyle \to E} \qquad\qquad \dfrac{\begin{matrix}[A]^{(i)}\\ \vdots \\ B\end{matrix}}{A \to B}\,{\scriptstyle \to I,i}$$

Figure 11: natural deduction rules for conditionals

We have motivated two rules for proofs with conditionals. These rules are displayed in Figure 11. The first rule, *modus ponens*, or *conditional elimination* [$\to E$] allows us to step from a conditional and its antecedent to the consequent of the conditional. We call the conditional premise $A \to B$ the *major* premise of the [$\to E$] inference, and the antecedent A the *minor* premise of that inference. When we apply the inference [$\to E$], we combine two proofs: the proof of $A \to B$ and the proof of A. The new proof has as assumptions any assumptions made in the proof of $A \to B$ and also any assumptions made in the proof of A. The conclusion is B.

The major premise in a connective rule features that connective.

The second rule, *conditional introduction* [$\to I$], allows us to use a proof from A to B as a proof of $A \to B$. The assumption of A is discharged in this step. The proof of $A \to B$ has as its assumptions all of the assumptions used in the proof of B except for the instances of A that we discharge in this step. Its conclusion is $A \to B$.

Now we come to the first formal definition, giving an account of what counts as a proof in this natural deduction system for the language of conditionals.

DEFINITION 1.1 [PROOFS FOR CONDITIONALS] A proof is a *tree* consisting of formulas, some of which may be *bracketed*. The formula at the root of a proof is said to be its CONCLUSION. The unbracketed formulas at the leaves of the tree are the PREMISES of the proof.

- Any FORMULA A is a proof, with premise A and conclusion A. The formula A is not bracketed.
- If π_l is a proof, with conclusion $A \to B$ and π_r is a proof, with conclusion A, then these proofs may be combined, into the following proof,

$$\dfrac{\begin{array}{c}\vdots\,\pi_l\\ A \to B\end{array} \qquad \begin{array}{c}\vdots\,\pi_r\\ A\end{array}}{B}\,{\to}E$$

  which has conclusion B, and which has premises consisting of the premises of π_l together with the premises of π_r.

How do you choose the number for the label (i) on the discharged formula? Find the largest number labelling a discharge in the proof π, and then choose the next one.

- If π is a proof with conclusion B, then the following tree

$$\dfrac{\begin{array}{c}[A]^{(i)}\\ \vdots\,\pi\\ B\end{array}}{A \to B}\,{\to}I,i$$

  is a proof with conclusion $A \to B$. Its premises are the premises of the original proof π, except for the premise A which is now discharged. We indicate this discharge by *bracketing* it.
- Nothing else is a proof.

This is a recursive definition, in just the same manner as the recursive definition of the class FORMULA. We define atomic proofs (in this case, consisting of a single formula), and then show how new (larger) proofs can be built out of smaller proofs.

$$\dfrac{\dfrac{\dfrac{[B \to C]^{(2)} \qquad \dfrac{A \to B \quad [A]^{(1)}}{B}\,{\to}E}{C}\,{\to}E}{A \to C}\,{\to}I,1}{(B \to C) \to (A \to C)}\,{\to}I,2$$

SUFFIXING (DEDUCTION)

$$\dfrac{\dfrac{\dfrac{[A \to B]^{(1)} \quad [A]^{(2)}}{B}\,{\to}E}{(A \to B) \to B}\,{\to}I,1}{A \to ((A \to B) \to B)}\,{\to}I,2$$

ASSERTION (FORMULA)

$$\dfrac{\dfrac{\dfrac{\dfrac{[A \to B]^{(3)} \qquad \dfrac{[C \to A]^{(2)} \quad [C]^{(1)}}{A}\,{\to}E}{B}\,{\to}E}{C \to B}\,{\to}I,1}{(C \to A) \to (C \to B)}\,{\to}I,2}{(A \to B) \to ((C \to A) \to (C \to B))}\,{\to}I,3$$

PREFIXING (FORMULA)

Figure 12: THREE IMPLICATIONAL PROOFS

Figure 12 gives three proofs constructed using our rules. The first is a proof from $A \to B$ to $(B \to C) \to (A \to C)$. This is the inference of *suffixing*. (We "suffix" both A and B with $\to C$.) The other proofs conclude in formulas justified on the basis of *no* undischarged assumptions. It is worth your time to read through these proofs to make sure that you understand the way each proof is constructed. A good way to understand the shape of these proofs is to try writing them out from top-to-bottom, identifying the basic proofs you start with, and only adding the discharging brackets at the stage of the proof where the discharge occurs.

You can try a number of different strategies when making proofs for yourself without copying existing ones. For example, you might like to try your hand at constructing a proof to the conclusion that $B \to (A \to C)$ from the assumption $A \to (B \to C)$. Here are two strategies you could use to piece a proof together.

TOP–DOWN: You start with the assumptions and see what you can do with them. In this case, with $A \to (B \to C)$ you can, clearly, get $B \to C$, if you are prepared to assume A. And then, with the assumption of B we can deduce C. Now it is clear that we can get $B \to (A \to C)$ if we discharge our assumptions, A first, and then B.

In either case, this is the proof we construct:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to (B \to C) \quad [A]^{(1)}}{B \to C}{\scriptstyle \to E} \qquad [B]^{(2)}}{C}{\scriptstyle \to E}}{A \to C}{\scriptstyle \to I,1}}{B \to (A \to C)}{\scriptstyle \to I,2}$$

BOTTOM–UP: Start with the conclusion, and find what you could use to prove it. Notice that to prove $B \to (A \to C)$ you could prove $A \to C$ using B as an assumption. Then to prove $A \to C$ you could prove C using A as an assumption. So, our goal is now to prove C using A, B and $A \to (B \to C)$ as assumptions. But this is an easy pair of applications of [$\to$*E*].

» «

Before exploring some more of the formal and structural properties of this kind of proof, let's pause for a moment to consider in more detail we might interpret the components of these proof structures. It is one thing to specify a formal structure as representing a network of connections between judgements. It is another to have a view of what kinds of connections between judgements are modelled in such a structure. To be specific: What kind of act is making an assumption? What kind of act is discharging that assumption that has been made? There are different things that you can say about this, but one way to understand the making of assumptions in a proof is that when you *suppose* A in a proof, you (under the scope of that assumption) treat the judgement A as if it had been asserted. You enter "A" in the "asserted" scoresheet, and treat it as if it had been asserted, for the purposes of reasoning, without actually undergoing the commitments. This enables us to infer *from* that commitment without actually having to undertake the commitment. We can attend to the inferential transition between A and B independently of actually asserting A. Doing so gives us a way to distinguish different senses of inferring B from A. The strong sense is the sense in which we have already granted A: To infer B from A in that context tells us that A and B both hold—or it commits us to the even strong claim, B *because* A. This can be distinguished from the weak sense of inferring B from A *hypothetically*, which tells us merely that *if* A *then* B. This incurs no commitment to B (or to A), but gives us a way to make explicit the inferential commitment we incur. With interpretation of supposition in mind, we can interpret proof structures as follows.

- An identity proof A represents the act of supposing A. Its *conclusion* is A, the very content that is supposed. Its only *active supposition* is A.

- Given a proof of the conclusions $A \rightarrow B$ (in which the suppositions in X are active) and a proof of A (in which the suppositions in Y are active), we have two corresponding (possibly complex) acts, the act of inferring $A \rightarrow B$ from X and the act of inferring A from Y. The proof given by extending those two proofs by an [$\rightarrow$*E*] step, to conclude B represents the complex act of (a) inferring $A \rightarrow B$ from X, (b) inferring A from Y, and then (c) deducing B from $A \rightarrow B$ and A. The active suppositions of this complex deduction are given in X and Y.

- Given a proof of B, in which the suppositions in X and A are active, this corresponds to the act of deducing B from X together with A. We interpret the proof of $A \rightarrow B$ from X, found by discharging A from the active assumptions, as representing the complex act of (a) first deducing B from X and A, and then (b) concluding $A \rightarrow B$, from the deduction from A to B. Now the conclusion is $A \rightarrow B$ and the active suppositions are those in X.

» «

I have been intentionally unspecific when it comes how formulas are discharged formulas in proofs. In the examples in Figure 12 you will notice that at each step when a discharge occurs, one and only one formula is discharged. By this I do not mean that at each [$\rightarrow$*I*] step a formula A is discharged and a different formula B is not. I mean that in the proofs we have seen so far, at each [$\rightarrow$*I*] step, a single *instance* of the formula is discharged. Not all proofs are like this. Consider this proof from the assumption $A \rightarrow (A \rightarrow B)$ to the conclusion $A \rightarrow B$. At the final step of this proof, two instances of the assumption A are discharged at once.

$$\dfrac{\dfrac{\dfrac{A \rightarrow (A \rightarrow B) \quad [A]^{(1)}}{A \rightarrow B}{\scriptstyle \rightarrow E} \qquad [A]^{(1)}}{B}{\scriptstyle \rightarrow E}}{A \rightarrow B}{\scriptstyle \rightarrow I,1}$$

For this to count as a proof, we must read the rule [$\rightarrow$*I*] as licensing the discharge of *one or more instances* of a formula in the inference to the conditional. Once we think of the rule in this way, one further generalisation comes to mind: If we think of an [$\rightarrow I$] move as discharging a *collection* of instances of our assumption, someone of a generalising spirit will ask if that collection can be empty. Can we discharge an assumption that *isn't there*? If we can, then *this* counts as a proof:

“Yesterday upon the stair, I met a man who wasn't there. He wasn't there again today. I wish that man would go away.” — Hughes Mearns

$$\frac{A}{B \rightarrow A}{\scriptstyle \rightarrow I,1}$$

Here, we assume A, and then, we infer $B \rightarrow A$ discharging *all* of the active assumptions of B in the proof at this point. The collection of active assumptions of B is, of course, empty. No matter, they are all discharged, and we have our conclusion: $B \rightarrow A$.

You might think that this is silly: how can you discharge a nonexistent assumption? Nonetheless, discharging assumptions that are not there plays a role. To give you a taste of why, notice that the inference from A to $B \rightarrow A$ is *valid* if we read “$\rightarrow$” as the material conditional of standard two-valued classical propositional logic. In a pluralist spirit we will investigate different policies for discharging formulas.

For more in a “pluralist spirit” see my work with Jc Beall [12, 13, 194].

DEFINITION 1.2 [DISCHARGE POLICY] A DISCHARGE POLICY may either allow or disallow *duplicate* discharge (more than one instance of a formula at once) or *vacuous* discharge (*zero* instances of a formula in a discharge step). Here are the names for the four discharge policies:

		VACUOUS	
		YES	NO
DUPLICATES	YES	*Standard*	*Relevant*
	NO	*“Affine”*	*Linear*

I am not happy with the label “affine,” but that's what the literature has given us. Does anyone have any better ideas for this? “Standard” is not “classical” because it suffices for intuitionistic logic in this context, not classical logic. It's not “intuitionistic” because “intuitionistic” is difficult to pronounce, and it is not *distinctively* intuitionistic. As we shall see later, it's the shape of proof and not the discharge policy that gives us intuitionistic implicational logic.

The “standard” discharge policy is to allow both vacuous and duplicate discharge.

There are reasons to explore each of the different policies. As I indicated above, you might think vacuous discharge does not make much sense. However, we can say more than that: it seems downright *mistaken* if we are to understand a judgement of the form $A \rightarrow B$ to record the claim that B may be inferred *from* A. If A is not used in the inference to B, then we hardly have reason to think that B follows from A in this sense. So, if you are after a conditional which is *relevant* in this way, you would be interested in discharge policies that ban vacuous discharge [3, 4, 184].

There are also reasons to ban duplicate discharge: Victor Pambuccian has found an interesting example of doing without duplicate discharge in early 20th Century geometry [157]. He traces cases where geometers took care to keep track of the number of times a postulate was used in a proof. So, they draw a distinction between $A \rightarrow (A \rightarrow B)$ and $A \rightarrow B$. The judgement that $A \rightarrow (A \rightarrow B)$ records the fact that B can be deduced from two uses of A. $A \rightarrow B$ records that B can be deduced from A used only once. More recently, work in *fuzzy logic* [20, 96, 143] motivates keeping track of the number of times premises are used. If a conditional $A \to B$ fails to be true to the degree that A is truer than B, then $A \to (A \to B)$ may be truer than $A \to B$.

Consider the claim I'll call (α)—*If (α) is true, then I am a monkey's uncle.*

Finally, for some [11, 146, 173, 191], Curry's Paradox motivates banning indiscriminate duplicate discharge. If we have a claim A which both implies $A \to B$ and is implied by it then we can reason as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{[A]^{(1)}}{A \to B}\,\dagger \quad [A]^{(1)}}{B}\,{\scriptstyle \to E}}{A \to B}\,{\scriptstyle \to I,1} \qquad \dfrac{\dfrac{\dfrac{\dfrac{[A]^{(2)}}{A \to B}\,\dagger \quad [A]^{(2)}}{B}\,{\scriptstyle \to E}}{A \to B}\,{\scriptstyle \to I,2}}{A}\,\dagger}{B}\,{\scriptstyle \to E}$$

Where we have used '†' to mark the steps where we have gone from A to $A \to B$ or back. Notice that this is a proof of B from *no premises at all*! So, if we have a claim A which is equivalent to $A \to B$, and if we allow vacuous discharge, then we can derive B.

DEFINITION 1.3 [KINDS OF PROOFS] A proof in which every discharge is *linear* is a *linear proof*. Similarly, a proof in which every discharge is *relevant* is a *relevant proof*, a proof in which every discharge is *affine* is an *affine proof*. If a proof has some duplicate discharge and some vacuous discharge, it is at least a *standard proof*.

We will *generalise* the notion of an argument later, in a number of directions. But this notion of argument is suited to the kind of proof we are considering here.

Proofs underwrite *arguments*. If we have a proof from a collection X of assumptions to a conclusion A, then the argument $X \therefore A$ is *valid* by the light of the rules we have used. So, in this section, we will think of *arguments* as structures involving a collection of assumptions and a single conclusion. But what kind of thing is that collection X? It isn't a *set*, because the number of premises makes a difference: (The example here involves linear discharge policies. We will see later that even when we allow for duplicate discharge, there is a sense in which the number of occurrences of a formula in the premises might still matter.) There is a linear proof from $A \to (A \to B), A, A$ to B:

$$\dfrac{\dfrac{A \to (A \to B) \quad A}{A \to B}\,{\scriptstyle \to E} \qquad A}{B}\,{\scriptstyle \to E}$$

We shall see later that there is *no* linear proof from $A \to (A \to B), A$ to B. (If we ban duplicate discharge, then the number of assumptions in a proof matters.) The *collection* appropriate for our analysis at this stage is what is called a *multiset*, because we want to pay attention to the number of times we make an assumption in an argument.

DEFINITION 1.4 [MULTISET] Given a class X of objects (such as the class FORMULA), a *multiset* M of objects from X is a special kind of collection of elements of X. For each x in X, there is a natural number $o_M(x)$, the number of OCCURRENCES of the object x in the multiset M. The number $o_M(x)$ is sometimes said to be the DEGREE to which x is a member of M. The multiset M is FINITE if $o_M(x) > 0$ for only finitely many objects x. The multiset M is identical to the multiset M' if and only if $o_M(x) = o_{M'}(x)$ for every x in X.

Multisets may be presented in lists, in much the same way that sets can. For example, $[1, 2, 2]$ is the finite multiset containing 1 only once and 2 twice. $[1, 2, 2] = [2, 1, 2]$, but $[1, 2, 2] \neq [1, 1, 2]$. We shall only consider finite multisets of *formulas*, and not multisets that contain other multisets as members. This means that we can do without the brackets and write our multisets as lists. We will write "A, B, B, C" for the finite multiset containing B twice and A and C once. The empty multiset, to which everything is a member to degree zero, is $[\,]$.

If you like, you could define a multiset of formulas to *be* the occurrence function $o_M(x)$ function $o_M : \text{FORMULA} \rightarrow \omega$. Then $o_1 = o_2$ when $o_1(A) = o_2(A)$ for each formula A. $o(A)$ is the number of times A is in the multiset o.

DEFINITION 1.5 [COMPARING MULTISETS] When M and M' are multisets and $o_M(x) \leq o_{M'}(x)$ for each x in X, we say that M is a SUB-MULTISET of M', and M' is a SUPER-MULTISET of M.

The GROUND of the multiset M is the set of all objects that are members of M to a non-zero degree. So, for example, the ground of the multiset A, B, B, C is the set $\{A, B, C\}$.

We use finite multisets as a part of a discriminating analysis of proofs and arguments. (An even more discriminating analysis will consider premises to be structured in *lists*, according to which A, B differs from B, A. You can examine this in Exercise 24 on page 46.) We have no need to consider *infinite* multisets in this section, as multisets represent the premise collections in arguments, and it is quite natural to consider only arguments with finitely many premises, since proofs, as we have defined them feature only finitely many assumptions. So, we will consider arguments in the following way.

DEFINITION 1.6 [ARGUMENT] An ARGUMENT $X \therefore A$ is a structure consisting of a finite multiset X of formulas as its *premises*, and a single formula A as its *conclusion*. The premise multiset X may be empty. An argument $X \therefore A$ is *standardly valid* if and only if there is some proof with undischarged assumptions forming the multiset X, and with the conclusion A. It is *relevantly valid* if and only if there is a relevant proof from the multiset X of premises to A, and so on.

John Slaney has joked that the empty multiset $[\,]$ should be distinguished from the empty set $\emptyset$, since *nothing* is a member of $\emptyset$, but *everything* is a member of $[\,]$ *zero times*.

Here are some features of validity.

LEMMA 1.7 [VALIDITY FACTS] *Let v-validity be any of linear, relevant, affine or standard validity.*

1. $A \therefore A$ *is v-valid.*
2. $X, A \therefore B$ *is v-valid if and only if* $X \therefore A \rightarrow B$ *is v-valid.*
3. *If* $X, A \therefore B$ *and* $Y \therefore A$ *are both v-valid, so is* $X, Y \therefore B$.
4. *If* $X \therefore B$ *is affine or standardly valid, so is* $X, A \therefore B$.
5. *If* $X, A, A \therefore B$ *is relevantly or standardly valid, so is* $X, A \therefore B$.

Proof: (1) is given by the proof consisting of A as premise and conclusion.

For (2), take a proof π from X, A to B, and in a single step $\rightarrow I$, discharge the (single instance of) A to construct the proof of $A \rightarrow B$ from X. Conversely, if you have a proof from X to $A \rightarrow B$, add a (single) premise A and apply $\rightarrow E$ to derive B. In both cases here, if the original proofs satisfy a constraint (vacuous or multiple discharge) so do the new proofs.

For (3), take a proof from X, A to B, but replace the instance of assumption of A indicated in the premises, and replace this with the *proof* from Y to A. The result is a proof, from X, Y to B as desired. This proof satisfies the constraints satisfied by both of the original proofs.

For (4), if we have a proof π from X to B, we extend it as follows

$$\dfrac{\dfrac{\begin{array}{c}X\\ \vdots\,\pi\\ B\end{array}}{A \rightarrow B}{\scriptstyle\rightarrow I} \qquad A}{B}{\scriptstyle\rightarrow E}$$

to construct a proof from to B involving the new premise A, as well as the original premises X. The $\rightarrow I$ step requires a vacuous discharge.

Finally (5): if we have a proof π from X, A, A to B (that is, a proof with X and *two* instances of A as premises to derive the conclusion B) we discharge the two instances of A to derive $A \rightarrow B$ and then reinstate a single instance of A to as a premise to derive B again.

$$\dfrac{\dfrac{\begin{array}{c}X, [A, A]^{(i)}\\ \vdots\,\pi\\ B\end{array}}{A \rightarrow B}{\scriptstyle\rightarrow I,i} \qquad A}{B}{\scriptstyle\rightarrow E}$$

■

Now, having established these facts, we might focus all our attention on the distinction between those arguments that are valid and those that are not, to attend to facts about validity such as those we have just proved. That would be to ignore the distinctive features of proof theory. We care not only *that* an argument is proved, but *how* it is proved. For each of these facts about validity, we showed not only the bare existential fact (for example, if there is a proof from X, A to B, then there is a proof from X to $A \rightarrow B$) but the stronger and more specific fact (if there is a proof from X, A to B then from this proof we construct the proof from X to $A \rightarrow B$ in a uniform way). This is the power of proof theory. We focus on proofs, not only as a certificate for the validity of an argument, but as a structure worth attention in its own right.

» «

It is often a straightforward matter to show that an argument is valid. Find a proof from the premises to the conclusion, and you are done. It seems more difficult to show that an argument is not valid. According to the literal reading of this definition, if an argument is not valid there is no proof from the premises to the conclusion. So, the direct way to show that an argument is invalid is to show that it has no proof from the premises to the conclusion. There are infinitely many proofs. It would take forever to through all of the proofs and check that none of them are proofs from X to A in order to convince yourself that the argument from X to A is not valid. To show that the argument is not valid, that there is no proof from X to A, some subtlety is called for. We will end this section by looking at how we might summon up the skill we need.

One subtlety would be to change the terms of discussion entirely, and introduce a totally new concept. If you could show that all valid arguments have some special property – and one that is easy to detect when present and when absent – then you could show that an argument is invalid by showing it lacks that special property. How this might manage to work depends on the special property. We shall look at one of these properties in Chapter 3 when we show that all valid arguments *preserve truth* in all *models*. Then to show that an argument is invalid, you could provide a model in which truth is *not* preserved from the premises to the conclusion. If all valid arguments are truth-in-a-model-preserving, then such a model would count as a counterexample to the validity of your argument.

In this chapter, on the other hand, we will not go beyond the conceptual bounds of proof theory itself. We will find instead a way to show that an argument is invalid, using an analysis of the structure of proofs. The collection of *all* proofs is too large to survey. From premises X and conclusion A, the collection of *direct* proofs – those that go straight from X to A without any detours down byways or highways – should be more tractable. If we could show that there are not many *direct* proofs from a given collection of premises to a conclusion, then we might be able to exploit this fact to show that for a given set of premises and a conclusion there are *no* direct proofs from X to A. If, in addition, you were to how that any proof from a premise set to a conclusion could somehow be converted into a direct proof from the same premises to that conclusion, then you would have shown that there is no proof from X to A.

Happily, this technique works. To show how it works we need to understand what it is for a proof to be have no detours. These proofs which head straight from the premises to the conclusion without detours are so important that they have their own name. They are called *normal*.

I *think* that the terminology 'normal' comes from Prawitz [169], though the idea comes from Gentzen.

1.3 | NORMAL PROOFS

It is best to introduce normal proofs by contrasting them with *non*-normal proofs. Non-normal proofs are not difficult to find. Suppose you want to show that the following argument is valid

$$p \to q \therefore p \to ((q \to r) \to r)$$

On page 8 You might note first that we have already seen an argument which takes us from $p \to q$ to $(q \to r) \to (p \to r)$. This is SUFFIXING.

$$
\begin{prooftree}
\AxiomC{$[q \to r]^{(2)}$}
\AxiomC{$p \to q$}
\AxiomC{$[p]^{(1)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{q}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{r}
\RightLabel{$\scriptstyle \to I,1$}
\UnaryInfC{$p \to r$}
\RightLabel{$\scriptstyle \to I,2$}
\UnaryInfC{$(q \to r) \to (p \to r)$}
\end{prooftree}
$$

So, we have $p \to q \therefore (q \to r) \to (p \to r)$. But we also have the general principle *permuting* antecedents: $A \to (B \to C) \therefore B \to (A \to C)$.

$$
\begin{prooftree}
\AxiomC{$A \to (B \to C)$}
\AxiomC{$[A]^{(3)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{$B \to C$}
\AxiomC{$[B]^{(4)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{C}
\RightLabel{$\scriptstyle \to I,3$}
\UnaryInfC{$A \to C$}
\RightLabel{$\scriptstyle \to I,4$}
\UnaryInfC{$B \to (A \to C)$}
\end{prooftree}
$$

We can apply this in the case where $A = (q \to r)$, $B = p$ and $C = r$ to get $(q \to r) \to (p \to r) \therefore p \to ((q \to r) \to r)$. We then chain reasoning together, to get us from $p \to q$ to $p \to ((q \to r) \to r)$, which we wanted. But take a look at the whole proof:

$$
\begin{prooftree}
\AxiomC{$[q \to r]^{(2)}$}
\AxiomC{$p \to q$}
\AxiomC{$[p]^{(1)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{q}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{r}
\RightLabel{$\scriptstyle \to I,1$}
\UnaryInfC{$p \to r$}
\RightLabel{$\scriptstyle \to I,2$}
\UnaryInfC{$(q \to r) \to (p \to r)$}
\AxiomC{$[q \to r]^{(3)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{$p \to r$}
\AxiomC{$[p]^{(4)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{r}
\RightLabel{$\scriptstyle \to I,3$}
\UnaryInfC{$(q \to r) \to r$}
\RightLabel{$\scriptstyle \to I,4$}
\UnaryInfC{$p \to ((q \to r) \to r)$}
\end{prooftree}
$$

This proof is circuitous. It gets us from our premise $(p \to q)$ to our conclusion $(p \to ((q \to r) \to r))$, but it does it in a roundabout way. We break down the conditionals $p \to q$, $q \to r$ to construct $(q \to r) \to (p \to r)$ halfway through the proof, only to break that down again (deducing r on its own, for a second time) to build the required conclusion. This is most dramatic around the intermediate conclusion $p \to ((q \to r) \to r)$ which is built up *from* $p \to r$ only to be used to justify $p \to r$ at the next step. We may eliminate this redundancy by cutting out the intermediate formula $p \to ((q \to r) \to r)$ like this:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[q \to r]^{(3)} \qquad \dfrac{p \to q \quad [p]^{(1)}}{q}{\scriptstyle \to E}}{r}{\scriptstyle \to E}}{p \to r}{\scriptstyle \to I,1} \qquad [p]^{(4)}}{r}{\scriptstyle \to E}}{(q \to r) \to r}{\scriptstyle \to I,3}}{p \to ((q \to r) \to r)}{\scriptstyle \to I,4}$$

The resulting proof is a lot simpler already. But now the $p \to r$ is constructed from r only to be broken up immediately to return r. We can delete the redundant $p \to r$ in the same way.

$$\dfrac{\dfrac{\dfrac{[q \to r]^{(3)} \qquad \dfrac{p \to q \quad [p]^{(4)}}{q}{\scriptstyle \to E}}{r}{\scriptstyle \to E}}{(q \to r) \to r}{\scriptstyle \to I,3}}{p \to ((q \to r) \to r)}{\scriptstyle \to I,4}$$

This proof takes us directly from its premise to its conclusion, through no extraneous formulas. Every formula used in this proof is either found in the premise, or in the conclusion. This wasn't true in the original, roundabout proof. We say this new proof is *normal*, the original proof was not.

This is a general phenomenon. Take a proof ending in [$\to E$]: it goes from A to B by way of a sub-proof π_1, and then we discharge A to conclude $A \to B$. Imagine that at the very next step, we use a different proof – say π_2 – with conclusion A to deduce B by means of an implication elimination. This proof contains a redundant step. Instead of taking the detour through the formula $A \to B$, we could use the proof π_1 of B, but instead of taking A as an *assumption*, we could use the proof of A we have at hand, namely π_2. The before-and-after comparison is this:

$$\text{BEFORE:} \quad \dfrac{\dfrac{\begin{array}{c}[A]^{(i)} \\ \vdots\, \pi_1 \\ B\end{array}}{A \to B}{\scriptstyle \to I,i} \qquad \begin{array}{c}\vdots\, \pi_2 \\ A\end{array}}{B}{\scriptstyle \to E} \qquad \text{AFTER:} \quad \begin{array}{c}\vdots\, \pi_2 \\ A \\ \vdots\, \pi_1 \\ B\end{array}$$

The result is a proof of B from the same premises as our original proof. The premises are the premises of π_1 (other than the instances of A that were discharged in the other proof) together with the premises of π_2. This new proof does not go through the formula $A \to B$, so it is, in a sense, simpler than the original.

Well … there are some subtleties with counting, as usual with proofs. If the discharge of A was vacuous, then we have nowhere to plug in the new proof π_2, so π_2, and its premises, don't appear in the final proof. On the other hand, if a number of duplicates of A were discharged, then the new proof will contain that many copies of π_2, and hence, that many copies of the premises of π_2.

Let's make this discussion more concrete, by considering an example where π_1 has two instances of A in the premise list. The original proof containing the introduction and then elimination of $A \to B$ is

$$\dfrac{\dfrac{\dfrac{\dfrac{A \to (A \to B) \quad [A]^{(1)}}{A \to B}{\scriptstyle \to E} \quad [A]^{(1)}}{B}{\scriptstyle \to E}}{A \to B}{\scriptstyle \to I,1} \quad \dfrac{(A \to A) \to A \quad \dfrac{[A]^{(2)}}{A \to A}{\scriptstyle \to I,2}}{A}{\scriptstyle \to E}}{B}{\scriptstyle \to E}$$

We can cut out the [$\to I$/$\to E$] pair (we call such pairs INDIRECT PAIRS) using the technique described above, we place a copy of the inference to A at *both* places that the A is discharged (with label 1). The result is this proof, which does not make that detour.

$$\dfrac{\dfrac{A \to (A \to B) \quad \dfrac{(A \to A) \to A \quad \dfrac{[A]^{(2)}}{A \to A}{\scriptstyle \to I,2}}{A}{\scriptstyle \to E}}{A \to B}{\scriptstyle \to E} \quad \dfrac{(A \to A) \to A \quad \dfrac{[A]^{(2)}}{A \to A}{\scriptstyle \to I,2}}{A}{\scriptstyle \to E}}{B}{\scriptstyle \to E}$$

which is a proof from the same premises ($A \to (A \to B)$ and $(A \to A) \to A$) to the same conclusion B, except for multiplicity. In this proof the premise $(A \to A) \to A$ is used twice instead of once. (Notice too that the label '2' is used twice. We could relabel one subproof to $A \to A$ to use a different label, but there is no ambiguity here because the two proofs to $A \to A$ do not overlap. Our convention for labelling is merely that at the time we get to an [$\to I$] label, the numerical tag is unique in the proof *above* that step.)

We have motivated the concept of normality. Here is the definition:

DEFINITION 1.8 [NORMAL PROOF] A proof is NORMAL if and only if the concluding formula $A \to B$ introduced in an [$\to I$] step is not then immediately used as the major premise of an [$\to E$] step.

DEFINITION 1.9 [INDIRECT PAIR; DETOUR FORMULA] If a formula $A \to B$ introduced in an [$\to I$] step in a proof is also the major premise of a following [$\to E$] step in that proof, then we shall call this pair of inferences an INDIRECT PAIR and we will call the instance $A \to B$ in the middle of this indirect pair a DETOUR FORMULA in that proof.

So, a normal proof is one without any indirect pairs. It has no detour formulas.

Normality is not only important for proving that an argument is invalid by showing that it has no normal proofs. The claim that every valid argument has a normal proof could well be *vital*. If we think of the rules for conditionals as somehow *defining* the connective, then proving something by means of a roundabout [$\rightarrow$*I*/$\rightarrow$*E*] step that you *cannot* prove without it would seem to be illicit. If the conditional is *defined* by way of its rules then it seems that the things one can prove *from* a conditional ought to be merely the things one can prove from whatever it was you used to *introduce* the conditional. If we could prove more from a conditional $A \rightarrow B$ than one could prove on the basis on the information used to *introduce* the conditional, then we are conjuring new arguments out of thin air.

For this reason, many have thought that being able to convert non-normal proofs to normal proofs is not only desirable, it is critical if the proof system is to be properly logical. We will not continue in this philosophical vein here. We will take up this topic in a later section, after we understand the behaviour of normal proofs a little better. Let us return to the study of normal proofs.

Normal proofs are, intuitively at least, proofs without a kind of redundancy. It turns out that avoiding this kind of redundancy in a proof means that you must avoid another kind of redundancy too. A normal proof from X to A may use only a very restricted repertoire of formulas. It will contain only the *subformulas* of X and A.

DEFINITION 1.10 [SUBFORMULAS AND PARSE TREES] The PARSE TREE for an atom is that atom itself. The PARSE TREE for a conditional $A \rightarrow B$ is the tree containing $A \rightarrow B$ at the root, connected to the parse tree for A and the parse tree for B. The SUBFORMULAS of a formula A are those formulas found in A's parse tree. We let $\mathrm{sf}(A)$ be the set of all subformulas of A, so $\mathrm{sf}(p) = \{p\}$, and $\mathrm{sf}(A \rightarrow B) = \{A \rightarrow B\} \cup \mathrm{sf}(A) \cup \mathrm{sf}(B)$. To generalise, when X is a multiset of formulas, we will write $\mathrm{sf}(X)$ for the set of subformulas of each formula in X.

Here is the parse tree for $(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow p)$:

$$\dfrac{\dfrac{p \quad q}{p \rightarrow q} \qquad \dfrac{\dfrac{q \quad r}{q \rightarrow r} \quad p}{(q \rightarrow r) \rightarrow p}}{(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow p)}$$

So, $\mathrm{sf}((p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow p)) = \{(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow p), p \rightarrow q, p, q, (q \rightarrow r) \rightarrow p, q \rightarrow r, r\}$.

We may prove the following theorem.

THEOREM 1.11 [THE SUBFORMULA THEOREM] *If π is a normal proof from the premises X to the conclusion A, then π contains only formulas in* $\mathrm{sf}(X, A)$.

Notice that this is *not* the case for non-normal proofs. Consider the following circuitous proof from A to A.

$$\dfrac{\dfrac{[A]^{(1)}}{A \to A}{\scriptstyle \to I,1} \quad A}{A}{\scriptstyle \to E}$$

Here $A \to A$ is in the proof, but it is not a subformula of the premise (A) or the conclusion (also A).

The subformula property for normal proofs goes some way to reassure us that a normal proof is *direct*. A normal proof from X to A cannot stray so far away from the premises and the conclusion so as to incorporate material outside X and A. This fact goes some way to defend the notion that validity is *analytic* in a strong sense. The validity of an argument is grounded a proof where the consitituents of that proof are found by *analysing* the premises and the conclusion into their constituents. Here is how the subformula theorem is proved.

Proof: We look carefully at how proofs are constructed. If π is a normal proof, then it is constructed in exactly the same way as all proofs are, but the fact that the proof is normal gives us some useful information. By the definition of proofs, π either is a lone assumption, or π ends in an application of [$\to I$], or it ends in an application of [$\to E$]. Assumptions are the basic building blocks of proofs. We will show that assumption-only proofs have the subformula property, and then, also show on the assumption that the proofs we have on had have the subformula property, then the normal proofs we construct from them also have the property. Then it will follow that all normal proofs have the subformula property, because all of the normal proofs can be generated in this way.

Notice that the subproofs of normal proofs are normal. If a subproof of a proof contains an indirect pair, then so does the larger proof.

ASSUMPTION: A sole assumption, considered as a proof, satisfies the subformula property. The assumption A is the only constituent of the proof and it is both a premise and the conclusion.

INTRODUCTION: In the case of [$\to I$], π is constructed from another normal proof π' from X to B, with the new step added on (and with the discharge of a number – possibly zero – of assumptions). π is a proof from X' to $A \to B$, where X' is X with the deletion of some number of instances of A. Since π' is normal, we may assume that every formula in π' is in $\mathrm{sf}(X, B)$. Notice that $\mathrm{sf}(X', A \to B)$ contains every element of $\mathrm{sf}(X, B)$, since X differs only from X' by the deletion of some instances of A. So, every formula in π (namely, those formulas in π', together with $A \to B$) is in $\mathrm{sf}(X', A \to B)$ as desired.

ELIMINATION: In the case of [$\to E$], π is constructed out of *two* normal proofs: one (call it π_1) to the conclusion of a conditional $A \to B$ from premises X, and the other (call it π_2) to the conclusion of the antecedent of that conditional A from premises Y. Both π_1 and π_2 are normal, so we may assume that each formula in π_1 is in $\mathrm{sf}(X, A \to B)$ and each formula in π_2 is in $\mathrm{sf}(Y, A)$. We wish to show that every formula in π is in $\mathrm{sf}(X, Y, B)$. This seems difficult ($A \to B$ is in the proof—where can it be found inside X, Y or B?), but we also have some more information: π_1 cannot end in the *introduction* of the conditional $A \to B$. So, π_1 is either the assumption $A \to B$ itself (in which case $Y = A \to B$, and clearly in this case each formula in π *is* in $\mathrm{sf}(X, A \to B, B)$) or π_1 ends in a [$\to E$] step. But if π_1 ends in an [$\to E$] step, the major premise of that inference is a formula of the form $C \to (A \to B)$. So π_1 contains the formula $C \to (A \to B)$, so *whatever* list Y is, $C \to (A \to B) \in \mathrm{sf}(Y, A)$, and so, $A \to B \in \mathrm{sf}(Y)$. In this case too, every formula in π is in $\mathrm{sf}(X, Y, B)$, as desired.

This completes the proof of our theorem. Every normal proof is constructed from assumptions by introduction and elimination steps in this way. The subformula property is preserved through each step of the construction. ■

Normal proofs are handy to work with. Even though an argument might have very many proofs, it will have many fewer normal proofs. We can exploit this fact when searching for proofs.

EXAMPLE 1.12 [NO NORMAL PROOFS] There is no normal proof from p to q. There is no normal relevant proof from $p \to r$ to $p \to (q \to r)$.

Proof: Normal proofs from p to q (if there are any) contain only formulas in $\mathrm{sf}(p, q)$: that is, they contain only p and q. That means they contain no [$\to I$] or [$\to E$] steps, since they contain no conditionals at all. It follows that any such proof must consist solely of an assumption. As a result, the proof cannot have a premise p that differs from the conclusion q. There is no normal proof from p to q.

For the second example, if there is a normal proof of $p \to (q \to r)$, from $p \to r$, it must end in an $\to I$ step, from a normal (relevant) proof from $p \to r$ and p to $q \to r$. Similarly, this proof must also end in an $\to I$ step, from a normal (relevant) proof from $p \to r$, p and q to r. Now, what normal relevant proofs can be found from $p \to r$, p and q to r? There are none. Any such proof would have to use q as a premise somewhere, but since it is normal, it contains only subformulas of $p \to r$, p, q and r—namely those formulas themselves. There is no formula involving q other than q itself on that list, so there is nowhere for q to go. It cannot be used, so it will not be a premise in the proof. There is no normal relevant proof from the premises $p \to r$, p and q to the conclusion r. ■

These facts are interesting enough. It would be more productive, however, to show that there is no proof at all from p to q, and no relevant proof from $p \to r$ to $p \to (q \to r)$. We can do this if we have some way of showing that if we have a proof for some argument, we have a normal proof for that argument.

So, we now work our way towards the following theorem:

THEOREM 1.13 [NORMALISATION THEOREM] *A proof π from X to A reduces in some number of steps to a normal proof π' from X' to A.*

If π is linear, so is π', and $X = X'$. If π is affine, so is π', and X' is a sub-multiset of X. If π is relevant, then so is π', and X' covers the same ground as X, and is a super-multiset of X. If π is standard, then so is π', and X' covers no more ground than X.

$[1, 2, 2, 3]$ covers the same ground as–and is a super-multiset of–$[1, 2, 3]$. And $[2, 2, 3, 3]$ covers no more ground than $[1, 2, 3]$.

Notice how the premise multiset of the normal proof is related to the premise multiset of the original proof. If we allow duplicate discharge, then the premise multiset may contain formulas to a greater degree than in the original proof, but the normal proof will not contain any premises that weren't in the original proof. If we allow vacuous discharge, then the normal proof might contain fewer premises than the original proof.

The normalisation theorem mentions the notion of *reduction*, so let us first define it.

DEFINITION 1.14 [REDUCTION] A proof π *reduces* to π' (shorthand: $\pi \rightsquigarrow \pi'$) if some indirect pair in π is eliminated, to result in π'.

$$\begin{array}{c}\dfrac{\dfrac{\begin{array}{c}[A]^{(i)}\\ \vdots\ \pi_1\\ B\end{array}}{A \to B}{\scriptstyle \to I, i} \qquad \begin{array}{c}\vdots\ \pi_2\\ A\end{array}}{B}{\scriptstyle \to E}\\ \vdots\\ C\end{array} \qquad \rightsquigarrow \qquad \begin{array}{c}\vdots\ \pi_2\\ A\\ \vdots\ \pi_1\\ B\\ \vdots\\ C\end{array}$$

If there is no π' such that $\pi \rightsquigarrow \pi'$, then π' is normal. If $\pi_0 \rightsquigarrow \pi_2 \rightsquigarrow \cdots \rightsquigarrow \pi_n$ we write "$\pi_0 \rightsquigarrow_* \pi_n$" and we say that π_0 reduces to π_n in a number of steps. We aim to show that for any proof π, there is some normal π^* such that $\pi \rightsquigarrow_* \pi^*$.

We allow that $\pi \rightsquigarrow_* \pi$. A proof 'reduces' to itself in *zero* steps.

The only difficult part in proving the normalisation theorem is showing that the process reduction can terminate in a normal proof. In the case where we do not allow duplicate discharge, there is no difficulty at all.

Proof [Theorem 1.13: *linear and affine cases*]: If π is a linear proof, or is an affine proof, then whenever you pick an indirect pair and normalise it, the result is a shorter proof. At most one copy of the proof π_2 for A is inserted into the proof π_1. (Perhaps no substitution is made in the case of an affine proof, if a vacuous discharge was made.) Proofs have some finite size, so this process cannot go on indefinitely. Keep deleting indirect pairs until there are no pairs left to delete. The result is a normal proof to the conclusion A. The premises X remain undisturbed, except in the affine case, where we may have lost premises along the way. (An assumption from π_2 might disappear if we did not need to make the substitution.) In this case, the premise multiset X' from the normal proof is a *sub*-multiset of X, as desired. ■

If we allow duplicate discharge, however, we cannot be sure that in normalising we go from a larger to a smaller proof. The example on page 18 goes from a proof with 11 formulas to another proof with 11 formulas. In some cases a reduction step can take us from a smaller proof to a properly larger proof. Sometimes, the result is *much* larger. So size alone is no guarantee that the process terminates.

To gain some understanding of the general process of transforming a non-normal proof into a normal one, we must find some other measure that decreases as normalisation progresses. If this measure has a least value then we can be sure that the process will stop. The appropriate measure in this case will not be too difficult to find. Let's look at a part of the process of normalisation: the complexity of the formula that is normalised.

Well, the process stops if the measures are ordered appropriately—so that there's no *infinitely descending chain*.

DEFINITION 1.15 [COMPLEXITY] A formula's *complexity* is the number of connectives in that formula. In this case, it is the number of instances of '$\to$' in the formula.

The crucial features of complexity are that each formula has a finite complexity, and that the proper subformulas of a formula each have a lower complexity than the original formula. This means that complexity is a good measure for an induction, like the size of a proof.

Now, suppose we have a proof containing just one indirect pair, introducing and eliminating $A \to B$, and suppose that otherwise, π_1 (the proof of B from A) and π_2 (the proof of A) are normal.

$$\text{BEFORE: } \dfrac{\dfrac{\begin{matrix}[A]^{(i)} \\ \vdots\, \pi_1 \\ B\end{matrix}}{A \to B}{\scriptstyle \to I, i} \qquad \begin{matrix}\vdots\, \pi_2 \\ A\end{matrix}}{B}{\scriptstyle \to E} \qquad \text{AFTER: } \begin{matrix}\vdots\, \pi_2 \\ A \\ \vdots\, \pi_1 \\ B\end{matrix}$$

This the new proof need not be necessarily normal, even though π_1 and π_2 are. The new proof is non-normal if π_2 ends in the introduction of A and π_1 starts off with the elimination of A. Notice, however, that the non-normality of the new proof is, somehow, *smaller*. There is no non-normality with respect to $A \to B$ or any other formula as complex as that. The potential non-normality is with respect to a subformula A. This result would still hold if the proofs π_1 and π_2 weren't normal themselves, but when they might have [$\to I/\!\to E$] pairs for formulas less complex than $A \to B$. If $A \to B$ is the most complex detour formula in the original proof, then the new proof has a *smaller* most complex detour formula.

DEFINITION 1.16 [NON-NORMALITY] The *non-normality measure* of a proof is a sequence $\langle c_1, c_2, \ldots, c_n\rangle$ of numbers such that c_i is the number of indirect pairs of formulas of complexity i. The sequence for a proof stops at the last non-zero value. Sequences are ordered with their last number as most significant. That is, $\langle c_1, \ldots, c_n\rangle > \langle d_1, \ldots, d_m\rangle$ if and only if $n > m$, or if $n = m$, when $c_n > d_n$, or if $c_n = d_n$, when $\langle c_1, \ldots, c_{n-1}\rangle > \langle d_1, \ldots, d_{n-1}\rangle$.

Non-normality measures satisfy the finite descending chain condition. Starting at any particular measure, you cannot find an infinite descending chain of measures below it. Of course, there are infinitely many measures smaller than $\langle 0, 1\rangle$ (in this case, $\langle 0\rangle, \langle 1\rangle, \langle 2\rangle, \ldots$). However, to form a *descending sequence* from $\langle 0, 1\rangle$ you must choose one of these as your next measure. Say you choose $\langle 500\rangle$. From that, you have only finitely many (500, in this case) steps until $\langle\rangle$ and the sequence stops. This generalises. From the sequence $\langle c_1, \ldots, c_n\rangle$, you lower c_n until it gets to zero. Then you look at the index for $n - 1$, which might have grown enormously. Nonetheless, it is some finite number, and now you must reduce this value. And so on, until you reach the last quantity, and from there, the empty sequence $\langle\rangle$. Here is an example sequence using this ordering $\langle 3, 2, 30\rangle > \langle 2, 8, 23\rangle > \langle 1, 47, 15\rangle > \langle 138, 478\rangle > \cdots > \langle 1, 3088\rangle > \langle 314159\rangle > \cdots > \langle 1\rangle > \langle\rangle$.

LEMMA 1.17 [NON-NORMALITY REDUCTION] *Any a proof with an indirect pair reduces in one step to some proof with a lower measure of non-normality.*

Proof: Choose a detour formula in π of greatest complexity (say n), such that its proof contains no other detour formulas of complexity n. Normalise that proof. The result is a proof π' with fewer detour formulas of complexity n (and perhaps many more of $n - 1$, etc.). So, it has a lower non-normality measure. ∎

Now we have a proof of our normalisation theorem.

Proof [of Theorem 1.13: *for the relevant and standard cases*]: Start with π, a proof that isn't normal, and use Lemma 1.17 to choose a proof π' with a lower measure of non-normality. If π' is normal, we're done. If it isn't, continue the process. There is no infinite descending chain of non-normality measures, so this process will stop at some point, and the result is a normal proof. ∎

Every proof may be transformed into a normal proof. If there is a linear proof from X to A then there is a normal linear proof from X to A. Linear proofs are satisfying and strict in this manner. If we allow vacuous discharge or duplicate discharge, matters are not so straightforward. For example, there is a non-normal standard proof from p, q to p:

$$\dfrac{\dfrac{p}{q \to p}{\scriptstyle \to I,1} \qquad q}{p}{\scriptstyle \to E}$$

but there is no normal standard proof from exactly these premises to the same conclusion, since any normal proof from atomic premises to an atomic conclusion must be an assumption alone. We have a normal proof from p to p (it is very short!), but there is no normal proof from p to p that involves q as an extra premise.

Similarly, there is a relevant proof from $p \to (p \to q), p$ to q, but it is non-normal.

$$\dfrac{\dfrac{\dfrac{\dfrac{p \to (p \to q) \qquad [p]^{(1)}}{p \to q}{\scriptstyle \to E} \qquad [p]^{(1)}}{q}{\scriptstyle \to E}}{p \to q}{\scriptstyle \to I,1} \qquad p}{q}{\scriptstyle \to E}$$

There is no normal relevant proof from $p \to (p \to q), p$ to q. Any normal relevant proof from $p \to (p \to q)$ and p to q must use [$\to E$] to deduce $p \to q$, and then the only other possible move is either [$\to I$] (in which case we return to $p \to (p \to q)$ none the wiser) or we perform another [$\to E$] with another assumption p to deduce q, and we are done. Alas, we have claimed two undischarged assumptions of p. In the non-linear cases, the transformation from a non-normal to a normal proof does not preserve the number of times a premise is used in the proof.

1.4 | STRONG NORMALISATION AND TERMS

It is very tempting to view normalisation as a process of reducing a proof down to its essence, of unwinding detours, and making explicit the essential logical connections made in the proof between the premises and the conclusion. The result of normalising a proof π from X to A shows the connections made from X to A in that proof π, without the need to bring in the extraneous information in any detours that may have been used in π. Another analogy is that the complex non-normal proof is *evaluated* into its normal form, in the same way that a numerical term like $5 + (2 \times (7 + 3))$ is evaluated into *its* normal form, the numeral 25.

This passage is the hardest part of Chapter 1. Feel free to skip over the proofs of theorems in this section, until page 38 on first reading.

If this is the case, then the process of normalisation should give us two distinct “answers” for the underlying structure of the one proof. Can two different reduction sequences for a single proof result in *different* normal proofs? To investigate this, we need to pay attention to the different processes of reduction we can take when reducing a proof. To do that, we'll introduce a new notion of reduction:

DEFINITION 1.18 [PARALLEL REDUCTION] A proof π *parallel reduces* to π' if some number of indirect pairs in π are eliminated in parallel. We write “$\pi \overset{\rightsquigarrow}{\rightsquigarrow} \pi'$.”

For example, consider the proof with the following two detour formulas marked:

$$
\begin{prooftree}
\AxiomC{$A \to (A \to B)$}
\AxiomC{$[A]^{(1)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{$A \to B$}
\AxiomC{$[A]^{(1)}$}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{B}
\RightLabel{$\scriptstyle \to I,1$}
\UnaryInfC{$A \to B$}
\AxiomC{$[A]^{(2)}$}
\RightLabel{$\scriptstyle \to I,2$}
\UnaryInfC{$A \to A$}
\AxiomC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{B}
\end{prooftree}
$$

To process them we can take them in any order. Eliminating the $A \to B$, we have

$$
\begin{prooftree}
\AxiomC{$A \to (A \to B)$}
\AxiomC{$[A]^{(2)}$}
\RightLabel{$\scriptstyle \to I,2$}
\UnaryInfC{$A \to A$}
\AxiomC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{$A \to B$}
\AxiomC{$[A]^{(2)}$}
\RightLabel{$\scriptstyle \to I,2$}
\UnaryInfC{$A \to A$}
\AxiomC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{B}
\end{prooftree}
$$

which now has two copies of the $A \to A$ to be reduced. However, these copies do not overlap in scope (they cannot, as they are duplicated in the place of assumptions discharged in an eliminated $\to I$ rule) so they can be processed together. The result is the proof

$$
\begin{prooftree}
\AxiomC{$A \to (A \to B)$}
\AxiomC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{$A \to B$}
\AxiomC{A}
\RightLabel{$\scriptstyle \to E$}
\BinaryInfC{B}
\end{prooftree}
$$

You can check that if you had processed the formulas to be eliminated in the other order, the result would have been the same.

Can you see why this is called the diamond property?

LEMMA 1.19 [DIAMOND PROPERTY FOR $\overset{\rightsquigarrow}{\rightsquigarrow}$] *If $\pi \overset{\rightsquigarrow}{\rightsquigarrow} \pi_1$ and $\pi \overset{\rightsquigarrow}{\rightsquigarrow} \pi_2$ then there is some proof π' where $\pi_1 \overset{\rightsquigarrow}{\rightsquigarrow} \pi'$ and $\pi_2 \overset{\rightsquigarrow}{\rightsquigarrow} \pi'$.*

They may have multiplied, if they occurred in a proof part duplicated in the reduction step. But some may have vanished, too, if they were in a part of the proof that disappeared during reduction.

Proof: Take the detour formulas in the proof π that are eliminated in either the move to π_1 or the move to π_2. 'Colour' them in π, and transform the proof to π_1. Some of the coloured formulas may remain. Do the same in the move from π to π_2. The result are two proofs π_1 and π_2 in which some formulas may be coloured. The proof π' is found by parallel reducing either collection of formulas in π_1 or π_2. ∎

THEOREM 1.20 [ONLY ONE NORMAL FORM] *Given any proof π, if $\pi \rightsquigarrow_* \pi'$ then if $\pi \rightsquigarrow_* \pi''$, it must be that $\pi' = \pi''$. That is, any sequence of reduction steps from π that terminates in a normal form must terminate in a* unique *normal form.*

Proof: Suppose that $\pi \rightsquigarrow_* \pi'$, and $\pi \rightsquigarrow_* \pi''$. It follows that we have two reduction sequences

$$\pi \overset{\rightsquigarrow}{\rightsquigarrow} \pi_1' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_2' \overset{\rightsquigarrow}{\rightsquigarrow} \cdots \overset{\rightsquigarrow}{\rightsquigarrow} \pi_n' \overset{\rightsquigarrow}{\rightsquigarrow} \pi'$$

$$\pi \overset{\rightsquigarrow}{\rightsquigarrow} \pi_1'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_2'' \overset{\rightsquigarrow}{\rightsquigarrow} \cdots \overset{\rightsquigarrow}{\rightsquigarrow} \pi_m'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi''$$

By the diamond property, we have a $\pi_{1,1}$ where $\pi_1' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_{1,1}$ and $\pi_1'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_{1,1}$. Then $\pi_1'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_{1,1}$ and $\pi_1'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_2''$ so by the diamond property there is some $\pi_{2,1}$ where $\pi_2'' \overset{\rightsquigarrow}{\rightsquigarrow} \pi_{2,1}$ and $\pi_{1,1} \overset{\rightsquigarrow}{\rightsquigarrow} \pi_{2,1}$. Continue in this vein, guided by the picture below:

$$\begin{array}{ccccccccc}
\pi & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_1' & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_2' & \overset{\rightsquigarrow}{\rightsquigarrow} & \cdots & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_n' \\
\wr\wr & & \wr\wr & & \wr\wr & & & & \wr\wr \\
\pi_1'' & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{1,1} & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{1,2} & \overset{\rightsquigarrow}{\rightsquigarrow} & \cdots & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{1,n} \\
\wr\wr & & \wr\wr & & \wr\wr & & & & \wr\wr \\
\pi_2'' & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{2,1} & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{2,2} & \overset{\rightsquigarrow}{\rightsquigarrow} & \cdots & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{2,n} \\
\wr\wr & & \wr\wr & & \wr\wr & & & & \wr\wr \\
\vdots & & \vdots & & \vdots & & & & \vdots \\
\wr\wr & & \wr\wr & & \wr\wr & & & & \wr\wr \\
\pi_m'' & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{m,1} & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi_{m,2} & \overset{\rightsquigarrow}{\rightsquigarrow} & \cdots & \overset{\rightsquigarrow}{\rightsquigarrow} & \pi^*
\end{array}$$

to find the desired proof π^*. So, if π_n' and π_n'' are *normal* they must be identical. ∎

So, sequences of reductions from π cannot terminate in two different proofs. A normal form for a proof is unique.

This result goes a lot of the way towards justifying the idea that normalisation corresponds to evaluating the underlying essence of a proof. The normal form is well defined an unique. But this leaves us with a remaining question. We have seen that for each proof π there is some process to evaluate its normal form π^*, and further, the proof of the previous theorem shows us that *any* finite sequence of reductions from π can be extended to eventually reach π^*. Does it follow that any process of reductions from π terminates in its normal form π^*? That is: are our proofs *strongly normalising*?

DEFINITION 1.21 [STRONGLY NORMALISING] A proof π is strongly normalising (under a reduction relation $\rightsquigarrow$) if and only if there is no infinite reduction sequence starting from π.

This does not follow from weak normalisation (there is some reduction to a normal form) and the diamond property, which gives us unique normal form theorem. This is straightforward to see, because a "reduction" process which allows us to run reduction steps backwards as well as forwards if the proof is not already normal, would still allow for weak normalisation, would still have the diamond property and would have a unique normal form. But it would not be strongly normalising. (We could go on forever reducing one detour only to put it back, forever.) Is there any guarantee that our reduction process process will always terminate?

A naive approach would be to define some measure on proofs which always reduces under any reduction step. This seems hopeless, because anything like the measure we have already defined can increase, rather than decrease, under reductions. (Take a proof with a small detour formula $A \to B$ where the assumption A is discharged a number of times in the proof of the major premise $A \to B$, and in which there are larger detour formulas in the proof of the minor premise A. This proof is duplicated in the reduction, and the measure of the new proof could rise significantly, as we have eliminated a small detour formula at the cost of many large detour formulas.)

We will prove that every proof is strongly normalising under the relation $\rightsquigarrow$ of deleting detour formulas. To assist in talking about this, we need to make a few more definitions. First, the *reduction tree*.

DEFINITION 1.22 [REDUCTION TREE] The reduction tree (under $\rightsquigarrow$) of a proof π is the tree whose branches are the reduction sequences on the relation $\rightsquigarrow$. So, from the root π we reach any proof accessible in one $\rightsquigarrow$ step from π. From each π' where $\pi \rightsquigarrow \pi'$, we branch similarly. Each node has only finitely many successors as there are only finitely many detour formulas in a proof. For each proof π, $\nu(\pi)$ is the size of its reduction tree.

LEMMA 1.23 [THE SIZE OF REDUCTION TREES] *A strongly normalising proof has a finite reduction tree. It follows that not only is every reduction path finite, but there is a* longest *reduction path.*

It's true that every finitely branching tree with finite branches is finite. But is it *obvious* that it's true?

Proof: This is a corollary of König's Lemma, which states that every tree in which the number of immediate descendants of a node is finite (it is finitely *branching*), and in which every branch is finitely long, is itself *finite*. Since the reduction tree for a strongly normalising proof is finitely branching, and each branch has a finite length, it follows that any strongly normalising proof not only has only finite reduction paths, it also has a *longest* reduction path. ■

Now to prove that every proof is strongly normalising. To do this, we define a new property that proofs can have: of being **red**. It will turn out that all **red** proofs are strongly normalising. It will also turn out that all proofs are **red**.

The term '**red**' should bring to mind 'reducible.' This formulation of strong normalisation is originally due to William Tait [236]. I am following the presentation of Jean-Yves Girard [87, 88].

DEFINITION 1.24 [**red** PROOFS] We define a new predicate '**red**' applying to proofs in the following way.

» A proof of an atomic formula is **red** if and only if it is strongly normalising.

» A proof π of an implication formula $A \rightarrow B$ is **red** if and only if whenever π' is a **red** proof of A, then the proof

$$\dfrac{\begin{array}{c}\vdots\,\pi \\ A \rightarrow B\end{array} \quad \begin{array}{c}\vdots\,\pi' \\ A\end{array}}{B}$$

is a **red** proof of type B.

We will have cause to talk often of the proof found by extending a proof π of $A \rightarrow B$ and a proof π' of A to form the proof of B by adding an $\rightarrow E$ step. We will write '$(\pi\,\pi')$' to denote this proof. If you like, you can think of it as the application of the proof π to the proof π'.

Now, our aim will be twofold: to show that every **red** proof is strongly normalising, and to show that every proof is **red**. We start by proving the following crucial lemma:

LEMMA 1.25 [PROPERTIES OF **red** PROOFS] *For any proof π, the following three conditions hold:*

- C1 *If π is **red** then π is strongly normalisable.*
- C2 *If π is **red** and π reduces to π' in one step, then π' is **red** too.*
- C3 *If π is a proof not ending in $\rightarrow$I, and whenever we eliminate one indirect pair in π we have a **red** proof, then π is **red** too.*

Proof: We prove this result by induction on the formula proved by π. We start with proofs of atomic formulas.

- C1 Any **red** proof of an atomic formula is strongly normalising, by the definition of '**red**'.
- C2 If π is strongly normalising, then so is any proof to which π reduces.
- C3 π does not end in $\rightarrow I$ as it is a proof of an atomic formula. If whenever $\pi \rightsquigarrow \pi'$ and π' is **red**, since π' is a proof of an atomic formula, it is strongly normalising. Since *any* reduction path through π must travel through one such proof π', each such path through π terminates. So, π is **red**.

Now we prove the results for a proof π of $A \rightarrow B$, under the assumption that C1, C2 and C3 they hold for proofs of A and proofs of B. We can then conclude that they hold of *all* proofs, by induction on the complexity of the formula proved.

- C1 If π is a **red** proof of $A \rightarrow B$, consider the proof

$$\sigma : \quad \dfrac{\begin{array}{c}\vdots\,\pi \\ A \rightarrow B\end{array} \quad A}{B}$$

The assumption A is a normal proof of its conclusion A not ending in $\rightarrow I$, so c3 applies and it is **red**. So, by the definition of **red** proofs of implication formulas, σ is a **red** proof of B. Condition c1 tells us that **red** proofs of B are strongly normalising, so any reduction sequence for σ must terminate. It follows that any reduction sequence for π must terminate too, since if we had a non-terminating reduction sequence for π, we could apply the same reductions to the proof σ. But since σ is strongly normalising, this cannot happen. It follows that π is strongly normalising too.

c2 Suppose that π reduces in one step to a proof π'. Given that π is **red**, we wish to show that π' is **red** too. Since π' is a proof of $A \rightarrow B$, we want to show that for any **red** proof π'' of A, the proof $(\pi'\,\pi'')$ is **red**. But this proof is **red** since the **red** proof $(\pi\,\pi'')$ reduces to $(\pi'\,\pi'')$ in one step (by reducing π to π'), and c2 applies to proofs of B.

c3 Suppose that π does not end in $[\rightarrow I]$, and suppose that all of the proofs reached from π in one step are **red**. Let σ be a **red** proof of A. We wish to show that the proof $(\pi\,\sigma)$ is **red**. By c1 for the formula A, we know that σ is strongly normalising. So, we may reason by induction on the length of the longest reduction path for σ. If σ is normal (with path of length 0), then $(\pi\,\sigma)$ reduces in one step only to $(\pi'\,\sigma)$, with π' one step from π. But π' is **red** so $(\pi'\,\sigma)$ is too.

On the other hand, suppose σ is not yet normal, but the result holds for all σ' with shorter reduction paths than σ. So, suppose τ reduces to $(\pi\,\sigma')$ with σ' one step from σ. σ' is **red** by the induction hypothesis c2 for A, and σ' has a shorter reduction path, so the induction hypothesis for σ' tells us that $(\pi\,\sigma')$ is **red**.

There is no other possibility for reduction as π does not end in $\rightarrow I$, so reductions must occur wholly in π or wholly in σ, and not in the last step of $(\pi\,\sigma)$.

This completes the proof by induction. The conditions c1, c2 and c3 hold of every proof. ∎

Now we prove one more crucial lemma.

LEMMA 1.26 [**red** PROOFS ENDING IN $[\rightarrow I]$] *If for each* **red** *proof σ of A, the proof*

$$\pi(\sigma):\quad \begin{array}{c} \vdots\ \sigma \\ A \\ \vdots\ \pi \\ B \end{array}$$

is **red**, *then so is the proof*

$$\tau:\quad \dfrac{\begin{array}{c} [A] \\ \vdots\ \pi \\ B \end{array}}{A \rightarrow B}\,{\rightarrow}I$$

Proof: We show that the $(\tau\,\sigma)$ is **red** whenever σ is **red**. This will suffice to show that the proof τ is **red**, by the definition of the predicate '**red**' for proofs of $A \rightarrow B$. We will show that every proof resulting from $(\tau\,\sigma)$ in one step is **red**, and we will reason by induction on the sum of the sizes of the reduction trees of π and σ. There are three cases:

» $(\tau\,\sigma) \rightsquigarrow \pi(\sigma)$. In this case, $\pi(\sigma)$ is **red** by the hypothesis of the proof.

» $(\tau\,\sigma) \rightsquigarrow (\tau'\,\sigma)$. In this case the sum of the size of the reduction trees of τ' and σ is smaller, and we may appeal to the induction hypothesis.

» $(\tau\,\sigma) \rightsquigarrow (\tau\,\sigma')$. In this case the sum of the size of the reduction trees is τ and σ' smaller, and we may appeal to the induction hypothesis. ∎

We are set to prove our major theorem:

THEOREM 1.27 [ALL PROOFS ARE **red**] *Every proof π is* **red**.

To do this, we'll approach it by induction, as follows:

LEMMA 1.28 [**red** PROOFS BY INDUCTION] *For each proof π with assumptions $A_1, \ldots, A_n$, and for any* **red** *proofs $\sigma_1, \ldots, \sigma_n$ of the formulas $A_1, \ldots, A_n$ respectively, the proof $\pi(\sigma_1, \ldots, \sigma_n)$ in which each assumption A_i is replaced by the proof σ_i is* **red**.

Proof: We prove this by induction on the construction of the proof.

» If π is an assumption A_1, the claim is a tautology (if σ_1 is **red**, then σ_1 is **red**).

» If π ends in [$\rightarrow$*E*], and is $(\pi_1\,\pi_2)$, then by the induction hypothesis $\pi_1(\sigma_1, \ldots, \sigma_n)$ and $\pi_2(\sigma_1, \ldots, \sigma_n)$ are **red**. Since $\pi_1(\sigma_1, \ldots, \sigma_n)$ has type $A \rightarrow B$ the definition of **red**ness tells us that when ever it is applied to a **red** proof the result is also **red**. Therefore, the proof $(\pi_1(\sigma_1, \ldots, \sigma_n)\,\pi_2(\sigma_1, \ldots, \sigma_n))$ is **red**, but this proof is simply $\pi(\sigma_1, \ldots, \sigma_n)$.

» If π ends in an application of [$\rightarrow$*I*], then this case is dealt with by Lemma 1.26: if π is a proof of $A \rightarrow B$ ending in $\rightarrow$*E*, then we may assume that π', the proof of B from A inside π is **red**, so by Lemma 1.26, the result π is **red** too.

It follows that *every* proof is **red**. ∎

It follows also that every proof is strongly normalising, since all **red** proofs are strongly normalising.

» «

It is very tempting to think of proofs as *processes* or *functions* that convert the information presented in the premises into the information in the conclusion. This is doubly tempting when you look at the notation for implication. In $\to E$ we apply something which converts A to B (a function from A to B?) to something which delivers you A (from premises) into something which delivers you B. In $\to I$ if we can produce B (when supplied with A, at least in the presence of other resources—the other premises) then we can (in the context of the other resources at least) convert As into Bs at will.

Let's make this talk a little more precise, by making *explicit* this kind of *function*-talk. It will give us a new vocabulary to talk of proofs.

We start with simple notation to talk about functions. The idea is straightforward. Consider numbers, and addition. If you have a number, you can add 2 to it, and the result is another number. If you like, if x is a number then

$$x + 2$$

is another number. Now, suppose we don't want to talk about a particular number, like $5+2$ or $7+2$ or $x+2$ for any choice of x, but we want to talk about the *operation* or of adding two. There is a sense in which just writing "$x + 2$" should be enough to tell someone what we mean. It is relatively clear that we are treating the "x" as a marker for the input of the function, and "$x + 2$" is the output. The *function* is the output as it varies for different values of the input. Sometimes leaving the variables there is not so useful. Consider the subtraction

$$x - y$$

You can think of this as the function that takes the input value x and takes away y. Or you can think of it as the function that takes the input value y and subtracts it from x. or you can think of it as the function that takes two input values x and y, and takes the second away from the first. Which do we mean? When we apply this function to the input value 5, what is the result? For this reason, we have a way of making explicit the different distinctions: it is the λ-notation, due to Alonzo Church [39]. The function that takes the input value x and returns $x + 2$ is denoted

$$\lambda x.(x + 2)$$

The function taking the input value y and subtracts it from x is

$$\lambda y.(x - y)$$

The function that takes *two* inputs and subtracts the second from the first is

$$\lambda x.\lambda y.(x - y)$$

Notice how this function works. If you feed it the input 5, you get the output $\lambda y.(5 - y)$. We can write *application* of a function to its input by way of juxtaposition. The result is that

$$(\lambda x.\lambda y.(x - y)\, 5)$$

evaluates to the result $\lambda y.(5-y)$. This is the function that subtracts y from 5. When you feed *this* function the input 2 (i.e., you evaluate $(\lambda y.(5-y)\,2)$) the result is $5-2$. In other words, 3. So, functions can have other functions as outputs.

Now, suppose you have a function f that takes two inputs y and z, and we wish to consider what happens when you apply f to a pair where the first value is the repeated as the second value. (If f is $\lambda x.\lambda y.(x-y)$ and the input value is a number, then the result should be 0.) We can do this by applying f to the value x twice, to get $((f\,x)\,x)$. But this is not a function, it is the result of applying f to x and x. If you consider this as a function of x you get

$$\lambda x.((f\,x)\,x)$$

This is the function that takes x and feeds it *twice* into f. But just as functions can create other functions as *outputs*, there is no reason not to make functions take other functions as *inputs*. The process here was competely general—we knew nothing specific about f— so the function

$$\lambda y.\lambda x.((y\,x)\,x)$$

takes an input y, and returns the function $\lambda x.((y\,x)\,x)$. This function takes an input x, and then applies y to x and then applies the result to x again. When you feed it a function, it returns the *diagonal* of that function.

Draw the function as a table of values for each pair of inputs, and you will see why this is called the '*diagonal*.'

Now, sometimes this construction does not work. Suppose we feed our diagonal function $\lambda y.\lambda x.((y\,x)\,x)$ an input that is not a function, or that is a function that does not expect two inputs? (That is, it is not a function that returns another function.) In that case, we may not get a sensible output. One response is to bite the bullet and say that everything is a function, and that we can apply anything to anything else. We won't take that approach here, as something becomes very interesting if we consider what happens if we consider variables (the x and y in the expression $\lambda y.\lambda x.((y\,x)\,x)$) to be *typed*. We could consider y to only take inputs which are functions of the right kind. That is, y is a function that expects values of some kind (let's say, of type A), and when given a value, returns a function. In fact, the function it returns has to be a function that expects values of the very same kind (also type A). The *result* is an object (perhaps a function) of some kind or other (say, type B). In other words, we can say that the variable y takes values of type $A \to (A \to B)$. Then we expect the variable x to take values of type A. We'll write these facts as follows:

This is the *untyped* λ-calculus.

$$y : A \to (A \to B) \qquad x : A$$

Now, we may put these two things together, to say derive the type of the result of applying the function y to the input value x.

$$\frac{y : A \to (A \to B) \quad x : A}{(y\,x) : A \to B}$$

Applying the result to x again, we get

$$\dfrac{\dfrac{y : A \to (A \to B) \quad x : A}{(y\,x) : A \to B} \qquad x : A}{((y\,x)\,x) : B}$$

Then when we abstract away the particular choice of the input value x, we have this

$$\dfrac{\dfrac{\dfrac{y : A \to (A \to B) \quad [x : A]}{(y\,x) : A \to B} \qquad [x : A]}{((y\,x)\,x) : B}}{\lambda x.((y\,x)\,x) : A \to B}$$

and abstracting away the choice of y, we have

$$\dfrac{\dfrac{\dfrac{\dfrac{[y : A \to (A \to B)] \quad [x : A]}{(y\,x) : A \to B} \qquad [x : A]}{((y\,x)\,x) : B}}{\lambda x.((y\,x)\,x) : A \to B}}{\lambda y.\lambda x.((y\,x)\,x) : (A \to (A \to B)) \to (A \to B)}$$

so the diagonal function $\lambda y.\lambda x.((y\,x)\,x)$ has type $(A \to (A \to B)) \to (A \to B)$. It takes functions of type $A \to (A \to B)$ as input and returns an output of type $A \to B$.

Does that process look like something you have already seen?

We may use these λ-terms to represent proofs. Here are the definitions. We will first think of formulas as *types*.

$$\textsc{type} ::= \textsc{atom} \mid (\textsc{type} \to \textsc{type})$$

Then, given the class of types, we can construct terms for each type.

DEFINITION 1.29 [TYPED SIMPLE λ-TERMS] The class of typed simple λ-terms is defined as follows:

- » For each type A, there is an infinite supply of variables x^A, y^A, z^A, w^A, x_1^A, x_2^A, etc.
- » If M is a term of type $A \to B$ and N is a term of type A, then $(M\,N)$ is a term of type B.
- » If M is a term of type B then $\lambda x^A.M$ is a term of type $A \to B$.

These formation rules for types may be represented in ways familiar to those of us who care for proofs. See Figure 13.

Sometimes we write variables without superscripts, and leave the typing of the variable understood from the context. It is simpler to write $\lambda y.\lambda x.((y\,x)\,x)$ instead of $\lambda y^{A\to(A\to B)}.\lambda x^A((y^{A\to(A\to B)}\,x^A)\,x^A)$.

$$\frac{M : A \to B \quad N : A}{(M\,N) : B}\,{\to}E \qquad \frac{\begin{array}{c}[x : A]^{(i)} \\ \vdots \\ M : B\end{array}}{\lambda x.M : A \to B}\,{\to}I,i$$

Figure 13: RULES FOR λ-TERMS

Not everything that *looks* like a typed λ-term actually is. Consider the term

$$\lambda x.(x\,x)$$

There is no such simple typed λ-term. Were there such a term, then x would have to both have type $A \to B$ and type A. But as things stand now, a variable can have only one type. Not every λ-term is a *typed* λ-term.

Now, it is clear that typed λ-terms stand in some interesting relationship to proofs. From any typed λ-term we can reconstruct a unique proof.Take $\lambda x.\lambda y.(y\,x)$, where y has type $p \to q$ and x has type p. We can rewrite the unique formation pedigree of the term as a tree.

$$\cfrac{\cfrac{\cfrac{[y : p \to q] \quad [x : p]}{(y\,x) : q}}{\lambda y.(y\,x) : (p \to q) \to q}}{\lambda x.\lambda y.(y\,x) : p \to ((p \to q) \to q)}$$

and once we erase the terms, we have a proof of $p \to ((p \to q) \to q)$. The term is a compact, linear representation of the proof which is presented as a tree.

The mapping from terms to proofs is many-to-one. Each typed term constructs a single proof, but there are many different terms for the one proof. Consider the proofs

$$\frac{p \to q \quad p}{q} \qquad \frac{p \to (q \to r) \quad p}{(q \to r)}$$

we can label them as follows

$$\frac{x : p \to q \quad y : p}{(xy) : q} \qquad \frac{z : p \to (q \to r) \quad y : p}{(zy) : q \to r}$$

we could combine them into the proof

$$\cfrac{\cfrac{z : p \to (q \to r) \quad y : p}{(zy) : q \to r} \quad \cfrac{x : p \to q \quad y : p}{(xy) : q}}{(zy)(xy) : r}$$

but if we wished to discharge just *one* of the instances of p, we would have to have chosen a different term for one of the two subproofs. We could have chosen the variable w for the first p, and used the following term:

$$\dfrac{\dfrac{\dfrac{z : p \to (q \to r) \quad [w : p]}{(zw) : q \to r} \quad \dfrac{x : p \to q \quad y : p}{(xy) : q}}{(zw)(xy) : r}}{\lambda w.(zw)(xy) : p \to r}$$

So, the choice of variables allows us a great deal of choice in the construction of a term for a proof. The choice of variables both does *not* matter (who cares if we replace x^A by y^A) and *does* matter (when it comes to discharge an assumption, the formulas discharged are exactly those labelled by the particular free variable bound by λ at that stage).

DEFINITION 1.30 [FROM TERMS TO PROOFS AND BACK] For every typed term M (of type A), we find $\textsc{proof}(M)$ (of the formula A) as follows:

» $\textsc{proof}(x^A)$ is the identity proof A.

» If $\textsc{proof}(M^{A \to B})$ is the proof π_1 of $A \to B$ and $\textsc{proof}(N^A)$ is the proof π_2 of A, then extend them with one [$\to$*E*] step into the proof $\textsc{proof}(MN^B)$ of B.

» If $\textsc{proof}(M^B)$ is a proof π of B and x^A is a variable of type A, then construct the proof $\textsc{proof}((\lambda x.M)^{A \to B})$ of type $A \to B$ as follows: Extend the proof π by discharging each premise in π of type A labelled with the variable x^A.

Conversely, for any proof π, we find the set $\textsc{terms}(\pi)$ as follows:

» $\textsc{terms}(A)$ is the set of variables of type A. (Note that the term is an unbound variable, whose type is the only assumption in the proof.)

» If π_l is a proof of $A \to B$, and M (of type $A \to B$) is a member of $\textsc{terms}(\pi_l)$, and N (of type A) is a member of $\textsc{terms}(\pi_r)$, then (MN) (which is of type B) is a member of $\textsc{terms}(\pi)$, where π is the proof found by extending π_l and π_r by the [$\to$*E*] step. (Note that if the unbound variables in M have types corresponding to the assumptions in π_l and those in N have types corresponding to the assumptions in π_r, then the unbound variables in (MN) have types corresponding to the variables in π.)

» Suppose π is a proof of B, and we extend π into the proof π' by discharging some set (possibly empty) of instances of the formula A, to derive $A \to B$ using [$\to$*I*]. Then M is a member of $\textsc{terms}(\pi)$ for which a variable x labels *all* and *only* those assumptions A that are discharged in this [$\to$*I*] step, then $\lambda x.M$ is a member of $\textsc{terms}(\pi')$. (Notice that the free variables in $\lambda x.M$ correspond to the remaining active assumptions in π'.)

THEOREM 1.31 [TERMS ARE PROOFS ARE TERMS] *If* $M \in \textsc{terms}(\pi)$ *then* $\pi = \textsc{proof}(M)$. *Conversely,* $M' \in \textsc{terms}(\textsc{proof}(M))$ *if and only if* M' *is a relabelling of* M.

Proof: For the first part, we proceed by induction on the proof π. If π is an atomic proof, then since $\textsc{terms}(A)$ is the set of variables of type A, and since $\textsc{proof}(x^A)$ is the identity proof A, we have the base case of the induction. If π is composed of two proofs, π_l of $A \to B$, and π_r of A, joined by an $[\to E]$ step, then M is in $\textsc{terms}(\pi)$ if and only if $M = (N_1 N_2)$ where $N_1 \in \textsc{terms}(\pi_l)$ and $N_2 \in \textsc{terms}(\pi_r)$. But by the induction hypothesis, if $N_1 \in \textsc{terms}(\pi_l)$ and $N_2 \in \textsc{terms}(\pi_r)$, then $\pi_l = \textsc{proof}(N_1)$ and $\pi_r = \textsc{proof}(N_2)$, and as a result, $\pi = \textsc{proof}(M)$, as desired.

Finally, if π is a proof of B, extended to the proof π' of $A \to B$ by discharging some (possibly empty) set of instances of A, then if M is in $\textsc{terms}(\pi)$ if and only if $M = \lambda x.N$, $N \in \textsc{terms}(\pi')$, and x labels those (and only those) instances of A discharged in π. By the induction hypothesis, $\pi' = \textsc{proof}(N)$. It follows that $\pi = \textsc{proof}(\lambda x.N)$, since x labels all and only the formulas discharged in the step from π' to π.

For the second part of the proof, if $M' \in \textsc{terms}(\textsc{proof}(M))$, then if M is a variable, $\textsc{proof}(M)$ is an identity proof of some formula A, and $\textsc{terms}(\textsc{proof}(M))$ is a variable with type A, so the base case of our hypothesis is proved. Suppose the hypothesis holds for terms simpler than our term M. If M is an application term $(N_1 N_2)$, then $\textsc{proof}(N_1 N_2)$ ends in $[\to E]$, and the two subproofs are $\textsc{proof}(N_1)$ and (N_2) respectively. By hypothesis, $\textsc{term}(\textsc{proof}(N_1))$ is some relabelling of N_1 and $\textsc{term}(\textsc{proof}(N_2))$ is some relabelling of N_2, so $\textsc{term}(\textsc{proof}(N_1 N_2))$ may only be relabelling of $(N_1 N_2)$ as well. Similarly, if M is an abstraction term $\lambda x.N$, then $\textsc{proof}(\lambda x.N)$ ends in $[\to I]$ to prove some conditional $A \to B$, and $\textsc{proof}(N)$ is a proof of B, in which some (possibly empty) collection of instances of A are about to be discharged. By hypothesis, $\textsc{term}(\textsc{proof}(N))$ is a relabelling of N, so $\textsc{term}(\textsc{proof}(\lambda x.N))$ can only be a relabelling of $\lambda x.N$. ∎

The following theorem shows that the λ-terms of different kinds of proofs have different features.

THEOREM 1.32 [DISCHARGE CONDITIONS AND TERMS] M *is a* linear λ*-term* (*a term of some* linear *proof*) *iff each* λ *expression in* M *binds exactly one variable.* M *is a* relevant λ*-term* (*a term of a* relevant *proof*) *iff each* λ *expression in* M *binds at* least *one variable.* M *is a an* affine λ*-term* (*a term of some* affine *proof*) *iff each* λ *expression binds at* most *one variable.*

Proof: Check the definition of $\textsc{proof}(M)$. If M satisfies the conditions on variable binding, $\textsc{proof}(M)$ satisfies the corresponding discharge conditions. Conversely, if π satisfies a discharge condition, the terms in $\textsc{term}(\pi)$ are the corresponding kinds of λ-term. ∎

The most interesting connection between proofs and λ-terms is not simply this pair of mappings. It is the connection between *normalisation* and *evaluation*. We have seen how the application of a function, like $\lambda x.((y\,x)\,x)$ to an input like M is found by removing the lambda binder, and substituting the term M for each variable x that was bound by the binder. In this case, we get $((y\,M)\,M)$.

DEFINITION 1.33 [β REDUCTION] The term $\lambda x.M\,N$ is said to immediately β-reduce to the term $M[x := N]$ found by substituting the term N for each free occurrence of x in M.

Furthermore, M β-reduces in one step to M' if and only if some subterm N inside M immediately β-reduces to N' and $M' = M[N := N']$. A term M is said to β-reduce to M^* if there is some chain $M = M_1, \cdots M_n = M^*$ where each M_i β-reduces in one step to M_{i+1}.

Consider what this means for *proofs*. The term $(\lambda x.M\,N)$ immediately β-reduces to $M[x := N]$. Representing this transformation as a proof, we have

$$\dfrac{\dfrac{\begin{matrix}[x:A]\\ \vdots\,\pi_l\\ M:B\end{matrix}}{\lambda x.M:A\to B}\qquad \begin{matrix}\vdots\,\pi_r\\ N:A\end{matrix}}{(\lambda x.M\,N):B} \quad\Longrightarrow^{\beta}\quad \begin{matrix}\vdots\,\pi_r\\ N:A\\ \vdots\,\pi_l\\ M[x:=N]:B\end{matrix}$$

and β-reduction corresponds to normalisation. This fact leads immediately to the following theorem.

THEOREM 1.34 [NORMALISATION AND β-REDUCTION] *A proof* PROOF(N) *is normal if and only if the term* N *does not* β*-reduce to another term. If* N β*-reduces to* N' *then a normalisation process sends* PROOF(N) *to* PROOF(N').

This natural reading of normalisation as function application, and the easy way that we think of $(\lambda x.M\,N)$ as *being identical to* $M[x := N]$ leads some to make the following claim:

If π and π' *normalise* to the same proof,
then π and π' are *really* the same proof.

We will discuss proposals for the identity of proofs in a later section.

1.5 | HISTORY

Gentzen's technique for natural deduction is not the only way to represent this kind of reasoning, with introduction and elimination rules for connectives. Independently of Gentzen, the Polish logician, Stanisław Jaśkowski constructed a closely related, but different system for presenting proofs in a natural deduction style. In Jaśkowski's system, a proof is a *structured list* of formulas. Each formula in the list is either a *supposition*, or it follows from earlier formulas in the list by means of the rule of *modus ponens* (conditional elimination), or it is proved by *conditionalisation*. To prove something by conditionalisation you first make a supposition of the antecedent: at this point you start a *box*. The contents of a box constitute a proof, so if you want to use a formula from outside the box, you may *repeat* a formula into the inside. A conditionalisation step allows you to exit the box, discharging the supposition you made upon entry. Boxes can be nested, as follows:

1.	$A \to (A \to B)$	Supposition
2.	A	Supposition
3.	$A \to (A \to B)$	1, Repeat
4.	$A \to B$	2, 3, Modus Ponens
5.	B	2, 4, Modus Ponens
6.	$A \to B$	2–5, Conditionalisation
7.	$(A \to (A \to B)) \to (A \to B)$	1–6, Conditionalisation

This nesting of boxes, and repeating or reiteration of formulas to enter boxes, is the distinctive feature of Jaśkowski's system. Notice that we could prove the formula $(A \to (A \to B)) \to (A \to B)$ without using a duplicate discharge. The formula A is used twice as a minor premise in a Modus Ponens inference (on line 4, and on line 5), and it is then discharged at line 6. In a Gentzen proof of the same formula, the assumption A would have to be made twice.

Jaśkowski proofs also straightforwardly incorporate the effects of a vacuous discharge in a Gentzen proof. We can prove $A \to (B \to A)$ using the rules as they stand, without making any special plea for a vacuous discharge:

1.	A	Supposition
2.	B	Supposition
3.	A	1, Repeat
4.	$B \to A$	2–3, Conditionalisation
5.	$A \to (B \to A)$	1–4, Conditionalisation

The formula B is supposed, and it is not *used* in the proof that follows. The formula A on line 4 occurs *after* the formula B on line 3, in the subproof, but it is harder to see that it is inferred *from* that$B$. Conditionalisation, in Jaśkowski's system, colludes with reiteration to allow the effect of vacuous discharge. It appears that the "fine control" over inferential connections between formulas in proofs in a Gentzen proof is somewhat obscured in the *linearisation* of a Jaśkowski proof. The fact that one formula occurs *after* another says nothing about how that formula is inferentially connected to its forbear.

Jaśkowski's account of proof was modified in presentation by Frederic Fitch (boxes become assumption *lines* to the left, and hence become somewhat simpler to draw and to typeset). Fitch's natural deduction system gained quite some popularity in undergraduate education in logic in the 1960s and following decades in the United States [72]. Further forms of natural deduction are given in Patrick Suppes *Introduction to Logic* [235] and Edward Lemmon's text *Beginning Logic* [129]. Lemmon's and Suppes' accounts of natural deduction are similar to Fitch's, except that they do without the need to reiterate by *breaking the box*.

1	(1)	$A \rightarrow (A \rightarrow B)$	Assumption
2	(2)	A	Assumption
1,2	(3)	$A \rightarrow B$	1, 2, Modus Ponens
1,2	(4)	B	2,3, Modus Ponens
1	(5)	$A \rightarrow B$	2, 4, Conditionalisation
	(6)	B	1, 5, Conditionalisation

Now, line numbers are joined by *assumption numbers*: each formula is tagged with the line number of each assumption upon which that formula depends. The rules for the conditional are straightforward: If $A \rightarrow B$ depends on the assumptions X and A depends on the assumptions Y, then you can derive B, depending on the assumptions X, Y. (You should ask yourself if X, Y is the *set* union of the *sets* X and Y, or the *multiset* union of the *multisets* X and Y. For Lemmon, the assumption collections are *sets*.) For conditionalisation, if B depends on X, A, then you can derive $A \rightarrow B$ on the basis of X alone. As you can see, vacuous discharge is harder to motivate, as the rules stand now. If we attempt to use the strategy of the Jaśkowski proof, we are soon stuck:

1	(1)	A	Assumption
2	(2)	B	Assumption
$\vdots$	(3)	$\vdots$	

There is no way to attach the assumption number "2" to the formula A. The linear presentation is now explicitly *detached* from the inferential connections between formulas by way of the assumption numbers. Now the assumption numbers tell you all you need to know about the provenance of formulas. In Lemmon's own system, you *can* prove the formula $A \rightarrow (B \rightarrow A)$ but only, as it happens, by taking a detour through conjunction or some other connective.

1	(1)	A	Assumption
2	(2)	B	Assumption
1,2	(3)	$A \wedge B$	1,2, Conjunction intro
1,2	(4)	A	3, Conjunction elim
1	(5)	$B \rightarrow A$	2,4, Conditionalisation
	(6)	$A \rightarrow (B \rightarrow A)$	1,5, Conditionalisation

This seems quite unsatisfactory, as it breaks the normalisation property. (The formula $A \rightarrow (B \rightarrow A)$ is proved only by a non-normal proof—in this case, a proof in which a conjunction is introduced and then immediately eliminated.) Normalisation can be restored to Lemmon's system, but at the cost of the introduction of a new rule, the rule of *weakening*, which says that if A depends on assumptions X, then we can infer A depending on assumptions X together with another formula.

For more information on the history of natural deduction, consult Jeffrey Pelletier's article [161].

Notice that the lines in a Lemmon proof don't just contain *formulas* (or formulas tagged a line number and information about how the formula was deduced). They are *pairs*, consisting of a formula, and the formulas upon which the formula depends. In a Gentzen proof this information is implicit in the structure of the proof. (The formulas upon which a formula depends in a Gentzen proof are the leaves in the tree above that formula that are undischarged at the moment that this formula is derived.) This feature of Lemmon's system was not original to him. The idea of making completely explicit the assumptions upon which a formula depends had also occurred to Gentzen, and this insight is our topic for the next section.

» «

Linear, relevant and affine implication have a long history. Relevant implication burst on the scene through the work of Alan Anderson and Nuel Belnap in the 1960s and 1970s [3, 4], though it had precursors in the work of the Russian logician, I. E. Orlov in the 1920s [56, 155]. The idea of a proof in which conditionals could only be introduced if the assumption for discharge was genuinely *used* is indeed one of the motivations for relevant implication in the Anderson–Belnap tradition. However, *other* motivating concerns played a role in the development of relevant logics. For other work on relevant logic, the work of Dunn [60, 62], Routley and Meyer [207], Read [184] and Mares [137] are all useful. Linear logic arose much more centrally out of proof-theoretical concerns in the work of the proof-theorist Jean-Yves Girard in the 1980s [86, 88]. (Girard, also, named the logic including weakening but avoiding contraction *affine* logic [86, page ?].) A helpful introduction to linear logic is the text of Troelstra [239]. Affine logic, as a distinct system, goes back at least to the early 1980s in the work of Grišin [92], and is discussed extensively in an influential paper by Ono and Komori from 1985 [154]. Affine implication is quite close to the implication in Łukasiewicz's infinitely valued logic—which is slightly stronger, but shares the property of rejecting all *contraction*-related principles [192]. These logics are all *substructural* logics [57, 158, 193].

The definition of normality used here due to Prawitz [169], though noemalisation is present in Fitch's *Symbolic Logic* [72, p. 115ff], and glimpses of the idea are present in Gentzen's original work [81].

The λ-calculus is due to Alonzo Church [39], and the study of λ-calculi has found many different applications in logic, computer science, type theory and related fields [8, 99, 218]. The correspondence between formulas/proofs and types/terms is known as the Curry–Howard correspondence [46, 112].

1.6 | EXERCISES

Working through these exercises will help you understand the material. As with all logic exercises, if you want to deepen your understanding of these techniques, you should attempt the exercises until they are no longer difficult. So, attempt each of the different kinds of *basic* exercises, until you know you can do them. Then move on to the *intermediate*

I am not altogether confident about the division of the exercises into "basic," "intermediate," and "advanced." I'd appreciate your feedback on whether some exercises are too easy or too difficult for their categories.

exercises, and so on. (The *project* exercises are not the kind of thing that can be completed in one sitting.)

BASIC EXERCISES

Q1 Which of the following formulas have proofs with no premises?

Formula 4 is *Peirce's Law*. It is a two-valued classical logic tautology.

1 : $p \to (p \to p)$
2 : $p \to (q \to q)$
3 : $((p \to p) \to p) \to p$
4 : $((p \to q) \to p) \to p$
5 : $((q \to q) \to p) \to p$
6 : $((p \to q) \to q) \to p$
7 : $p \to (q \to (q \to p))$
8 : $(p \to q) \to (p \to (p \to q))$
9 : $((q \to p) \to p) \to ((p \to q) \to q)$
10 : $(p \to q) \to ((q \to p) \to (p \to p))$
11 : $(p \to q) \to ((q \to p) \to (p \to q))$
12 : $(p \to q) \to ((p \to (q \to r)) \to (p \to r))$
13 : $(q \to p) \to ((p \to q) \to ((q \to p) \to (p \to q)))$
14 : $((p \to p) \to p) \to ((p \to p) \to ((p \to p) \to p))$
15 : $(p_1 \to p_2) \to ((q \to (p_2 \to r)) \to (q \to (p_1 \to r)))$

For each formula that can be proved, find a proof that complies with the strictest discharge policy possible.

Q2 Annotate your proofs from Exercise 1 with λ-terms.

Q3 Construct a proof from $q \to r$ to $(q \to (p \to p)) \to (q \to r)$ using vacuous discharge. Then construct a proof of $q \to (p \to p)$ (also using vacuous discharge). Combine the two proofs, using [$\to$*E*] to deduce $q \to r$. Normalise the proof you find. Then annotate each proof with λ-terms, and explain the β reductions of the terms corresponding to the normalisation.

Then construct a proof from $(p \to r) \to ((p \to r) \to q))$ to $(p \to r) \to q$ using duplicate discharge. Then construct a proof from $p \to (q \to r)$ and $p \to q$ to $p \to r$ (also using duplicate discharge). Combine the two proofs, using [$\to$*E*] to deduce q. Normalise the proof you find. Then annotate each proof with λ-terms, and explain the β reductions of the terms corresponding to the normalisation.

Q4 Find types and proofs for each of the following terms.

1 : $\lambda x.\lambda y.x$
2 : $\lambda x.\lambda y.\lambda z.((xz)(yz))$
3 : $\lambda x.\lambda y.\lambda z.(x(yz))$
4 : $\lambda x.\lambda y.(yx)$
5 : $\lambda x.\lambda y.((yx)x)$

Which of the proofs are linear, which are relevant and which are affine?

Q5 Show that there is no normal relevant proof of these formulas.

1 : $p \to (q \to p)$

$2 : (p \to q) \to (p \to (r \to q))$
$3 : p \to (p \to p)$

Q6 Show that there is no normal affine proof of these formulas.

$1 : (p \to q) \to ((p \to (q \to r)) \to (p \to r))$
$2 : (p \to (p \to q)) \to (p \to q)$

Q7 Show that there is no normal proof of these formulas.

$1 : ((p \to q) \to p) \to p$
$2 : ((p \to q) \to q) \to ((q \to p) \to p)$

Q8 Find a formula that can has both a relevant proof and an affine proof, but no linear proof.

INTERMEDIATE EXERCISES

Q9 Consider the following "truth tables."

GD3

$\to$	t	n	f
t	t	n	f
n	t	t	f
f	t	t	t

Ł3

$\to$	t	n	f
t	t	n	f
n	t	t	f
f	t	t	t

RM3

$\to$	t	b	f
t	t	f	f
n	t	b	f
f	t	t	t

BN4

$\to$	t	b	n	f
t	t	f	n	f
b	t	b	n	f
n	t	n	t	n
f	t	t	t	t

A GD3 tautology is a formula that receives the value t in every GD3 valuation. An Ł3 tautology is a formula that receives the value t in every Ł3 valuation. Show that every formula with a standard proof is a GD3 tautology. Show that every formula with an affine proof is an Ł3 tautology.

A RM3 tautology is a formula that receives either of the values t and b in every RM3 valuation. A BN4 tautology is a formula that receives either of the values t and b in every BN4 valuation. Show that every formula with a relevant proof is an RM3 tautology. Show that every formula with a linear proof is a BN4 tautology.

Q10 Consider proofs that have paired steps of the form [$\to E/\to I$]. That is, a conditional is eliminated only to be introduced again. The proof has a sub-proof of the form of this proof fragment:

$$\dfrac{\dfrac{A \to B \quad [A]^{(i)}}{B}{\scriptstyle \to E}}{A \to B}{\scriptstyle \to I, i}$$

These proofs contain redundancies too, but they may well be normal. Call a proof with a pair like this CIRCUITOUS. Show that all circuitous proofs may be transformed into non-circuitous proofs with the same premises and conclusion.

Q11 In Exercise 5 you showed that there is no normal relevant proof of $p \to (p \to p)$. By normalisation, it follows that there is no relevant proof (normal or not) of $p \rightarrow (p \rightarrow p)$. Use this fact to explain why it is more natural to consider relevant arguments with *multisets* of premises and not just *sets* of premises. (HINT: is the argument from p, p to p relevantly valid?)

Q12 You might think that "if … then …" is a slender foundation upon which to build an account of logical consequence. Remarkably, there is rather a lot that you can do with implication alone, as these next questions ask you to explore.

First, define $A \hat{\vee} B$ as follows: $A \hat{\vee} B ::= (A \rightarrow B) \rightarrow B$. In what way is "$\hat{\vee}$" like *disjunction*? What usual features of disjunction are not had by $\hat{\vee}$? (Pay attention to the behaviour of $\hat{\vee}$ with respect to different discharge policies for implication.)

Q13 Think about what it would take to have introduction and elimination rules for $\hat{\vee}$ that do not involve the conditional connective $\rightarrow$. Can you do this?

Q14 Now consider *negation*. Given an ATOM p, define the p-negation $\neg_p A$ to be $A \rightarrow p$. In what way is "$\neg_p$" like negation? What usual features of negation are not had by $\neg_p$ defined in this way? (Pay attention to the behaviour of $\neg$ with respect to different discharge policies for implication.)

Q15 Provide introduction and elimination rules for $\neg_p$ that do not involve the conditional connective $\rightarrow$.

Q16 You have probably noticed that the inference from $\neg_p \neg_p A$ to A is not, in general, valid. Define a *new* language CFORMULA inside FORMULA as follows:

$$\text{CFORMULA} ::= \neg_p \neg_p \text{ATOM} \mid (\text{CFORMULA} \rightarrow \text{CFORMULA})$$

Show that $\neg_p \neg_p A \therefore A$ and $A \therefore \neg_p \neg_p A$ are valid when A is a CFORMULA.

Q17 Now define $A \dot{\wedge} B$ to be $\neg_p (A \rightarrow \neg_p B)$, and $A \dot{\vee} B$ to be $\neg_p A \rightarrow B$. In what way are $A \dot{\wedge} B$ and $A \dot{\vee} B$ like conjunction and disjunction of A and B respectively? (Consider the difference between when A and B are FORMULAS and when they are CFORMULAS.)

Q18 Show that if there is a normal relevant proof of $A \rightarrow B$ then there is an ATOM occurring in both A and B.

Q19 Show that if we have two conditional connectives $\rightarrow_1$ and $\rightarrow_2$ defined using different discharge policies, then the conditionals collapse, in the sense that we can construct proofs from $A \rightarrow_1 B$ to $A \rightarrow_2 B$ and *vice versa*.

Q20 Explain the significance of the result of Exercise 19.

Q21 Add rules the obvious introduction rules for a *conjunction* connective $\otimes$ as follows:

$$\frac{A \quad B}{A \otimes B}\,{\scriptstyle \otimes I}$$

Show that if we have the following two $\otimes E$ rules:

$$\frac{A \otimes B}{A}\,{\scriptstyle \otimes E_1} \qquad\qquad \frac{A \otimes B}{B}\,{\scriptstyle \otimes E_2}$$

we may simulate the behaviour of vacuous discharge. Show, then, that we may normalise proofs involving these rules (by showing how to eliminate all indirect pairs, including $\otimes I/\otimes E$ pairs).

ADVANCED EXERCISES

Q22 Another demonstration of the subformula property for normal proofs uses the notion of a *track* in a proof.

DEFINITION 1.35 [TRACK] A sequence $A_0, \dots, A_n$ of formula instances in the proof π is a *track* of length $n + 1$ in the proof π if and only if

- A_0 is a *leaf* in the proof tree.
- Each A_{i+1} is immediately below A_i.
- For each $i < n$, A_i is not a minor premise of an application of $[{\to}E]$.

A track whose terminus A_n is the conclusion of the proof π is said to be a TRACK OF ORDER 0. If we have a track t whose terminus A_n is the minor premise of an application of $[{\to}E]$ whose conclusion is in a track of order n, we say that t is a TRACK OF ORDER $n + 1$.

The following annotated proof gives an example of tracks.

$$\cfrac{\cfrac{\heartsuit [B \to C]^{(2)} \qquad \cfrac{\cfrac{\spadesuit\, A \to ((D \to D) \to B) \quad \diamondsuit [A]^{(2)}}{\spadesuit\, (D \to D) \to B}\,{\scriptstyle \to E} \qquad \cfrac{\clubsuit [D]^{(1)}}{\clubsuit\, D \to D}\,{\scriptstyle \to I,1}}{\spadesuit\, B}\,{\scriptstyle \to E}}{\cfrac{\heartsuit\, C}{\heartsuit\, A \to C}\,{\scriptstyle \to I,2}}\,{\scriptstyle \to E}}{\heartsuit\, (B \to C) \to (A \to C)}\,{\scriptstyle \to I,3}$$

(Don't let the fact that this proof has one track of each order 0, 1, 2 and 3 make you think that proofs can't have more than one track of the same order. Look at this example—

$$\frac{\dfrac{A \to (B \to C) \quad A}{B \to C} \qquad B}{C}$$

—it has two tracks of order 1.) The formulas labelled with $\heartsuit$ form one track, starting with $B \to C$ and ending at the conclusion of the proof. Since this track ends at the conclusion of the proof, it is a track of order 0. The track consisting of $\spadesuit$ formulas starts at $A \to ((D \to D) \to B)$ and ends at B. It is a track of order 1, since its final formula is the minor premise in the [$\rightarrow E$] whose conclusion is C, in the ♡ track of order 0. Similarly, the ♢ track is order 2 and the ♣ track has order 3.

For this exercise, prove the following lemma by induction on the construction of a proof.

LEMMA 1.36 *In every proof, every formula is in one and only one track, and each track has one and only one order.*

Then prove this lemma.

LEMMA 1.37 *Let* $t : A_0, \dots, A_n$ *be a track in a normal proof. Then*

a) *The rules applied within the track consist of a sequence (possibly empty) of* [$\rightarrow$E] *steps and then a sequence (possibly empty) of* [$\rightarrow$I] *steps.*

b) *Every formula* A_i *in* t *is a subformula of* A_0 *or of* A_n.

Now prove the subformula theorem, using these lemmas.

Q23 Consider the result of Exercise 19. Show how you might define a natural deduction system containing (say) both a linear and a standard conditional, in which there is *no* collapse. That is, construct a system of natural deduction proofs in which there are two conditional connectives: $\rightarrow_l$ for linear conditionals, and $\rightarrow_s$ for standard conditionals, such that whenever an argument is valid for a linear conditional, it is (in some appropraite sense) valid in the system you design (when $\rightarrow$ is translated as $\rightarrow_l$) and whenever an argument is valid for a standard conditional, it is (in some appropriate sense) valid in the system you design (when $\rightarrow$ is translated as $\rightarrow_s$). What mixed inferences (those using both $\rightarrow_l$ and $\rightarrow_s$) are valid in your system?

Q24 Suppose we have a new discharge policy that is "stricter than linear." The *ordered* discharge policy allows you to discharge only the *rightmost* assumption at any one time. It is best paired with a strict version of [$\rightarrow E$] according to which the major premise ($A \rightarrow B$) is on the left, and the minor premise (A) is on the right. What is the resulting logic *like*? Does it have the normalisation property?

Q25 Take the logic of Exercise 24, and extend it with *another* connective $\leftarrow$, with the rule [$\leftarrow E$] in which the major premise ($B \leftarrow A$) is on the *right*, and the minor premise (A) is on the *left*, and [$\leftarrow I$], in which the *leftmost* assumption is discharged. Examine the connections between $\rightarrow$ and $\leftarrow$. Does normalisation work for *these* proofs? This is *Lambek's* logic for syntactic types [125, 126, 147, 148].

Q26 Show that there is a way to be even *stricter* than the discharge policy of Exercise 24. What is the *strictest* discharge policy for $\rightarrow I$, that will result in a system which normalises, provided that $\rightarrow E$ (in which the major premise is leftmost) is the only other rule for implication.

Q27 Consider the introduction rule for $\otimes$ given in Exercise 21. Construct an appropriate *elimination* rule for fusion which does not allow the simulation of vacuous (or duplicate) discharge, and for which proofs normalise.

Q28 Identify two proofs where one can be reduced to the other by way of the elimination of *circuitous* steps (see Exercise 10). Characterise the identities this provides among λ-terms. Can this kind if identification be maintained along with β-reduction?

PROJECT

Q29 Thoroughly and systematically explain and evaluate the considerations for choosing one discharge policy over another. This will involve looking at the different *uses* to which one might put a system of natural deduction, and then, relative to a use, what one might say in favour of a different policy.

SEQUENT CALCULUS 2

In this chapter we will look at a different way of thinking about proof and consequence: Gentzen's *sequent calculus*. The core idea is straightforward. We want to know what follows from what, so we will keep a track of facts of consequence: facts we will record in the following form:

$$A \vdash B$$

One can read "$A \vdash B$" in a number of ways. You can say that B follows from A, or that A entails B, that the argument from A to B is valid, or that asserting A clashes with denying B, or – and this is the understanding most appropriate for us – that there is a proof from A to B. The symbol used between A and B is sometimes called the TURNSTILE.

"Scorning a turnstile wheel at her reverend helm, she sported there a tiller; and that tiller was in one mass, curiously carved from the long narrow lower jaw of her hereditary foe. The helmsman who steered by that tiller in a tempest, felt like the Tartar, when he holds back his fiery steed by clutching its jaw. A noble craft, but somehow a most melancholy! All noble things are touched with that." — Herman Melville, *Moby Dick*.

Once we have this notion of consequence, we can ask ourselves what properties consequence has. There are many different ways you could answer this question. The focus of this section will be a particular technique, originally due to Gerhard Gentzen. We can think of consequence – relative to a particular *language* – like this: when we want to know about the relation of consequence, we first consider each different kind of formula in the language. To make the discussion concrete, let's consider a very simple language: the language of propositional logic with only two connectives, *conjunction* $\wedge$ and *disjunction* $\vee$. That is, we will now look at formulas expressed in a language with the following grammar:

$$\text{FORMULA} ::= \text{ATOM} \mid (\text{FORMULA} \wedge \text{FORMULA}) \mid (\text{FORMULA} \vee \text{FORMULA})$$

To characterise consequence relations, we need to characterise how consequence works on the *atoms* of the language, and then how the addition of $\wedge$ and $\vee$ expands the repertoire of facts about consequence. To do this, we need to know when $A \vdash B$ when A is a conjunction, or when A is a disjunction, and when B is a conjunction, or when B is a disjunction. In other words, for each connective, we need to know when we can prove something *from* a formula featuring that connective, and when we can prove a formula featuring that connective. Another way of putting it is that we wish to know how a connective behave on the left of the turnstile, and how it behaves on the right.

In a sequent system, we will have rules concerning statements about consequence—and these statements are the *sequents* at the heart of the system. Because we can make false claims as well as true ones, we will use the following *bent* turnstile for the general case of a sequent

$$A \succ B$$

I follow Lloyd Humberstone, who, as far as I am aware, introduced this convention for sequents [114]. Gentzen used the arrow, which we have already used for the object-language conditional.

and we reserve the straight turnstile $A \vdash B$ for when we wish to explicitly claim that the sequent $A \succ B$ is *derivable*. In what follows, $p \succ p \wedge q$

is a perfectly good sequent, though it will not be a derivable one (for $p \wedge q$ does not follow from p), so we will not have $p \vdash p \wedge q$.

This is a *formal* account of consequence. We look only at the form of propositions and not their content. For atomic propositions (those with no internal form) there is nothing upon which we could pin a claim to consequence. Thus we never have $p \vdash q$ where p and q are different atoms, while $p \vdash p$ for all atoms p.

The answers for our language seem straightforward. For atomic formulas, p and q, the sequent $p \succ q$ is derivable only if p and q are the *same* atom: so we have $p \vdash p$ for each atom p. For conjunction, we can say that if $A \succ B$ and $A \succ C$ are derivable, then so is $A \succ B \wedge C$. That's how we can infer *to* a conjunction. Inferring *from* a conjunction is also straightforward. We can say that $A \wedge B \succ C$ when $A \succ C$, or when $B \succ C$. For disjunction, we can reason similarly. We can say $A \vee B \succ C$ when $A \succ C$ and $B \succ C$. We can say $A \succ B \vee C$ when $A \succ B$, or when $A \succ C$. This is *inclusive* disjunction, not exclusive disjunction.

You can think of these definitions as adding new material (in this case, conjunction and disjunction) to a pre-existing language. Think of the inferential repertoire of the basic language as settled (in our discussion this is *very* basic, just the atoms), and the connective rules are "definitional" extensions of the basic language. These thoughts are the raw materials for the development of an account of proof and logical consequence in general.

2.1 | DERIVATIONS

Like natural deduction proofs, derivations involving sequents are trees. The structure is as before:

$$\frac{\dfrac{\bullet \quad \bullet}{\bullet} \qquad \dfrac{\dfrac{\bullet \quad \bullet}{\bullet}}{\bullet}}{\bullet}$$

Where each position on the tree follows from those above it. In a tree, the *order* of the branches does not matter. These are two different ways to present the *same* tree:

$$\frac{A \quad B}{C} \qquad \frac{B \quad A}{C}$$

In this case, the tree structure is at the one and the same time *simpler* and *more complicated* than the tree structure of natural deduction proofs. They are simpler, in that there is no discharge. They are more complicated, in that trees are not trees of formulas. They are trees consisting of *sequents*. As a result, we will call these structures DERIVATIONS instead of PROOFS. The distinction is simple. For us, a proof is a structure in which the *formulas* are connected by inferential relations in a tree-like structure. A proof will go *from* some formulas *to* other formulas, *via* yet other formulas. Our structures involving sequents are quite different. The last sequent in a tree (the *endsequent*) is itself a statement of consequence, with its own antecedent and consequent (or premise and conclusion, if you prefer.) The tree *derivation* shows you why (or perhaps how) you can infer from the antecedent to the consequent. The rules for constructing sequent derivations are found in Figure 21.

I say "tree-like" since we will see different structures in later chapters.

$$p \succ p \; Id \qquad \frac{L \succ C \quad C \succ R}{L \succ R}\,Cut$$

$$\frac{A \succ R}{A \land B \succ R}\land L_1 \qquad \frac{A \succ R}{B \land A \succ R}\land L_2 \qquad \frac{L \succ A \quad L \succ B}{L \succ A \land B}\land R$$

$$\frac{A \succ R \quad B \succ R}{A \lor B \succ R}\lor L \qquad \frac{L \succ A}{L \succ A \lor B}\lor R_1 \qquad \frac{L \succ A}{L \succ B \lor A}\lor R_2$$

Figure 21: a simple sequent system

definition 2.1 [simple sequent derivation] If the leaves of a tree are instances of the *Id* rule, and if its transitions from node to node are instances of the other rules in Figure 21, then the tree is said to be a simple sequent derivation.

We must read these rules completely literally. Do not presume any properties of conjunction or disjunction other than those that can be demonstrated on the basis of the rules. We will take these rules as *constituting* the behaviour of the connectives $\land$ and $\lor$.

example 2.2 [example sequent derivations] In this section, we will look at a few sequent derivations, demonstrating some simple properties of conjunction, disjunction, and the consequence relation.

The first derivations show some commutative and associative properties of conjunction and disjunction. Here is the conjunction case, with derivations to the effect that $p \land q \vdash q \land p$, and that $p \land (q \land r) \vdash (p \land q) \land r$.

$$\frac{\dfrac{q \succ q}{p \land q \succ q}\land L_2 \quad \dfrac{p \succ p}{p \land q \succ p}\land L_1}{p \land q \succ q \land p}\land R \qquad \frac{\dfrac{\dfrac{p \succ p}{p \land (q \land r) \succ p}\land L_1 \quad \dfrac{\dfrac{q \succ q}{q \land r \succ q}\land L_1}{p \land (q \land r) \succ q}\land L_2}{p \land (q \land r) \succ p \land q}\land R \quad \dfrac{\dfrac{r \succ r}{q \land r \succ r}\land L_2}{p \land (q \land r) \succ r}\land L_2}{p \land (q \land r) \succ (p \land q) \land r}\land R$$

Next come the cases for disjunction. The first derivation below is for the commutativity of disjunction, and the second is its associativity.

$$\frac{\dfrac{p \succ p}{p \succ q \lor p}\lor R_1 \quad \dfrac{q \succ q}{q \succ p \lor q}\lor R_2}{p \lor q \succ q \lor p}\lor L \qquad \frac{\dfrac{\dfrac{p \succ p}{p \succ p \lor (q \lor r)}\lor R_1 \quad \dfrac{\dfrac{q \succ q}{q \succ q \lor r}\lor R_1}{q \succ p \lor (q \lor r)}\lor R_2}{p \lor q \succ p \lor (q \lor r)}\lor L \quad \dfrac{\dfrac{r \succ r}{r \succ q \lor r}\lor R_2}{r \succ p \lor (q \lor r)}\lor R_2}{(p \lor q) \lor r \succ p \lor (q \lor r)}\lor L$$

It is important to notice that these are not derivations of the commutativity or associativity of conjunction or disjunction in *general*. They only show the commutativity and associativity of conjunction and disjunction of *atomic* formulas. These are not derivations of $A \land B \succ B \land A$ (for example) since $A \succ A$ is not an axiom if A is a complex formula. We will see more on this in the next section.

Exercise 15 on page 115 asks you to make this duality precise.

You can see that the disjunction derivations have the same structure as those for conjunction. You can convert any derivation into another (its *dual*) by swapping conjunction and disjunction, and swapping the left-hand side of the sequent with the right-hand side. Here are some more examples of duality between derivations. The first is the dual of the second, and the third is the dual of the fourth.

$$\dfrac{p \succ p \quad p \succ p}{p \lor p \succ p}{\scriptstyle \lor L} \qquad \dfrac{p \succ p \quad p \succ p}{p \succ p \land p}{\scriptstyle \land R} \qquad \dfrac{p \succ p \quad \dfrac{p \succ p}{p \land q \succ p}{\scriptstyle \land L_1}}{p \lor (p \land q) \succ p}{\scriptstyle \lor L} \qquad \dfrac{p \succ p \quad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R}}{p \succ p \land (p \lor q)}{\scriptstyle \land R}$$

You can use derivations you have at hand, like these, as components of other derivations. One way to do this is to use the *Cut* rule.

$$\dfrac{\dfrac{p \succ p \quad \dfrac{p \succ p}{p \land q \succ p}{\scriptstyle \land L_1}}{p \lor (p \land q) \succ p}{\scriptstyle \lor L} \quad \dfrac{p \succ p \quad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R}}{p \succ p \land (p \lor q)}{\scriptstyle \land R}}{p \lor (p \land q) \succ p \land (p \lor q)}{\scriptstyle Cut}$$

Notice, too, that each of these derivations we've seen so far move from less complex formulas at the top to more complex formulas, at the bottom. Reading from bottom to top, you can see the formulas decomposing into their constituent parts. This isn't the case for all sequent derivations. Derivations that use the *Cut* rule can include new (more complex) material in the process of deduction. Here is an example:

$$\dfrac{\dfrac{\dfrac{p \succ p}{p \succ q \lor p}{\scriptstyle \lor R_1} \quad \dfrac{q \succ q}{q \succ q \lor p}{\scriptstyle \lor R_2}}{p \lor q \succ q \lor p}{\scriptstyle \lor L} \quad \dfrac{\dfrac{q \succ q}{q \succ p \lor q}{\scriptstyle \lor R_1} \quad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_2}}{q \lor p \succ p \lor q}{\scriptstyle \lor L}}{p \lor q \succ p \lor q}{\scriptstyle Cut}$$

We call the concluding sequent of a derivation the "ENDSEQUENT."

This derivation is a complicated way to deduce $p \lor q \succ p \lor q$, and it includes $q \lor p$, which is not a subformula of any formula in the final sequent of the derivation. Reading from bottom to top, the *Cut* step can introduce new formulas into the derivation.

2.2 | IDENTITY & CUT CAN BE ELIMINATED

The two distinctive rules in our proof system are *Id* and *Cut*. These rules are not about any particular kind of formula—they are *structural*, governing the behaviour of derivations, no matter *what* the nature of the formulas flanking the turnstiles. In this section we will look at the distinctive behaviour of *Id* and of *Cut*. We start with *Id*.

IDENTITY

This derivation of $p \lor q \succ p \lor q$ is *a* derivation of an identity (a sequent of the form $A \succ A$). There is a more systematic way to show that $p \lor q \succ p \lor q$, and any identity sequent. Here is a derivation of the sequent without *Cut*, and its dual, for conjunction.

$$\dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1} \qquad \dfrac{q \succ q}{q \succ p \lor q}{\scriptstyle \lor R_2}}{p \lor q \succ p \lor q}{\scriptstyle \lor L} \qquad\qquad \dfrac{\dfrac{p \succ p}{p \land q \succ p}{\scriptstyle \land L_1} \qquad \dfrac{q \succ q}{p \land q \succ q}{\scriptstyle \land L_2}}{p \land q \succ p \land q}{\scriptstyle \land R}$$

We can piece together these little derivations in order to derive any sequent of the form $A \succ A$. For example, here is the start of derivation of $p \land (q \lor (r_1 \land r_2)) \succ p \land (q \lor (r_1 \land r_2))$.

$$\dfrac{\dfrac{p \succ p}{p \land (q \lor (r_1 \land r_2)) \succ p}{\scriptstyle \land L_1} \qquad \dfrac{q \lor (r_1 \land r_2) \succ q \lor (r_1 \land r_2)}{p \land (q \lor (r_1 \land r_2)) \succ q \lor (r_1 \land r_2)}{\scriptstyle \land L_2}}{p \land (q \lor (r_1 \land r_2)) \succ p \land (q \lor (r_1 \land r_2))}{\scriptstyle \land R}$$

It's not a complete derivation yet, as one leaf $q \lor (r_1 \land r_2) \succ q \lor (r_1 \land r_2)$ is not an axiom. However, we can add the derivation for it.

$$\dfrac{\dfrac{p \succ p}{p \land (q \lor (r_1 \land r_2)) \succ p} \qquad \dfrac{\dfrac{\dfrac{q \succ q}{q \succ q \lor (r_1 \land r_2)}{\scriptstyle \lor R_1} \qquad \dfrac{\dfrac{\dfrac{r_1 \succ r_1}{r_1 \land r_2 \succ r_1}{\scriptstyle \land L_1} \qquad \dfrac{r_2 \succ r_2}{r_1 \land r_2 \succ r_2}{\scriptstyle \land L_2}}{r_1 \land r_2 \succ r_1 \land r_2}{\scriptstyle \land R}}{r_1 \land r_2 \succ q \lor (r_1 \land r_2)}{\scriptstyle \lor R_2}}{q \lor (r_1 \land r_2) \succ q \lor (r_1 \land r_2)}{\scriptstyle \lor L}}{p \land (q \lor (r_1 \land r_2)) \succ q \lor (r_1 \land r_2)}}{p \land (q \lor (r_1 \land r_2)) \succ p \land (q \lor (r_1 \land r_2))}$$

The derivation of $q \lor (r_1 \land r_2) \succ q \lor (r_1 \land r_2)$ itself contains a smaller identity derivation, for $r_1 \land r_2 \succ r_1 \land r_2$. The derivation displayed here uses shading to indicate the way the derivations are nested together. This result is general, and it is worth a theorem of its own.

THEOREM 2.3 [IDENTITY DERIVATIONS] *For each formula A, the sequent $A \succ A$ has a derivation. A derivation for $A \succ A$ may be systematically constructed from the identity derivations for the subformulas of A.*

Proof: We define Id_A, the IDENTITY DERIVATION FOR A by induction on the construction of A, as follows. Id_p is the axiom $p \succ p$. For complex formulas, we have

$$Id_{A \lor B} : \dfrac{\dfrac{Id_A}{A \succ A \lor B}{\scriptstyle \lor R_1} \qquad \dfrac{Id_B}{B \succ A \lor B}{\scriptstyle \lor R_2}}{A \lor B \succ A \lor B}{\scriptstyle \lor L} \qquad Id_{A \land B} : \dfrac{\dfrac{Id_A}{A \land B \succ A}{\scriptstyle \land L_1} \qquad \dfrac{Id_B}{A \land B \succ B}{\scriptstyle \land L_2}}{A \land B \succ A \land B}{\scriptstyle \land R}$$

$\blacksquare$

We say that $A \succ A$ is DERIVABLE in the sequent system. If we think of *Id* as a degenerate *rule* (a rule with no premise), then its generalisation, Id_A, is a *derivable rule*.

It might seem *crazy* to have a proof of identity, like $A \succ A$ where A is a complex formula. Why don't we take Id_A as an axiom? There are a few different reasons we might like to consider for taking Id_A as derivable instead of one of the primitive axioms of the system.

THE SYSTEM IS SIMPLE: In an axiomatic theory, it is always preferable to minimise the number of primitive assumptions. Here, it's clear that Id_A is derivable, so there is no need for it to be an axiom. A system with fewer axioms is preferable to one with more, for the reason that we have reduced derivations to a smaller set of primitive notions.

These are part of a general story, to be explored throughout this book, of what it is to be a logical constant. These sorts of considerations have a long history [95].

THE RULES ARE SEPARABLE: In the system without Id_A as an axiom, when we consider a sequent like $L \succ R$ in order to know whether it it is derived (in the absence of *Cut*, at least), we can ask only two distinct questions: concerning L and where *it* came from, and concerning R and *its* origin. Consider L. If it is complex perhaps $L \succ R$ is derivable by means of a left rule like [$\wedge$L] or [$\vee$L]. On the other hand, if R is complex, then perhaps the sequent is derivable by means of a right rule, like [$\wedge$R] or [$\vee$R]. If both are primitive, then $L \succ R$ is derivable by identity only. And that is it! You check the left, check the right, and there's no other possibility. There is no other condition under which the sequent is derivable. In the presence of Id_A, one would have to check if $L = R$ as well as the left and right rules. In the absence of Id_A, the identity condition arises only when both L and R are simple formulas. We appeal to identity only when the other rules do not apply.

THE SYSTEM PROVIDES A CONSTRAINT: Following on from this, in the absence of a general identity axiom, the burden on deriving identity is passed over to the connective rules. Allowing derivations of identity statements is a hurdle over which a connective rule might be able to jump, or over which it might *fail*. As we shall see later, this is provides a constraint we can use to sort out "good" definitions from "bad" ones. Given that the left and right rules for conjunction and disjunction tell you how the connectives are to be introduced, it would seem that the rules are defective (or at the very least, *incomplete*) if they don't allow the derivation of each instance of *Id*. We will make much more of this when we consider other connectives. However, before we make more of the philosophical motivations and implications of this constraint, we will add another possible constraint on connective rules, this time to do with the other rule in our system, *Cut*.

CUT

Some of the nice properties of a sequent system are as a matter of fact, the nice features of derivations that are constructed without the *Cut* rule.

Derivations constructed without *Cut* satisfy the subformula property.

THEOREM 2.4 [SUBFORMULA PROPERTY] *If δ is a sequent derivation not containing Cut, then the formulas in δ are all subformulas of the formulas in the endsequent of δ.*

Proof: You can see this merely by looking at the rules. Each rule except for *Cut* has the subformula property. ∎

Notice how much simpler this proof is than the proof of Theorem 1.11.

A derivation is said to be CUT-FREE if it does not contain an instance of the *Cut* rule. Doing without *Cut* is good for some things, and bad for others. In the system of proof we're studying in this section, sequents have *very many* more proofs with *Cut* than without it.

EXAMPLE 2.5 [DERIVATIONS WITH OR WITHOUT CUT] $p \succ p \lor q$ has only one *Cut*-free derivation, it has infinitely many derivations using *Cut*.

You can see that there is only one *Cut*-free derivation with $p \succ p \lor q$ as the endsequent. The only possible last inference in such a derivation is $[\lor R]$, and the only possible premise for that inference is $p \succ p$. This completes that proof.

On the other hand, there are very many different last inferences in a derivation featuring *Cut*. The most trivial example of this is the following derivation:

$$\dfrac{p \succ p \qquad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{p \succ p \lor q}{\scriptstyle Cut}$$

which contains the *Cut*-free derivation of $p \succ p \lor q$ inside it. We can nest the cuts with the identity sequent $p \succ p$ as deeply as we like.

$$\dfrac{p \succ p \qquad \dfrac{p \succ p \qquad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{p \succ p \lor q}{\scriptstyle Cut}}{p \succ p \lor q}{\scriptstyle Cut} \qquad \dfrac{p \succ p \qquad \dfrac{p \succ p \qquad \dfrac{p \succ p \qquad \dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{p \succ p \lor q}{\scriptstyle Cut}}{p \succ p \lor q}{\scriptstyle Cut}}{p \succ p \lor q}{\scriptstyle Cut} \qquad \cdots$$

However, we can construct an never ending supply of different derivations of our sequent, and we involve different material in the derivation. For absolutely any formula A you wish to choose, we can implicate A (an "innocent bystander") in the derivation of $p \succ p \lor q$ as follows:

$$\dfrac{\dfrac{p \succ p}{p \succ p \lor (q \land A)}{\scriptstyle \lor R_1} \qquad \dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1} \qquad \dfrac{\dfrac{q \succ q}{q \land A \succ q}{\scriptstyle \land L_1}}{q \land A \succ p \lor q}{\scriptstyle \lor R_2}}{p \lor (q \land A) \succ p \lor q}{\scriptstyle \lor L}}{p \succ p \lor q}{\scriptstyle Cut}$$

Well, it's doing *work*, in that $p \vee (q \wedge A)$ is, for many choices for A, genuinely intermediate between p and $p \vee q$. However, A is doing the kind of work that could be done by *any* formula. Choosing different values for A makes no difference to the shape of the derivation. A is doing the kind of work that doesn't require special qualifications.

Of course, there is an important sense in which the *Cut* formula $p \vee (q \wedge A)$ is doing no genuine work in this derivation. It is merely repeating the left formula p or the right formula q.

So, using *Cut* makes the search for derivations rather difficult. There are very many more *possible* derivations of a sequent, and many more actual derivations. The search space is much more constrained if we are looking for *Cut*-free derivations instead. Constructing derivations, on the other hand, is easier if we are permitted to use *Cut*. We have very many more options for constructing a derivation, since we are able to pass through formulas "intermediate" between the desired antecedent and consequent.

Do we *need* to use *Cut*? Is there anything derivable with *Cut* that cannot be derived without? Take a derivation involving *Cut*, like this:

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$p \wedge (q \wedge r) \succ p$}
\AxiomC{$q \succ q$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$q \wedge r \succ q$}
\RightLabel{$\wedge L_2$}
\UnaryInfC{$p \wedge (q \wedge r) \succ q$}
\RightLabel{$\wedge R$}
\BinaryInfC{$p \wedge (q \wedge r) \succ p \wedge q$}
\AxiomC{$q \succ q$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$p \wedge q \succ q$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \wedge q \succ q \vee r$}
\RightLabel{\textit{Cut}}
\BinaryInfC{$p \wedge (q \wedge r) \succ q \vee r$}
\end{prooftree}
$$

The systematic technique I am using will be revealed in detail very soon.

This sequent $p \wedge (q \wedge r) \succ q \vee r$ did not have to be derived using *Cut*. We can *eliminate* the *Cut*-step from the derivation in a systematic way by showing that whenever we use a *Cut* in a derivation we could have either done without it, or used it *earlier*. For example in the last inference here, we did not need to leave the *Cut* until the last step. We could have *Cut* on the sequent $p \wedge q \succ q$ and left the inference to $q \vee r$ until later:

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$p \wedge (q \wedge r) \succ p$}
\AxiomC{$q \succ q$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$q \wedge r \succ q$}
\RightLabel{$\wedge L_2$}
\UnaryInfC{$p \wedge (q \wedge r) \succ q$}
\RightLabel{$\wedge R$}
\BinaryInfC{$p \wedge (q \wedge r) \succ p \wedge q$}
\AxiomC{$q \succ q$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$p \wedge q \succ q$}
\RightLabel{\textit{Cut}}
\BinaryInfC{$p \wedge (q \wedge r) \succ q$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \wedge (q \wedge r) \succ q \vee r$}
\end{prooftree}
$$

The similarity with non-normal proofs as discussed in the previous section is *not* an accident.

Now the *Cut* takes place on the conjunction $p \wedge q$, which is introduced immediately before the application of the *Cut*. Notice that in this case we use the *Cut* to get us to $p \wedge (q \wedge r) \succ$, which is one of the sequents already seen in the derivation! This derivation repeats itself. (Do not be deceived, however. It is not a *general* phenomenon among proofs involving *Cut* that they repeat themselves. The original proof did not repeat any sequents except for the axiom $q \succ q$.)

No, the interesting feature of this new proof is that before the *Cut*, the *Cut* formula is introduced on the right in the derivation of left sequent $p \wedge (q \wedge r) \succ p \wedge q$, and it is introduced on the left in the derivation of the right sequent $p \wedge q \succ q$.

Notice that in general, if we have a *Cut* applied to a conjunction which is introduced on both sides of the step, we have a shorter route to $L \succ R$. We can sidestep the move through $A \land B$ to *Cut* on the formula A, since we have $L \succ A$ and $A \succ R$.

$$\dfrac{\dfrac{L \succ A \quad L \succ B}{L \succ A \land B}{\scriptstyle \land R} \quad \dfrac{A \succ R}{A \land B \succ R}{\scriptstyle \land L_1}}{L \succ R}{\scriptstyle Cut}$$

In our example we do the same: We *Cut* with $p \land (q \land r) \succ q$ on the left and $q \succ q$ on the right, to get the first proof below in which the *Cut* moves *further* up the derivation. Clearly, however, this *Cut* is redundant, as cutting on an identity sequent does nothing. We could eliminate that step, without cost.

$$\dfrac{\dfrac{\dfrac{\dfrac{q \succ q}{q \land r \succ q}{\scriptstyle \land L_1}}{p \land (q \land r) \succ q}{\scriptstyle \land L_2} \quad q \succ q}{p \land (q \land r) \succ q}{\scriptstyle Cut}}{p \land (q \land r) \succ q \lor r}{\scriptstyle \lor R_1} \qquad \dfrac{\dfrac{\dfrac{q \succ q}{q \land r \succ q}{\scriptstyle \land L_1}}{p \land (q \land r) \succ q}{\scriptstyle \land L_2}}{p \land (q \land r) \succ q \lor r}{\scriptstyle \lor R_1}$$

We have a *Cut*-free derivation of our concluding sequent.

As I hinted before, this technique is a general one. We may use exactly the same method to convert *any* derivation using *Cut* into a derivation without it. To do this, we will make explicit a number of the concepts we saw in this example.

DEFINITION 2.6 [ACTIVE AND PASSIVE FORMULAS] The formulas L and R in each inference in Figure 21 are said to be *passive* in the inference (they "do nothing" in the step from top to bottom), while the other formulas are *active*.

A formula is active in a step in a derivation if that formula is either introduced or eliminated. The active formulas in the connective rules are the *principal* formula (the conjunction or disjunction introduced, below the line) or the *constituents* from which the principal formula is constructed. The active formulas in a *Cut* step are the two instances of the *Cut*-formula, present above the line, but absent below the line.

DEFINITION 2.7 [DEPTH OF AN INFERENCE] The DEPTH of an inference in a derivation δ is the number of nodes in the sub-derivation of δ in which that inference is the last step, minus one. In other words, it is the number of sequents above the conclusion of that inference.

Now we can proceed to present the technique for eliminating *Cut*s from a derivation. First we show that *Cut*s may be moved upward. Then we show that this process will terminate in a *Cut*-free derivation. This first lemma is the bulk of the procedure for eliminating *Cut*s from derivations.

LEMMA 2.8 [CUT-DEPTH REDUCTION] *Given a derivation δ of $A \succ C$, whose final inference is Cut, but which is otherwise Cut-free, and in which that inference has a depth of n, we can transform δ another derivation δ' of $A \succ C$ which is Cut-free, or in which each Cut step has a depth less than n.*

Proof: Our derivation δ contains two subderivations: δ_l ending in $A \succ B$ and δ_r ending in $B \succ C$. These subderivations are *Cut*-free.

$$\dfrac{\begin{matrix}\vdots\,\delta_l \\ A \succ B\end{matrix} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A \succ C}$$

To find our new derivation, we look at the two instances of the *Cut*-formula B and its roles in the final inference in δ_l and in δ_r. We have the following two cases: either B is passive in one or other of these inferences, or it is not.

CASE 1: THE CUT-FORMULA IS PASSIVE IN EITHER INFERENCE Suppose that the formula B is *passive* in the last inference in δ_l or passive in the last inference in δ_r. For example, if δ_l ends in $[\wedge L_1]$, then we may push the *Cut* above it like this:

The $\wedge L_2$ case is the same, except for the choice of A_2 instead of A_1.

$$\text{BEFORE: } \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l' \\ A_1 \succ B\end{matrix}}{A_1 \wedge A_2 \succ B}{\scriptstyle \wedge L_1} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A_1 \wedge A_2 \succ C}{\scriptstyle Cut} \qquad \text{AFTER: } \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l' \\ A_1 \succ B\end{matrix} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A_1 \succ C}{\scriptstyle Cut}}{A_1 \wedge A_2 \succ C}{\scriptstyle \wedge L_1}$$

The resulting derivation has a *Cut*-depth lower by one. If, on the other hand, δ_l ends in $[\vee L]$, we may push the *Cut* above that $[\vee L]$ step. The result is a derivation in which we have duplicated the *Cut*, but we have reduced the *Cut*-depth more significantly, as the effect of δ_l is split between the two cuts.

$$\text{BEFORE: } \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l^1 \\ A_1 \succ B\end{matrix} \quad \begin{matrix}\vdots\,\delta_l^2 \\ A_2 \succ B\end{matrix}}{A_1 \vee A_2 \succ B}{\scriptstyle \vee L} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A_1 \vee A_2 \succ C}{\scriptstyle Cut} \qquad \text{AFTER: } \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l^1 \\ A_1 \succ B\end{matrix} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A_1 \succ C}{\scriptstyle Cut} \quad \dfrac{\begin{matrix}\vdots\,\delta_l^2 \\ A_2 \succ B\end{matrix} \quad \begin{matrix}\vdots\,\delta_r \\ B \succ C\end{matrix}}{A_2 \succ C}{\scriptstyle Cut}}{A_1 \vee A_2 \succ C}{\scriptstyle \vee L}$$

The other two ways in which the *Cut* formula could be passive are when δ_2 ends in $[\vee R]$ or $[\wedge R]$. The technique for these is identical to the examples we have seen. The *Cut* passes over $[\vee R]$ trivially, and it passes over $[\wedge R]$ by splitting into two cuts. In every instance, the depth is reduced.

CASE 2: THE CUT-FORMULA IS ACTIVE In the remaining case, the *Cut*-formula formula B may be assumed to be active in the last inference in both δ_l and in δ_r, because we have dealt with the case in which it is passive in either inference. What we do now depends on the form of the formula B. In each case, the structure of the formula B determines the final rule in both δ_l and δ_r.

CASE 2A: THE CUT-FORMULA IS ATOMIC If the *Cut*-formula is an atom, then the only inference in which an atomic formula is active in the conclusion is *Id*. In this case, the *Cut* is redundant.

$$\text{BEFORE:}\quad \dfrac{p \succ p \quad p \succ p}{p \succ p}\,{\scriptstyle Cut} \qquad \text{AFTER:}\quad p \succ p$$

CASE 2B: THE CUT-FORMULA IS A CONJUNCTION If the *Cut*-formula is a conjunction $B_1 \land B_2$, then the only inferences in which a conjunction is active in the conclusion are $[\land R]$ and $[\land L]$. Let us suppose that in the inference $[\land L]$, we have inferred the sequent $B_1 \land B_2 \succ C$ from the premise sequent $B_1 \succ C$. In this case, it is clear that we could have *Cut* on B_1 instead of the conjunction $B_1 \land B_2$, and the *Cut* is shallower.

The choice for $[\land L_2]$ instead of $[\land L_1]$ involves choosing B_2 instead of B_1.

$$\text{BEFORE:}\quad \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l^1 \\ A \succ B_1\end{matrix} \quad \begin{matrix}\vdots\,\delta_l^2 \\ A \succ B_2\end{matrix}}{A \succ B_1 \land B_2}\,{\scriptstyle \land R} \quad \dfrac{\begin{matrix}\vdots\,\delta_r' \\ B_1 \succ C\end{matrix}}{B_1 \land B_2 \succ C}\,{\scriptstyle \land L_1}}{A \succ C}\,{\scriptstyle Cut} \qquad \text{AFTER:}\quad \dfrac{\begin{matrix}\vdots\,\delta_l^1 \\ A \succ B_1\end{matrix} \quad \begin{matrix}\vdots\,\delta_r' \\ B_1 \succ C\end{matrix}}{A \succ C}\,{\scriptstyle Cut}$$

CASE 2C: THE CUT-FORMULA IS A DISJUNCTION The case for disjunction is similar. If the *Cut*-formula is a disjunction $B_1 \lor B_2$, then the only inferences in which a conjunction is active in the conclusion are $\lor R$ and $\lor L$. Let's suppose that in $\lor R$ the disjunction $B_1 \lor B_2$ is introduced in an inference from B_1. In this case, it is clear that we could have *Cut* on B_1 instead of the disjunction $B_1 \lor B_2$, with a shallower *Cut*.

$$\text{BEFORE:}\quad \dfrac{\dfrac{\begin{matrix}\vdots\,\delta_l' \\ A \succ B_1\end{matrix}}{A \succ B_1 \lor B_2}\,{\scriptstyle \lor R_1} \quad \dfrac{\begin{matrix}\vdots\,\delta_r^1 \\ B_1 \succ C\end{matrix} \quad \begin{matrix}\vdots\,\delta_r^2 \\ B_2 \succ C\end{matrix}}{B_1 \lor B_2 \succ C}\,{\scriptstyle \lor L}}{A \succ C}\,{\scriptstyle Cut} \qquad \text{AFTER:}\quad \dfrac{\begin{matrix}\vdots\,\delta_l' \\ A \succ B_1\end{matrix} \quad \begin{matrix}\vdots\,\delta_r^1 \\ B_1 \succ C\end{matrix}}{A \succ C}\,{\scriptstyle Cut}$$

In every case, then, we have traded in a derivation for a derivation either without *Cut* or with a shallower cut. ■

The process of reducing *Cut*-depth cannot continue indefinitely, since the starting *Cut*-depth of any derivation is finite. At some point we find a derivation of our sequent $A \succ C$ with a *Cut*-depth of *zero*: We find a derivation of $A \succ C$ without a *Cut*. That is,

THEOREM 2.9 [CUT ELIMINATION] *If a sequent is derivable with Cut, it is derivable without Cut.*

Proof: Given a derivation of a sequent $A \succ C$, take a *Cut* with no *Cut*s above it. This *Cut* has some depth, say n. Use the lemma to find a derivation with lower *Cut*-depth. Continue until there is no *Cut* remaining in this part of the derivation. (The depth of each *Cut* decreases, so this process cannot continue indefinitely.) Keep selecting cuts in the original derivation and eliminate them one-by-one. Since there are only finitely many cuts, this process terminates. The result is a *Cut*-free derivation. ■

This result is extremely powerful, and it has a number of fruitful consequences for our understanding of logical consequence, which we will consider in Section 2.4, but before that, we will extend our result to a richer language, putting together what natural deduction and the sequent calculus.

2.3 | COMPLEX SEQUENTS

Simple sequents are straightforward. They are simple – perhaps they are too simple to be a comprehensive analysis of the logical relationships between judgements involving conjunction and disjunction, let alone other connectives (like conditionals or negation). Staying with conjunction and disjunction for a moment: consider the sequent

$$p \land (q \lor r) \succ (p \land q) \lor (p \land r)$$

Is that sequent *valid*? It is not too hard to show that it has no *Cut*-free derivation.

EXAMPLE 2.10 [DISTRIBUTION IS NOT DERIVABLE] The sequent $p \land (q \lor r) \succ (p \land q) \lor r$ is not derivable.

Proof: Any *Cut*-free derivation of $p \land (q \lor r) \succ (p \land q) \lor r$ must end in either a $\land L$ step or a $\lor R$ step. Consider the two cases:

CASE 1: THE DERIVATION ENDS WITH [$\land$L]: Then we infer our sequent from either $p \succ (p \land q) \lor r$, or from $q \lor r \succ (p \land q) \lor r$. Neither of these are derivable. As you can see, $p \succ (p \land q) \lor r$ is derivable only, using $\lor R$ from either $p \succ p \land q$ or from $p \succ r$. The latter is not derivable (it is not an axiom, and it cannot be inferred from *anywhere*) and the former is derivable only when $p \succ q$ is—and it isn't. Similarly, $q \lor r \succ (p \land q) \lor r$ is derivable only when $q \succ (p \land q) \lor r$ is derivable, and this is only derivable when either $q \succ p \land q$ or when $q \succ r$ are derivable, and as before, neither of *these* are derivable either.

CASE 2: THE DERIVATION ENDS WITH [$\lor$R]: Then we infer our sequent from either $p \land (q \lor r) \succ p \land q$ or from $p \land (q \lor r) \succ r$. By dual reasoning, neither of these sequents are derivable. So, $p \land (q \lor r) \succ (p \land q) \lor r$ has no *Cut*-free derivation, and by Theorem 2.9 it has no derivation at all. ∎

Reflecting on this sequent, though, it seems that on one account of proof involving disjunction and conjunction, it should be derivable. After all, can't we reason like this?

> Suppose $p \land (q \lor r)$. It follows that p, and that $q \lor r$. So we have two cases: Case (1) q holds, so by p we have $p \land q$, and therefore $(p \land q) \lor (p \land r)$. Case (2) r holds, so by p we have $p \land r$, and therefore $(p \land q) \lor (p \land r)$. So in either case we have $(p \land q) \lor (p \land r)$.

This looks like perfectly reasonable reasoning, and it's reasoning that cannot be reflected in our simple sequent system. The reason is that at the point we wish to split into two cases (on the basis of $q \lor r$), we want to *also* use the information that p holds. Our simple sequents have no space for that. If we expand them a little bit, to allow for more than one formula on the left, we could represent the reasoning like this:

Be careful: I have not yet defined the rules for conjunction or disjunction that this derivation is using.

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p \quad q \succ q}{p, q \succ p \land q}}{p, q \succ (p \land q) \lor (p \land r)} \quad \dfrac{\dfrac{p \succ p \quad r \succ r}{p, r \succ p \land r}}{p, r \succ (p \land q) \lor (p \land r)}}{p, q \lor r \succ (p \land q) \lor (p \land r)}}{p \land (q \lor r) \succ (p \land q) \lor (p \land r)}$$

Now, a sequent has the shape

$$X \succ B$$

where X is a collection of formulas, and B is a single formula. The flexibility of sequents like this allows for us to represent more reasoning. The sequent $X \succ B$ can be seen as making the claim that the conclusion B follows from the premises X. Once we allow sequents to have this structure, we should revisit our rules for conjunction and disjunction to see how they should behave in this new setting. Recall the simple sequent system rules in Figure 21 on page 51. These rules tell us the behaviour of identity sequents, the form of the *Cut* rules, and how to introduce a conjunction or a disjunction on the left hand side of a sequent, and on the right. We want to do the same in our new setting, where sequents can have multiple premises. Here is one option, where we take the old rules, and simply replace all parameters on the left (instances of L) with collections, and allow for extra formulas on the left where the rules allow. The result is in Figure 22. This is the simplest and most straightforward

$$p \succ p \; {\scriptstyle Id} \qquad \dfrac{X \succ C \quad Y, C \succ R}{X, Y \succ R}\, {\scriptstyle Cut}$$

$$\dfrac{X, A \succ R}{X, A \land B \succ R}\, {\scriptstyle \land L_1} \qquad \dfrac{X, A \succ R}{X, B \land A \succ R}\, {\scriptstyle \land L_2} \qquad \dfrac{X \succ A \qquad X \succ B}{X \succ A \land B}\, {\scriptstyle \land R}$$

$$\dfrac{X, A \succ R \qquad X, B \succ R}{X, A \lor B \succ R}\, {\scriptstyle \lor L} \qquad \dfrac{X \succ A}{X \succ A \lor B}\, {\scriptstyle \lor R_1} \qquad \dfrac{X \succ A}{X \succ B \lor A}\, {\scriptstyle \lor R_2}$$

Figure 22: LATTICE RULES FOR MULTIPLE PREMISE SEQUENTS

extension of the simple sequent proof system to allow for sequents with multiple premises. Notice that each simple sequent derivation counts as a derivation in this new system, because we have just modified our rules to allow for more cases: now we allow more than one formula on the left of a sequent. As a result, in this new sequent system, identity sequents (of the form $A \succ A$) remain derivable, with the same derivations as before. Theorem 2.3 applies to this system of sequents. The subformula property holds, too, as each rule (other than *Cut*) still has the property that the formulas in the premise sequents of a rule remain (at least, as subformulas) in the concluding sequent of that rule.

The process for the elimination of the *Cut* rule in derivations is, however, more complicated. We cannot simply take the existing proof of the *Cut* elimination theorem and state that it applies here, too, for now, we have more flexibility in our derivations, and more ways that *Cut* can interact with a derivation. We will need to do more work to show how *Cut* can be eliminated from these derivations.

Before we do that, however, let's do a little more to explore the behaviour of these sequents. Notice the intuitive justification of the distribution sequent $p \land (q \lor r) \succ (p \land q) \lor (p \land r)$ on page 61. In that derivation, we derive a sequent $p, q \succ p \land q$ on the way to proving our desired sequent. No doubt, $p, q \succ p \land q$ seems like a plausibly valid sequent (if p and q hold, so does $p \land q$). Can we derive it using the rules of our sequent system, as laid out in Figure 22?

It turns out that we can't, not as they stand. The sequents $p \succ p$ and $q \succ q$ are both instances of the identity axiom. We can derive $p \land q \succ p$ and $p \land q \succ q$ using the $[\land L]$ rules, but there is no way to derive $p, q \land p \land q$. To derive $X \succ p \land q$ using the $[\land R]$ rule we need to first derive $X \succ p$ and $X \succ q$, so to use $[\land R]$ to derive $p, q \succ p \land q$, we need to derive $p, q \succ p$ and $p, q \succ q$. But these are not axiomatic sequents. We can prove p from p, but not from p together with q.

This should remind you of the discussion of vacuous discharge in Chapter 1. It turns out that exactly the same phenomenon can be observed here with sequent derivations. We have some more options available to us, concerning how to use multiple premises in our sequents. These rules govern how the *structures* of our sequents behave, and as a result, they are called *structural rules*.

Contrast structural rules with the *connective* rules, which govern the behaviour of judgements featuring particular components, like conjunction, disjunction or the conditional.

» «

How should we derive $p, q \succ p$? The justification seems to not depend on any particular behaviour of q (the q could be *any* proposition at all), and there is nothing special in the behaviour of the sequent $p \succ p$ here, other than the fact that we've already derived *it*. There is a general principle at play. If we have derived a sequent $X \succ R$, then we could have the *weaker* sequent $X, A \succ R$, for any formula A. If X and A hold, so does R, because X holds. (We do not need to appeal to A in the justification of R. It stands unused.) In fact, the conclusion of this move of weakening is not general enough. The weakened-in item here need not be another formula. It could well be a collection of formulas all of its own. So, we have the general form of the structural rule of weakening:

$$\dfrac{X \succ A}{X, Y \succ A}\,K$$

The label, *K*, comes from Schönfinkel's *Combinatory Logic* [46, 47, 215]. Now, we can go much further in our derivation of the distribution sequent:

You can remember it like this: **K** for wea**K**ening, but Schönfinkel took K to stand for *Konstanzfunktion*.

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p, q \succ p}\,K \quad \dfrac{q \succ q}{p, q \succ q}\,K}{p, q \succ p \land q}\,\land R}{p, q \succ (p \land q) \lor (p \land r)}\,\lor R \quad \dfrac{\dfrac{\dfrac{p \succ p}{p, r \succ p}\,K \quad \dfrac{r \succ r}{p, r \succ p}\,K}{p, r \succ p \land r}\,\land R}{p, r \succ (p \land q) \lor (p \land r)}\,\lor R}{p, q \lor r \succ (p \land q) \lor (p \land r)}\,\lor L$$

However, that is not enough to justify $p \land (q \lor r) \succ (p \land q) \lor (p \land r)$, as our rules stand. We can continue like this

$$\dfrac{\dfrac{p, q \lor r \succ (p \land q) \lor (p \land r)}{p \land (q \lor r), q \lor r \succ (p \land q) \lor (p \land r)}\,\land L}{p \land (q \lor r), p \land (q \lor r) \succ (p \land q) \lor (p \land r)}\,\land L$$

and we are nearly at our desired conclusion. We just need some way to coalesce the two appeals to $p \land (q \lor r)$ into one. We want to appeal to the structural rule of *contraction*.

$$\dfrac{X, Y, Y \succ R}{X, Y \succ R}\,W$$

The '**W**' isn't due to Schönfinkel—this comes from Haskell Curry [46]. No, I will not attempt to justify the choice of letter by making some reference to cont**W**action. Curry's justification was the natural association of the letter with repetition.

The full derivation of distribution, using *Weakening* and *Contraction*, goes as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p, q \succ p}\,K \quad \dfrac{q \succ q}{p, q \succ q}\,K}{p, q \succ p \land q}\,\land R}{p, q \succ (p \land q) \lor (p \land r)}\,\lor R \quad \dfrac{\dfrac{\dfrac{p \succ p}{p, r \succ p}\,K \quad \dfrac{r \succ r}{p, r \succ p}\,K}{p, r \succ p \land r}\,\land R}{p, r \succ (p \land q) \lor (p \land r)}\,\lor R}{p, q \lor r \succ (p \land q) \lor (p \land r)}\,\lor L}{p \land (q \lor r), q \lor r \succ (p \land q) \lor (p \land r)}\,\land L}{\dfrac{p \land (q \lor r), p \land (q \lor r) \succ (p \land q) \lor (p \land r)}{p \land (q \lor r) \succ (p \land q) \lor (p \land r)}\,W}\,\land L$$

There are other possible structural rules you can explore [193], especially if you treat sequents as composed of lists or other structured collections of formulas. We will not explore those structural rules here. For us, the left hand side of a sequent will consist of a multiset of formulas: the sequent $A, B \succ C$ is literally the *same* sequent as $B, A \succ C$. Contraction and weakening will be enough structural rules for us to consider, just as vacuous discharge and duplicate discharge were enough for us to consider when it came to natural deduction.

If the left hand side of a sequent is a list of formulas, then $p, q \succ r$ is different from $q, p \succ r$.

» «

The parallels between structural rules and discharge policies are no coincidence. Given sequents that allow for more than one formula on the left, it is straightforward to give sequent rules for the conditional. Here is a natural [$\to R$] rule:

This can be motivated directly in terms of natural deduction proofs if you like. Given a proof from premises X together with A to B, if we discharge the A, we have constructed a proof from X to $A \to B$.

$$\frac{X, A \succ B}{X \succ A \to B}\;{\scriptstyle \to R}$$

What would a matching [$\to L$] rule look like? One constraint is that we would like to be able to derive $A \to B \succ A \to B$ from the sequents $A \succ A$ and $B \succ B$. So, we need to be able to derive $A \to B, A \succ B$ so we can go on like this:

$$\frac{A \to B, A \succ B}{A \to B \succ A \to B}\;{\scriptstyle \to R}$$

That much is sure. But think about this more generally: when should we be able to derive $A \to B, Z \succ R$ for arbitrary Z and R? Well, if we could use some of the formulas in Z to derive A, and we could use the other formulas in Z, together with B to derive R, then we'd have enough to conclude R on the basis of Z and $A \to B$. Splitting the Z up into two parts, we get this rule:

$$\frac{X \succ A \qquad B, Y \succ R}{X, A \to B, Y \succ R}\;{\scriptstyle \to R}$$

This rule suffices to derive $A \to B, A \succ B$, since we have $A \succ A$ and $B \succ B$ (here X is A, Y is nothing at all, and R is B). Now we can see the connection between weakening and vacuous discharge. Here is a derivation of $p \succ q \to p$, paired with a natural deduction proof using vacuous discharge:

$$\frac{\dfrac{p \succ p}{p, q \succ p}\;{\scriptstyle K}}{p \succ q \to p}\;{\scriptstyle \to R} \qquad\qquad \frac{p}{q \to p}\;{\scriptstyle \to I, 1}$$

Similarly, duplicate discharge fits naturally with contraction.

$$\frac{\dfrac{\dfrac{p \succ p \qquad \dfrac{p \succ p \quad q \succ q}{p \to q, p \succ q}\;{\scriptstyle \to L}}{p \to (p \to q), p, p \succ q}\;{\scriptstyle \to L}}{p \to (p \to q), p \succ q}\;{\scriptstyle W}}{p \to (p \to q) \succ p \to q}\;{\scriptstyle \to R} \qquad \frac{\dfrac{\dfrac{p \to (p \to q) \qquad [p]^{(1)}}{p \to q}\;{\scriptstyle \to E} \qquad [p]^{(1)}}{q}\;{\scriptstyle \to E}}{p \to q}\;{\scriptstyle \to I, 1}$$

The [$\to R$] rule has an interesting feature, not shared by any of the other rules in the system so far. The number of formulas on the left in the premise sequent $X, A \succ B$ is lowered by one in the conclusion sequent $X \succ A \to B$. This makes sense even if X is empty. Sequents, then, can have *empty* left hand sides. This corresponds to the move in a natural deduction proof when all premises have been discharged.

$$\dfrac{\dfrac{\dfrac{p \succ p \quad q \succ q}{p \to q, p \succ q}{\scriptstyle \to L}}{p \succ (p \to q) \to q}{\scriptstyle \to R}}{\succ p \to ((p \to q) \to q)}{\scriptstyle \to R} \qquad \dfrac{\dfrac{\dfrac{[p \to q]^{(1)} \quad [p]^{(2)}}{q}{\scriptstyle \to E}}{(p \to q) \to q}{\scriptstyle \to I,1}}{p \to ((p \to q) \to q)}{\scriptstyle \to I,2}$$

If the left hand side of a sequent can be empty, what about its right? If we generalise $A \succ B$ to $\succ B$, thinking of this sequent as signifying what can be derived with no active assumptions, then how should we think of $A \succ$, with an empty right? We know that if $\succ B$ is derivable, and $B \succ B'$ is derivable, then by *Cut*, we have $\succ B'$. Consequences of tautologies are tautologies. Applying this reasoning with *Cut* to empty conclusions, if $A \succ$ is derivable and $A' \succ A$ is also derivable, then so is $A' \succ$. The preservation goes the other way. The natural analogue is that if $A \succ$ is derivable, then A is a *contradiction*. Those things that entail contradictions are also contradictions. If the empty left hand side signifies truth on the basis of derivation alone, the empty right signifies falsity on the basis of derivation alone. This insight, and this slight expansion of the notion of a sequent gives us the scope to define *negation*. The natural rule for negation on the right is this:

$$\dfrac{X, A \succ}{X \succ \neg A}{\scriptstyle \neg R}$$

If X and A together are contradictory, then X suffices for $\neg A$, the negation of A. What is an appropriate rule for negation on the left of a sequent? The following rule:

$$\dfrac{X \succ A}{X, \neg A \succ}{\scriptstyle \neg L}$$

allows for the derivation of the identity sequent $\neg A \succ \neg A$, and seems well motivated.

Later in this section we'll see that it also allows for the elimination of *Cut*.

$$\dfrac{\dfrac{A \succ A}{\neg A, A \succ}{\scriptstyle \neg L}}{\neg A \succ \neg A}{\scriptstyle \neg R}$$

These rules, then, give us a logical system with a rich repertoire of propositional connectives: $\wedge, \vee, \to, \neg$. Here are some derivations showing some of the interactions between these connective rules. The first is a derivation of a principle connecting conditionals and conjunctions:

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p \quad q \succ q}{p \to q, p \succ q}{\scriptstyle \to L}}{(p \to q) \wedge (p \to r), p \succ q}{\scriptstyle \wedge L_1} \quad \dfrac{\dfrac{p \succ p \quad r \succ r}{p \to r, p \succ r}{\scriptstyle \to L}}{(p \to q) \wedge (p \to r), p \succ r}{\scriptstyle \wedge L_2}}{(p \to q) \wedge (p \to r), p \succ q \wedge r}{\scriptstyle \wedge R}}{(p \to q) \wedge (p \to r) \succ p \to (q \wedge r)}{\scriptstyle \to R}$$

Notice that this derivation does not use either contraction or weakening. Neither does this next derivation, connecting conditionals, conjunction and disjunction:

$$\dfrac{\dfrac{\dfrac{p \succ p \quad r \succ r}{p \to r, p \succ r}{\scriptstyle \to L}}{(p \to r) \land (q \to r), p \succ r}{\scriptstyle \land L_1} \qquad \dfrac{\dfrac{q \succ q \quad r \succ r}{q \to r, q \succ r}{\scriptstyle \to L}}{(p \to r) \land (q \to r), q \succ r}{\scriptstyle \land L_2}}{\dfrac{(p \to r) \land (q \to r), p \lor q \succ r}{(p \to r) \land (q \to r) \succ (p \lor q) \to r}{\scriptstyle \to R}}{\scriptstyle \lor L}$$

This next pair of derivations shows that two de Morgan laws hold in this sequent system—again, without making appeal to any structural rules.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{\neg(p \lor q), p \succ}{\scriptstyle \neg L}}{\neg(p \lor q) \succ \neg p}{\scriptstyle \neg R} \qquad \dfrac{\dfrac{\dfrac{q \succ q}{q \succ p \lor q}{\scriptstyle \lor R_2}}{\neg(p \lor q), q \succ}{\scriptstyle \neg L}}{\neg(p \lor q) \succ \neg q}{\scriptstyle \neg R}}{\neg(p \lor q) \succ \neg p \land \neg q}{\scriptstyle \land R}}{\succ \neg(p \lor q) \to (\neg p \land \neg q)}{\scriptstyle \to R}
\qquad
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\neg p, p \succ}{\scriptstyle \neg L}}{\neg p \land \neg q, p \succ}{\scriptstyle \land L_1} \qquad \dfrac{\dfrac{q \succ q}{\neg q, q \succ}{\scriptstyle \neg L}}{\neg p \land \neg q, q \succ}{\scriptstyle \land L_2}}{\neg p \land \neg q, p \lor q \succ}{\scriptstyle \lor L}}{\neg p \land \neg q \succ \neg(p \lor q)}{\scriptstyle \neg R}}{\succ (\neg p \land \neg q) \to \neg(p \lor q)}{\scriptstyle \to R}$$

For the other de Morgan laws, connecting the $\neg p \lor \neg q$ and $\neg(p \land q)$, one direction is derivable:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\neg p, p \succ}{\scriptstyle \neg L}}{\neg p, p \land q \succ}{\scriptstyle \land L_1} \qquad \dfrac{\dfrac{q \succ q}{\neg q, q \succ}{\scriptstyle \neg L}}{\neg q, p \land q \succ}{\scriptstyle \land L_2}}{\neg p \lor \neg q, p \land q \succ}{\scriptstyle \lor L}}{\neg p \lor \neg q \succ \neg(p \land q)}{\scriptstyle \neg R}}{\succ (\neg p \lor \neg q) \to \neg(p \land q)}{\scriptstyle \to R}$$

The other direction, however, cannot be derived. There is no way to derive $\neg(p \land q) \succ \neg p \lor \neg q$. (The rule [$\neg L$] cannot be applied, since the right hand side of the sequent is not empty, and there is no way to apply [$\lor R$], since neither disjunct $\neg p$ or $\neg q$ can be derived from $\neg(p \land q)$.)

The logic of negation in this sequent system is *not* classical two valued – or Boolean – logic. While we can derive $p \succ \neg\neg p$:

$$\dfrac{\dfrac{p \succ p}{p, \neg p \succ}{\scriptstyle \neg L}}{p \succ \neg\neg p}{\scriptstyle \neg R}$$

there is no way to derive the converse, $\neg\neg p \succ p$. Again, we cannot apply [$\neg L$] to derive $\neg\neg p \succ p$ since the left hand side is not empty. For the same sort of reason, we cannot derive $\succ p \lor \neg p$.

If we had some way to move the $\neg p$ over to the right hand side while keeping the p on the right hand side as well, we could derive $\neg\neg p \succ p$, as we will see soon.

We have seen the behaviour of the conditional varies with respect to the presence or absence of structural rules, in a way that parallels the use of different discharge policies in natural deduction. The influence of these structural rules extends beyond the behaviour of the conditional, to the other connectives. We have seen this already with the distribution of $\wedge$ over $\vee$: $p \wedge (q \vee r) \succ (p \wedge q) \vee (p \wedge r)$ can be derived using contraction and weakening, as seen on page 63.

Here are two more negation principles derivable using contraction, are not derivable without it.

$$\dfrac{\dfrac{\dfrac{p \succ p \qquad \dfrac{p \succ p}{\neg p, p \succ}{\scriptstyle \neg L}}{p \to \neg p, p, p \succ}{\scriptstyle \to L}}{p \to \neg p, p \succ}{\scriptstyle W}}{p \to \neg p \succ \neg p}{\scriptstyle \neg R} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p, \neg p \succ}{\scriptstyle \neg L}}{p, p \wedge \neg p \succ}{\scriptstyle \wedge L_2}}{p \wedge \neg p, p \wedge \neg p \succ}{\scriptstyle \wedge L_1}}{p \wedge \neg p \succ}{\scriptstyle W}}{\succ \neg(p \wedge \neg p)}{\scriptstyle \neg R}$$

The principle to the effect that the falsity of a conditional entails the falsity of its consequent can be derived using weakening, as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{q, p \succ p}{\scriptstyle K}}{q \succ p \to q}{\scriptstyle \to R}}{\neg(p \to q), q \succ}{\scriptstyle \neg L}}{\neg(p \to q) \succ \neg q}{\scriptstyle \neg R}$$

The shift to sequents with a collection of formulas on the left hand side has brought the question of structural rules to the fore. Do repeated formulas matter? This is the question of contraction. Does the addition of an extra formula break validity? This is the question of weakening.

With the introduction of the possibility of empty right hand sides in sequents, the question of weakening arises for the right hand side, too. If we have a derivation of $X \succ$ (that is, if the premises X are inconsistent), can move to the sequent $X \succ A$, adding the conclusion A where there was none before? This is the structural rule of weakening on the *right*. This means that we now have three different structural rules.

$$\dfrac{X, Y, Y \succ R}{X, Y \succ R}{\scriptstyle W} \qquad \dfrac{X \succ R}{X, Y \succ R}{\scriptstyle KL} \qquad \dfrac{X \succ}{X \succ R}{\scriptstyle KR}$$

With weakening on the right, contradictions entail arbitrary conclusions:

$$\dfrac{\dfrac{p \succ p}{p, \neg p \succ}{\scriptstyle \neg L}}{p, \neg p \succ q}{\scriptstyle KR}$$

We'll *prove* this in the next chapter, when we will see how sequents and derivations can be related to models. We'll show that the derivable sequents in this system match exactly those that hold in every Kripke model for intuitionistic logic.

The proof system with the full complement of structural rules is a sequent system for intuitionistic propositional logic [48, 103]. If we drop the rule of weakening on the right, we get *Minimal logic*, if we also drop the rule of weakening on the left, we get *Intuitionistic Relevant logic* (without distribution), and the system with out contraction or weakening is *Intuitionistic Linear logic*.

Or, *nearly*. We haven't quite modelled the whole of intuitionistic linear logic. We have missed out one connective, known in the linear logic and relevant logic traditions under various guises, such as *multiplicative conjunction* or *fusion*. Here is another way we could give left and right rules for a conjunction-like connective:

$$\dfrac{X, A, B \succ R}{X, A \otimes B \succ R}\,{\otimes L} \qquad \dfrac{X \succ A \quad Y \succ B}{X, Y \succ A \otimes B}\,{\otimes R}$$

Here, the connective $\otimes$ corresponds directly to the comma of premise combination in sequents. In the left rule, this is immediate. We trade in a comma on the left for a fusion. To derive the fusion of two formulas A and B, we derive A from X, and B from Y and combine the premises X and Y with another comma.

We can see immediately that the rules allow us to derive identity sequents for $\otimes$:

$$\dfrac{\dfrac{A \succ A \quad B \succ B}{A, B \succ A \otimes B}\,{\otimes R}}{A \otimes B \succ A \otimes B}\,{\otimes L}$$

and $\otimes$ is connected intimately with the conditional. The sequent $A \otimes B \succ C$ holds if and only if $A \succ B \to C$, as we can see using this pair of derivations:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\\ A \succ B \to C\end{array} \quad \dfrac{B \succ B \quad C \succ C}{B \to C, B \succ C}\,{\to L}}{A, B \succ C}\,{\mathit{Cut}}}{A \otimes B \succ C}\,{\otimes L} \qquad \dfrac{\dfrac{\dfrac{A \succ A \quad B \succ B}{A, B \succ A \otimes B}\,{\otimes R} \quad \begin{array}{c}\vdots\\ A \otimes B \succ C\end{array}}{A, B \succ C}\,{\mathit{Cut}}}{A \succ B \to C}\,{\to R}$$

The connective $\otimes$ becomes salient in relevant and linear logic because in the context of intuitionistic or minimal logic – that is, in the presence of contraction and weakening (on the left) – $\otimes$ is another way of expressing the standard conjunction $\wedge$.

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p \quad q \succ q}{p, q \succ p \otimes q}\,{\otimes R}}{p, p \wedge q \succ p \otimes q}\,{\wedge L_2}}{p \wedge q, p \wedge q \succ p \otimes q}\,{\wedge L_1}}{p \wedge q \succ p \otimes q}\,{W} \qquad \dfrac{\dfrac{\dfrac{p \succ p}{p, q \succ p}\,{KL} \quad \dfrac{q \succ q}{p, q \succ q}\,{KL}}{p, q \succ p \wedge q}\,{\wedge R}}{p \otimes q \succ p \wedge q}\,{\otimes L}$$

In the presence of contraction, the lattice conjunction $\wedge$ is *at least as strong as* fusion $\otimes$. In the presence of weakening, fusion is *at least as* *strong as* lattice conjunction. So, in the presence of both contraction and weakening, the two coincide in strength, and it is straightforward to use the one to do the work of the other.

To round out the family of different possible connective rules, it will be worth our while to turn to another kind of concept that can be defined by way of left and right sequent rules: the propositional *constant*. One motivation for introducing propositional constants into our language is the connection between negation and the conditional. If you attend to the negation and the conditional rules, you will see some similarities between them. Consider the right rules, to start:

$$\frac{X, A \succ B}{X \succ A \to B}{\scriptstyle \to R} \qquad \frac{X, A \succ}{X \succ \neg A}{\scriptstyle \neg R}$$

Negation looks like the special case of a conditional where instead of a consequent, we have the empty right hand side of the sequent. The parallel is preserved when we consider the left rules.

$$\frac{X \succ A \quad B, Y \succ R}{X, A \to B, Y \succ R}{\scriptstyle \to L} \qquad \frac{X \succ A}{X, \neg A \succ}{\scriptstyle \neg L}$$

The negation rule is what one would get from the conditional rule where the consequent formula B somehow stands in for the empty right hand side. (Make Y and R empty, and let B be a formula that does the same job as the empty right hand side of the sequent, and the result is the [$\neg L$] rule. So, if we had a formula – call it f – that were governed by this pair of rules

$$f \succ \;{\scriptstyle fL} \qquad \frac{X \succ}{X \succ f}{\scriptstyle fR}$$

then it turns out that $\neg p$ is equivalent to $p \to f$.

$$\cfrac{\cfrac{p \succ p \quad f \succ}{p \to f, p \succ}{\scriptstyle \to L}}{p \to f \succ \neg p}{\scriptstyle \neg R} \qquad \cfrac{\cfrac{\cfrac{p \succ p}{\neg p, p \succ}{\scriptstyle \neg L}}{\neg p, p \succ f}{\scriptstyle fR}}{\neg p \succ p \to f}{\scriptstyle \to R}$$

There is a sense in which the formula f is *false*—hence the choice of letter. It is a propositional constant, whose interpretation is fixed by its definition, unlike the other atoms, p, q, etc.

Just as we can have a formula corresponding to the behaviour of the empty right hand side of a sequent, we can have a formula which corresponds to the empty left hand side: t

$$\frac{X \succ R}{X, t \succ R}{\scriptstyle tL} \qquad \succ t \;{\scriptstyle tR}$$

Notice that for both f and for t, the identity sequent is derivable from the left and right rules, as for the other connectives, though in this case these do not 'connect' or combine other propositions.

$$\frac{f \succ}{f \succ f}{\scriptstyle fR} \qquad \frac{\succ t}{t \succ t}{\scriptstyle tL}$$

In the presence of weakening on the right [KR], we can derive $f \succ R$ from the axiom $f \succ$ – given [KR] f entails all propositions. In the presence of weakening on the left [KL], we can derive $X \succ t$ from the axiom $\succ t$ – given [KL] t is entailed by anything and everything. These facts does not hold in the absence of weakening. The role of the empty left hand side and the empty right hand side is separable from the role of the strongest and weakest propositions. We can introduce rules for *those* concepts, too, in a straightforward way:

To make sure you understand the difference between t and $\top$ and f and $\bot$, it's a useful exercise to show that t is an identity for $\otimes$, in that $t \otimes p$ is equivalent to p, while $\top$ is an identity for $\wedge$, in that $\top \wedge p$ is equivalent to p. Similarly, $\bot$ is an identity for $\vee$. What about f? Is there a connective for which f is an identity?

$$X, \bot \succ R \;{\scriptstyle \bot L} \qquad X \succ \top \;{\scriptstyle \top R}$$

Here there is no need to have a left rule for $\top$ or a right rule for $\bot$. The identity sequent $\top \succ \top$ is an instance of [$\top R$], and $\bot \succ \bot$ is an instance of [$\bot L$].

$$p \succ p \;{\scriptstyle Id} \qquad \frac{X \succ C \quad Y, C \succ R}{X, Y \succ R}\,{\scriptstyle Cut}$$

$$\frac{X, Y, Y \succ R}{X, Y \succ R}\,{\scriptstyle W} \qquad \frac{X \succ R}{X, Y \succ R}\,{\scriptstyle KL} \qquad \frac{X \succ}{X \succ R}\,{\scriptstyle KR}$$

$$\frac{X, A \succ R}{X, A \wedge B \succ R}\,{\scriptstyle \wedge L_1} \qquad \frac{X, A \succ R}{X, B \wedge A \succ R}\,{\scriptstyle \wedge L_2} \qquad \frac{X \succ A \quad X \succ B}{X \succ A \wedge B}\,{\scriptstyle \wedge R}$$

$$\frac{X, A \succ R \quad X, B \succ R}{X, A \vee B \succ R}\,{\scriptstyle \vee L} \qquad \frac{X \succ A}{X \succ A \vee B}\,{\scriptstyle \vee R_1} \qquad \frac{X \succ A}{X \succ B \vee A}\,{\scriptstyle \vee R_2}$$

$$\frac{X, A, B \succ R}{X, A \otimes B \succ R}\,{\scriptstyle \otimes L} \qquad \frac{X \succ A \quad Y \succ B}{X, Y \succ A \otimes B}\,{\scriptstyle \otimes R}$$

$$\frac{X \succ A \quad B, Y \succ R}{X, A \to B, Y \succ R}\,{\scriptstyle \to L} \qquad \frac{X, A \succ B}{X \succ A \to B}\,{\scriptstyle \to R}$$

$$\frac{X \succ A}{X, \neg A \succ}\,{\scriptstyle \neg L} \qquad \frac{X, A \succ}{X \succ \neg A}\,{\scriptstyle \neg R}$$

$$\frac{X \succ R}{X, t \succ R}\,{\scriptstyle tL} \qquad \succ t \;{\scriptstyle tR} \qquad X \succ \top \;{\scriptstyle \top R} \qquad f \succ \;{\scriptstyle fL} \qquad \frac{X \succ}{X \succ f}\,{\scriptstyle fR} \qquad X, \bot \succ R \;{\scriptstyle \bot L}$$

Figure 23: RULES FOR MULTIPLE PREMISE SEQUENTS

That is a great many connectives and rules! To help you keep stock, Figure 23 contains a summary of all of the sequent rules we have seen so far for each of these connectives. This family of rules (taken as a whole, altogether) defines a proof system for intuitionistic propositional logic. However, the rules give us more than that. They are a toolkit, which can be used to define a number of different proof systems. One axis of variation is the choice of structural rules: you can choose from any of the structural rules, independently of the others (giving eight different possibilities) and for each of the nine logical concepts which have been given rules ($\wedge, \vee, \otimes, \rightarrow, \neg, t, f, \top, \bot$) you have the choice of including that concept or not. That gives us $8 \times 2^9 = 4096$ different systems of rules, of varying degrees of expressive power and logical strength.

Actually, it defines intuitionistic propositional logic with two different conjunction connectives, $\wedge$ and $\otimes$, which are provably equivalent.

Not all of these 4096 systems differ in expressive power, of course. If we have negation at hand, and t, then $\neg t$ can be shown to be logically equivalent to f, so even if we do not include the f rules, we can get their effect using $\neg t$. But beware, you cannot, in general, define t as $\neg f$, and nor can you define $\bot$ as $\neg\top$, in the absence of weakening on the right.

» «

For all of the variety of these 4096 different systems, none of them give us a derivation of the following classically valid sequents:

$$\succ p \vee \neg p \qquad \neg\neg p \succ p \qquad \neg(p \wedge q) \succ \neg p \vee \neg q \qquad (p \rightarrow q) \rightarrow p \succ p$$

One way to strengthen this would be, for example, to allow for an inference rule that warrants the inference from $X \succ \neg\neg A$ to $X \succ A$ – the elimination of a double negation. We could, for example, derive Peirce's Law using this principle, together with contraction and weakening:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\neg p, p \succ}\,{\scriptstyle \neg L}}{\neg p, p \succ q}\,{\scriptstyle KR}}{\neg p \succ p \rightarrow q}\,{\scriptstyle \rightarrow R} \qquad p \succ p}{(p \rightarrow q) \rightarrow p, \neg p \succ p}\,{\scriptstyle \rightarrow L}}{(p \rightarrow q) \rightarrow p, \neg p, \neg p \succ}\,{\scriptstyle \neg L}}{(p \rightarrow q) \rightarrow p, \neg p \succ}\,{\scriptstyle W}}{(p \rightarrow q) \rightarrow p \succ \neg\neg p}\,{\scriptstyle \neg R}}{(p \rightarrow q) \rightarrow p \succ p}\,{\scriptstyle DNE}$$

But there is something quite unsatisfying in this derivation. It passes through negation, when that concept isn't used in the end sequent. There should be some explanation for why $(p \rightarrow q) \rightarrow p$ classically entails p that doesn't appeal to negation, and which uses the rules for the conditional alone. Looking at the behaviour of $\neg p$ in the derivation, you can see that we use it purely to provide some way to apply contraction, duplicating that $\neg p$ (reading from bottom to top), giving us two copies, one for each premise of the [$\rightarrow$*L*] rule we need to apply. There is a sense in which a p on the right of a sequent (or a $\neg\neg p$) is like a $\neg p$ on the left. If we could keep the conclusion p on the right, and allow *it* to be duplicated there, we could avoid the whole charade of disguising the p-on-the-right as a $\neg p$-on-the-left. We could do this:

$$\dfrac{\dfrac{\dfrac{p \succ p}{p \succ q, p}\,{\scriptstyle KR}}{\succ p \to q, p}\,{\scriptstyle \to R} \qquad p \succ p}{\dfrac{(p \to q) \to p \succ p, p}{(p \to q) \to p \succ p}\,{\scriptstyle W}}\,{\scriptstyle \to L}$$

which exposes the central structure of the derivation, and trades in any $\neg p$-on-the-left for a p-on-the-right. Now the connective rules in this derivation involve the connective in the sequent itself: the derivation is much simpler.

But what does a sequent – like $\succ p \to q, p$ – with more than one formula on the right *mean*? The individual formulas on the right are the different *cases* that are being considered. In a general sequent $A, B \succ C, D$, we have that the two cases C, D follow from the premises A, B. In other words, if *all* of A and B are hold, then *some* of C and D do too. So, in our derivation, weakening on the right tells us that $p \succ q, p$ – if p holds, then at least one of q and p hold. Then we discharge the assumption p, and conclude that at least one of $p \to q$ and p hold. And indeed, in classical two-valued logic, this is correct: either $p \to q$ holds (if p is false) or p does (otherwise).

The move from multiple premise sequents $X \succ R$ representing proofs from premises X to a conclusion R to multiple premise and multiple conclusion sequents $X \succ Y$ representing proofs from a range of premises to a range of concluding *cases* was one of Gerhard Gentzen's great insights in proof theory in the 1930s [81, 82]. Sequents with this structure provide an elegant and natural proof system for classical logic.

Here is another derivation, showing how extra positions on the right of a sequent give us exactly the space we need to derive the formerly underivable sequents $\neg\neg p \succ p$ and $\neg(p \wedge q) \succ \neg p \vee \neg q$.

$$\dfrac{\dfrac{p \succ p}{\succ \neg p, p}\,{\scriptstyle \neg R}}{\neg\neg p \succ p}\,{\scriptstyle \neg L} \qquad \dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\succ p, \neg p}\,{\scriptstyle \neg R}}{\succ p, \neg p \vee \neg q}\,{\scriptstyle \vee R_1} \qquad \dfrac{\dfrac{q \succ q}{\succ q, \neg q}\,{\scriptstyle \neg R}}{\succ q, \neg p \vee \neg q}\,{\scriptstyle \vee R_2}}{\succ p \wedge q, \neg p \vee \neg q}\,{\scriptstyle \wedge R}}{\neg(p \wedge q) \succ \neg p \vee \neg q}\,{\scriptstyle \neg L}$$

These derivations use rules for the connectives which generalise the multiple premise rules to allow for multiple formulas in the conclusion of a sequent. Here are the negation rules, in their generality:

$$\dfrac{X \succ A, Y}{X, \neg A \succ Y}\,{\scriptstyle \neg L} \qquad \dfrac{X, A \succ Y}{X \succ \neg A, Y}\,{\scriptstyle \neg R}$$

Here X and Y are completely arbitrary collections of formulas. X may be empty, or many formulas. So may Y. The rules tell us that if I have can derive A (as one of the active cases), then I can ensure that it must be one of the other cases that hold if I add $\neg A$ to the stock of my premises. (Notice that this works equally well in the case of *no* other cases: then if A follows from X, then X and $\neg A$ are inconsistent). That is the $[\neg L]$ rule. The $[\neg R]$ rule tells us that if from X and A I can prove Y (a single formula, or a range of cases, or even no conclusion at all—if the premises are inconsistent), then from X alone, the cases that follow are Y or $\neg A$. This is eminently understandable classical reasoning.

The same can be said for the rules for the other connectives. Now that we have space for more than one formula on the right, we can add an intensional (multiplicative) *disjunction* $\oplus$ – called *fission* – to parallel our intensional conjunction $\otimes$, *fusion*. The rules for fission correspond to the rules for fusion, swapping left and right:

$$\frac{X, A \succ Y \quad X', B \succ Y'}{X, X', A \oplus B \succ Y, Y'}\,{\oplus L} \qquad \frac{X \succ A, B, Y}{X \succ A \oplus B, Y}\,{\oplus R}$$

We introduce $A \oplus B$ on the right by converting the two cases A and B into the one case $A \oplus B$. (A natural way to understand 'or'.) To derive some cases (Y and Y') *from* $A \oplus B$ (with other premises), you derive some cases (say Y) from A (with some of the other premises) and the other cases (say Y') from B (with the remaining premises).

Another variation in the rules for sequents with multiple premises and multiple conclusions is in the structural rules. In sequents with single or no conclusions, we had the option of weakening on the right as well as weakening on the left, and the left structural rule is independent from the right one. In multiple conclusion sequents, this changes. This change is most clear in the presence of negation. If we have weakening on the left, then weakening on the right follows, given negation. After all, if I want to weaken A into the right of the sequent $X \succ Y$, I can weaken $\neg A$ in on the left instead, and convert this to an A on the right, using a *Cut*.

For example, in minimal logic, we have weakening on the left without weakening on the right.

$$\dfrac{\dfrac{A \succ A}{\succ A, \neg A}\,{\neg R} \qquad \dfrac{\begin{matrix}\vdots\\ X \succ Y\end{matrix}}{X, \neg A \succ Y}\,{K}}{X \succ A, Y}\,{Cut}$$

Even without the presence of negation, the need to have weakening on the right if we have weakening on the left is made clear in the process of eliminating *Cut*s from derivations. Consider a derivation in which a sequent $X, A \succ Y$ – where that A has been weakened in – is *Cut* with another sequent: $X' \succ A, Y'$.

$$\dfrac{\begin{matrix}\vdots\\ X' \succ A, Y'\end{matrix} \qquad \dfrac{\begin{matrix}\vdots\\ X \succ Y\end{matrix}}{X, A \succ Y}\,{K}}{X, X' \succ Y, Y'}\,{Cut}$$

We have moved from the sequent $X \succ Y$ to the weaker sequent $X, X' \succ Y, Y'$ where we have weakened on both sides. In sequent systems of this structure, it is hard separate weakening on one side of a sequent from weakening on the other. The same holds for contraction, for exactly the same reasons. For example, in the presence of negation and *Cut*, contraction on the left leads to contraction on the right.

$$\dfrac{\dfrac{A \succ A}{\succ A, \neg A}{\scriptstyle \neg R} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\vdots}{X \succ A, A, Y}}{X, \neg A \succ A, Y}{\scriptstyle \neg R}}{X, \neg A, \neg A \succ Y}{\scriptstyle \neg R}}{X, \neg A \succ Y}{\scriptstyle W}}{X \succ A, Y}{\scriptstyle Cut}$$

For this reason, in these sequents we will not separate structural rules out into left and right versions. Rather, there is a single rule of weakening, that allows weakening in collections of formulas *simultaneously* on the right and left, and similarly for contraction.

This will make commuting a *Cut* over a structural rule much more straightforward. A *Cut* after a weakening converts into a weakening – not a series of weakenings – after a *Cut*, and a *Cut* after a contraction converts into a contraction – not a series of contractions – after a *Cut*.

The full complement of rules for multiple premise / multiple conclusion sequents is found in Figure 24.

» «

DEFINITION 2.11 [SEQUENT SYSTEMS] A sequent system is given by a selection of rules from Figures 23 or 24.

An *intuitionistic* or *multiple premise* sequent system has derivations involving sequents of the form $X \succ R$ where X is a multiset of formulas and R is either a single formula, or empty. The structural rules are *Id*, *Cut*, and any choice from among [*W*], [*KR*] and [*KL*], and the connective rules are the rules for any connective among ($\wedge$, $\vee$, $\otimes$, $\rightarrow$, $\neg$, t, $\top$, f, $\bot$). The sequent system for *intuitionistic logic* has all structural rules and all connectives.

For intuitionist and classical logic, it is typical to leave out some connectives which are definable in terms of the others. For example, in intuitionist logic and in classical logic, f is equivalent to $\neg t$, and $\wedge$ is equivalent to $\otimes$. In classical logic, $\neg$ and $\wedge$ (for example), suffice to define *all* the connectives. However, for the sake of simplicity, and to show no favouritism between connectives, we will treat all of these concepts as independent, and we will allow the choice of any of them as primitive.

A *classical* or *multiple premise, multiple conclusion* sequent system has derivations involving sequents of the form $X \succ Y$ where X and Y are multisets of formulas. The structural rules are *Id*, *Cut*, and any choice from among [*W*], [*K*], and the connective rules are the rules for any connective among ($\wedge$, $\vee$, $\otimes$, $\oplus$, $\rightarrow$, $\neg$, t, $\top$, f, $\bot$). The sequent system for *classical logic* has all structural rules and all connectives.

THEOREM 2.12 [IDENTITY DERIVATIONS] *In any sequent system as defined in Definition 2.11, for any formula A, the identity sequent $A \succ A$ may be derived.*

Proof: For each formula A in the language, we define the identity derivation Id_A inductively, using the clauses in Figure 25. Notice that a derivation involving a concept (a connective or a propositional concept) uses only the sequent structure necessary for the rules involving that concept. So, for example, the identity derivation $Id_{p \wedge (q \vee r)}$ uses simple sequents

$$p \succ p \; Id \qquad \frac{X \succ C, Y \quad X', C \succ Y'}{X, X' \succ Y, Y'} \; Cut$$

$$\frac{X, X', X' \succ Y, Y', Y'}{X, X' \succ Y, Y'} \; W \qquad \frac{X \succ Y}{X, X' \succ Y, Y'} \; K$$

$$\frac{X, A \succ Y}{X, A \land B \succ Y} \land L_1 \qquad \frac{X, A \succ Y}{X, B \land A \succ Y} \land L_2 \qquad \frac{X \succ A, Y \qquad X \succ B, Y}{X \succ A \land B, Y} \land R$$

$$\frac{X, A \succ Y \qquad X, B \succ Y}{X, A \lor B \succ Y} \lor L \qquad \frac{X \succ A, Y}{X \succ A \lor B, Y} \lor R_1 \qquad \frac{X \succ A, Y}{X \succ B \lor A, Y} \lor R_2$$

$$\frac{X, A, B \succ Y}{X, A \otimes B \succ Y} \otimes L \qquad \frac{X \succ A, Y \quad X' \succ B, Y'}{X, X' \succ A \otimes B, Y, Y'} \otimes R$$

$$\frac{X, A \succ Y \quad X', B \succ Y'}{X, X', A \oplus B \succ Y, Y'} \oplus L \qquad \frac{X \succ A, B, Y}{X \succ A \oplus B, Y} \oplus R$$

$$\frac{X \succ A, Y \quad B, X' \succ Y'}{X, A \to B, X' \succ Y, Y'} \to L \qquad \frac{X, A \succ B, Y}{X \succ A \to B, Y} \to R$$

$$\frac{X \succ A, Y}{X, \neg A \succ Y} \neg L \qquad \frac{X, A \succ Y}{X \succ \neg A, Y} \neg R$$

$$\frac{X \succ Y}{X, t \succ Y} \; tL \qquad \succ t \; tR \qquad X \succ \top, Y \; \top R$$

$$f \succ \; fL \qquad \frac{X \succ Y}{X \succ f, Y} \; fR \qquad X, \bot \succ Y \; \bot L$$

Figure 24: rules for multiple conclusion sequents

$$Id_p: \quad p \succ p$$

$$Id_{A\land B}: \quad \cfrac{\cfrac{Id_A}{A \land B \succ A}\,{\scriptstyle \land L_1} \qquad \cfrac{Id_B}{A \land B \succ B}\,{\scriptstyle \land L_2}}{A \land B \succ A \land B}\,{\scriptstyle \land R} \qquad Id_{A\lor B}: \quad \cfrac{\cfrac{Id_A}{A \succ A \lor B}\,{\scriptstyle \lor R_1} \qquad \cfrac{Id_B}{B \succ A \lor B}\,{\scriptstyle \lor R_2}}{A \lor B \succ A \lor B}\,{\scriptstyle \lor L}$$

$$Id_{A\otimes B}: \quad \cfrac{\cfrac{Id_A \qquad Id_B}{A, B \succ A \otimes B}\,{\scriptstyle \otimes R}}{A \otimes B \succ A \otimes B}\,{\scriptstyle \otimes L} \qquad Id_{A\oplus B}: \quad \cfrac{\cfrac{Id_A \qquad Id_B}{A \oplus B \succ A, B}\,{\scriptstyle \oplus L}}{A \oplus B \succ A \oplus B}\,{\scriptstyle \oplus R}$$

$$Id_{A\to B}: \quad \cfrac{\cfrac{Id_A \qquad Id_B}{A \to B, A \succ B}\,{\scriptstyle \to L}}{A \to B \succ A \to B}\,{\scriptstyle \to R} \qquad Id_{\neg A}: \quad \cfrac{\cfrac{Id_A}{A, \neg A \succ}\,{\scriptstyle \neg L}}{\neg A \succ \neg A}\,{\scriptstyle \neg R}$$

$$Id_t: \quad \cfrac{\succ t}{t \succ t}\,{\scriptstyle tL} \qquad Id_\top: \quad \top \succ \top\,{\scriptstyle \top R} \qquad Id_f: \quad \cfrac{f \succ}{f \succ f}\,{\scriptstyle fR} \qquad Id_\bot: \quad \bot \succ \bot\,{\scriptstyle \bot L}$$

Figure 25: IDENTITY DERIVATIONS FOR COMPLEX SEQUENTS

only, and is a derivation in any sequent system with conjunction and disjunction in the vocabulary. The identity derivation for $\oplus$ requires multiple conclusions, so that is not a derivation in multiple premise sequent systems, but the connective $\oplus$ cannot be defined in such systems. The derivations for $\to$, $\neg$ and $\otimes$, require multiple premises, but none of the rest do. The derivations for f and $\neg$ require empty right hand sides, and the derivation for t requires an empty left hand side. ∎

Now, before going on to show that in each of our sequent systems, the *Cut* rule may be eliminated, generalising Theorem 2.9—let us pause and examine some of the behaviour of each of the connectives we have defined.

$$\land \quad \lor \quad \otimes \quad \oplus \quad \to \quad \neg$$

Take $\neg$, to start. It is *order inverting* in each of our logical systems, in the following sense. In any system, if $A_1 \succ A_2$ is derivable, so is $\neg A_2 \succ \neg A_1$. The proof is a slight generalisation of the proof of the identity derivation for negations:

$$\cfrac{\cfrac{A_1 \succ A_2}{A_1, \neg A_2 \succ}\,{\scriptstyle \neg L}}{\neg A_2 \succ \neg A_1}\,{\scriptstyle \neg R}$$

We call this property *order inverting* because as we go from the *stronger* A_1 to the *weaker* A_2, negation flips the order: we go from the *weaker* $\neg A_1$ to the *stronger* $\neg A_2$. Conjunction has two positions (the left conjunct, and the right conjunct), and these positions are not order inverting: they are order *preserving*. If $A_1 \succ A_2$ is derivable, then so are $A_1 \land B \succ A_2 \land B$, and $B \land A_1 \succ B \land A_2$:

$$\dfrac{\dfrac{A_1 \succ A_2}{A_1 \land B \succ A_2}{\scriptstyle \land L_1} \quad \dfrac{B \succ B}{A_1 \land B \succ B}{\scriptstyle \land L_2}}{A_1 \land B \succ A_2 \land B}{\scriptstyle \land R} \qquad \dfrac{\dfrac{B \succ B}{B \land A_1 \succ B}{\scriptstyle \land L_1} \quad \dfrac{A_1 \succ A_2}{B \land A_1 \succ A_2}{\scriptstyle \land L_1}}{B \land A_1 \succ B \land A_2}{\scriptstyle \land R}$$

In the same way, each position in $\lor$, $\otimes$, $\oplus$ is *order preserving*. This leaves the conditional. This is the most interesting connective of our family, for one position in the conditional (the antecedent) is order inverting, and the other (the consequent) is order preserving.

Verifying this is an exercise left to you.

$$\dfrac{\dfrac{A_1 \succ A_2 \quad B \succ B}{A_2 \to B, A_1 \succ B}{\scriptstyle \to L}}{A_2 \to B \succ A_1 \to B}{\scriptstyle \to R} \qquad \dfrac{\dfrac{B \succ B \quad A_1 \succ A_2}{B \to A_1, B \succ A_2}{\scriptstyle \to L}}{B \to A_1 \succ B \to A_2}{\scriptstyle \to R}$$

We also call an order inverting position inside a connective *negative* and an order preserving position inside a connective *positive*. This can generalise when connectives are nested in the following way:

DEFINITION 2.13 [POSITIVE AND NEGATIVE OCCURRENCES OF ATOMS] In any formula A in which the atom p occurs, we classify that occurrence as POSITIVE or NEGATIVE in the following way:

- p occurs positively in the atom p itself.
- p occurs positively in $A \land B$, $A \lor B$, $A \otimes B$, $A \oplus B$, iff it occurs positively in A or in B.
- p occurs negatively in $A \land B$, $A \lor B$, $A \otimes B$, $A \oplus B$, iff it occurs negatively in A or in B.
- p occurs positively in $\neg A$ iff it occurs negatively in A.
- p occurs negatively in $\neg A$ iff it occurs positively in A.
- p occurs positively in $A \to B$ iff it occurs negatively in A or positively in B.
- p occurs negatively in $A \to B$ iff it occurs positively in A or negatively in B.

It is a good exercise to prove the following theorem:

If we had defined a biconditional connective $\leftrightarrow$ where $A \leftrightarrow B$ behaves like $(A \to B) \land (B \to A)$, then would p be positive or negative in $p \leftrightarrow q$?

THEOREM 2.14 *If* $B_1 \succ B_2$ *is derivable and the indicated occurrence* p *is positive in* $A(p)$, *then* $A(B_1) \succ A(B_2)$ *is derivable. If that occurrence is negative, then* $A(B_2) \succ A(B_1)$ *is derivable.*

» «

The last result for this section is generalising Theorem 2.9 for our sequent systems. We want to show that *Cut* is eliminable in each of our sequent systems. The strategy for the proof is exactly the same as the proof for simple sequents, but the details are more involved because there are more positions in our sequents on which rules can operate. The proof will proceed by showing how a *Cut* in a derivation, like this

$$\dfrac{\begin{array}{c}\vdots\,\delta_l\\ X \succ C, Y\end{array} \qquad \begin{array}{c}\vdots\,\delta_r\\ X', C \succ Y'\end{array}}{X, X' \succ Y, Y'}\ \mathit{Cut}$$

can be made *simpler* in some way. Exactly which way depends on how the *Cut*-formula C behaves in the left and right derivations δ_l and δ_r. If C is a *passive* formula in the last inference of either δ_l or δ_r, the strategy is to pass the *Cut* upwards beyond that inference. If C is *active* in the last inference of δ_l and δ_r, we convert the *Cut* on C into *Cuts* on smaller formulas. The added complexity for our complex sequents is the first case, not the second. Showing how *Cut* formulas that are active in both premises of the *Cut* can be reduced is a matter of inspecting the left and right rules for each connective. The complexity arises out of the many more different ways that a *Cut* formula can be *passive* in an inference step.

How to treat the *Cut* depends, as before, on the behaviour of A in δ and δ', but now the structural rules and the more complex sequents give us many more options for ways in which a formula can be passive in a step in a derivation. Consider this example:

$$\dfrac{\begin{array}{c}\vdots\,\delta_l\\ X \succ C\end{array} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_l\\ X', C, C \succ R\end{array}}{X, C \succ R}\ W}{X, X' \succ R}\ \mathit{Cut}$$

If we are to push this *Cut* over the contraction, the result is not a *Cut* on a simpler formula (a *Cut* with a smaller *rank*), nor is it a *Cut* closer to the leaves of the derivation tree (a *Cut* with a smaller *depth*). The result is *two Cuts*, with the same rank as before, and one of which has the same depth as the original *Cut*.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_l\\ X \succ C\end{array} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_l\\ X \succ C\end{array} \qquad \begin{array}{c}\vdots\,\delta_r\\ X', C, C \succ R\end{array}}{X, X' \succ R}\ \mathit{Cut}}{X, X', X' \succ R}\ \mathit{Cut}}{X, X' \succ R}\ W$$

It looks like we have made things worse by pushing the *Cut* upwards. Things are no better (and perhaps worse) in the case of sequents with multiple conclusions. Here, we could have contractions on both premises of a *Cut* step:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_l\\ X \succ C, C, Y\end{array}}{X \succ C, Y}\ W \qquad \dfrac{\begin{array}{c}\vdots\,\delta_r\\ X', C, C \succ Y'\end{array}}{X', C \succ Y'}\ W}{X, X' \succ Y, Y'}\ \mathit{Cut}$$

Now, to push a *Cut* up the derivation, there would be wild proliferation, no matter what order we try to take. The process for pushing *Cut*s up a derivation will not be quite as straightforward as in the case of simple sequents.

There are a number of options. Gentzen's original technique was to generalise. Instead of eliminating *Cut*, we introduce a more *general* inference rule, *Mix*

$$\frac{X \succ C, \ldots, C, Y \quad X', C, \ldots, C \succ Y'}{X, X' \succ Y, Y'}\ \textit{Mix}$$

in which any number of instances of the *Mix*-formula may be removed. Then a *Mix* after a contraction (on either the left or the right) of a sequent can be traded in for a *Mix* processing more formulas before that contraction. That is an insightful idea, but eliminating *Mix* has other complications – such as the complication of processing the *Mix* when one *Mix* formula is active – and we will not follow Gentzen's lead.

Another approach to eliminate *Cut* is to rewrite the rules of the sequent system in order to get rid of *Contraction* completely. We rearrange the rules to allow us to prove "contraction elimination" theorem, to the effect that if the premise sequent of a contraction step is derivable, then so is its conclusion, *without using contraction* [65, 153]. This involves rewriting each of the connective rules in such a way as to embed enough contraction into the rules that the separate rule is no longer required. For example, consider the conditional left rule:

$$\frac{X, A \succ Y \quad X', B \succ Y'}{X, A \to B, X' \succ Y, Y'}\ {\to}L$$

Suppose some formula appears both in X and X', and we wish to contract it into one instance after we make this inference. Instead of doing this in an explicit step of contraction, we could eliminate the requirement by placing everything that occurs either in X or X' in *both* premises, and do the same for Y and Y', for we may need to contract formulas in the right, too. The result would be a variant of our inference rule:

$$\frac{X, A \succ Y \quad X, B \succ Y}{X, A \to B \succ Y}\ {\to}L'$$

Now there is no need to contract formulas that end up duplicated in the concluding sequent by way of appearing in both premise sequents. But now, we have made both premise sequents *fatter*, by stuffing them with whatever appears in the endsequent of the inference, apart from the active formula. This is no problem if weakening is one of the structural rules, and it is in systems with contraction and weakening that this technique works best. Once the rules have been converted, the elimination of *Cut* is simpler, in that it never needs to be commuted over a separate contraction inference.

There are more details to take care of, especially in intuitionist logic. See exercise ?? for the details.

We won't follow the lead of this approach, either, because our goal is an understanding of the process of *Cut* elimination that is relatively uniform between systems in which the structural rule is present and those in which it is absent – without modifying the connective rules. The target is an understanding of *Cut* elimination for any sequent system, as given in Definition 2.11. Tailoring the connective rules to incorporate the structural rules results is a fine approach for particular sequent systems in which those structural rules are present, but it makes the general approach to *Cut* elimination less transparent. The most difficult case for us will be systems in which contraction is a structural rule but weakening is absent. The techniques of internalising contraction in the connective rules, as used in G3 systems, do not apply in this case. We must look elsewhere to understand how to take control of contraction and *Cut*.

[Reference/explanation to be added.]

Another, more recent approach to managing contraction and *Cut* is due to Katalin Bimbó [22, 23]. She has shown that if we keep track of the *contraction count* for an instance of *Cut* in a derivation (the number of applications of contraction above the *Cut*), then we can find a measure which always does decrease as the *Cut* makes its upward journey.

...and in others, later in this book.

The approach we will take to eliminating *Cut* in these systems is originally due to Haskell Curry [45], and has been systematised and generalised by Nuel Belnap in his magisterial work on Display Logic [16], and further generalised by the current author [193]. The crucial idea is to understand the process of *Cut* elimination by keeping track of the *ancestors* of the *Cut* formula in a derivation — in either the left premise of the *Cut* step or the right — which form a *tree* above the *Cut* formula, and we perform the *Cut* instead at each of the *leaves* of that tree. At those steps the *Cut* formula is active in one premise of the *Cut*. Then inspect the passive ancestors of the *Cut* formula in the other premise, and commute each *Cut* up to the leaves of *that* tree. The result is a derivation in which all of these *Cuts* now involve active formulas in both premises. And these can then be eliminated by reducing the *Cuts* to *Cuts* on simpler formulas. Let's illustrate this process with a concrete example. Here is a derivation with a single *Cut*, where the *Cut* formula, $\neg p$, is passive in both inferences leading up to the *Cut*.

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \succ p \vee q$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p \vee q, \boxed{\neg p}$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg(p \vee q) \succ \boxed{\neg p}$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \boxed{\neg p}$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \boxed{\neg p}$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$\neg p \wedge r \succ \boxed{\neg p}$}
\RightLabel{$\oplus L$}
\BinaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \boxed{\neg p}, \boxed{\neg p}$}
\RightLabel{W}
\UnaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \boxed{\neg p}$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\RightLabel{$\neg L$}
\UnaryInfC{$\boxed{\neg p} \succ \neg p$}
\AxiomC{$\boxed{\neg p} \succ \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\boxed{\neg p}, \boxed{\neg p} \succ \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\boxed{\neg p} \succ \neg p \otimes \top$}
\RightLabel{Cut}
\BinaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top$}
\end{prooftree}
$$

This derivation uses a large complement of concepts: $\wedge, \vee, \otimes, \oplus, \neg, \top$. The derivation doesn't use weakening, so fusion and fission differ in strength from lattice conjunction and disjunction.

In the derivation, the coloured boxes show the tree of ancestors of the *Cut* formula. The orange boxes trace the ancestry of the $\neg p$ as used in the left premise of the *Cut*, while the blue boxes, trace its ancestry as used in the right premise.

Notice that the leaves of the orange tree both end in [$\neg R$] steps, while one leaf of the blue tree introduces $\neg p$ in a [$\neg L$] step, and the other introduces $\neg p$ in a $[\top R]$ axiom, where the $\neg p$ is passive. In the process of *Cut* reduction, we shift the *Cut* to the leaves of one of these trees.

Let choose the orange tree to process first. (We could have chosen the blue tree. The result at the end might be different, but it would still result in a *Cut*-free derivation, whichever choice we made.) The result of commuting the *Cut* up the orange tree of passive ancestors is that we *Cut* the indicated sequents at the leaves of the orange tree of ancestors with the other *Cut* premise and we replace the $\neg p$ instances in that tree with the result of making the *Cut*, leaving the rest of the derivation undisturbed. Here is the result:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{\succ p \lor q, \neg p}{\scriptstyle \neg R} \quad \dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R}}{\neg p \succ \neg p}{\scriptstyle \neg L} \quad \neg p \succ \top}{\neg p, \neg p \succ \neg p \otimes \top}{\scriptstyle \otimes R}}{\neg p \succ \neg p \otimes \top}{\scriptstyle W}}{\succ p \lor q, \neg p \otimes \top}{\scriptstyle Cut}}{\neg(p \lor q) \succ \neg p \otimes \top}{\scriptstyle \neg L} \quad \dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R} \quad \dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R}}{\neg p \succ \neg p}{\scriptstyle \neg L} \quad \neg p \succ \top}{\neg p, \neg p \succ \neg p \otimes \top}{\scriptstyle \otimes R}}{\neg p \succ \neg p \otimes \top}{\scriptstyle W}}{\succ p, \neg p \otimes \top}{\scriptstyle Cut}}{\neg p \succ \neg p \otimes \top}{\scriptstyle \neg L}}{\neg p \land r \succ \neg p \otimes \top}{\scriptstyle \land L_1}}{\neg(p \lor q) \oplus (\neg p \land r) \succ \neg p \otimes \top, \neg p \otimes \top}{\scriptstyle \oplus L}}{\neg(p \lor q) \oplus (\neg p \land r) \succ \neg p \otimes \top}{\scriptstyle W}$$

We have a tree deriving the required sequent, and the *Cut* has been driven up to the leaves of the orange tree. This means, in this case, that in the left premise of each *Cut* inference, the *Cut* formula is active. We can do the same, passing each *Cut* up to the leaves of the blue tree of passive instances of the *Cut* formula in consequent position.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{\succ p \lor q, \neg p}{\scriptstyle \neg R} \quad \dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R}}{\neg p \succ \neg p}{\scriptstyle \neg L}}{\succ p \lor q, \neg p}{\scriptstyle Cut} \quad \dfrac{\dfrac{\dfrac{p \succ p}{p \succ p \lor q}{\scriptstyle \lor R_1}}{\succ p \lor q, \neg p}{\scriptstyle \neg R} \quad \neg p \succ \top}{\succ p \lor q, \top}{\scriptstyle Cut}}{\succ p \lor q, p \lor q, \neg p \otimes \top}{\scriptstyle \otimes R}}{\succ p \lor q, \neg p \otimes \top}{\scriptstyle W}}{\neg(p \lor q) \succ \neg p \otimes \top}{\scriptstyle \neg L} \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R} \quad \dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R}}{\neg p \succ \neg p}{\scriptstyle \neg L}}{\succ p, \neg p}{\scriptstyle Cut} \quad \dfrac{\dfrac{p \succ p}{\succ p, \neg p}{\scriptstyle \neg R} \quad \neg p \succ \top}{\succ p, \top}{\scriptstyle Cut}}{\succ p, p, \neg p \otimes \top}{\scriptstyle \otimes R}}{\succ p, \neg p \otimes \top}{\scriptstyle W}}{\neg p \succ \neg p \otimes \top}{\scriptstyle \neg L}}{\neg p \land r \succ \neg p \otimes \top}{\scriptstyle \land L_1}}{\neg(p \lor q) \oplus (\neg p \land r) \succ \neg p \otimes \top, \neg p \otimes \top}{\scriptstyle \oplus L}}{\neg(p \lor q) \oplus (\neg p \land r) \succ \neg p \otimes \top}{\scriptstyle W}$$

Here, the *Cuts* have been pushed up to the tops of both trees of ancestors of the *Cut* formula. Now each *Cut* formula is either either active in its premise, or passive in an axiom.

The *Cut* steps featuring the axiomatic sequent $\neg p \succ \top$ may be immediately simplified. The $\neg p$ is passive in the axiom, and the conclusion of the *Cut* is another instance of the axiom, so we can simplify the derivation by replacing the premise with its conclusion at no cost, like this:

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \succ p \vee q$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p \vee q, \neg p$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \neg p$}
\RightLabel{Cut}
\BinaryInfC{$\succ p \vee q, \neg p$}
\AxiomC{$\succ p \vee q, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p \vee q, p \vee q, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p \vee q, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg(p \vee q) \succ \neg p \otimes \top$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \neg p$}
\RightLabel{Cut}
\BinaryInfC{$\succ p, \neg p$}
\AxiomC{$\succ p, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p, p, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \neg p \otimes \top$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$\neg p \wedge r \succ \neg p \otimes \top$}
\RightLabel{$\oplus L$}
\BinaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top$}
\end{prooftree}
$$

In the remaining *Cuts*, the *Cut* formula is active in both premises. We make this transformation:

$$
\begin{prooftree}
\AxiomC{$\vdots\ \delta_l$}
\noLine
\UnaryInfC{$X, C \succ Y$}
\RightLabel{$\neg R$}
\UnaryInfC{$X \succ \neg C, Y$}
\AxiomC{$\vdots\ \delta_r$}
\noLine
\UnaryInfC{$X' \succ C, Y'$}
\RightLabel{$\neg L$}
\UnaryInfC{$X', \neg C \succ Y'$}
\RightLabel{Cut}
\BinaryInfC{$X, X' \succ Y, Y'$}
\end{prooftree}
\quad \rightsquigarrow \quad
\begin{prooftree}
\AxiomC{$\vdots\ \delta_r$}
\noLine
\UnaryInfC{$X' \succ C, Y'$}
\AxiomC{$\vdots\ \delta_l$}
\noLine
\UnaryInfC{$X, C \succ Y$}
\RightLabel{Cut}
\BinaryInfC{$X, X' \succ Y, Y'$}
\end{prooftree}
$$

to reduce the degree of those *Cuts*, replacing each *Cut* on a negation, $\neg p$, by a *Cut* on p.

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\AxiomC{$p \succ p$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \succ p \vee q$}
\RightLabel{Cut}
\BinaryInfC{$\succ p \vee q, \neg p$}
\AxiomC{$\succ p \vee q, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p \vee q, p \vee q, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p \vee q, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg(p \vee q) \succ \neg p \otimes \top$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\AxiomC{$p \succ p$}
\RightLabel{Cut}
\BinaryInfC{$\succ p, \neg p$}
\AxiomC{$\succ p, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p, p, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \neg p \otimes \top$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$\neg p \wedge r \succ \neg p \otimes \top$}
\RightLabel{$\oplus L$}
\BinaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top$}
\end{prooftree}
$$

In the right *Cut* in the derivation, the right premise is an identity sequent, and a *Cut* on an identity is trivial (its other premise is its conclusion), so that can be deleted. In the *Cut* step on the left of the derivation, the *Cut* formula is passive in both sides. If we commute the *Cut* up in either direction, the result is a *Cut* free derivation of $\succ p \vee q, \neg q$ (the order of the $\vee R$ and $\neg R$ steps in that derivation depend on whether we push the *Cut* up the left branch or the right branch first). Pushing it up the left branch first, the result is this:

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\RightLabel{$\vee R_1$}
\UnaryInfC{$p \succ p \vee q$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p \vee q, \neg p$}
\AxiomC{$\succ p \vee q, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p \vee q, p \vee q, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p \vee q, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg(p \vee q) \succ \neg p \otimes \top$}
\AxiomC{$p \succ p$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ p, \neg p$}
\AxiomC{$\succ p, \top$}
\RightLabel{$\otimes R$}
\BinaryInfC{$\succ p, p, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\succ p, \neg p \otimes \top$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg p \succ \neg p \otimes \top$}
\RightLabel{$\wedge L_1$}
\UnaryInfC{$\neg p \wedge r \succ \neg p \otimes \top$}
\RightLabel{$\oplus L$}
\BinaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top, \neg p \otimes \top$}
\RightLabel{W}
\UnaryInfC{$\neg(p \vee q) \oplus (\neg p \wedge r) \succ \neg p \otimes \top$}
\end{prooftree}
$$

a derivation with no *Cut*s. Notice that the final derivation uses the same connective rules – introducing the same formulas – as were in the original derivation, but in a different order. The same structural rules were also used, but operating on *different* formulas. Had we pushed the original *Cut* up the right branch (to the leaves of the blue tree, the ancestors of the *Cut* formula in the right premise of the *Cut*), the resulting derivation would have been different, but we would also have found a *Cut*-free derivation.

Exercise: eliminate the *Cut* in the original derivation in just this way. What *Cut*-free derivation *do* you get?

This example illustrates four components of the general process of eliminating *Cut* from a derivation.

1. Defining the trees of ancestors of the *Cut* formulas in a derivation.
2. Transferring a *Cut* to the top of the tree of ancestors.
3. Eliminating a *Cut* when one instance of the *Cut* formula is introduced as a *passive* formula in a premise of the *Cut* (for example, in a weakening inference, or a [$\bot L$] or [$\top R$] axiom).
4. Replacing a *Cut* on a complex formula, *active* in both premises of that *Cut* with *Cut*s on subformulas of that formula, from the same premises.

To make these four components of the *Cut* elimination process precise, we need to extend our notion of *active* and *passive* formulas in inference rules in derivations from the case for simple sequents given Definition 2.6 on page 57. The generalisation is natural.

DEFINITION 2.15 [ACTIVE AND PASSIVE FORMULAS] Any formula appearing as a component of the multisets X, Y, R in the rules for multiple premise sequents in Figure 23, or in the multisets X, X′, Y, or Y′ in multiple conclusion sequents in Figure 24 are said to be PASSIVE in that inference step. The other formulas are said to be ACTIVE in those inferences.

The presentation of the rules in Figures 23 and 24 not only allows us to define active and passive formulas in each instance of a rule. It also helps

This idea is, in fact, rather subtle, when sequents are composed of multisets of formulas. Take this [$\otimes R$] inference:

$$\dfrac{p \succ p \quad p \succ \top}{p, p \succ p \otimes \top}\otimes R$$

Which occurrence of p in a premise of the inference is the ancestor of the *first* p in the sequent in the conclusion of the inference? This question, in fact, makes no sense, because the left hand side of the sequent $p, p \succ p \otimes p$ is a multiset and its members do not come in any order. The one formula p is a member of that multiset twice. We can depict the ancestry of formulas with colours like this:

$$\dfrac{\boxed{p} \succ p \quad \boxed{p} \succ \top}{\boxed{p}, \boxed{p} \succ p \otimes \top}\otimes R$$

but this does not pair the *first* member of the multiset $\boxed{p}, \boxed{p}$ with the p in the left hand side of the sequent $p \succ p$, for the multiset does not have a first member. This is exactly the same account of the ancestry of formulas as depicted here:

$$\dfrac{\boxed{p} \succ p \quad \boxed{p} \succ \top}{\boxed{p}, \boxed{p} \succ p \otimes \top}\otimes R$$

The two displayed derivations are two different representations of the one derivation, in which the formulas are written in a different order. There is no fact of the matter concerning whether the first instance of ‘p’ in the sequent $p, p \succ p \otimes \top$ is paired with LHS p in $p \succ p$ or the p in $p \succ \top$, because there is so such thing as *the first instance of ‘p’ in* $p, p \succ p \otimes \top$. The situation is fully described when we say that one p in the LHS of $p, p \succ p \otimes \top$ is matched with one premise, and the other p, with the other premise.

In later chapters we will revisit these notions when considering proof systems with different sequent structures. They require a new analysis.

us define the ancestors of a formula in a derivation, though we must be careful with the details.

DEFINITION 2.16 [PARENTS, ORPHANS, ANCESTRY] When D is a formula occurring passively in a premise sequent of an inference falling under some rule, it is a PARENT of a single occurrence of the same formula D in the concluding sequent of that rule, where both occurrences fall under the same multiset (or formula) variable (X, Y, R, etc.) in the schematic statement of the rule. For example, in any instance of the rule $\otimes R$

$$\dfrac{X \succ A, Y \quad X' \succ B, Y'}{X, X' \succ A \otimes B, Y, Y'}\otimes R$$

A formula D occurring inside the multiset Y in the right hand side of the premise $X \succ A, Y$ of the rule is a parent of one occurrence of the formula D occurring in the right hand side of the conclusion $X, X' \succ A \otimes B, Y, Y'$. There may be *other* instances of D occurring in that right hand side of the sequent, but they do not have our particular premise D as a parent. If that extra D in the conclusion is also a member of Y, then its parent is another D occurring in the Y in the premise. If that D is in Y', its parent is in the other premise sequent. If that D is, on the other hand, the formula $A \otimes B$, then it is an active formula in this inference, and it has no parent in the premises of the inference. Passive formulas in the conclusions of most of our inference rules have single parents, but rules in which side formulas are repeated in the premises (such as *Contraction*, $\wedge R$, and $\vee L$) bring us dual parentage. In an instance of *Contraction*:

$$\dfrac{X, X', X' \succ Y, Y', Y'}{X, X' \succ Y, Y'}W$$

a formula present in X' or in Y' in the conclusion has two parents in the premise of the rule, while formulas in X and Y in the conclusion have a single parent in the premise.

In some inferences (and in all axioms) formulas occurring passively do not have parents. The formulas in X' or Y' in a weakening inference

$$\dfrac{X \succ Y}{X, X' \succ Y, Y'}K$$

and the passive formulas in [$\bot L$] and [$\top R$] axioms have no parents. They are said to be ORPHANS.

The ANCESTRY of a formula occurring passively in an inference rule of some derivation is the following collection of occurrences of the same formula in that derivation: Its ancestry is *empty* if that instance is an orphan. If it is not an orphan, its ancestry is the collection consisting of its *parents*, together with the *ancestries* of those parents.

That completes the definitions of active and passive formulas, of parents, of orphans, and of ancestry. Given any proof system, a definition of these notions for that proof system will be called an *analysis*.

LEMMA 2.17 [ANCESTRY PRESERVES POSITION IN SEQUENTS] *The ancestors of a formula occurrence* A *in a derivation are on the same side of their sequents as the original occurrence* A.

Proof: If you inspect the inference rules in Figures 23 and 24, you see that the multisets X, X', Y, Y', and the formula R never swaps sides from premise to conclusion of an inference. So a parent–child relationship is always between formulas on the same side of a sequent. ∎

The process of *Cut* elimination involves pushing a *Cut* up the tree of ancestors of the *Cut* formula, to the leaves. These orphans are either *active* in the inferences leading up to the *Cut* or *passive*. In either case, the *Cut* can be eliminated entirely, or converted into *Cut*s on subformulas of the *Cut* formula.

Let's turn next to the results which allow us to transfer a *Cut* in a derivation up to the orphans in the ancestry of the *Cut* formula. To make the process precise, consider a single *Cut* step:

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \succ C, Y\end{array}\qquad \begin{array}{c}\vdots\,\delta_2\\ X', C \succ Y'\end{array}}{X, X' \succ Y, Y'}\,{\scriptstyle Cut}$$

If we are, for example, pushing the *Cut* up the tree of ancestors of C in δ_2, then the result of the process involve a *Cut* with δ_1 for each orphan at the leaves of the tree of ancestors, and the remaining Cs in the ancestry in δ_2 will be replaced by the result of the *Cut*. In general, this means that each non-orphan-C (in that ancestry in δ_2) will be replaced by each formula in X, and for good measure, for each such C we also add each formula in Y *on the other side of the sequent*. If we make the substitution all the way down the ancestry of the displayed C in the derivation δ_2 of $X', C \succ Y$, the result of the substitution is indeed $X, X' \succ Y, Y'$, which we want to conclude. This process of replacing C by a sequent is *sequent-substitution*, as defined here:

DEFINITION 2.18 [SEQUENT SUBSTITUTION FOR MULTIPLE CONCLUSIONS] The result of substituting the sequent $X \succ Y$ for the given instance of C in $X', C \succ Y'$, or in $X' \succ C, Y$, is the sequent $X', X \succ Y', Y$. To substitute a sequent for more than one instance of a formula in a sequent, perform that substitution once for each instance, so we add one copy of X and one copy of Y for each copy of C.

In single conclusion sequents, the definition sequent substitution needs more care, for there is less room for substitution in the RHS of a sequent. In this case, the *Cut* has the following shape:

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \succ C\end{array}\qquad \begin{array}{c}\vdots\,\delta_2\\ X', C \succ R\end{array}}{X, X' \succ R}\,{\scriptstyle Cut}$$

Here, the ancestry of C in δ_1 is very simple. It consists only of a single parent families until we reach an orphan, since the ancestry remains on the RHS of each sequent. The ancestry of the displayed C in δ_2 could be more complex. To substitute for C in $X', C \succ R$, we substitute a sequent of the form $X \succ$ for C, and the result is $X, X' \succ R$. If we substitute for C in $X \succ C$, we can substitute a sequent of the form $X' \succ R$, and the result is also $X, X' \succ R$.

DEFINITION 2.19 [SINGLE CONCLUSION SEQUENT SUBSTITUTION] The result of substituting $X \succ$ for C in $X', C \succ R$ is $X, X' \succ R$. The result of substituting $X \succ R$ for C in $X' \succ C$ is also $X, X' \succ R$. (To substitute a sequent for more than one instance of a formula in a sequent, perform the required substitution once for each instance of that formula.)

So, to pass the *Cut*

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \succ C, Y\end{array} \qquad \begin{array}{c}\vdots\,\delta_2\\ X', C \succ Y'\end{array}}{X, X' \succ Y, Y'}\;\textit{Cut}$$

up the ancestry of a formula in a derivation, we perform a sequent substitution through the ancestry of the cut formula along either the left or the right. So, for example, if we push the *Cut* up the derivation δ_2, we perform the *Cut* on each of the orphans instances of C in the ancestry in δ_2, and throughout the rest of the ancestry of C in δ_2 substitute the remainder of the *Cut* premise (here $X \succ Y$) for C. For this to actually produce a *derivation*, we need to show that substituting sequents for formulas in an ancestry does not invalidate any of the rules. That is, we will show that for any sequent rule

$$\dfrac{S_1 \quad \cdots \quad S_n}{S}$$

In our case, n is 0, 1 or 2, but it rules could have more premises in general

where some formulas C in S have parents in S_1 to S_n, then for any sequent $X \succ Y$, the result of substituting $X \succ Y$ for each instance of C in the ancestry is remains an instance of the same rule. That is, we need to show that our rules are *closed under substitution for parents and children*.

Were we to have a rule (sometimes called *Mingle*) like the converse of *Contraction*,

$$\dfrac{X, X' \succ Y, Y'}{X, X', X' \succ Y, Y', Y'}$$

then this Lemma would fail. The rule would not be closed under substitution of a sequent for a single formula in X' (or Y') in the conclusion. You can repair the defect by generalising the rule to

$$\dfrac{X, X' \succ Y, Y' \quad X, X'' \succ Y, Y''}{X, X', X'' \succ Y, Y', Y''}$$

LEMMA 2.20 [RULES ARE CLOSED UNDER SUBSTITUTION] *For each inference rule in Figures 23 and 24, and any formula* C *occurring passively in the conclusion of that rule, and any sequent* $U \succ V$ (*where this has the form* $U \succ R$ *in the case of single conclusion rules where* C *occurs in the* right *hand side of the conclusion, or* $U \succ$ *in the case of single conclusion rules where* C *occurs in the* left *hand side of the conclusion*), *the result of substituting* $U \succ V$ *for* C *and its parents in that inference remains an instance of the rule.*

Proof: This is verified by inspection of each of the rules. The multiple conclusion case is simplest. The passive formulas in the rules represented in Figure 24 are those occurring in X, X', Y or Y'. To substitute a sequent (say) $U \succ V$ for any such formula we need to find some place to slot in the extra formulas U and V. But the multisets X, Y, X', Y' are arbitrary and are as large as one likes. The crucial condition for the substitution is that no rules have a shape like

$$\frac{X \succ A}{X \succ A'}$$

for then, I could not always substitute $U \succ V$ for a formula in X, since there is nowhere for the formulas in V to go—the restriction on the right hand side to being a single formula blocks the substitution. Each of the rules in Figure 24 have places for arbitrary passive formulas on the left *and* the right, so replacing one formula in one such position with a family of others, both on the left and the right, results in another instance of the rule.

In the case of the contraction rule, if the formula being substituted for occurs once in the conclusion and twice in the premise, then the result of substituting $U \succ V$ for the contracted formula will have all of the formulas in U on the left and V on the right duplicated in the premise of the rule. This is why the rule of contraction has the form that it does

$$\frac{X, X', X' \succ Y, Y', Y'}{X, X' \succ Y, Y'}\ {\scriptstyle W}$$

so, for example, the result of substituting $U \succ V$ for C in

$$\frac{X, A, A \succ Y}{X, A \succ Y}\ {\scriptstyle W}$$

is the inference

$$\frac{X, U, U \succ Y, V, V}{X, U \succ Y, V}\ {\scriptstyle W}$$

which is, indeed, another instance of the contraction rule.

For the case of rules in multiple premise single conclusion sequents, in Figure 23, we need to show that at least the more restrictive substitutions are possible. The restrictions on substitution are required, because many rules for single conclusion sequents do not have space for passive formulas on the right. The negation rules, in particular, have this shape:

$$\frac{X \succ A}{X, \neg A \succ}\ {\scriptstyle \neg L} \qquad \frac{X, A \succ}{X \succ \neg A}\ {\scriptstyle \neg R}$$

and here, there are no formulas on the right at all in two of the sequents. Each of the connective right rules have no space for passive formulas on the right — the formula on the right is active in the rule. For these rules, the only formulas that could be substituted for are formulas on the left. And, in general, to substitute for C in $X, C \succ D$ we substitute a sequent $U \succ$, to get $X, U \succ D$. So, to substitute in inferences in which no formulas on the right are passive, we replace a single passive formula on the left by a multiset of formulas. But whenever a formula is passive on the left, the inference rule allows arbitrary multisets of passive formulas, so such a substitution is permissible. If a formula is passive on the right in a sequent (such as C in $X \succ C$) then we can substitute $U \succ R$ for *it*, to get $X, U \succ R$. For this to be acceptable, we need to check that any sequents in rules in which a formula on the right is passive (it is an R in the rule statements in Figure 23), then the rule also provides space for multiset of passive formulas on the left, available for substitution. This is the case for all such rules. (There is no rule with the shape $\succ R$ or $A \succ R$, where we are allowed arbitrary formulas on the right but not arbitrary multisets of passive formulas on the left.) So, each of our rules is general enough to allow for substitution, and the lemma is proved. ■

Here is how this lemma will be applied, in our quest to eliminate *Cut* from proofs.

LEMMA 2.21 [COMMUTING CUT WITH RULES] *If a rule* R *is closed under sequent substitution for any formula occurring passively in that rule, then the rule commutes with Cut in the following sense. If the Cut formula occurs on the* RHS *in the rule, then we can transform this proof fragment*:

$$\dfrac{\dfrac{X_1 \succ Y_1, m_1C \quad \cdots \quad X_n \succ Y_n, m_nC}{X \succ Y, C}\,R \qquad X', C \succ Y'}{X, X' \succ Y, Y'}\,\mathrm{Cut}$$

into this:

$$\dfrac{\dfrac{X_1 \succ Y_1, m_1C \quad X', C \succ Y'}{X_1, m_1X' \succ Y_1, m_1Y'}\,m_1\mathrm{Cuts} \quad \cdots \quad \dfrac{X_n \succ Y_n, m_nC \quad X', C \succ Y'}{X_n, m_nX' \succ Y_n, m_nY'}\,m_n\mathrm{Cuts}}{X, X' \succ Y, Y'}\,R$$

where mC *is* m *copies of the formula* C (m *may be zero*), *and* mX *is* m *copies of each formula in the multiset* X. *The dual transformation is made if the the Cut formula occurs on the* LHS *in the rule* R.

Proof: This follows immediately from the fact that the rule R is closed under sequent substitution. ■

Here is another fact about ancestry in derivations which will become important in understanding the behaviour of *Cut*s on orphan formulas.

LEMMA 2.22 [ACTIVE FORMULAS ARE SOLE ORPHANS] *If the formula* C *is an orphan and active in the conclusion of some inference rule, then it is the* only *orphan in that sequent if that inference has premises.*

This condition would not be satisfied by an elimination rule like this:

$$\dfrac{X \succ A \oplus B, Y}{X \succ A, B, Y}$$

in which both A and B are active and orphans in the conclusion.

Proof: Consider each of the left and right connective rules. The only orphans in the rules are the active formula introduced. All the other formulas in the conclusion of the rule have parents in some premise sequent. The exception to this are the rules $\top R$ and $\bot L$, which have no premises, so in any instance of these rules, each formula present is an orphan. In the case of *Weakening*, the formulas added by weakening are passive, not active. In the case of *Contraction*, or *Cut*, there are no orphans at all. ■

We can now consider the detail of how to commute a *Cut* up the ancestry of the cut formula, to meet the orphans in that ancestry.

LEMMA 2.23 [CONVERTING CUTS INTO ORPHAN ACTIVE CUTS] *Any sequent system in which* (1) *we have an analysis of active, passive formulas, and ancestry such that ancestry preserves positions in sequents;* (2) *each rule is closed under sequent substitution in passive positions;* (3) *active formulas are the sole orphans in each rule with one or more premises, then for any derivation ending in a Cut*

$$\dfrac{\dfrac{\vdots\,\delta_1}{X \succ C, Y} \qquad \dfrac{\vdots\,\delta_2}{X', C \succ Y'}}{X, X' \succ Y, Y'}\,\text{Cut}$$

(*and in which there are no other Cuts on* C) *can be systematically transformed into a derivation of the same conclusion in which the only Cuts on* C *are inferences in which the Cut formula* C *is an orphan and active in both of the premises of that Cut. The resulting derivation contains no inference rules not present in the original derivation, and in particular, if* δ_1 *and* δ_2 *are Cut-free, then the only Cuts in the new derivation are those Cuts on* C *in which both instances of* C *are orphans and active.*

Proof: Consider the tree of ancestors of the cut formula C in the left derivation δ_1. Substitute $X' \succ Y'$ for C for all non-orphan instances of C in the ancestry in δ_1. For each inference in which the substitution is made to premises and the conclusion, or in the case of rules such as $\bot L$ and $\top R$, which have no premises, the inference remains an instance of that rule, since the rules are closed under substitution of sequents for passive formulas in the conclusion. (And, if C was not an orphan in the final inference of δ_1, the result is now the conclusion of the *Cut* step, $X, X' \succ Y, Y'$.) For sequents in δ_1 where C occurs as an *orphan* in this ancestry, either it is introduced in a connective rule and it is the sole orphan, or it is a family of passive instances, or it is a $\top$ or $\bot$ in their distinctive rules, or it is one side of an *Id* axiom. In the first case (the sole orphan active formula in a rule with premises), replace the component

$$\dfrac{\vdots}{U \succ C, V}$$

in which the indicated C is active, by the following instance of *Cut*.

$$\dfrac{\dfrac{\vdots}{U \succ C, V} \qquad \dfrac{\vdots\,\delta_2}{X', C \succ Y'}}{U, X' \succ V, Y'}\,\textit{Cut}$$

The conclusion of this inference is the sequent $U \succ C, V$ where $X' \succ Y'$ has been substituted for the displayed C, so it leads appropriately into the rest of the derivation, where the substitution has occurred for non-orphan instances of C, including those which had this displayed C as a parent.

For sequents in which there are a number of orphan instances of C in the ancestry, these are either *all* passive, or C is $\top$ and this is the $\top R$ axiom, or C is $\bot$ and this is the $\bot L$ axiom. In the case of all orphan instances of C being passive in the inference rule, this means that C is introduced by weakening (its parents are not in the premises, and it is not active), so the rule is also closed under substitution of $X' \succ Y'$ for C, so we may replace the conclusion of this rule by the desired substitution.

The remaining case is where C is either $\top$ and the sequent in which C is an orphan is a [$\top R$] axiom or it is $\bot$ and the sequent is a [$\bot L$] axiom. Consider the case for $\top$. Our sequent has the form

$$X \succ \top, \ldots, \top, Y$$

where one $\top$ is active and the others passive, and we wish to conclude

$$X, X', \ldots, X' \succ Y, Y', \ldots, Y'$$

substituting $X' \succ Y'$ for each instance of $\top$. We have a derivation δ_2:

$$\begin{array}{c} \vdots\, \delta_2 \\ X', \top \succ Y' \end{array}$$

In this derivation, $\top$ is passive *everywhere,* since there is no inference rule in which $\top$ is active on the left of a sequent. As a result, we can substitute $X, X', \ldots, X' \succ Y, Y', \ldots, Y'$ (with one fewer instances of X' and Y') for the ancestry of that $\top$ in δ_2. The result is a derivation of the required sequent.

So, we have pushed the *Cut* up the ancestry of C in δ_1. The same technique allows us to push the remaining *Cut*s up the ancestry of C in each instance of δ_2, and the result is a derivation in which the remaining *Cut*s on C feature C as active in both premises of the *Cut*. ■

The remaining component of the elimination of *Cut*s is reducing the complexity of *Cut* formulas.

LEMMA 2.24 [REDUCTION OF RANK FOR CUT FORMULAS] *In any derivation*

$$\dfrac{\begin{array}{c} \vdots\, \delta_1 \\ X \succ C, Y \end{array} \qquad \begin{array}{c} \vdots\, \delta_2 \\ X', C \succ Y' \end{array}}{X, X' \succ Y, Y'}\,{\scriptstyle Cut}$$

where C *is active in the final inferences of* δ_1 *and* δ_2, *this Cut can be replaced by Cuts on subformulas of* C, *or by no Cuts at all.*

Proof: For *identity* sequents, the *Cut* reduction is immediate. A *Cut*

$$\dfrac{p \succ p \quad p \succ p}{p \succ p}\,{\scriptstyle Cut}$$

can be replaced by the axiom $p \succ p$. Then for non-identity sequents, we replace *Cut*s on a case-by-case basis. For multiple premise and single conclusion sequents, use the reductions in Figure 26. For multiple conclusion sequents, use the reductions in Figure 27. ■

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1} \quad \overset{\vdots\,\delta_l^2}{X \succ C_2}}{X \succ C_1 \wedge C_2}\,{\wedge R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X', C_i \succ R}}{X', C_1 \wedge C_2 \succ R}\,{\wedge L_i}}{X, X' \succ R}\,\mathit{Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l^i}{X \succ C_i} \quad \overset{\vdots\,\delta_r}{X', C_i \succ R}}{X, X' \succ R}\,\mathit{Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X \succ C_i}}{X \succ C_1 \vee C_2}\,{\vee R_i} \qquad \dfrac{\overset{\vdots\,\delta_r^1}{X', C_1 \succ R} \quad \overset{\vdots\,\delta_r^2}{X', C_2 \succ R}}{X', C_1 \vee C_2 \succ R}\,{\vee L}}{X, X' \succ R}\,\mathit{Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l}{X \succ C_i} \quad \overset{\vdots\,\delta_r^i}{X', C_i \succ R}}{X, X' \succ R}\,\mathit{Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1} \quad \overset{\vdots\,\delta_l^2}{X' \succ C_2}}{X, X' \succ C_1 \otimes C_2}\,{\otimes R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X'', C_1, C_2 \succ R}}{X'', C_1 \otimes C_2 \succ R}\,{\otimes L}}{X, X', X'' \succ R}\,\mathit{Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1} \quad \dfrac{\overset{\vdots\,\delta_l^2}{X' \succ C_2} \quad \overset{\vdots\,\delta_r}{X'', C_1, C_2 \succ R}}{X', X'', C_1 \succ R}\,\mathit{Cut}}{X, X', X'' \succ R}\,\mathit{Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X, C_1 \succ C_2}}{X \succ C_1 \to C_2}\,{\to R} \qquad \dfrac{\overset{\vdots\,\delta_r^1}{X' \succ C_1} \quad \overset{\vdots\,\delta_r^2}{X'', C_2 \succ R}}{X', X'', C_1 \to C_2 \succ R}\,{\to L}}{X, X', X'' \succ R}\,\mathit{Cut} \quad \rightsquigarrow \quad \dfrac{\dfrac{\overset{\vdots\,\delta_r^1}{X' \succ C_1} \quad \overset{\vdots\,\delta_l}{X, C_1 \succ C_2}}{X, X' \succ C_2}\,\mathit{Cut} \quad \overset{\vdots\,\delta_r^2}{X', C_2 \succ R}}{X, X', X'' \succ R}\,\mathit{Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X, C \succ}}{X \succ \neg C}\,{\neg R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X' \succ C}}{X', \neg C \succ}\,{\neg L}}{X, X' \succ}\,\mathit{Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_r}{X' \succ C} \quad \overset{\vdots\,\delta_l}{X, C \succ}}{X, X' \succ}\,\mathit{Cut}$$

$$\dfrac{\succ t \qquad \dfrac{\overset{\vdots\,\delta}{X \succ R}}{X, t \succ R}\,{tL}}{X \succ R}\,\mathit{Cut} \quad \rightsquigarrow \quad \overset{\vdots\,\delta}{X \succ R} \qquad\qquad \dfrac{\dfrac{\overset{\vdots\,\delta}{X \succ}}{X \succ f}\,{fR} \qquad f \succ}{X \succ}\,\mathit{Cut} \quad \rightsquigarrow \quad \overset{\vdots\,\delta}{X \succ}$$

Figure 26: ACTIVE FORMULA CUT REDUCTIONS: MULTIPLE PREMISE

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1, Y} \quad \overset{\vdots\,\delta_l^2}{X \succ C_2, Y}}{X \succ C_1 \wedge C_2, Y}{\scriptstyle \wedge R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X', C_i \succ Y'}}{X', C_1 \wedge C_2 \succ Y'}{\scriptstyle \wedge L_i}}{X, X' \succ Y, Y'}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l^i}{X \succ C_i, Y} \quad \overset{\vdots\,\delta_r}{X', C_i \succ Y'}}{X, X' \succ Y, Y'}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X \succ C_i, Y}}{X \succ C_1 \vee C_2, Y}{\scriptstyle \vee R_i} \qquad \dfrac{\overset{\vdots\,\delta_r^1}{X', C_1 \succ Y'} \quad \overset{\vdots\,\delta_r^2}{X', C_2 \succ Y'}}{X', C_1 \vee C_2 \succ Y'}{\scriptstyle \vee L}}{X, X' \succ Y, Y'}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l}{X \succ C_i, Y} \quad \overset{\vdots\,\delta_r^i}{X', C_i \succ Y'}}{X, X' \succ Y, Y'}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1, Y} \quad \overset{\vdots\,\delta_l^2}{X' \succ C_2, Y'}}{X, X' \succ C_1 \otimes C_2, Y, Y'}{\scriptstyle \otimes R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X'', C_1, C_2 \succ Y''}}{X'', C_1 \otimes C_2 \succ Y''}{\scriptstyle \otimes L}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_l^1}{X \succ C_1, Y} \quad \dfrac{\overset{\vdots\,\delta_l^2}{X' \succ C_2, Y'} \quad \overset{\vdots\,\delta_r}{X'', C_1, C_2 \succ Y''}}{X', X'', C_1 \succ Y', Y''}{\scriptstyle Cut}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X \succ C_1, C_2, Y}}{X \succ C_1 \oplus C_2, Y}{\scriptstyle \oplus R} \qquad \dfrac{\overset{\vdots\,\delta_r^1}{X', C_1 \succ Y'} \quad \overset{\vdots\,\delta_r^2}{X'', C_2 \succ Y''}}{X', X'', C_1 \oplus C_2 \succ Y', Y''}{\scriptstyle \oplus L}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\dfrac{\overset{\vdots\,\delta_l}{X \succ C_1, C_2, Y} \quad \overset{\vdots\,\delta_r^1}{X', C_1 \succ Y'}}{X, X', C_2 \succ Y, Y'}{\scriptstyle Cut} \quad \overset{\vdots\,\delta_r^2}{X', C_2 \succ Y'}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X, C_1 \succ C_2, Y}}{X \succ C_1 \to C_2, Y}{\scriptstyle \to R} \qquad \dfrac{\overset{\vdots\,\delta_r^1}{X' \succ C_1, Y'} \quad \overset{\vdots\,\delta_r^2}{X'', C_2 \succ Y''}}{X', X'', C_1 \to C_2 \succ Y', Y''}{\scriptstyle \to L}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\dfrac{\overset{\vdots\,\delta_r^1}{X' \succ C_1, Y'} \quad \overset{\vdots\,\delta_l}{X, C_1 \succ C_2, Y}}{X, X' \succ C_2, Y, Y'}{\scriptstyle Cut} \quad \overset{\vdots\,\delta_r^2}{X', C_2 \succ Y'}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{\overset{\vdots\,\delta_l}{X, C \succ Y}}{X \succ \neg C, Y}{\scriptstyle \neg R} \qquad \dfrac{\overset{\vdots\,\delta_r}{X' \succ C, Y'}}{X', \neg C \succ Y'}{\scriptstyle \neg L}}{X, X' \succ Y, Y'}{\scriptstyle Cut} \quad \rightsquigarrow \quad \dfrac{\overset{\vdots\,\delta_r}{X' \succ C, Y'} \quad \overset{\vdots\,\delta_l}{X, C \succ Y}}{X, X' \succ Y, Y'}{\scriptstyle Cut}$$

$$\dfrac{\succ t \qquad \dfrac{\overset{\vdots\,\delta}{X \succ Y}}{X, t \succ Y}{\scriptstyle tL}}{X \succ Y}{\scriptstyle Cut} \quad \rightsquigarrow \quad \overset{\vdots\,\delta}{X \succ Y} \qquad \dfrac{\dfrac{\overset{\vdots\,\delta}{X \succ Y}}{X \succ f, Y}{\scriptstyle fR} \qquad f \succ}{X \succ Y}{\scriptstyle Cut} \quad \rightsquigarrow \quad \overset{\vdots\,\delta}{X \succ Y}$$

Figure 27: ACTIVE FORMULA CUT REDUCTIONS: MULTIPLE CONCLUSION

Putting these results together, we show how *Cut* can be eliminated from derivations in any of our sequent systems.

THEOREM 2.25 [CUT ELIMINATION IN SEQUENT SYSTEMS] *Any derivation δ in a sequent system may be systematically transformed into a derivation of the same sequent in which no Cuts are used.*

Proof: We prove this by induction on the *Cut* complexity of the derivation, where this complexity is the sequence $\langle c_0, c_1, c_2, \ldots, c_m \rangle$ where the derivation has c_i *Cut*s of rank i, and no other *Cut*s. Cut complexity is ordered as usual, with higher ranks more significant than lower. Select a *Cut* in δ where there are no *Cut*s *above*. Use the process of the previous two lemmas to push that *Cut* up to orphans (temporarily blowing up the *Cut* measure by possibly duplicating it past contractions) and then replacing the *Cut* with *Cut*s on subformulas of C, *lowering* the *Cut* complexity. This process possibly duplicates material in the derivations, but since this derivation *up* to our *Cut* contains no *Cut*s, this does not increase the *Cut* measure by duplicating other *Cut*s. The result is a derivation with lower *Cut* complexity. Continue the process, and since there is no infinitely descending sequence of *Cut* complexity, the process terminates in a *Cut* free derivation. ■

Recall the definition of the non-normality measure in Definition 1.16 on page 23.

We have shown that in each of our sequent systems, if a sequent $X \succ Y$ has a derivation, then it has a *Cut*-free derivation—and furthermore, the a *Cut*-free derivation can be found by *eliminating* the *Cut*s from the original derivation. This is a result which is rich in significance. In the next section, we will explore some of its consequences.

2.4 | CONSEQUENCES OF CUT ELIMINATION

A core consequence of *Cut* elimination is the subformula property.

THEOREM 2.26 [SUBFORMULA PROPERTY] *If δ is a Cut-free derivation of a sequent $X \succ Y$, then δ contains only subformulas of formulas in the endsequent $X \succ Y$.*

Proof: For any inference falling under a rule (other than *Cut*) in any of our sequent systems, the formulas in the premise sequents are subformulas of formulas in the concluding sequent. So, since any derivation is a tree of sequents structured in accordance with the rules, for any sequent in that tree, only subformulas of formulas in a given sequent can occur above that sequent in the tree. In particular, all formulas in a *Cut*-free derivation of $X \succ Y$ are subformulas of formulas in $X \succ Y$. ■

This result holds for any of our sequent systems, so it holds for classical logic, intuitionistic logic, minimal logic, non-distributive relevant logic, linear logic, etc., and for fragments of these logics in which we have rules for only part of the traditional vocabulary. The phenomenon is robust.

In particular, this result shows that our presentation of these logical systems is appropriately *modular*. For example, Peirce's Law

$$\succ ((p \to q) \to p) \to p$$

is derivable in classical logic. This means it has a *Cut* free derivation, and in particular, it has a derivation in which the only rules that apply act on subformulas of $((p \to q) \to p) \to p$. In particular, we do not need to use any rules except for [$\to$*L*] and [$\to$*R*], so we do not need to appeal to negation, conjunction, or any other logical concept. The rules for the conditional encapsulate the semantics of the conditional in a way that needs no supplementation by the rules for any other connective. This can be stated, formally, in the following theorem:

THEOREM 2.27 [CONSERVATIVE EXTENSION] *If we extend a sequent system $\mathfrak{S}$ for some subset of our family of connectives, by adding the left and rules for other connectives in our family to form sequent system $\mathfrak{S}'$, this addition is* conservative, *in the sense that system $\mathfrak{S}'$ can derive no new sequents $X \succ Y$ where X and Y are taken from the language of $\mathfrak{S}$.*

Proof: Take a sequent $X \succ Y$ from the language of $\mathfrak{S}$, and which is derivable in $\mathfrak{S}'$. Take some *Cut*-free derivation of the sequent. The connective rules in this derivation apply only to subformulas of formulas in $X \succ Y$, and so, are rules from the system $\mathfrak{S}$. So this sequent could already have been derived in $\mathfrak{S}$. ■

This result means that the addition of new logical concepts gives us new concepts to express, and new ways to prove things – even in new ways to prove things from our old vocabulary – but it does not change the landscape of what can be derived in that old vocabulary. The significance of this result will be one of the central topics of the middle part of the book.

The subformula property is significant if you think of sequent rules for a connective as presenting the meaning of that connective. Consider the derivability of Peirce's Law. Not only does the separability of the system ensure that Peirce's Law holds in virtue of the rules [$\to$*L*] and [$\to$*R*] governing the conditional. The subformula property assures us that the sequent is derivable in virtue of the instances of those rules applying to subformulas of the formula itself. Peirce's Law $\succ ((p \to q) \to p) \to p$ holds in virtue of the semantic properties of $(p \to q) \to p$, $p \to q$, p and q. There is a profound sense in which the sequent is *analytic*. A *Cut*-free derivation of a sequent $X \succ Y$ shows that the sequent holds in virtue of an *analysis* of the sequent into its components. The sequent holds not in virtue of some relations that the components hold to *other* judgements, but in terms of the internal relationships between those components, and the *Cut*-free derivation of a sequent gives an analysis of the sequent into its components that suffices to establish that the sequent holds.

Excursus: This is not to say that all valid sequents have one and only one such analysis. The sequent calculus allows for sequents to hold for different reasons. The sequent $p \wedge q \succ p \vee q$, for example, has the following

Cut-free derivations:

$$\dfrac{\dfrac{p \succ p}{p \wedge q \succ p}{\scriptstyle \wedge L_1}}{p \wedge q \succ p \vee q}{\scriptstyle \vee R_2} \qquad \dfrac{\dfrac{q \succ q}{p \wedge q \succ q}{\scriptstyle \wedge L_2}}{p \wedge q \succ p \vee q}{\scriptstyle \vee R_2}$$

where according to the first, the sequent holds in virtue of the p, shared between $p \wedge q$ and $p \vee q$, and according to the second, the sequent holds in virtue of the shared q. *End of Excursus*

Another consequence of the *Cut*-elimination theorem is the *decidability* of logical consequence in our languages. This is easiest to see in the case of simple sequents.

THEOREM 2.28 [SIMPLE SEQUENT DECIDABILITY] *There is an* algorithm *for determining whether or not a simple sequent* $A \succ B$ *is valid.*

To determine whether or not $A \succ B$ has a simple sequent derivation, we use the notion of a sequent's *possible ancestry*.

DEFINITION 2.29 [POSSIBLE ANCESTRY] Given any sequent, its POSSIBLE PARENTS are each of the sequents from which it could have been derived, using any rule other than *Cut*. That is, if the sequent $A \succ B$ is the concluding sequent in an instance of a rule, for which $C \succ D$ is a premise, then $C \succ D$ is one of the possible parents of $A \succ B$. The POSSIBLE ANCESTRY of the sequent $A \succ B$ is the tree with $A \succ B$ as its root, with links to each possible parent $C \succ D$, and then each of these sequents is further connected to *its* possible ancestry.

Figure 28 depicts the possible ancestry of $q \vee r \succ (p \wedge q) \vee r$.

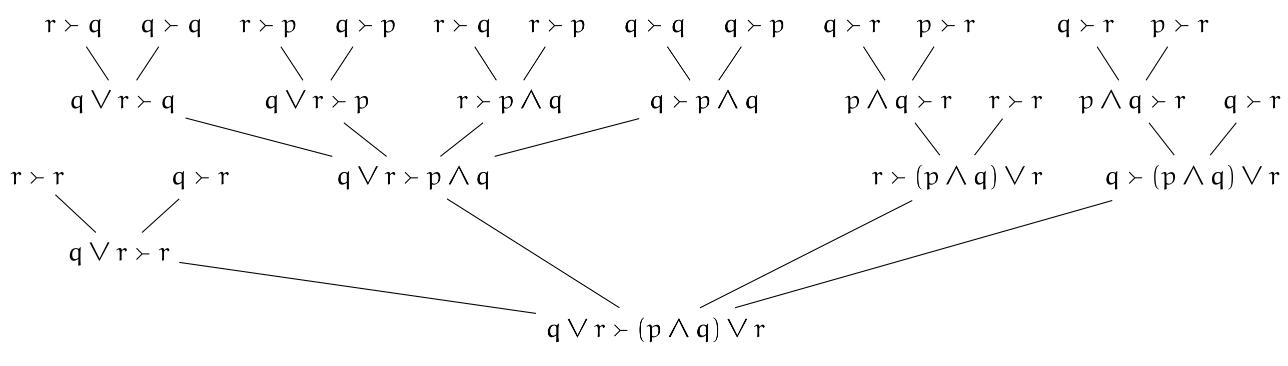


Figure 28: THE POSSIBLE ANCESTRY OF $q \vee r \succ (p \wedge q) \vee r$

LEMMA 2.30 [POSSIBLE ANCESTRY IS FINITE] *The possible ancestry of any sequent* $A \succ B$ *in a simple sequent system contains only finitely many nodes.*

Proof: We prove this by induction on the complexity of $A \succ B$. A sequent $p \succ q$ consisting of atoms has *no* possible parents, and so its ancestry is the trivial tree consisting of $p \succ q$ itself, and is finite.

Take any sequent $A \succ B$, and suppose that the hypothesis holds for simpler sequents. Inspecting the rules (see Figure 21, on page 51), we can see that each of its possible parents are simpler sequents, and, in addition, there are only finitely many possible parents. The hypothesis holds for each possible parents (their possible ancestries are finite) so the possible ancestry of $A \succ B$ is finite as well. ▪

Now we can use the possible ancestry of a sequent in order to find a derivation for that sequent—if it has one.

Proof: Given the possible ancestry of the sequent $A \succ B$, start at the leaves, and mark any leaf of the form $p \succ p$ as derivable, and delete (cross out) other leaf as underivable. We have marked the derivable sequents and deleted the underivable ones. Let's call this process *processing* the leaves. The marked sequents have derivations, and the deleted sequents do not. Let's suppose that all of the potential parents of the sequent $C \succ D$ have been processed, and we explain what it is to process $C \succ D$. Given a sequent $C \succ D$, examine its possible parents of $C \succ D$ which have not been deleted, and check, for each rule that can derive $C \succ D$, whether the premises of that rule have survived (been marked, not deleted). If so, mark $C \succ D$ as derivable, since it can be derived using any of the derivations of the parents, and the rule under which any of the surviving parents fall. If not enough parents have survived, the sequent $C \succ D$ has no derivation is deleted. This completes the process. ▪

We have already seen one example of this at the beginning of the previous section, when we saw that distribution $(p \wedge (q \vee r) \succ (p \wedge q) \vee (p \wedge r))$ is not derivable (see page 60), though now we can describe this process in terms of the possible ancestry of $p \wedge (q \vee r) \succ (p \wedge q) \vee (p \wedge r)$. This tree has $p \wedge (q \vee r) \succ (p \wedge q) \vee (p \wedge r)$ at its root, and this has four possible parents: $p \succ (p \wedge q) \vee (p \wedge r)$ and $q \vee r \succ (p \wedge q) \vee (p \wedge r)$ on the one hand, and $p \wedge (q \vee r) \succ p \wedge q$ and $p \wedge (q \vee r) \succ p \wedge r$ on the other. Each of these would suffice as sole parents (the relevant rules are [$\wedge$*L*] and [$\vee$*R*], which have single premises), but none of these are marked in *their* possible ancestries (see page 60 for the details), and as a result, $p \wedge (q \vee r) \succ (p \wedge q) \vee (p \wedge r)$ is not derivable.

Searching for derivations in the naïve manner described by this theorem is not as efficient as we can be: we don't need to search for *all* possible derivations of a sequent if we know about some of the special properties of the rules of the system. For example, consider the sequent $A \vee B \succ C \wedge D$ (where A, B, C and D are possibly complex statements). This is derivable in two ways (a) from $A \succ C \wedge D$ and $B \succ C \wedge D$ by [$\vee$*L*] or (b) from $A \vee B \succ C$ and $A \vee B \succ D$ by [$\wedge$*R*]. Instead of searching *both* of these possibilites, we may notice that *either* choice would be enough to search for a derivation, since the rules [$\vee$*L*] and [$\wedge$*R*] 'lose no information' in an important sense.

DEFINITION 2.31 [INVERTIBILITY] A sequent rule of the form

$$\frac{S_1 \ \cdots \ S_n}{S}$$

is *invertible* if and only if whenever the sequent S is derivable, so are the sequents $S_1, \ldots, S_n$.

THEOREM 2.32 [INVERTIBLE SEQUENT RULES] *The rules* [$\lor$L] *and* [$\land$R] *are invertible, but the rules* [$\lor$R] *and* [$\land$L] *are not.*

Proof: Consider [$\lor$*L*]. If $A \lor B \succ C$ is derivable, then since we have a derivation of $A \succ A \lor B$ (by [$\lor$*R*]), a use of *Cut* shows us that $A \succ C$ is derivable. Similarly, since we have a derivation of $B \succ A \lor B$, the sequent $B \succ C$ is derivable too. So, from the conclusion $A \lor B \succ C$ of a [$\lor$*L*] inference, we may derive the premises. The case for [$\land$*R*] is completely analogous.

For [$\land$*L*], on the other hand, we have a derivation of $p \land q \succ p$, but no derivation of the premise $q \succ p$, so this rule is not invertible. Similarly, $p \succ q \lor p$ is derivable, but $p \succ q$ is not. ■

It follows that when searching for a derivation of a sequent, instead of searching through its *entire* possible ancestry, if it may be derived from an invertible rule, we can look to the premises of *that* rule, and ignore the other branches of its ancestry.

EXAMPLE 2.33 [DERIVATION SEARCH] The sequent $(p \land q) \lor (q \land r) \succ (p \lor r) \land p$ is not derivable. By the invertibility of [$\lor$*L*], it *is* derivable only if (a) $p \land q \succ (p \lor r) \land p$ and (b) $q \land r \succ (p \lor r) \land p$ are both derivable. Here is the possible ancestry for $p \land q \succ (p \lor r) \land p$, where we, at the first step, appeal to the invertible rule [$\land$*R*], and we start from the top and strike out any underivable sequents.

$$\dfrac{\dfrac{\dfrac{p \succ p \quad \cancel{p \succ r}}{p \succ p \lor r} \quad \dfrac{\cancel{q \succ r} \quad \cancel{q \succ r}}{\cancel{q \succ p \lor r}} \quad \dfrac{p \succ p \quad \cancel{q \succ p}}{p \land q \succ p} \quad \dfrac{\cancel{p \succ r} \quad \cancel{q \succ r}}{\cancel{p \land q \succ r}}}{p \land q \succ p \lor r} \qquad \dfrac{p \succ p \quad \cancel{q \succ p}}{p \land q \succ p}}{p \land q \succ (p \lor r) \land p}$$

The sequent at the root survives. It is derivable. The other required premise, for our target sequent, $q \land r \succ (p \lor r) \land p$, is less fortunate.

$$\dfrac{\dfrac{\dfrac{r \succ r \quad \cancel{q \succ r}}{q \land r \succ r} \quad \dfrac{\cancel{r \succ p} \quad \cancel{q \succ p}}{\cancel{q \land r \succ p}} \quad \dfrac{r \succ r \quad \cancel{r \succ p}}{r \succ p \lor r} \quad \dfrac{\cancel{q \succ r} \quad \cancel{q \succ p}}{\cancel{q \succ p \lor r}}}{q \land r \succ p \lor r} \qquad \dfrac{\cancel{q \succ p} \quad \cancel{r \succ p}}{\cancel{q \land r \succ p}}}{\cancel{q \land r \succ (p \lor r) \land p}}$$

This sequent is not derivable, because $q \land r \succ p$ is underivable.

» «

This result can be generalised to apply to any of our complex sequent systems, but in the case of some of these systems, we need to do more work. Given a system $\mathfrak{S}$, and a sequent $X \succ Y$ we produce its possible ancestry, and use this to determine whether there the sequent has a derivation. In the case of systems without the contraction rule, this result is no more complex than for simple sequent systems.

THEOREM 2.34 [CONTRACTION FREE SYSTEMS ARE DECIDABLE] *Any of our sequent systems (multiple conclusion or single conclusion) without the contraction rule is decidable. In particular, given any sequent* $X \succ Y$*, its possible ancestry is finite.*

Proof: If you inspect every rule other than contraction and *Cut*, you can see that (a) the number of formulas in each premise sequent of a rule is less than or equal to the number of formulas in the conclusion sequent, and also, (b) each formula in a premise sequent must be a *subformula* of a formula in the concluding sequent, and (c) for each sequent, only a finite number of rules (other than *Cut*) could produce that sequent as a conclusion. So, the possible ancestry (defined as before) for a sequent $X \succ Y$ is a finitely branching tree, in which the sequents along a branch reduce in complexity, and each sequent contains only subformulas of sequents lower down in the branch. So, each branch has only a finite length. By König's Lemma, the possible ancestry is finite.

König's Lemma: A finitely branching tree in which every branch has finite length is itself finite.

We use the possible ancestry to check for the existence of a proof as before. Starting with the leaves of the tree (the sequents containing only atoms), we check for derivability immediately, marking those that survive, and crossing out those that are not axioms. Then, for nodes lower down in the tree, once all its parents have been processed, mark a sequent as derivable if any of the premises of a rule under which it falls have survived. If not enough possible parents have survived, and there are no premises upon which to derive the sequent, strike it out. Continue until the tree is complete. ■

For systems *with* the contraction rule, things get more complicated, for with contraction, we can derive *smaller* sequents from *larger* sequents. This means that the possible ancestry for a sequent is no longer finite. In the case of systems with contraction and weakening, this is especially egregious. We can arbitrarily lengthen any derivation with moves like this: Replace a sequent $X, A \succ Y$ occurring in a derivation with this *weakening–contraction* two-step:

$$\dfrac{\dfrac{\begin{matrix}\vdots\\ X, A \succ Y\end{matrix}}{X, A, A \succ Y}\,K}{\begin{matrix}X, A \succ Y\\ \vdots\end{matrix}}\,W$$

to extend the derivation by two sequents. This can be repeated *ad libitum*. So, we need some way to limit the search for derivations. Clearly, when we search for derivations for a sequent, we want to limit the search space so we don't end up chasing our own tail. At least we should restrict our search to *concise* derivations.

DEFINITION 2.35 [CONCISE DERIVATIONS] A derivation δ is CONCISE iff each branch of the tree of sequents contains each sequent only once.

THEOREM 2.36 [DERIVATIONS CAN BE MADE CONCISE] *If a sequent* $X \succ Y$ *has a derivation (in some system), it has a concise derivation as a sub-tree of that derivation.*

Proof: Take a derivation δ of $X \succ Y$. If there is some brach where a sequent $U \succ V$ is repeated, delete all steps in the branch between the first occurrence of the sequent in this branch and the last (and one instance of $U \succ V$), including deleting any other branches of the tree which branch into this intermediate segment. The result is a smaller derivation of $X \succ Y$. If there are still repeated sequents in branches, continue the process. It cannot continue forever, as the derivation is smaller at each stage of the process. The end of the process is a concise derivation of $X \succ Y$ which found inside the original derivation δ. ∎

Searching for *concise* derivations cuts down on the search space. This step alone is not enough, though, for contraction is an insidious rule. When we ask ourselves whether the sequent $X, A \succ Y$ is derivable, perhaps it was derived from $X, A, A \succ Y$. And where was this derived? Perhaps from $X, A, A, A \succ Y$, and so on. If we search for concise derivations using contraction, the search space is still very large. Consider a derivation where contraction needs to be used a *lot*. The sequent $p \succ p \otimes (p \otimes (p \otimes p))$ is derivable using contraction.

$$\dfrac{\dfrac{\dfrac{p \succ p \quad \dfrac{p \succ p \quad \dfrac{p \succ p \quad p \succ p}{p, p \succ p \otimes p}\otimes R}{p, p, p \succ p \otimes (p \otimes p)}\otimes R}{p, p, p, p \succ p \otimes (p \otimes (p \otimes p))}\otimes R}{p, p \succ p \otimes (p \otimes (p \otimes p))}W}{p \succ p \otimes (p \otimes (p \otimes p))}W$$

In this derivation, we derive $p, p, p, p \succ p \otimes (p \otimes (p \otimes p))$, and then two steps of contraction reduce the four instances of p to one. When looking for a derivation of $p \succ p \otimes (p \otimes (p \otimes p))$, we can find one when we go through the more complex sequent $p, p, p, p \succ p \otimes (p \otimes (p \otimes p))$, having duplicated the p three times.

Why two steps? In the first, we contract two instances of p, p into one. In the second, two instances of p are contracted into one.

Why is so much contraction *needed* in a derivation of $p \succ p \otimes (p \otimes (p \otimes p))$? They are required here because repetitions of p are introduced by each [$\otimes R$] step, and we need an [$\otimes R$] step for each $\otimes$ in the right of the sequent. To derive $p \succ p \otimes (p \otimes (p \otimes p))$ without using *Cut*, the

An un-serious option is a contraction on the *right*, but why would you try that?

only serious options are a contraction on the left or [$\otimes R$]. The [$\otimes R$] step cannot be immediately applied, since we have only one p to go around (the possible pairs of parents are $p \succ p$ and $\succ p \otimes (p \otimes p)$ or $\succ p$ and $p \succ p \otimes (p \otimes p)$, and in either case, at least one possible parent is underivable). So, a contraction it must be. But we don't *need* to apply contraction again to go to the four-way repetition of p, for the sequent $p, p \succ p \otimes (p \otimes (p \otimes p))$ is derivable by way of the two parents $p \succ p$ and $p \succ p \otimes (p \otimes p)$. The $p \succ p$ is an axiom, and again, we can derive $p \succ p \otimes (p \otimes p)$ by contraction from $p, p \succ p \otimes (p \otimes p)$, apply [$\otimes R$] to split again, and so on. The resulting derivation is different:

$$
\begin{prooftree}
\AxiomC{$p \succ p$}
\AxiomC{$p \succ p$}
\AxiomC{$p \succ p$}
\AxiomC{$p \succ p$}
\RightLabel{$\otimes R$}
\BinaryInfC{$p, p \succ p \otimes p$}
\RightLabel{W}
\UnaryInfC{$p \succ p \otimes p$}
\RightLabel{$\otimes R$}
\BinaryInfC{$p, p \succ p \otimes (p \otimes p)$}
\RightLabel{W}
\UnaryInfC{$p \succ p \otimes (p \otimes p)$}
\RightLabel{$\otimes R$}
\BinaryInfC{$p, p \succ p \otimes (p \otimes (p \otimes p))$}
\RightLabel{W}
\UnaryInfC{$p \succ p \otimes (p \otimes (p \otimes p))$}
\end{prooftree}
$$

We apply contraction immediately after the duplication is incurred in the [$\otimes R$] step. And this generalises to the other rules. Contraction is required to the extent that different premise sequents in our rules pile up copies of formulas. If the rules are applied repeatedly, the piles of copies can be sizeable. However, if we "clean up" as we go, contracting formulas as the first occur together (when we indeed want to contract them), instead of delaying the process for later, the process is manageable. In fact, we can make the contraction step a part of the connective rule [$\otimes R$], like this: We could reformulate [$\otimes R$] to have the following shape:

There is an analogy somewhere in the vicinity about housework, or dealing with email, or other small but numerous recurring tasks, but I will not pause to make it.

$$
\frac{X \succ A, Y \quad X' \succ B, Y'}{[X, X'] \succ [[A \otimes B], Y, Y']}\;\otimes R'
$$

where $[X, X']$ is some multiset formed from X, X', allowing for (but not requiring) any formula in X, X' to be contracted *once*, and $[[A \otimes B], Y, Y']$ is a multiset formed from $A \otimes B, Y, Y'$, allowing for (but not requiring) any formula in Y, Y' to be contracted *once*, and allowing (but not requiring) $A \otimes B$ to be contracted *once* or *twice*. This means that any new repetitions introduced in the output of this rule could be dealt with on the spot. Why would a contraction be required? Perhaps because a formula was supplied to the conclusion *both* from the left premise, and from the right premise, whereas I need only one in the resulting sequent. So contract the formula in X or X', or Y or Y', in this step. Or the formula $A \otimes B$ might already be in Y or in Y' – or in both. In that case, we can also contract also as needed. Clearly, if something is derivable using the old sequent rule [$\otimes R$], it is derivable using this new rule (nothing *forces* us to use contraction), and if something is derivable using the new rule [$\otimes R'$], and we have contraction available, we can derive it using the old rule too. Now the derivation of $p \succ p \otimes (p \otimes (p \otimes p))$ can be rewritten:

$$\dfrac{p \succ p \qquad \dfrac{p \succ p \qquad \dfrac{p \succ p \quad p \succ p}{p \succ p \otimes p}{\scriptstyle \otimes R'}}{p \succ p \otimes (p \otimes p)}{\scriptstyle \otimes R'}}{p \succ p \otimes (p \otimes (p \otimes p))}{\scriptstyle \otimes R'}$$

and we contract the duplicate ps in the left as we apply $[\otimes R']$. No explicit contraction step is required. This is the genius of allowing contraction *inside* the rules.

We can do the same for *all* our rules. For any system $\mathfrak{S}$ including contraction, we call the system $\mathfrak{S}^W$ the system with contraction internalised into the rules.

DEFINITION 2.37 [CONTRACTED MULTISETS] Given multisets X, Y and Z, a multiset M is said to be an multiset of kind $[[X], Y], Z$ if and only if its occurrences satisfy

$$\begin{array}{rcccl} \max(1, o_{X,Y,Z}(x) - 2) & \le & o_M(x) & \le & o_{X,Y,Z}(x) \quad \text{for } x \in X \\ \max(1, o_{X,Y,Z}(y) - 1) & \le & o_M(y) & \le & o_{X,Y,Z}(y) \quad \text{for } x \in Y \text{ and } x \notin X \\ & & o_M(z) & = & o_{X,Y,Z}(z) \quad \text{otherwise} \end{array}$$

That is, we allow for repeated members of X to be reduced by 2 (with a floor of 1 – you can only eliminiate copies when you have one copy left), and repeated members of Y (except for those in X) to be reduced by 1 (again, with a floor of 1), and members of Z (other than those occuring in X, Y) are be unchanged.

In this definition, any of X, Y, Z can be empty. For example, when X is empty, we have $[Y], Z$ (allowing repeats to be reduced by one in Y), and if Z is empty, we have $[[X], Y]$ (allowing for two repeats to be eliminated in X and one in Y), and in our notation we allow for the brackets to be in other orders: in other words, $[Y, [X]] = [[X], Y]$ and $[[X], Y], Z = [Y, [X]], Z = Z, [Y, [X]]$, etc.

EXAMPLE 2.38 The multiset p, q, r is of kind $[p, q], p, q, r$, as is p, p, q, r and p, q, q, r. The multisets of kind $[[p], p, q, r], p, r$ are

$$p, p, p, q, r, r \quad p, p, p, q, r \quad p, p, q, r, r \quad p, p, q, r \quad p, q, r, r \quad p, q, r$$

Given this notation for contracted multisets, we can specify the rules sequent systems with contraction folded into the rules. These rules are in Figure 29. Many of the rules are specified using contracted multisets in the conclusions. In these cases, the rules have more than one possible conclusion. You get a different instance of the rule for each different choice of a contracted multiset for the left hand side and right hand side of the sequent in the conclusion. For example, given the premise sequent $\neg p \succ p, \neg p$, these are both instances of $[\neg L']$, where the introduced formula is $\neg p$.

$$\dfrac{\neg p \succ p, \neg p}{\neg p, \neg p \succ \neg p}{\scriptstyle \neg L'} \qquad \dfrac{\neg p \succ p, \neg p}{\neg p \succ \neg p}{\scriptstyle \neg L'}$$

$$p \succ p \; Id \qquad \frac{X \succ Y}{X, X' \succ Y, Y'} \; K$$

$$\frac{X, A \succ Y}{[A \wedge B], X \succ Y} \wedge L_1' \qquad \frac{X, B \succ Y}{[A \wedge B], X \succ Y} \wedge L_2' \qquad \frac{X \succ A, Y \qquad X \succ B, Y}{X \succ [A \wedge B], Y} \wedge R'$$

$$\frac{X, A \succ Y \qquad X, B \succ Y}{[A \vee B], X \succ Y} \vee L' \qquad \frac{X \succ A, Y}{X \succ [A \vee B], Y} \vee R_1' \qquad \frac{X \succ B, Y}{X \succ [A \vee B], Y} \vee R_2'$$

$$\frac{X, A, B \succ Y}{X, [A \otimes B] \succ Y} \otimes L' \qquad \frac{X \succ A, Y \quad X' \succ B, Y'}{[X, X'] \succ [[A \otimes B], Y, Y']} \otimes R'$$

$$\frac{X, A \succ Y \quad X', B \succ Y'}{[X, X', [A \oplus B]] \succ [Y, Y']} \oplus L' \qquad \frac{X \succ A, B, Y}{X \succ [A \oplus B], Y} \oplus R'$$

$$\frac{X \succ A, Y \quad B, X' \succ Y'}{[[A \to B], X, X'] \succ [Y, Y']} \to L' \qquad \frac{X, A \succ B, Y}{X \succ [A \to B], Y} \to R'$$

$$\frac{X \succ A, Y}{X, [\neg A] \succ Y} \neg L' \qquad \frac{X, A \succ Y}{X \succ [\neg A], Y} \neg R'$$

$$\frac{X \succ Y}{X, t \succ Y} \; tL \qquad \succ t \; tR \qquad X \succ \top, Y \; \top R$$

$$f \succ \; fL \qquad \frac{X \succ Y}{X \succ f, Y} \; fR \qquad X, \bot \succ Y \; \bot L$$

Figure 29: SEQUENT RULES WITH IMPLICIT CONTRACTION

In the first case, there is no contraction applied, in the second, one instance of $\neg p$ is contracted on the left (as allowed in the specification of $[\neg L']$).

A crucial feature of these rules is that while contractions are applied, they are not applied *too much*.

LEMMA 2.39 [FINITELY MANY POSSIBLE PARENTS] *Each sequent* $X \succ Y$ *has only finitely many possible parents in the sequent rules with implicit contraction.*

Proof: For each rule under which the conclusion might fall, there are only finitely many ways to split the conclusion up in order to find possible premises. Take, $[\otimes R']$ for example, and suppose the conclusion sequent has the shape $U \succ A \otimes B, V$. The possible premises have the form $X \succ A, Y$ and $X' \succ B, Y'$ for different choices of X, X', Y, Y', where $U \subseteq X \cup X'$, $X \subseteq U$ and $X' \subseteq U$, and $V \backslash (A \otimes B) \subseteq Y \cup Y'$, $Y \subseteq V$ and $Y' \subseteq V$. Given that U and V are finite multisets, there are finitely many choices for U and V in this case, and in the same way, in every other. ∎

EXAMPLE 2.40 Given the sequent $p, q \succ r \otimes s$, the possible pairs of parents are:

You can see why I won't try drawing a tree for possible ancestry of *any* sequent in this system. It's just *too* large.

$p \succ r$ and $q \succ s$,	$p \succ s$ and $q \succ r$,
$p, q \succ r$ and $q \succ s$,	$p, q \succ s$ and $q \succ r$,
$q \succ r$ and $p, q \succ s$,	$q \succ s$ and $p, q \succ r$,
$p, q \succ r$ and $p, q \succ s$,	
$p \succ r, r \otimes s$ and $q \succ s$,	$p \succ s$ and $q \succ r, r \otimes s$,
$p, q \succ r, r \otimes s$ and $q \succ s$,	$p, q \succ s$ and $q \succ r, r \otimes s$.
$q \succ r, r \otimes s$ and $p, q \succ s$,	$q \succ s$ and $p, q \succ r, r \otimes s$,
$p, q \succ r, r \otimes s$ and $p, q \succ s$,	
$p \succ r$ and $q \succ s, r \otimes s$,	$p \succ s, r \otimes s$ and $q \succ r$,
$p, q \succ r$ and $q \succ s, r \otimes s$,	$p, q \succ s, r \otimes s$ and $q \succ r$,
$q \succ r$ and $p, q \succ s, r \otimes s$,	$q \succ s$ and $p, q \succ r$,
$p, q \succ r$ and $p, q \succ s, r \otimes s$,	
$p \succ r, r \otimes s$ and $q \succ s, r \otimes s$,	$p \succ s, r \otimes s$ and $q \succ r, r \otimes s$,
$p, q \succ r, r \otimes s$ and $q \succ s, r \otimes s$,	$p, q \succ s, r \otimes s$ and $q \succ r, r \otimes s$,
$q \succ r, r \otimes s$ and $p, q \succ s, r \otimes s$,	$q \succ s, r \otimes s$ and $p, q \succ r, r \otimes s$,
$p, q \succ r, r \otimes s$ and $p, q \succ s, r \otimes s$.	

There are 28 pairs of possible parents for this one sequent, following the $[\otimes R']$ rule. That is certainly much more than for rules where we *haven't* included contraction, but it is still finite.

It follows from this that the tree for possible ancestry is *finitely branching*. From any sequent you can trace only a finite number of possible parents. However, it is not the case that the possible ancestry of a sequent must be finite. Here is a fragment of the possible ancestry of $p \succ p \otimes p, p$, using our rules:

$$\dfrac{\dfrac{p \succ p \otimes p, p \qquad p \succ p \otimes p, p}{p \succ p \otimes p}\ {\otimes R'}}{p \succ p \otimes p, p}\ K$$

It follows that the *full* possible ancestry of $p \succ p \otimes p, p$ is infinite, for one of the ways we can derive $p \succ p \otimes p, p$ is from a pair of derivations of $p \succ p \otimes p, p$. We can stack derivations to an arbitrary depth – there is no limit on how long they might be. Clearly, when searching for a derivation for $X \succ Y$, if we find $X \succ Y$ in its own ancestry, we don't need to pursue *that* branch any further. There is no need to chase our own tail in the search for a derivation for $X \succ Y$. If all derivations of $X \succ Y$ went through earlier derivations of $X \succ Y$, this sequent wouldn't have *any* derivations.

But there are more ways to chase your tail than going around in a circle. We might also go around in an ever increasing *spiral*. Perhaps in my search for a derivation of $X, X' \succ Y, Y'$, I find that I could have done it by way of a derivation for $X, X', X' \succ Y, Y', Y'$. (We have already eliminated an *explicit* appeal to contraction in our rules, but we may still be able to mimic it through the contraction implicit in our connective rules.) We want to avoid having to derive $X, X' \succ Y, Y'$ through a derivation of a more complex sequent $X, X', X' \succ Y, Y', Y'$. To avoid this, we wish to use a stronger restriction on derivations than concision (Definition 2.35).

DEFINITION 2.41 [SUCCINCT DERIVATIONS] A derivation δ is SUCCINCT iff no branch of the tree that contains a sequent $X, X' \succ Y, Y'$ earlier contains a sequent $X, X', X' \succ Y, Y', Y'$ from which it could have been contracted.

To show that we can avoid searching for derivations that fail to be succinct, we prove the following lemma:

Anderson and Belnap call this Curry's Lemma, after Curry's 1950 proof of the result for classical and intuitionistic sequent systems [44].

LEMMA 2.42 [CURRY'S LEMMA] *In any system $\mathfrak{S}^W$, if a sequent $X, X', X' \succ Y, Y', Y'$ has a derivation with height n, then $X, X' \succ Y, Y'$ has a derivation with height $\leq n$.*

Proof: This is a straightforward induction on the length of the derivation of $X, X', X' \succ Y, Y', Y'$. If the sequent $X, X', X' \succ Y, Y', Y'$ is an axiom, its contraction $X, X' \succ Y, Y'$ is also an axiom. Suppose the result holds for derivations with height less than m and that $X, X', X' \succ Y, Y', Y'$ has a derivation of height m. If the formulas both occurrences of X' and Y' are all passive in the last step of the derivation, then the contraction could have occurred at the premise of the derivation, unless some components occurred in one premise and the others, in the other premise of the inference step. In that case, those premises may be contracted in this inference step. The only remaining case to consider is when one formula in X' or in Y' must is active in the final step of the conclusion. In that case, we are permitted to contract *that* instance as well in this inference step. This completes the proof. ■

So, we can restrict our attention to succinct derivations.

THEOREM 2.43 [DERIVATIONS CAN BE MADE SUCCINCT] *In any system* $\mathfrak{S}^W$, *if a sequent* $X \succ Y$ *has a derivation, it also has a succinct derivation.*

Proof: Given any branch of the derivation including a sequent $U, U', U' \succ V, V', V'$ and later, its contraction $U, U' \succ V, V'$, by Curry's Lemma, the derivation of $U, U', U' \succ V, V', V'$ can be transformed into a derivation of $U, U' \succ V, V'$ of no greater height. Replace the larger derivation of $U, U' \succ V, V'$ with this new, smaller derivation. Continue the process in our derivation, until all failures of succinctness are dealt with. ■

If I have a sequent $X \succ Y$ and I want to check it for derivability, I need only search the succinct possible ancestry. When I consider the sequents that could occur in a succinct derivation for $X \succ Y$, I need not worry about an unending sequence of larger and larger sequents, tracing contraction steps in reverse.

DEFINITION 2.44 [SUCCINCT POSSIBLE ANCESTRY] Given any sequent, its SUCCINCT POSSIBLE ANCESTRY is the tree constructed in the following way: start with the sequent itself (the root of the tree), and add a branch to the roots of the trees consisting of the succinct possible ancestry of the possible parents of that sequent. This tree is succinct (no branch contains a sequent and then, later a sequent from which is contracted) except concerning the root itself, which is new. Prune the tree by lopping off any branch at the point at which it contains a sequent from which the sequent at the root of the tree. The resulting tree is now the succinct possible ancestry of the starting sequent.

We have already shown that this tree is finitely branching. It remains to show that each branch is finite. This result was first proved by Saul Kripke in the late 1950s for the case of the sequent calculus for the implicational fragment of the relevant logic R [122], and so this result has his name:

LEMMA 2.45 [KRIPKE'S LEMMA] *Given a set* $\mathcal{S}$ *of sequents, each of which are comprised of formulas from some given finite set, and none of which is a contraction of any other sequent in the set,* $\mathcal{S}$ *is finite.*

Robert K. Meyer noticed, years later [145], that Kripke's Lemma follows from Dickson's Lemma, a result in number theory [53].

LEMMA 2.46 [DICKSON'S LEMMA] *For any infinite set* S *of* n*-tuples of natural numbers, there are at least two tuples* $(a_1, \ldots, a_n)$ *and* $(b_1, \ldots, b_n)$ *where* $a_i \leq b_i$ *for each* i.

That Kripke's Lemma follows from Dickson's is immediate.

Proof: Given any set $\mathcal{S}$ of sequents, comprised from formulas from some given finite set, partition it into finitely many classes, where each two sequents $X \succ Y$ and $X' \succ Y'$ are in the same class if and only if X and X' contain the same formulas (with possibly different repetitions), and the same holds for Y and Y'. (There are only finitely many such classes since there are only finitely many formulas out of which each sequent can be constructed.) For each such class, apply Dickson's Lemma in the following way: if the class is contains sequents made up from formulas in $A_1, \ldots, A_n$ on the left and $B_1, \ldots, B_m$ on the right, assign the tuple $(j_1, \ldots, j_n, k_1, \ldots, k_m)$ to the sequent containing j_i repetitions of A_i and k_i repetitions of B_i. There are infinitely many tuples if and only if there are infinitely many sequents in this class. By Dickson's Lemma, if there are infinitely many tuples, then we have some pair of tuples $(j_1, \ldots, j_n, k_1, \ldots, k_m)$ and $(j'_1, \ldots, j'_n, k'_1, \ldots, k'_m)$ where $j_i \leq j'_i$ and $k_i \leq k'_i$ for each i, but this means that the sequent corresponding to the first tuple is a contraction of the sequent corresponding to the second. Given that there are only finitely many equivalence classes, the only way that the entire set could be infinite is if one class (at least), was infinite. This completes the proof of Kripke's Lemma. ■

So, to prove Kripke's Lemma, we can simply refer to the general result of Dickson's Lemma and leave things at that. However, it is relatively straightforward to prove Dickson's Lemma, so if you are interested in a little bit of number theory, here is a proof:

Proof: A straightforward proof of Dickson's Lemma uses the fact that any infinite sequence $n_0, n_1, \ldots, n_i, \ldots$ of natural numbers has some *non-decreasing* infinite subsequence – that is, there is a selection of indices $i_0 < i_1 < i_2 < \cdots$ such that $n_{i_0} \leq n_{i_1} \leq n_{i_2} \leq \cdots$ – the subsequence $n_{i_0}, n_{i_1}, n_{i_2}, \ldots$ is never decreasing.

Why is there always such a subsequence? If the sequence is bounded above, then only finitely many numbers occur in the sequence, so at least one number occurs infinitely many times. Pick the constant subsequence consisting of one such number. On the other hand, if the sequence is unbounded, then define the sequence by setting $n_{i_0} = n_0$, and given n_{i_j}, for $n_{i_{j+1}}$, select the next item in the original sequence larger than n_{i_j}. Since the sequence is unbounded, there is always such a number. In either case, we have a non-descending infinite subsequence of our original sequence.

This reasoning is not at all constructive. It gives you no insight into what to do if you have not yet been able to verify that the original sequence is bounded above – and how *could* you verify this if I merely give you the sequence one item at a time? Consider the case where I feed you the sequence consisting of $1, 2, \ldots, n$ (for some *very* large number n) and only then continue with $0, 0, 0, \ldots$ The mathematics of providing an algorithm for *selecting* a subsequence is subtle [32, 159]. Here is one way to do it. Simply write down the original sequence, one item at a time. Whenever you write down a new item n_i in the sequence, simply cross out any numbers that you have written that are *larger* than n_i. The *result* (the numbers that are never crossed out) will form non-descending sequence. Showing that this sequence is infinitely long is the tricky part!

Now, consider the our set S of n-tuples, and represent it as a sequence S_0, in some arbitrarily chosen order. We can define the sequence S_1 as the infinite subsequence of S_0 in which the *first* element of each tuple never decreases from one tuple to the next. There is always such an infinite subsequence, applying our lemma to the sequence consisting of the first element of each tuple. Continue for each position in the tuples in the sequences. That is, given that S_i defined so that the first 1 to i elements of each tuple are non-decreasing from one item to the next in the infinite sequence, define S_{i+1} as the infinite subsequence of the sequence S_i of tuples where the $(i+1)$st element of each tuple never decreases from one tuple to the next. The *final* sequence S_n is an infinite series of n-tuples where each tuple is dominated by each later tuple, so the first tuple $(a_1, \ldots, a_n)$ and second tuple $(b_1, \ldots, b_n)$ in the list are such that $a_i \leq b_i$ for each i, and Dickson's Lemma is proved. ■

This completes all the components we need to show that all of our sequent systems are decidable.

THEOREM 2.47 [ALL SEQUENT SYSTEMS ARE DECIDABLE] *Any of our sequent systems (multiple conclusion or single conclusion). In particular, given any sequent* $X \succ Y$, *its possible ancestry (its* succinct *possible ancestry, in the case of systems with contraction) is finite.*

Proof: The proof takes the same shape as the proof for Theorem 2.34 on page 98, except in the presence of contraction, we use the *succinct* possible ancestry, terminating branches instead of adding nodes which are merely expansions of sequents we have already seen in the tree.

By Kripke's Lemma, the succinct possible ancestry is finite. Passing from the leaves to the root in the manner of the proof of Theorem 2.34, we have an algorithm for determining, for each sequent in the tree, whether it is derivable or not. ■

This result shows that a our sequent systems have the wherewithal to give us an algorithm for determining derivability. Producing a succinct possible ancestry for a sequent is not the most *efficient* way to test for derivability. There are many techniques for making derivation search more tractable, and the discipline of automated theorem proving is thriving [21, 50, 73, 238].

» «

The elimination of *Cut* is useful for more than just limiting the search for derivations. The fact that any derivable sequent has a *Cut*-free derivation has other consequences. One consequence is the fact of *interpolation*.

THEOREM 2.48 [INTERPOLATION FOR SIMPLE SEQUENTS] *If* $A \succ B$ *is derivable in the simple sequent system, then there is a formula* C *containing only atoms present in both* A *and* B *such that* $A \succ C$ *and* $C \succ B$ *are derivable.*

This result tells us that if the sequent $A \succ B$ is derivable then that consequence "factors through" a statement in the vocabulary shared between A and B. This means that the consequence $A \succ B$ not only relies only upon the material in A and B and nothing *else* (that is due to the availability of a *Cut*-free derivation) but also in some sense the derivation 'factors through' the material in common between A and B. The result is a straightforward consequence of the *Cut*-elimination theorem. A *Cut*-free derivation of $A \succ B$ provides us with an interpolant.

Proof: We prove this by induction on the construction of the derivation of $A \succ B$. We keep track of the interpolant with these rules:

$$p \succ_p p \; \mathit{Id}$$

$$\dfrac{A \succ_C R}{A \land B \succ_C R}\land L_1 \qquad \dfrac{A \succ_C R}{B \land A \succ_C R}\land L_2 \qquad \dfrac{L \succ_{C_1} A \qquad L \succ_{C_2} B}{L \succ_{C_1 \land C_2} A \land B}\land R$$

$$\dfrac{A \succ_{C_1} R \qquad B \succ_{C_2} R}{A \lor B \succ_{C_1 \lor C_2} R}\lor L \qquad \dfrac{L \succ_C A}{L \succ_C A \lor B}\lor R_1 \qquad \dfrac{L \succ_C A}{L \succ_C B \lor A}\lor R_2$$

We show by induction on the length of the derivation that if we have a derivation of $L \succ_C R$ then $L \succ C$ and $C \succ R$ and the atoms in C present in both L and in R. These properties are satisfied by the atomic sequent $p \succ_p p$, and it is straightforward to verify them for each of the rules. ∎

EXAMPLE 2.49 [A DERIVATION WITH AN INTERPOLANT] Take the sequent $p \land (q \lor (r_1 \land r_2)) \succ (q \lor r_1) \land (p \lor r_2)$. We may annotate a *Cut*-free derivation of it as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{q \succ_q q}{q \succ_q q \lor r}\lor R \qquad \dfrac{r_1 \succ_{r_1} r_1}{r_1 \land r_2 \succ_{r_1} r_1}\land L}{q \lor (r_1 \land r_2) \succ_{q \lor r_1} q \lor r_1}\lor L}{p \land (q \lor (r_1 \land r_2)) \succ_{q \lor r_1} q \lor r_1}\land L \qquad \dfrac{\dfrac{p \succ p}{p \succ_p p \lor r_2}\lor R}{p \land (q \lor (r_1 \land r_2)) \succ_p p \lor r_2}\land L}{p \land (q \lor (r_1 \land r_2)) \succ_{(q \lor r_1) \land p} (q \lor r_1) \land (p \lor r_2)}\land R$$

Notice that the interpolant $(q \lor r_1) \land p$ does not contain r_2, even though r_2 is present in both the antecedent and the consequent of the sequent. This tells us that r_2 is doing no 'work' in this derivation. Since we have

$$p \land (q \lor (r_1 \land r_2)) \succ (q \lor r_1) \land p, \qquad (q \lor r_1) \land p \succ (q \lor r_1) \land (p \lor r_2)$$

We can replace the r_2 in either derivation with another statement – say r_3 – preserving the structure of each derivation. We get the more general fact:

$$p \land (q \lor (r_1 \land r_2)) \succ (q \lor r_1) \land (p \lor r_3)$$

We can extend interpolation to complex sequent systems, too, though this takes a little more work, since in these derivations, formulas can switch sides in sequents. To prove interpolation, we will prove a stronger hypothesis, according to which the *splitting* of the sequent may be independent of the division between left and right. Here is the target result:

THEOREM 2.50 [SPLITTING FOR COMPLEX SEQUENTS] *For any of our sequent systems, given any derivable sequent* $X, X' \succ Y, Y'$, *we can find a formula* I *where* (1) $X \succ I, Y$ *and* $X', I \succ Y'$ *are derivable, and* (2) I *is a formula whose atoms occur both in* $X \cup Y$ *and* $X' \cup Y'$.

To prove this, we will make use of a family of *split sequent* rules. These rules generalise the rules of the sequent calculus, to operate on *pairs* of multisets of formulas on the left and right, and each sequent will be subscripted with an interpolating formula. So, split sequents have this form:

$$X; X' \succ_I Y; Y'$$

and the intended interpretation is that $X \succ I, Y$ and $X', I \succ Y'$ are both derivable – and that I a formula whose atoms occur both in $X \cup Y$ and $X' \cup Y'$. Figure 210 gives the axioms and rules for the splitting sequent system.

A derivation in the split sequent system is a tree of sequents, starting with axioms, and developed according to the rules, in the usual fashion.

lemma 2.51 [splitting sequents interpolate] *If* $X; X' \succ_I Y; Y'$ *can be derived in the split sequent system with some choice of structural rules (from* K *and* W*), then* $X \succ I, Y$ *and* $X', I \succ Y'$ *are both derivable in the sequent system with those structural rules, and furthermore,* I *is formed from atoms which are shared between* $X \cup Y$ *and* $X' \cup Y'$.

Proof: This is a straightforward induction on the length of the split sequent derivation. Space does not permit checking each of the rules here (there are many), but here is an indicative sample to show how to perform the verifications.

identities: We have $p; \succ_p ; p$ since $p \succ p$ and $p \succ p$ are derivable. (p is in the shared vocabulary of p and p), and we have $p; \succ \bot, p;$ since $p \succ \bot, p$, $\bot \succ$ are both derivable. ($\bot$ has no atoms, so it is in every vocabulary.) Similarly, we have $; p \succ_{\neg p} p;$ since $\succ \neg p, p$, and $p, \neg p \succ$ are both derivable. ($\neg p$ has atoms from the shared vocabulary of p and p.) And finally, we have $; p \succ_\top ; p$ since $\succ \top$ and $p, \top \succ p$ are both derivable. ($\top$ has no atoms, so it is in every vocabulary.)

weakening and contraction: Weakening and contraction do not modify the interpolating formula. The premise of the weakening rule is $X_1; X_1' \succ_I Y_1; Y_1'$, so we have derivations for $X_1 \succ I, Y_1$ and $X_1', I \succ Y_1'$ (and I is the vocabulary shared between $X_1 \cup Y_1$ and $X_1' \cup Y_1'$.) By weakening these underlying sequents, we have derivations for $X_1, X_2 \succ I, Y_1, Y_2$ and $X_1', X_2', I \succ Y_1', Y_2'$ (and I remains in the shared vocabulary), so we indeed have the conclusion of the splitting rule for weakening: $X_{12}; X_{12}' \succ_I Y_{12}; Y_{12}'$. The verification for contraction has exactly the asme form.

lattice connectives: Let's check the $[\wedge L]$ rules, and $[\wedge R]$ rules. For $[\wedge L;]$, if we have $A, X; X' \succ_I Y; Y'$ we have $A, X \succ I, Y$ and hence $A \wedge B, X \succ I, Y$, and since $X' \succ I, Y'$, we have $A \wedge B, X; X' \succ_I Y; Y'$. (Since the interpolant doesn't change, it remains in the shared vocabulary.) For $[;\wedge L]$, if we have $X; X', A \succ_I Y; Y'$, then we have $X \succ I, Y$, and $X', A, I \succ Y'$, and hence $X', A \wedge B, I \succ Y'$, and so, $X; X', A \wedge B \succ_I Y; Y'$. (And again, the interpolant doesn't change, so it remains in the shared vocabulary.)

$$p; \succ_{p} ;p \;{\scriptstyle Id} \qquad p; \succ_{\bot} p; \;{\scriptstyle Id} \qquad ;p \succ_{\neg p} p; \;{\scriptstyle Id} \qquad ;p \succ_{\top} ;p \;{\scriptstyle Id}$$

$$X;X' \succ_{\bot} \top, Y;Y' \;{\scriptstyle \top R;} \qquad X;X' \succ_{\top} Y;Y', \top \;{\scriptstyle ;\top R} \qquad \bot, X;X' \succ_{\bot} Y;Y' \;{\scriptstyle \bot L;} \qquad X;X', \bot \succ_{\top} Y;Y' \;{\scriptstyle ;\bot L}$$

$$\frac{A, X;X' \succ_{I} Y;Y'}{A \land B, X;X' \succ_{I} Y;Y'}\;{\scriptstyle \land L_1;} \qquad \frac{X;X', A \succ_{I} Y;Y'}{X;X', A \land B \succ_{I} Y;Y'}\;{\scriptstyle ;\land L_1} \qquad \frac{B, X;X' \succ_{I} Y;Y'}{A \land B, X;X' \succ_{I} Y;Y'}\;{\scriptstyle \land L_2;} \qquad \frac{X;X', B \succ_{I} Y;Y'}{X;X', A \land B \succ_{I} Y;Y'}\;{\scriptstyle ;\land L_2}$$

$$\frac{X;X' \succ_{I} A, Y;Y' \quad X;X' \succ_{J} B, Y;Y'}{X;X' \succ_{I \lor J} A \land B, Y;Y'}\;{\scriptstyle \land R;} \qquad \frac{X;X' \succ_{I} Y;Y', A \quad X;X' \succ_{J} Y;Y', B}{X;X' \succ_{I \land J} Y;Y', A \land B}\;{\scriptstyle ;\land R}$$

$$\frac{A, X;X' \succ_{I} Y;Y' \quad B, X;X' \succ_{J} Y;Y'}{A \lor B, X;X' \succ_{I \lor J} Y;Y'}\;{\scriptstyle \lor L;} \qquad \frac{X;X', A \succ_{I} Y;Y' \quad X;X', A \succ_{J} Y;Y'}{X;X', A \lor B \succ_{I \land J} Y;Y'}\;{\scriptstyle ;\lor L}$$

$$\frac{X;X' \succ_{I} A, Y;Y'}{X;X' \succ_{I} A \lor B, Y;Y'}\;{\scriptstyle \lor R_1;} \qquad \frac{X;X' \succ_{I} Y;Y', A}{X;X' \succ_{I} Y;Y';A \lor B}\;{\scriptstyle ;\lor R_1} \qquad \frac{X;X' \succ_{I} B, Y;Y'}{X;X' \succ_{I} A \lor B, Y;Y'}\;{\scriptstyle \lor R_2;} \qquad \frac{X \succ_{I} Y;Y', B}{X \succ_{I} Y;Y', A \lor B}\;{\scriptstyle ;\lor R_2}$$

$$\frac{A, B, X;X' \succ_{I} Y;Y'}{A \otimes B, X;X' \succ_{I} Y;Y'}\;{\scriptstyle \otimes L;} \qquad \frac{X;X', A, B \succ_{I} Y;Y'}{X;X', A \otimes B \succ_{I} Y;Y'}\;{\scriptstyle ;\otimes L}$$

$$\frac{X_1;X_1' \succ_{I} A, Y_1;Y_1' \quad X_2;X_2' \succ_{J} B, Y_2;Y_2'}{X_{1,2};X_{1,2}' \succ_{I \oplus J} A \otimes B, Y_{1,2};Y_{1,2}'}\;{\scriptstyle \otimes R;} \qquad \frac{X_1;X_1' \succ_{I} Y_1;Y_1', A \quad X_2;X_2' \succ_{J} Y_2;Y_2', B}{X_{1,2};X_{1,2}' \succ_{I \otimes J} Y_{1,2};Y_{1,2}', A \otimes B}\;{\scriptstyle ;\otimes R}$$

$$\frac{A, X_1;X_1' \succ_{I} Y_1;Y_1' \quad B, X_2;X_2' \succ_{J} Y_2;Y_2'}{A \oplus B, X_{1,2};X_{1,2}' \succ_{I \oplus J} Y_{1,2};Y_{1,2}'}\;{\scriptstyle \oplus L;} \qquad \frac{X_1;X_1', A \succ_{I} Y_1;Y_1' \quad X_2;X_2', B \succ_{J} Y_2;Y_2'}{X_{1,2};X_{1,2}', A \oplus B \succ_{I \otimes J} Y_{1,2};Y_{1,2}'}\;{\scriptstyle ;\oplus L}$$

$$\frac{X;X' \succ_{I} A, B, Y;Y'}{X;X' \succ_{I} A \oplus B, Y;Y'}\;{\scriptstyle \oplus R;} \qquad \frac{X;X' \succ_{I} Y;Y', A, B}{X;X' \succ_{I} Y;Y', A \oplus B}\;{\scriptstyle ;\oplus R}$$

$$\frac{X_1;X_1' \succ_{I} A, Y_1;Y_1' \quad B, X_2;X_2' \succ_{J} Y_2;Y_2'}{A \to B, X_{1,2};X_{1,2}' \succ_{I \oplus J} Y_{1,2};Y_{1,2}'}\;{\scriptstyle \to L;} \qquad \frac{X_1;X_1' \succ_{I} Y_1;Y_1', A \quad X_2;X_2', B \succ_{J} Y_2;Y_2'}{X_{1,2};X_{1,2}', A \to B \succ_{I \otimes J} Y_{1,2};Y_{1,2}'}\;{\scriptstyle ;\to L}$$

$$\frac{A, X;X' \succ_{I} B, Y;Y'}{X;X' \succ_{I} A \to B, Y;Y'}\;{\scriptstyle \to R;} \qquad \frac{X;X', A \succ_{I} Y;Y', A}{X;X' \succ_{I} Y;Y', A \to B}\;{\scriptstyle ;\to R}$$

$$\frac{X;X' \succ_{I} A, Y;Y'}{\neg A, X;X' \succ_{I} Y;Y'}\;{\scriptstyle \neg L;} \qquad \frac{X;X' \succ_{I} Y;Y', A}{X;X', \neg A \succ_{I} Y;Y'}\;{\scriptstyle ;\neg L} \qquad \frac{A, X;X' \succ_{I} Y;Y'}{X;X' \succ_{I} \neg A, Y;Y'}\;{\scriptstyle \neg R;} \qquad \frac{X;X', A \succ_{I} Y;Y'}{X;X' \succ_{I} Y;Y', \neg A}\;{\scriptstyle ;\neg R}$$

$$\frac{X;X' \succ_{I} Y, Y'}{t, X;X' \succ_{I} Y, Y'}\;{\scriptstyle tL;} \qquad \frac{X;X' \succ_{I} Y, Y'}{X;X', t \succ_{I} Y, Y'}\;{\scriptstyle ;tL} \qquad \succ_{f} t; \;{\scriptstyle tR;} \qquad \succ_{t} ;t \;{\scriptstyle ;tR}$$

$$f; \succ_{f} \;{\scriptstyle fL;} \qquad ;f \succ_{t} \;{\scriptstyle ;fL} \qquad \frac{X;X' \succ_{I} Y;Y'}{X;X' \succ_{I} f, Y;Y'}\;{\scriptstyle fR;} \qquad \frac{X;X' \succ_{I} Y;Y'}{X;X' \succ_{I} Y;Y', f}\;{\scriptstyle ;fR}$$

Figure 210: SPLITTING RULES FOR CONNECTIVES

$$\frac{X_1; X_1' \succ_I Y_1; Y_1'}{X_1, X_2; X_1', X_2' \succ_I Y_1, Y_2; Y_1', Y_2'}\,K \qquad \frac{X_1, X_2, X_2; X_1', X_2', X_2' \succ_I Y_1, Y_2, Y_2; Y_1', Y_2', Y_2'}{X_1, X_2; X_1', X_2' \succ_I Y_1, Y_2; Y_1', Y_2'}\,W$$

Figure 211: split structural rules

For [$\land R$;], if we have $X; X' \succ_I A, Y; Y'$ and $X; X' \succ_J B, Y; Y'$, then we can reason as follows:

$$\frac{\dfrac{X \succ I, A, Y}{X \succ I \lor J, A, Y}\,\lor R \quad \dfrac{X \succ J, B, Y}{X \succ I \lor J, B, Y}\,\lor R}{X \succ I \lor J, A \land B, Y}\,\land R \qquad \frac{X', I \succ Y' \quad X', J \succ Y'}{X', I \lor J \succ Y'}\,\lor L$$

so we have $X \succ I \lor J, A \land B, Y$ and $X', I \lor J \succ Y'$ and hence $X; X' \succ_{I \lor J} A \land B, Y; Y'$ is derivable. And since I is in the vocabulary shared between A, X, Y and X', Y', and J is in the vocabulary shared between B, X, Y and X', Y', it follows that $I \lor J$ is in the vocabulary shared between $A \lor B, X, Y$ and X', Y', Z'.

For [$\land R$;], if we have $X; X' \succ_I Y; Y', A$ and $X; X' \succ_J Y; Y', B$, then we can reason as follows:

$$\frac{X \succ I, Y \quad X \succ J, Y}{X \succ I \land J, Y}\,\land R \qquad \frac{\dfrac{X', I \succ Y', A}{X', I \land J \succ Y', A}\,\land L \quad \dfrac{X', J \succ Y', B}{X', I \land J \succ Y', B}\,\land L}{X', I \land J \succ Y', A \land B}\,\land R$$

so we have $X; X' \succ_{I \lor J} Y; Y', A \land B$ as desired. And since I is in the vocabulary shared between A, X, Y and X', Y', and J is in the vocabulary shared between B, X, Y and X', Y', it follows that $I \land J$ is in the vocabulary shared between $A \lor B, X, Y$ and X', Y', Z'.

multiplicative connectives: We'll check the conditional rules. (Fission and fusion are similar.) For [$\to L$;], if we have derivations of $X_1 \succ I, A, Y_1$ and $B, X_2 \succ J, Y_2$ and $X_1', I \succ Y_1'$ and $X_2', J \succ Y_2'$, then we can reason as follows:

$$\frac{\dfrac{X_1 \succ I, A, Y_1 \quad B, X_2 \succ J, Y_2}{A \to B, X_{12} \succ I, J, Y_{12}}\,\to L}{A \to B, X_1, X_2 \succ I \oplus J, Y_{12}}\,\oplus R \qquad \frac{X_1', I \succ Y_1' \quad X_2', J \succ Y_2'}{X_{12}', I \oplus J \succ Y_{12}'}\,\oplus L$$

and similarly, for [;$\to L$], if we have derivations of $X_1 \succ I, Y_1$, $X_2 \succ J, Y_2$, $X_1', I \succ A, Y_1'$ and $B, X_2', J \succ Y_2'$, we have:

$$\frac{X_1 \succ I, Y_1 \quad X_2 \succ J, Y_2}{X_{12} \succ I \otimes Y_{12}}\,\otimes R \qquad \frac{\dfrac{X_1', I \succ A, Y_1' \quad B, X_2', J \succ Y_2'}{A \to B, X_{12}', I, J \succ Y_{12}'}\,\to L}{A \to B, X_{12}', I \otimes J \succ Y_{12}'}\,\otimes L$$

and the right conditional rules are similarly verified, except the interpolating formula is constant, because we are not combining premise sequents. For [$\rightarrow R$;], if we have derivations for $A, X \succ I, B, Y$ and $X', I \succ Y'$ then to verify the split sequent $X; X' \succ_I A \rightarrow B, Y'; Y'$, we derive $X \succ I, A \rightarrow B, Y'$ and we are done. The same goes for the [;$\rightarrow R$] rule.

NEGATION: The negation rules work on the same principle as the [$\rightarrow R$] rules. The interpolating formula I remains constant as the formula A converted into $\neg A$ remains on the same side of the splitting as it shifts over the turnstile. Here is the case for [$\neg L$;], and the others are identical in form.

$$\dfrac{X \succ I, A, Y}{X, \neg A \succ I, Y}{\scriptstyle \neg L} \qquad X', I \succ Y'$$

■

UNITS: For [fL;], we have derivations of $f \succ f$ (the left splitting) and $f \succ$ (the right). For [; fL], we have $\succ t$ (the left splitting) and $f, t \succ$ (the right). For the [fR] rules, the interpolant is constant from premise to conclusion, and the rule merely inserts an extra f on the right hand side of a sequent (on one side of the splitting or the other), and that is an instance of the [fR] rule of the underlying sequent calculus. (The verification for the t rules the the same form.)

» «

More consequences of *Cut*-elimination and the admissibility of the identity rules Id_A will be considered as the book goes on. Exercises 8–14 ask you to consider different possible connective rules, some of which will admit of *Cut*-elimination and *Id*-admissibility when added, and others of which which will not. In Chapter 4 we will look at reasons why this might help us demarcate definitions of a kind of *properly logical* concept from those which are not logical in that sense.

2.5 | HISTORY

The idea of taking the essence of conjunction and disjunction to be expressed in these sequent rules is to take conjunction and disjunction to form what is known as a *lattice*. A lattice is on *ordered* structure in which we have for every pair of objects a *greatest lower bound* and a *least upper bound*. A *greatest lower bound* of x and y is something below both x and y but which is greatest among such things. A *least upper bound* of x and y is something above both x and y but which is the least among such things. Among statements, taking $\succ$ to be the ordering, $A \wedge B$ is the greatest lower bound of A and of B (since $A \wedge B \succ A$ and $A \wedge B \succ B$, and if $C \succ A$ and $C \succ B$ then $C \succ A \wedge B$) and $A \vee B$ is their least upper bound (for dual reasons).

Here, $\succ$ is the ordering. If $A \succ B$, think of A as occuring 'below' B in the ordering from stronger to weaker.

Well, you need to *squint*, and take A and A' to be *the same* if $A \succ A'$ and $A' \succ A$ to make $A \wedge B$ *the unique* greatest lower bound. If it helps, don't think of the *sentence* A but the *proposition*, where two logically equivalent sentences express the same proposition.

Lattices are wonderful structures, which may be applied in many different ways, not only to logic, but in many other domains as well. Davey and Priestley's *Introduction to Lattices and Order* [49] is an excellent way into the literature on lattices. The concept of a lattice dates to the late 19th Century in the work of Charles S. Peirce and Ernst Schröder, who independently generalised Boole's algebra of propositional logic. Richard Dedekind's work on 'ideals' in algebraic number theory was an independent mathematical motivation for the concept. Work in the area found a focus in the groundbreaking series of papers by Garrett Birkhoff, culminating in a the book *Lattice Theory* [24]. For more of the history, and for a comprehensive state of play for lattice theory and its many applications, George Grätzer's 1978 *General Lattice Theory* [89], and especially its 2003 Second Edition [90] is a good port of call.

We will not study much algebra in this book. However, algebraic technqiues find a very natural home in the study of logical systems. Helena Rasiowa's 1974 *An Algebraic Approach to Non-Classical Logics* [183] was the first look at lattices and other structures as models of a wide range of different systems. For a good guide to why this technique is important, and what it can do, you cannot go past J. Michael Dunn and Gary Hardegree's *Algebraic Methods in Philosophical Logic* [61].

Well, we won't study algebra *explicitly*. Algebraic considerations and sensibilities underly *much* of what will go on. But that will almost always stay under the surface.

The idea of studying derivations consisting of sequents, rather than proofs from premises to conclusions, is entirely due to Gentzen, in his groundbreaking work in proof theory. His motivation was to extend his results on normalisation from what we called the standard natural deduction system to classical logic as well as intuitionistic logic [81, 82]. To do this, it was fruitful to step back from proofs from premises X to a conclusion A to consider statements of the form '$X \succ A$,' making explicit at each step on which premises X the conclusion A depends. Then as we will see in the next chapter, normalisation 'corresponds' in some sense to the elimination of *Cut*s in a derivation. One of Gentzen's great insights was that sequents could be generalised to the form $X \succ Y$ to provide a uniform treatment of traditional Boolean classical logic. We will make much of this connection in the next chapter.

But for more connectives than just the conditional.

Gentzen didn't use the turnstile. His notation was '$\Gamma \rightarrow \mathfrak{A}$'. We use the arrow for a conditional, and the turnstile for a sequent separator.

Gentzen's own sequent calculi did not define lattice logic. They were proof systems for traditional intuitionistic and classical logic, in which the distribution of conjunction over disjunction—that is, $A \wedge (B \vee C) \succ (A \wedge B) \vee (A \wedge C)$—is valid. I have chosen to start with simple sequents for lattice logic for two reasons. First, it makes the procedure for the elimination of *Cut*s much more simple. There are fewer cases to consider and the essential shape of the argument is laid bare with fewer inessential details. Second, once we see the technique applied again and again, it will hopefully reinforce the thought that it is very general indeed. Sequents were introduced as a way of looking at an underlying proof structure. As a pluralist, I take it that there is more than one sort of underlying proof structure to examine, and so, sequents may take more than one sort of shape. Much work has been done recently on *why* Gentzen chose the rules he did for his sequent calculi. I have found papers by Jan von Plato [166, 167] most helpful. Gentzen's papers are available in his collected works [83], and a biography of Gentzen, whose life was cut short in the Second World War, has recently been written [142, 141].

2.6 | EXERCISES

BASIC EXERCISES

Q1 Find a derivation for $p \succ p \land (p \lor q)$ and a derivation for $p \lor (p \land q) \succ p$. Then find a *Cut*-free derivation for $p \lor (p \land q) \succ p \land (p \lor q)$ and compare it with the derivaion you get by joining the two original derivations with a *Cut*.

Q2 Show that there is no *Cut*-free derivation of the following sequents

$$
\begin{aligned}
1 &: p \lor (q \land r) \succ p \land (q \lor r) \\
2 &: p \land (q \lor r) \succ (p \land q) \lor r \\
3 &: p \land (q \lor (p \land r)) \succ (p \land q) \lor (p \land r)
\end{aligned}
$$

Q3 Suppose that there is a derivation of $A \succ B$. Let $C(A)$ be a formula containing A as a subformula, and let $C(B)$ be that formula with the subformula A replaced by B. Show that there is a derivation of $C(A) \succ C(B)$. Furthermore, show that a derivation of $C(A) \succ C(B)$ may be systematically constructed from the derivation of $A \succ B$ together with the context $C(-)$ (the shape of the formula $C(A)$ with a 'hole' in the place of the subformula A).

Q4 Find a derivation of $p \land (q \land r) \succ (p \land q) \land r$. Find a derivation of $(p \land q) \land r \succ p \land (q \land r)$. Put these two derivations together, with a *Cut*, to show that $p \land (q \land r) \succ p \land (q \land r)$. Then eliminate the cuts from this derivation. What do you get?

Q5 Do the same thing with derivations of $p \succ (p \land q) \lor p$ and $(p \land q) \lor p \succ p$. What is the result when you eliminate this cut?

Q6 Show that (1) $A \succ B \land C$ is derivable if and only if $A \succ B$ and $A \succ C$ is derivable, and that (2) $A \lor B \succ C$ is derivable if and only if $A \succ C$ and $B \succ C$ are derivable. Finally, (3) when is $A \lor B \succ C \land D$ derivable, in terms of the derivability relations between A, B, C and D.

Q7 Under what conditions do we have a derivation of $A \succ B$ when A contains only propositional atoms and *disjunctions* and B contains only propositional atoms and *conjunctions*.

Q8 Expand the system with the following rules for the propositional *constants* $\bot$ and $\top$.

$$A \succ \top \quad [\top R] \qquad\qquad \bot \succ A \quad [\bot L]$$

Show that *Cut* is eliminable from the new system. (You can think of $\bot$ and $\top$ as zero-place connectives. In fact, there is a sense in which $\top$ is a zero-place *conjunction* and $\bot$ is a zero-place *disjunction*. Can you see why?)

Q9 Show that simple sequents including $\top$ and $\bot$ are decidable, following Corollary 2.28 and the results of the previous question.

Q10 Show that every formula composed of just $\top$, $\bot$, $\land$ and $\lor$ is *equivalent* to either $\top$ or $\bot$. (What does this result remind you of?)

Q11 Prove the interpolation theorem (Corollary 2.48) for derivations involving $\wedge, \vee, \top$ and $\bot$.

Q12 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \succ R}{A \text{ tonk } B \succ R}\,\text{tonk}\,L \qquad \frac{L \succ B}{L \succ A \text{ tonk } B}\,\text{tonk}\,R$$

What new things can you derive using tonk? Can you derive A tonk $B \succ A$ tonk B? Is *Cut* eliminable for formulas involving tonk?

See Arthur Prior's "The Runabout Inference-Ticket" [176] for tonk's first appearance in print.

Q13 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \succ R}{A \text{ honk } B \succ R}\,\text{honk}\,L \qquad \frac{L \succ A \qquad L \succ B}{L \succ A \text{ honk } B}\,\text{honk}\,R$$

What new things can you derive using honk? Can you derive A honk $B \succ A$ honk B? Is *Cut* eliminable for formulas involving honk?

Q14 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \succ R \qquad B \succ R}{A \text{ plonk } B \succ R}\,\text{plonk}\,L \qquad \frac{L \succ B}{L \succ A \text{ plonk } B}\,\text{plonk}\,R$$

What new things can you derive using plonk? Can you derive A plonk $B \succ A$ plonk B? Is *Cut* eliminable for formulas involving plonk?

INTERMEDIATE EXERCISES

Q15 Give a formal, recursive definition of the *dual* of a sequent, and the *dual* of a derivation, in such a way that the dual of the sequent $p_1 \wedge (q_1 \vee r_1) \succ (p_2 \vee q_2) \wedge r_2$ is the sequent $(p_2 \wedge q_2) \vee r_2 \succ p_1 \vee (q_1 \wedge r_1)$. And then use this definition to prove the following theorem.

THEOREM 2.52 [DUALITY FOR DERIVATIONS] *A sequent* $A \succ B$ *is derivable if and only if its dual* $(A \succ B)^d$ *is derivable. Furthermore, the dual of the derivation of* $A \succ B$ *is a derivation of the dual of* $A \succ B$.

Q16 Even though the distribution sequent $p \wedge (q \vee r) \succ (p \wedge q) \vee r$ is not derivable (Example 2.10), some sequents of the form $A \wedge (B \vee C) \succ (A \wedge B) \vee C$ *are* derivable. Give an independent characterisation of the triples $\langle A, B, C\rangle$ such that $A \wedge (B \vee C) \succ (A \wedge B) \vee C$ is derivable.

Q17 Prove the invertibility result of Theorem 2.32 without appealing to the *Cut* rule or to *Cut*-elimination. (HINT: if a sequent $A \vee B \succ C$ has a derivation δ, consider the instances of $A \vee B$ 'leading to' the instance of $A \vee B$ in the conclusion. How does $A \vee B$ appear first in the derivation? Can you change the derivation in such a way as to make it derive $A \succ C$? Or to derive $B \succ C$ instead? Prove this, and a similar result for $\wedge L$.)

ADVANCED EXERCISES

Q18 Define a notion of reduction for simple sequent derivations parallel to the definition of reduction of natural deduction proofs in Chapter 1. Show that it is strongly normalising and that each derivation reduces to a unique *Cut*-free derivation.

Q19 Define *terms* corresponding to simple sequent derivations, in an analogy to the way that λ-terms correspond to natural deduction proofs for conditional formulas. For example, we may annotate each derivation with *terms* in the following way:

$$p \succ_x p \ \ \mathit{Id} \qquad \frac{L \succ_f A \quad A \succ_g R}{L \succ_{f \circ g} R}\ \mathit{Cut}$$

$$\frac{A \succ_f R}{A \wedge B \succ_{l[f]} R}\ {\wedge}L_1 \qquad \frac{B \succ_f R}{A \wedge B \succ_{r[f]} R}\ {\wedge}L_2 \qquad \frac{L \succ_f A \quad L \succ_g B}{L \succ_{f \| g} A \wedge B}\ {\wedge}R$$

where x is an atomic term (of type p), f and g are terms, $l[\ ]$ and $r[\ ]$ are one-place term constructors and $\|$ is a two-place term constructor (of a kind of parallel composition), and $\circ$ is a two-place term constructor (of *serial composition*). Define similar term constructors for the disjunction rules.

Then reducing a *Cut* will correspond to simplifying terms by eliminating serial composition. A *Cut* in which $A \wedge B$ is active will take the following form of reduction:

$$(f \| g) \circ l[h] \text{ reduces to } f \circ h \qquad (f \| g) \circ r[h] \text{ reduces to } g \circ h$$

Fill out all the other reduction rules for every other kind of step in the *Cut*-elimination argument.

Do these terms correspond to anything like computation? Do they have any other interpretation?

PROJECTS

Q20 Provide sequent formulations for logics intermediate between simple sequent logic and the logic of *distributive lattices* (in which $p \wedge (q \vee r) \succ (p \wedge q) \vee r$). Characterise *which* logics intermediate between lattice logic (the logic of simple sequents) and distributive lattice logic *have* sequent presentations, and which do not. (This requires making explicit what counts as a *logic* and what counts as a sequent presentation of a logic.)

FROM PROOFS TO MODELS 3

How do you show that an argument is *invalid*? It's one thing for there to be no proof from the premise P to the conclusion C, but it's another to *show* that. To show that an argument is valid, we can produce a proof from P to C. Is there some way to show that an argument is invalid in as direct and straightforward way as that? Simply failing to find a proof (and even *showing* that you'll never find a proof) is very illuminating. If an argument is invalid, is there something we can say about what makes it invalid, or is invalidity simply a negative notion, always and only to be understood as the absence of a proof?

According to one tradition in logic, if the argument from P to C is invalid, then this is vouchsafed by a *model*, which certifies that there is some way for P to be true and for C to be false. This is why C does not follow from P—if things were like *that*, then P would be true and C wouldn't be. So, C does not follow from P. In fact, on some ways of explaining validity, models are what defines validity. An argument is valid when there is no model where the premises are true and the conclusion is false. Invalidity is certified by the presence of a model. Validity, on this conception, is reduced to a negative notion, the absence of a counterexample.

In this book we have managed to go two whole chapters without conceiving of validity in this way, and without considering counterexamples to invalid arguments. We have taken validity to be the positive notion, not invalidity. The argument from P to C is valid if there is some proof from P to C. In this third chapter, we will pause to consider models, and to see how we can treat invalidity directly and as a positive notion of its own. However, instead of treating models as primary, as the starting point, we will conceive of models as subordinate to proofs and derivations. We will attempt to motivate models, and explore their properties, guided by their relationship to proofs.

To begin our journey, we will start at the beginning, with Aristotle, when he describes what is involved when premises do *not* form a syllogism. That is, when an inference with two premises, in syllogistic form, are not related in such a way as to generate a syllogism. Here is a passage from the *Prior Analytics*. The language here is very compressed, and there are a number of technical terms, but I will explain them.

> … if the first term belongs to all the middle, but the middle to none of the last term, there will be no syllogism in respect of the extremes; for nothing necessary follows from the terms being so related; for it is possible that the first should belong either to all or to none of the last, so that neither a particular nor a universal conclusion is necessary. But if there is no necessary consequence, there cannot be a syllogism by means of these premisses. As an example of a univer-

> sal affirmative relation between the extremes we may take the terms *animal, man, horse*; of a universal negative relation, the terms *animal, man, stone*. (*Prior Analytics*, Book I, Section 4)

The language is compressed, but the meaning is clear once things are spelled out. When Aristotle says "the first term belongs to all the middle, but the middle to none of the last term" the premises of the purported inference have the shape—All B is A; No B is C—where A is the first term, B is the middle term, and C is the last term. So, in "All B is A", the first term (A) belongs to all of the middle (B), and in "No B is C", the middle term (B) belongs to none of the last (C). Aristotle then claims that no conclusion connecting the first and the last terms (A and C) can be necessarily drawn from these premises because "neither a particular nor a universal conclusion is necessary." We cannot conclude *no* A is C, because of the case A: *animal,* B: *man,* C: *horse*. We would have

> All *men* are *animals*.
> No *men* are *horses*.
> So *no animals* are *horses*.

The premises are clearly true and the conclusion is clearly false. Similarly, we cannot conclude *all* A is C, because of the case A: *animal,* B: *man,* C: *stone*. We would have

> All *men* are *animals*.
> No *men* are *stone*.
> So *all animals* are *stone*.

Again, the premises are clearly true and the conclusion is clearly false. Again, we cannot conclude *some* A is C, for the same reason:

> All *men* are *animals*.
> No *men* are *stone*.
> So *some animals* are *stone*.

We show that an argument form fails to be a syllogism by providing an instance where we can grant the premises but resist the conclusion. This is a possible starting point for explaining what it is for an argument to be invalid, and this is the picture which inspires and informs the model theoretic tradition in logic.

» «

How can we incorporate this picture into the approach we have taken in our first two chapters, where proofs and derivations are primary? The crucial point of connection will be the notion of a *position*. If a derivation δ of the sequent $X \succ Y$ shows that the options Y follow from the premises X, then if there is no such derivation, there is some way of granting X as holding while also rejecting Y. So this is where we will start.

3.1 | positions and valuations

In the first two chapters, we started with systems of proof weaker than classical logic, and built up toward classical propositional logic. Here, it is more straightforward to start with classical logic, to move beyond it to other logical systems only when we have worked through the details of the simpler, classical case. After all, models for classical logic are most straightforward.

So, let us start with classical logic. We are dealing with sequents of the form $X \succ Y$, and the structural rules of contraction and weakening are present. We can think of a sequent $X \succ Y$ as making the claim that from the premises X, the cases Y follow. For parity with this positive view of proofs, we will introduce a new notation for the same structure, but understood negatively. We will think of a *position* $[X : Y]$ as making the claim that the sequent $X \succ Y$ is not derivable—or, to put that in another way, it is keeping open the possibility that *each* claim in X holds, while *no* claim in Y holds.

Recall, the sequent $A, B \succ C, D$ doesn't say that C and D both follow from A and from B. The claim that is being made is that if *both* A and B hold, then *either* C or D must follow.

definition 3.1 [positions and availability] A pair $[X : Y]$ of sets of formulas is called a position. If X and Y are finite, the position $[X : Y]$ is said to be available if the corresponding sequent $X \succ Y$ is not derivable. If X and Y are *infinite* (that is, if either X or Y contain infinitely many formulas), then $[X : Y]$ is available if for no finite $X' \subseteq X$ and $Y' \subseteq Y$ can we derive $X' \succ Y'$.

At this point we do not need to keep track of the number of repetitions of formulas, so we move from considering multisets to the simpler case of focussing on sets.

This definition allows for positions to be *large*, to contain very many formulas. Proofs and derivations are, by their very nature, finite. Not only are formulas finite, but a proof or derivation is a finite structure connecting a finite number of premises to a finite number of conclusions. It is a bridge connecting inputs to outputs, and a bridge we can traverse. A position, on the other hand, is a barrier, rendering the path from left to right impassable. If a position $[X : Y]$ is available, this means that there is no proof from any premises chosen from X to any conclusions chosen from Y. This makes perfect sense even when X and Y are infinite. For example, if we take two formulas A and B, where $[A : B]$ is an available position (so, $A \succ B$ is not derivable) then $[\mathit{to}(A) : \mathit{from}(B)]$ is also a available position, where $\mathit{to}(A)$ is defined as the set of all formulas C where $C \succ A$ is derivable, and $\mathit{from}(B)$ is the set of all formulas D where $B \succ D$ is derivable. This is a position, for we never have $X \succ Y$ for any sequent where $X \subseteq \mathit{to}(A)$ and $Y \subseteq \mathit{from}(B)$, since if we did (using *Cut*) we could derive $A \succ B$. However, the sets $\mathit{to}(A)$ and $\mathit{from}(B)$ are infinite.

An infinite path from A to B does not give us a way to actually *get* from A to B in any length of time, does it?

Because we consider positions involving infinitely many formulas, and the possibility of extending one position to another by adding formulas to the left or the right, we will introduce some terminology and notation to simplify things:

definition 3.2 [extending positions] We say that $[X' : Y']$ extends $[X : Y]$ when $X \subseteq X'$ and $Y \subseteq Y'$, and we gently abuse the notation '$\subseteq$' by writing '$[X : Y] \subseteq [X' : Y']$' when $[X' : Y']$ extends $[X : Y]$.

Recall that a position $[X : Y]$ is *available* when there is no derivation of any sequent $X' \succ Y'$ where $[X' : Y'] \subseteq [X : Y]$. In the case where X and Y are finite, this amounts to $X \succ Y$ having no derivation. We are only required to consider X' and Y' if either X or Y is infinite, for derivations contain no sequents with infinitely many formulas. They are always finite. However, it will be tedious to always move from a position $[X : Y]$ to a finite position $[X' : Y']$ it extends. Instead, we will make the following extension to our notation:

DEFINITION 3.3 [DERIVATIONS FOR GENERALISED SEQUENTS] When δ is a derivation *of* a sequent $X' \succ Y'$ (where X' and Y' are finite, as usual) we will say that it is also a derivation *for* any $X \rhd Y$ where $[X' : Y'] \subseteq [X : Y]$. So, in a regular sequent $X \succ Y$, X and Y are finite. For a GENERALISED SEQUENT $X \rhd Y$, we allow X or Y (or both) to be *infinite*.

This is the first point where we move from finitary proof theory to infinite sets. The relation $\vdash$ imposed on sets where $X \vdash Y$ iff there is a derivation for $X \rhd Y$ is an example of a *generalised consequence relation* (GCR), in Lloyd Humberstone's sense [114, Section 6.3].

So, a derivation *for* $X \rhd Y$ is a derivation of some sequent $X' \succ Y'$ where $X' \subseteq X$ and $Y' \subseteq Y$. This definition applies whether X and Y are finite or infinite.

Recall the motivation for considering positions. We are interested in understanding the structure of counterexamples to invalid arguments. If the position $[X : Y]$ is available, we take there to be some way to hold each member of X while resisting each member of Y. This possibility has a structure of its own. It makes sense to examine what is *true* or *false* relative to such a position. Clearly, each member of X is true according to $[X : Y]$, and each member of Y is false according to $[X : Y]$. But we can say more:

DEFINITION 3.4 [TRUTH AND FALSITY IN A POSITION] We say that A is TRUE according to $[X : Y]$ if there is a derivation for $X \rhd A, Y$, and A is FALSE according to $[X : Y]$ if there is a derivation for $X, A \rhd Y$.

If we can derive A, granting members of X as premises and allowing for each members of Y as alternate conclusions, then A is *true* according to $[X : Y]$. A may not have occurred explicitly in X, but ruling it in follows from the commitments of that position. If, granting A as a premise, along with other members of X, we can derive conclusions in Y, then A is *false* according to $[X : Y]$. A may not have occurred explicitly in Y, but ruling it out follows from the commitments of that position.

We will sometimes say that a formula true according to $[X : Y]$ is *verified* by $[X : Y]$, or *holds* in $[X : Y]$, etc, and similarly, formulas that are false according to $[X : Y]$ will sometimes be said to be *falsified* by $[X : Y]$. Notice first that it follows immediately that if $A \in X$, then A is true in $[X : Y]$, and similarly, if B is in Y, then B is false in $[X : Y]$.

I'll try to avoid saying that formulas false in $[X : Y]$ *fail* in $[X : Y]$, since this is too easy to confuse with 'failing to hold'. Many positions are incomplete. In $[p : q]$, for example, r is neither true nor false. It is natural to think of r as failing to hold in $[p : q]$ and failing to be false there, too.

A position is given by taking *some* things (those claims in X) to be true and others (those claims in Y) to be false. But typically, much more is true or false in a position than those formulas explicitly present in X and Y. For example, in the position $[p : q]$, the formula $p \land \neg q$ is true, since $p \succ p \land \neg q, q$ is easily derived:

$$\dfrac{p \succ p, q \qquad \dfrac{p, q \succ q}{p \succ \neg q, q}\,{\scriptstyle \neg R}}{p \succ p \land \neg q, q}\,{\scriptstyle \land R}$$

In this position, the formula $p \to q$ is false, as we can derive $p, p \to q \succ q$. (The derivation is obvious.)

Notice that if $[X : Y]$ is an available position, then no formula A is both true and false according to $[X : Y]$. (If $X' \succ A, Y'$ and $X'', A \succ Y''$ both hold, where $X', X'' \subseteq X$ and $Y', Y'' \subseteq Y$, then by *Cut*, $X', X'' \succ Y', Y''$ is derivable too, where $X' \cup X'' \subseteq X$ and $Y' \cup Y'' \subseteq Y$.)

Not only are truth and falsity incompatible in available positions. Truth and falsity as defined behave rather much like truth and falsity are thought to behave, when it comes to the truth conditions for connectives:

LEMMA 3.5 [TRUTH CONDITIONS AT POSITIONS] *For any position* $[X : Y]$,

- $A \land B$ *is true in* $[X : Y]$ *iff* A *and* B *are both true in* $[X : Y]$.
- $A \land B$ *is false in* $[X : Y]$ *if either* A *or* B *is false in* $[X : Y]$.
- $A \lor B$ *is true in* $[X : Y]$ *if either* A *or* B *is true in* $[X : Y]$.
- $A \lor B$ *is false in* $[X : Y]$ *iff* A *and* B *are both false in* $[X : Y]$.
- $A \to B$ *is true in* $[X : Y]$ *if either* A *is false or* B *is true in* $[X : Y]$.
- $A \to B$ *is false in* $[X : Y]$ *iff* A *is true and* B *is false in* $[X : Y]$.
- $\neg A$ *is true in* $[X : Y]$ *iff* A *is false in* $[X : Y]$.
- $\neg A$ *is false in* $[X : Y]$ *iff* A *is true in* $[X : Y]$.

Proof: Notice that three of the clauses (falsity for conjunctions, truth for disjunctions and for conditionals—the disjunctive clauses) are conditionals and not biconditionals.

The proof is a straightforward verification from the definitions. Here is the reasoning for conjunction. First, if A and B are both true in $[X : Y]$, then we have $X', X'' \subseteq X$ and $Y', Y'' \subseteq Y$ where $X' \succ A, Y'$ and $X'' \succ B, Y''$ are derivable, and so, we can derive $X', X'' \succ A \land B, Y', Y''$ by extending these derivations as follows:

$$\dfrac{\dfrac{X' \succ A, Y'}{X', X'' \succ A, Y', Y''}\,{\scriptstyle K} \qquad \dfrac{X'' \succ B, Y''}{X', X'' \succ B, Y', Y''}\,{\scriptstyle K}}{X', X'' \succ A \land B, Y', Y''}\,{\scriptstyle \land R}$$

Continually picking out the finite sets $X', X'' \subseteq X$ and $Y', Y'' \subseteq Y$ in such reasoning is rather tedious, and the bookkeeping obscures the essential core of the reasoning. That core can be presented as follows:

$$\dfrac{X \rhd A, Y \qquad X \rhd B, Y}{X \rhd A \land B, Y}\,{\scriptstyle \land R,K}$$

The derivations for $X \rhd A, Y$ and $X \rhd B, Y$ may indeed utilise different subsets of X and Y, but an instance of K suffices to coordinate them so that we can apply the $\wedge R$ rule to find our derivation for $X \rhd A \wedge B, Y$. In what follows, we will present the remaining constructions in this style.

Conversely, if we have derivations for $X \rhd A \wedge B, Y$, then we also have them for $X \rhd A, Y$ and $X \rhd B, Y$.

$$\dfrac{X \rhd A \wedge B, Y \qquad \dfrac{A \succ A}{A \wedge B \succ A}{\scriptstyle \wedge L}}{X \rhd A, Y}{\scriptstyle Cut}$$

$$\dfrac{X \rhd A \wedge B, Y \qquad \dfrac{B \succ B}{A \wedge B \succ B}{\scriptstyle \wedge L}}{X \rhd B, Y}{\scriptstyle Cut}$$

Furthermore, the falsity conditions (in one direction) for conjunction are straightforward. If we have a derivation for $X, A \rhd Y$, or for $X, B \rhd Y$, then we also have a derivation for $X, A \wedge B \rhd Y$.

$$\dfrac{X, A \rhd Y}{X, A \wedge B \rhd Y}{\scriptstyle \wedge L} \qquad \dfrac{X, B \rhd Y}{X, A \wedge B \rhd Y}{\scriptstyle \wedge L}$$

The reasoning for disjunction is precisely dual to that for conjunction, so I will leave that to the reader. The conditional is not much more difficult. If A is true and B is false in $[X : Y]$, then we have derivations for $X \rhd A, Y$ and $X, B \rhd Y$ are derivable, and so, we have a derivation for $X, A \to B \rhd Y$, as follows:

$$\dfrac{X \rhd A, Y \quad X, B \rhd Y}{X, A \to B \rhd Y}{\scriptstyle \to L}$$

Conversely, if we have a derivation for $X, A \to B \rhd Y$ is , then we also have derivations for $X \rhd A, Y$ and $X, B \succ Y$.

$$\dfrac{\dfrac{A \succ B, A}{\succ A \to B, A}{\scriptstyle \to R} \qquad X, A \to B \rhd Y}{X \rhd A, Y}{\scriptstyle Cut}$$

$$\dfrac{\dfrac{A, B \succ B}{B \succ A \to B}{\scriptstyle \to R} \qquad X, A \to B \rhd Y}{X, B \rhd Y}{\scriptstyle Cut}$$

Also, if we have a derivation for $X \rhd B, Y$ or for $X, A \rhd Y$, then we have a derivation for $X \rhd A \to B, Y$.

$$\dfrac{\dfrac{X \rhd B, Y}{X, A \rhd B, Y}{\scriptstyle K}}{X \rhd A \to B, Y}{\scriptstyle \to R} \qquad \dfrac{\dfrac{X, A \rhd Y}{X, A \rhd B, Y}{\scriptstyle K}}{X \rhd A \to B, Y}{\scriptstyle \to R}$$

For negation, we can proceed as follows. First, A is true at $[X : Y]$ iff $\neg A$ is false there:

$$\frac{X \rhd A, Y}{X, \neg A \rhd Y}\neg L \qquad \frac{\dfrac{A \succ A}{\succ A, \neg A}\neg R \quad X, \neg A \rhd Y}{X \rhd A, Y}\mathit{Cut}$$

And similarly, A is false at $[X : Y]$ iff $\neg A$ is true there:

$$\frac{X, A \rhd Y}{X \rhd \neg A, Y}\neg R \qquad \frac{X \rhd \neg A, Y \quad \dfrac{A \succ A}{A, \neg A \succ}\neg L}{X, A \rhd Y}\mathit{Cut}$$

■

So, positions look quite a lot like 'possibilities' or 'models' in some sense. They satisfy much of the traditional picture of truth and falsity conditions. However, we do not have biconditionals for the falsity conditions for conjunction or the truth conditions for disjunction or the conditional. There is no way to prove, for example, that a disjunction is true in a position if and only if one disjunction is true in that position. Take, for example, the position $[p \vee q :]$. The disjunction $p \vee q$ is true in this position, but neither p nor q is true in that position. We cannot derive $p \vee q \succ p$, and neither can we derive $p \vee q \succ q$. In the same way, conjunctions can be false without either conjunct being false (in the position $[p \vee q :]$, for example, $\neg p \wedge \neg q$ is false, but neither $\neg p$ nor $\neg q$ is false), and conditionals can be true without the antecedent being false or the consequent true (see $\neg p \vee q$ in the position $[p \vee q :]$ for an example). Arbitrary positions satisfy *much* of the traditional classical truth conditions for each the binary connectives $\wedge$, $\vee$ and $\rightarrow$, but not the whole package.

Only special positions satisfy the usual truth conditions. How special must a position be to do that job? These positions must be *total*, in a sense that I will explain. Since $X \succ Y, A \vee \neg A$ is derivable for each formula A, the formula $A \vee \neg A$ is true in every position. So, if we have a position that verifies at least one disjunct of each disjunction it verifies, that position must either make A true or make $\neg A$ true for each formula A. To do that, it must make A true or A false for each formula A. It must, in effect, be maximally opinionated, casting a verdict on every issue in the language.

DEFINITION 3.6 [INDISTINGUISHABLE POSITIONS] We will say that $[X : Y]$ and $[X' : Y']$ are INDISTINGUISHABLE as positions iff $[X : Y]$ and $[X' : Y']$ make the same formulas true, and make the same formulas false.

Different positions may be indistinguishable. For example, $[p :]$ is indistinguishable from $[: \neg p]$, since $p \succ A$ is derivable if and only if $\succ A, \neg p$ is derivable, and $p, B \succ$ is derivable if and only if $B \succ \neg p$ is derivable.

In other words, taking p to be true has the same effect as taking $\neg p$ to be false.

Indistinguishability interacts with truth and falsity in an obvious way. Something being true in a position is indistinguishable from it being *explicitly* taken to be true in that position.

LEMMA 3.7 [EXPANDING POSITIONS] *The position* $[X : Y]$ *makes* A *true* if and only if $[X : Y]$ *is indistinguishable from* $[X, A : Y]$, *and similarly,* $[X : Y]$ *makes* A *false iff* $[X : Y]$ *is indistinguishable from* $[X : A, Y]$.

Proof: Consider the case for truth. Suppose A is true in $[X : Y]$, so there is a derivation for $X \rhd A, Y$. To show that $[X, A : Y]$ is indistinguishable from $[X : Y]$, it suffices to show that if $[X, A : Y]$ makes B true, so does $[X : Y]$, and if $[X, A : Y]$ makes B false, so does $[X : Y]$, since the converses are immediate (by weakening). So suppose $[X, A : Y]$ makes B true. That means we have a derivation for $X, A \rhd B, Y$. Using *Cut* we can reason as follows:

$$\frac{X \rhd A, Y \quad X, A \rhd B, Y}{X \succ B, Y}\,\textit{Cut}$$

to show that B is true in $[X : Y]$. Similarly if $[X, A : Y]$ makes B false. That means we have a derivation for $X, A, B \rhd Y$. Again, using *Cut* we have:

$$\frac{X \rhd A, Y \quad X, A, B \rhd Y}{X, B \rhd Y}\,\textit{Cut}$$

and hence, B is false in $[X : Y]$. So, we have shown that if A is true in $[X : Y]$, then $[X : Y]$ is indistinguishable from $[X, A : Y]$. Conversely, if $[X : Y]$ is indistinguishable from $[X, A : Y]$, then since $A \succ A$ is derivable, it is immediate that A is true in $[X, A : Y]$, so it follows that A must be true in $[X : Y]$ too.

That completes the proof that $[X, A : Y]$ is indistinguishable from $[X : Y]$ iff A is true in $[X : Y]$. The proof for falsity has exactly the same structure and is left as an exercise. ■

DEFINITION 3.8 [THE CLOSURE OF A POSITION] Given a position $[X : Y]$, its CLOSURE $[[X : Y]]$ is defined as the position

$$\Big[\{A : A \text{ is true in } [X : Y]\} : \{B : B \text{ is false in } [X : Y]\}\Big]$$

It is immediate that $[[X : Y]]$ is indistinguishable from $[X : Y]$, and it is the *largest* position indistinguishable from $[X : Y]$. (To expand it further, we would need to add a formula on the left *not* true in $[X : Y]$, or a formula on the right not false in $[X : Y]$.) As befits a closure operator, the closure of a closure is itself: $[[[X : Y]]] = [[X : Y]]$.

For any position $[X : Y]$, its closure $[[X : Y]$ includes infinitely many formulas on each side, since $[X : Y]$ must make $p \vee \neg p$ true, and hence the left hand side of the closure contains $p \vee \neg p$, and also $\neg\neg(p \vee \neg p)$, $\neg\neg\neg\neg(p \vee \neg p)$, etc., and the right hand side contains the negations of each of these formulas.

So, let's return to the question of our special positions, those that verify a disjunct of each disjunction they verify, and falsify a conjunct of each conjunction they falsify. We have already seen that they must be maximally opinionated: for each formula A, they either make A true or make A false. It follows then, that if the position is available, its closure is special: it is a partition of the language.

DEFINITION 3.9 [PARTITION POSITIONS] The position $[X : Y]$ is a PARTITION of the language $\mathcal{L}$ if and only if $X \cup Y = \mathcal{L}$, and furthermore, if $X \cap Y = \emptyset$.

There are very many partition positions: any way of sorting out the language $\mathcal{L}$ into two sets (say, formulas with an even number of symbols; those with an odd number) will do. However, most of these partitions are not available. Recall, $[X : Y]$ is available if there is no derivation for $X \rhd Y$. Are there any *available* partition positions? The answer is yes—there are very many of them.

LEMMA 3.10 [POSITION EXTENSION] *For any available position* $[X : Y]$ *and any formula* A*, at least one of* $[X, A : Y]$ *and* $[X : A, Y]$ *is also available.*

The proof of this is an immediate application of the *Cut* rule.

Proof: If $[X, A : Y]$ and $[X : A, Y]$ were both unavailable, we'd have derivations for both $X \rhd A, Y$ and $X, A \rhd Y$. Applying *Cut* we would have:

$$\dfrac{X \rhd A, Y \quad X, A \rhd Y}{X \rhd Y}\,\textit{Cut}$$

so, contraposing, if $[X : Y]$ is available, one (at least) of $[X : A, Y]$ and $[X, A : Y]$ is also available. ∎

So, if we have an available position, we can walk through the entire language $\mathcal{L}$ a formula at a time, adding each formula either to the left of the position or to the right, maintaining availability.

THEOREM 3.11 [FINDING AVAILABLE PARTITION POSITIONS] *Every available position* $[X : Y]$ *is extended by some available partition position* $[X^\infty : Y^\infty]$.

Proof: Consider the branching tree of available positions extending $[X : Y]$ ordered by extension: the immediate descendents of $[X : Y]$ are those available positions $[X, A : Y]$ or $[X : B, Y]$ extending $[X : Y]$ by the addition of a single formula on the left or the right. The leaves of this tree are the limit positions, for (by the previous lemma) they alone cannot be further extended. There are many ways to construct a path through the tree, culminating in a leaf. One way to do so is to enumerate the language $\mathcal{L}$ in a sequence $A_0, A_1, A_2, \ldots$. Extend $[X : Y]$ to $[X : Y]_0$, choosing whichever of $[X, A_0 : Y]$ and $[X : A_0, Y]$ is available. In general, extend $[X : Y]_n$ to $[X : Y]_{n+1}$ by choosing one of the available positions extending $[X : Y]_n$ with the addition of A_{n+1} on the left or the right. Define $[X^\infty : Y^\infty]$ as the limit of this process: it includes the formula A_n on the left iff when it came up in the sequence, it was added to the left, and it includes that formula on the right iff it was added to the right when it was considered. The result is a partition by definition (all formulas are considered and included on the left or the right). The position is available for each strage along the way is available. There can be no finite $X' \subseteq X^\infty$ and $Y' \subseteq Y^\infty$ where $X' \succ Y'$ is derivable, for there is some stage n of the enumeration of the language where each formula in X' and Y' has been considered: $[X' : Y']$ must be included in $[X : Y]_n$ for that n, and by construction $[X : Y]_n$ is available, so $X' \succ Y'$ cannot be derived. ∎

So, any available position at all can be filled out into an available partition. For this reason, we will call available partition positions *limit* positions.

DEFINITION 3.12 [LIMIT POSITIONS] A position $[X : Y]$ is a LIMIT position if it is available and $[X : Y]$ is a partition of the language.

What are limit positions like? First, we can verify that they indeed satisfy all of the traditional two-valued classical truth conditions for formulas:

LEMMA 3.13 [TRUTH CONDITIONS AT LIMIT POSITIONS] *For any limit position* $[X : Y]$,

- $A \land B$ *is true in* $[X : Y]$ *iff* A *and* B *are both true in* $[X : Y]$.
- $A \land B$ *is false in* $[X : Y]$ *iff either* A *or* B *is false in* $[X : Y]$.
- $A \lor B$ *is true in* $[X : Y]$ *iff either* A *or* B *is true in* $[X : Y]$.
- $A \lor B$ *is false in* $[X : Y]$ *iff* A *and* B *are both false in* $[X : Y]$.
- $A \to B$ *is true in* $[X : Y]$ *iff either* A *is false or* B *is true in* $[X : Y]$.
- $A \to B$ *is false in* $[X : Y]$ *iff* A *is true and* B *is false in* $[X : Y]$.
- $\neg A$ *is true in* $[X : Y]$ *iff* A *is false in* $[X : Y]$.
- $\neg A$ *is false in* $[X : Y]$ *iff* A *is true in* $[X : Y]$.

Notice that these conditions are *exactly* the standard two valued truth conditions for the connectives. But we have not designed the sequent calculus with these truth conditions in mind. The two valued truth conditions, understood in this way, are a *consequence* of the rules for the connectives, not their *source*.

Proof: All but three of these conditions (the falsity conditions for conjunctions, truth conditions for disjunctions and conditionals) hold for all positions whatever. For partition positions $[X : Y]$ we have that for each formula A either A is true or A is false in $[X : Y]$, and for *available* positions, we have that A is not both true and false in $[X : Y]$. So, we can reason as follows: If $A \land B$ is false in $[X : Y]$ then $A \land B$ is not true at $[X : Y]$, and hence, either A is not true at $[X : Y]$ or B is not true at $[X : Y]$. In other words (since $[X : Y]$ is a partition position) A is false at $[X : Y]$ or B is false at $[X : Y]$. So, if $A \land B$ is false in $[X : Y]$, then either A is false in $[X : Y]$ or B is false there. Conversely, if A is false at $[X : Y]$, we have a derivation for some $X, A \rhd Y$. So, since $A \land B \succ A$ is derivable we have a derivation for $X, A \wedge B \succ Y$ by *Cut*, and $A \wedge B$ is false at $[X : Y]$. Similarly, if B is false at $[X : Y]$, so is $A \wedge B$.

The reasoning is exactly parallel for disjunctions and conditionals, so with that, we can declare the lemma proved. ■

Limit positions act exactly like *models* of classical logic, in the traditional two-valued sense. We can make this connection explicit, by showing how we can easily define each limit position.

LEMMA 3.14 [LIMIT POSITIONS AND VALUATIONS] *Any limit position* $[X : Y]$ *determines a* VALUATION FUNCTION $\nu_{[X:Y]}$ *assigning exactly one value of* true *or* false *to each atom in* $\mathcal{L}$, *defined by setting* $\nu_{[X:Y]}(p) = \mathsf{true}$ *iff* $p \in X$ *and* $\nu_{[X:Y]}(p) = \mathsf{false}$ *iff* $p \in Y$. *Conversely, for any valuation function* $\nu :$ ATOM $\to \{\mathsf{true}, \mathsf{false}\}$, *we can define the partition position* $[X_\nu : Y_\nu]$, *by setting* $X_\nu = \{A : \nu(A) = \mathsf{true}\}$ *and* $Y_\nu = \{B : \nu(B) = \mathsf{false}\}$, *where we define* ν *on the entire language in the usual way.*

Furthermore, for any valuation function ν, $\nu_{[X_\nu:Y_\nu]} = \nu$ *and for any limit position* $[X : Y]$, $[X_{\nu_{[X:Y]}} : Y_{\nu_{[X:Y]}}] = [X : Y]$.

A valuation function, defined in this way, can be understood as a *model* for a position $[X : Y]$ if it assigns each member of X the value true and assigns each member of Y the value false. In this circumstance, it also makes sense to think of ν as *refuting* the sequent $X \succ Y$, as it serves as a counterexample to the claim that the cases Y follow from the premises X. Now we may proceed to the proof of our lemma, connecting valuations and limit positions.

Proof: Take a limit position $[X : Y]$, and define $\nu_{[X:Y]}$ by setting $\nu_{[X:Y]}(p) =$ true iff $p \in X$, and $\nu_{[X:Y]}(p) =$ false iff $p \in Y$. This is, indeed a function, since $[X : Y]$ is a partition of $\mathcal{L}$. For any formula A in $\mathcal{L}$, we prove by induction on the structure of A that $A \in X$ if and only if $\nu_{[X:Y]}(A) =$ true, and $A \in Y$ if and only if $\nu_{[X:Y]}(A) =$ false. The result holds by definition if A is an atom. If A is a conjunction $B \wedge C$, and the result holds for B and C, then we can show that if $B \wedge C$ is in X, then by the truth conditions for partition positions, B and C are both in X. By the induction hypothesis, it follows that $\nu_{[X:Y]}(B) = \nu_{[X:Y]}(C) =$ true, and hence, by the definition of valuation functions, $\nu_{[X:Y]}(B \wedge C) =$ true as desired. Conversely, if $\nu_{[X:Y]}(B \wedge C) =$ true then $\nu_{[X:Y]}(B) = \nu_{[X:Y]}(C) =$ true, and by the induction hypothesis, $B, C \in X$, and by the truth conditions for conjunction in partition positions, $B \wedge C \in X$ too. The reasoning for the falsity conditions for conjunction, and for the other connectives work in exactly the same way. So we have shown that for any partition position $[X : Y]$, $\nu_{[X:Y]}$ is a valuation function, and $[X_{\nu_{[X:Y]}} : Y_{\nu_{[X:Y]}}] = [X : Y]$.

Now take a valuation function ν. It is immediate that $[X_\nu : Y_\nu]$ is a partition of the language, since every formula is assigned exactly one of the values true and false. To show that $[X_\nu : Y_\nu]$ is an *available* position, we need to show that there is no derivation for $X_\nu \succ Y_\nu$. To show this, we also proceed by induction, but it is an induction on the length of δ, our putative derivation for $X_\nu \succ Y_\nu$. Firstly, no *axiomatic* derivations are derivations for $X_v \rhd Y_v$, for our only axiom is *Identity*, and $[X_v : Y_v]$ is a partition, and no formula is both in X_v and Y_v. Now, suppose δ is a longer derivation, ending in some inference step, and the premises for that step are not derivations for $X_v \rhd Y_v$. We show that the conclusion is *also* not a derivation for $X_v \rhd Y_v$. To do this, it suffices to show that if the conclusion of a rule *is* extended by $[X_v : Y_v]$, then so is at least one of the premises. In other words, we need to show for each rule in the sequent calculus, if v refutes the conclusion (by assigning true to each formula on the left and false to each formula on the right), it also refutes on the of the premises. Consider *Cut*:

$$\frac{X \succ A, Y \quad X', A \succ Y'}{X, X' \succ Y, Y'}\ \mathit{Cut}$$

If v refutes $X, X' \succ Y, Y'$, then it either refutes $X \succ A, Y$ (if $v(A) = \text{false}$) or $X', A \succ Y'$ (if $v(A) = \text{true}$). Consider the conditional rules:

$$\frac{X \succ A, Y \quad X', B \succ Y'}{X, X', A \to B \succ Y, Y'}\ {\to}L \qquad \frac{X, A \succ B, Y}{X \succ A \to B, Y}\ {\to}R$$

For $\to$*L*, if v refutes $X, X', A \to B \succ Y, Y'$, it assigns $A \to B$ the value true. This means either $v(A) = \text{false}$, in which case v refutes $X \succ A, Y$, or $v(B) = \text{true}$, in which case v refutes $X', B \succ Y'$.

The rules for the other connectives proceed in the same way. So, for any derivation δ from axioms, no sequents in the derivation are refuted by v, since the axioms are not refuted, and any inference step *to* a refuted sequent must have started from some refuted sequent. It follows that $[X_v : Y_v]$ is indeed a limit position.

Is $v_{[X_v:Y_v]} = v$? It is immediate that it must be so. For any atom p where $v(p) = \text{true}$, we have $p \in X_v$. So, it follows that $v_{[X_v:Y_v]}(p) =$ true too. Similarly, if $v(p) = \text{false}$, we have $p \in Y_v$. So, it follows that $v_{[X_v:Y_v]}(p) = \text{false}$, so v and $v_{[X_v:Y_v]}$ agree as functions. ■

3.2 | SOUNDNESS AND COMPLETENESS

What we have shown, here, is a systematic and general soundness and completeness result relating derivations and valuations.

THEOREM 3.15 [SOUNDNESS AND COMPLETENESS FOR CLASSICAL LOGIC] *A sequent* $X \succ Y$ *has a derivation* δ *if and only if it has no refutation* v. *Equivalently, a sequent* $X \succ Y$ *has a refutation* v *if and only if it has no derivation* δ.

Proof: The *soundness* part of this result shows that a sequent never has both a derivation δ and a refutation v. We proceed as follows. Suppose $X \succ Y$ has a refutation v. Then it follows that $[X_v : Y_v]$ (formed by splitting the language into the formulas true according to v and those that are false according to v) is a partition, and by Lemma ??, it is *available*.

Since it is a limit position, there is no derivation for $X_v \rhd Y_v$, and as a result, there is no derivation of $X \succ Y$, since $X \subseteq X_v$ and $Y \subseteq Y_v$.

The *completeness* part of this result shows that a sequent always has either a derivation δ or a refutation v. To show this, suppose that $X \succ Y$ has no derivation. It follows $[X : Y]$ is a position, and so, by Lemma ??, it is extended by some limit position $[X^\infty : Y^\infty]$. The valuation $v_{[X^\infty : Y^\infty]}$ is a refutation of the sequent $X \succ Y$, since v assigns each member of X^∞ the value true (including each member of X) and each member of Y^∞ the value false (including each member of Y). ∎

[The significance of this result, on a purely technical level.]

The connection between positions and valuations runs even deeper than the correspondence between valuations and limit positions. Consider the following definition:

DEFINITION 3.16 [THE MODELLING RELATION] Let's write '$v \vDash [X : Y]$' when the valuation v models the position $[X : Y]$. That is, v assigns true to each member of X and false to each member of Y.

For each position $[X : Y]$ we can consider the set of valuations v where $v \vDash [X : Y]$. For some positions (for limit positions), there is only one such valuation. For others, there are many. The empty position $[\ :\ ]$ is modelled by *all* valuations, For some positions (those that are unavailable) there are none. On the other hand, for each set V of valuations, we can consider the positions $[X : Y]$ modelled by each of those valuations in V. That is, the set of all $[X : Y]$ where $v \vDash [X : Y]$ for each $v \in V$. But in this case, there is a *fundamental* such position:

$$[\{A : v(A) = \text{true for each } v \in V\} : \{B : v(B) = \text{false for each } v \in V\}]$$

This position pairs together all formulas true in all valuations in V with all formulas false in all valuations in V. The positions modelled by all valuations in V are exactly the positions extended by *this* position.

So, we have a mapping from positions to sets of valuations (the valuations modelling that position), and from sets of valuations back to positions (the position consisting of those formulas true in all those valuations, paired with those formulas false in those valuations).

$$\text{POSITIONS} \leftrightarrow \mathcal{P}(\text{VALUATIONS})$$

This two-way connection is an instance of an elegant connection known to algebraists as a *Galois connection*. Let's make this precise:

DEFINITION 3.17 [FROM POSITIONS TO SETS OF VALUATIONS & BACK] Given a position $[X : Y]$, define $\mathfrak{v}[X : Y]$ to be the set of all valuations v where $v \vDash [X : Y]$. So defined, $\mathfrak{v}$ is a function

$$\mathfrak{v} : \text{POSITIONS} \longrightarrow \mathcal{P}(\text{VALUATIONS})$$

Conversely, for any set V of valuations, the position $\mathfrak{p}(V)$ is defined as

$$[\{A : v(A) = \text{true for each } v \in V\} : \{B : v(B) = \text{false for each } v \in V\}]$$

and so defined, $\mathfrak{p}$ is a function

$$\mathfrak{p} : \mathcal{P}(\text{VALUATIONS}) \longrightarrow \text{POSITIONS}$$

The set POSITIONS is naturally partially ordered by extension (where $[X : Y] \subseteq [X' : Y']$ iff $X \subseteq X'$ and $Y \subseteq Y'$), and the set $\mathcal{P}(\text{VALUATIONS})$ is also partially ordered by the subset relation $\subseteq$. The functions $\mathfrak{v}$ and $\mathfrak{p}$ respect the ordering in the following way: if $[X : Y] \subseteq [X' : Y']$, then $\mathfrak{v}[X : Y] \supseteq \mathfrak{v}[X' : Y']$ (the larger the position, the fewer valuations modelling it) and if $V \subseteq V'$ then $\mathfrak{p}(V) \supseteq \mathfrak{p}(V')$ (the more valuations to satisfy, the smaller the position that can satisfy them all). The operations $\mathfrak{v}$ and $\mathfrak{p}$ are *order inverting*. However, the connection between $\mathfrak{v}$ and $\mathfrak{p}$ is more intimate than that. Suppose $[X : Y]$ is a position and V is a set of valuations. $V \supseteq \mathfrak{v}[X : Y]$ if and only if every valuation v in V, models $[X : Y]$. Now consider the position $\mathfrak{p}(V)$. We must have $[X : Y] \subseteq \mathfrak{p}(V)$, since at least the members of X are made true by every valuation in V and at least the members of Y are made false by every valuation in V. So, we have

$$V \subseteq \mathfrak{v}[X : Y] \text{ if and only if } [X : Y] \subseteq \mathfrak{p}(V)$$

This makes $\mathfrak{v}$ and $\mathfrak{p}$ an *antitone Galois connection*.

DEFINITION 3.18 If $F : A \to B$ and $G : B \to A$ are both maps on partially ordered sets $\langle A, \leq \rangle$ and $\langle B, \leq \rangle$, then they form an ANTITONE GALOIS CONNECTION if and only if for each $a \in A$ and $b \in B$

$$b \leq F(a) \text{ if and only if } a \leq G(b)$$

It is easy to see that for any antitone Galois connection, F, G, the maps $GF : A \to A$ and $FG : B \to B$ are closure operators: they satisfy $a \leq GF(a)$ and $b \leq FG(b)$, and $GFGF(a) = GF(a)$, and $FGFG(b) = FG(b)$. In our case, we have already seen the closure operator $\mathfrak{p}\mathfrak{v}$.

LEMMA 3.19 [$\mathfrak{p}\mathfrak{v}$ IS CLOSURE] *For any position* $[X : Y]$, *the position* $\mathfrak{p}(\mathfrak{v}[X : Y])$ *is the position's closure,* $[[X : Y]]$.

Proof: $\mathfrak{v}[X : Y]$ is the set of all valuations v where $v \vDash [X : Y]$. We need to show that if A is true (false) in $[X : Y]$, then A is assigned true (false) in all such valuations v, and if A is not true (false) in $[X : Y]$, then A assigned false (true) in at least one such valuation v.

Suppose A is true (false) in $[X : Y]$. Then we have some derivation in $X \rhd A, Y$ ($X, A \rhd Y$). It follows that any valuation v in $\mathfrak{v}[X : Y]$ must assign A true (false), by soundness, since the derivation for $X \rhd A, Y$ ($X, A \rhd Y$) would have a counterexample in the valuation v.

Conversely, suppose that A is not true (false) in $[X : Y]$. That means that there is no derivation in $X \rhd A, Y$ ($X, A \rhd Y$), and hence, $[X : A, Y]$ ($[X, A : Y]$) is an available position. By completeness, there is some valuation v modelling it, and this is a valuation in $\mathfrak{v}[X : Y]$ making A false (true).

It follows from this that $\mathfrak{p}(\mathfrak{v}[X : Y])$ is the closure $[[X : Y]]$ that we have already seen. ∎

What kind of closure operator is $\mathfrak{v}\mathfrak{p}$? Given a set V of valuations, what is the set $\mathfrak{v}(\mathfrak{p}(V))$? Are some sets of valuations not closed? Given a large enough language, indeed some sets of valuations are not closed. Take a language with an unending supply $p_0, p_1, p_2, \ldots$ of atoms, and consider the set V of valuations which make *at least* one atom p_i true. (So, it is the set of all valuations *except* the valuation assigning false to every atom). What is the position $\mathfrak{p}(V)$? What formulas are assigned true by every valuation in V? It turns out that only the tautologies are made true by all these valuations. If A is not a tautology, it can be made false by some valuation. Call that valuation ν_A, a falsifier for A. Now, ν_A may not be in our set V, but that is no problem: we can choose an atom p_i that does not occur in A and choose a valuation ν'_A that agrees with ν_A, except it assigns true to p_i. This means that ν'_A at least *is* in V, and since p_i is absent from A, ν'_A agrees with ν_A concerning A: they both declare it false. So, only the tautologies are made true by every valuation in V. Similarly, only the contradictions (formulas that cannot be made true) are made false by every valuation in V. The position $\mathfrak{p}(V)$ is the smallest closed position, $[T : C]$ consisting of the set T of tautologies and the set C of contradictions. But $\mathfrak{v}[T : C]$ is the set of *all* valuations, since even our missing valuation (the valuation making every atom false) takes tautologies to be true and contradictions false. So, our chosen set V of valuations is not closed.

» «

The maps $\mathfrak{p}$ and $\mathfrak{v}$ not only have interesting mathematical structures. They also point to a fundamental underlying duality between *proofs* and *models*. Mathematically speaking, we can think of either models or proofs as fundamental. The soundness and completeness theorems, as we have seen them, do not, in and of themselves, tell us whether we should take valuations as primitive—and so, soundness and completeness justifies the proof system as *agreeing* with two-valued valuations—or whether we should take derivations as primitive—and so, soundness and completeness justifies the model theory as *agreeing* with the sequent calculus and its account of proofs. While it is traditional to think of soundness and completeness as describing a proof system (it is sound when it can only derive sequents that are valid according to valuations, and it is complete when it can derive all such valid sequents), the mathematics of the theorem does not enforce that interpretation on us. All that soundness tells us is that there no sequent has both a derivation and a refutation. All that completeness tells us is that every sequent has either a derivation or a refutation. Nothing in *that* tells us that one or the other (derivations or refutations) are conceptually primary.

In this chapter, I have attempted to give a proof of the soundness and completeness theorem where the proof theoretical techniques are brought to the forefront. We have started from positions in order to define models. The existence of limit positions—and hence, of valuations—arises naturally out of the behaviour of the *Cut* rule. The existence of limit positions extending any available position—and this is a simple corollary of the *Cut* rule—is what drives the completeness theorem, and it does so, independently of the particulars of the logical vocabulary available in our language. It is true, I have written these last two sections presuming that the full complement of the classical vocabulary is present, but this played nothing but an expository role. All of these results hold for a languages with fewer connectives—in particular, they hold in the language not containing negation. The crucial feature here is the partition of the language, the limit position, not a maximal consistent set of sentences, as is usual in the metatheory of classical logic. Talk of maximal consistent sets of sentences *only* makes sense in the presence of negation (at least as a definable connective). In purely positive logic (say, logic with only $\wedge$ and $\vee$ as connectives, perhaps with $\top$ and $\bot$ along for good measure), the sequent $p \vee q \succ p \wedge q$ is not derivable: the position $[p \vee q : p \wedge q]$ is available. Consider limit positions extending this position. They must make one of p and q true, and they must make one of p and q false. Any valuation refuting this argument must make one of p and q true and the other false. The natural thing to do is to keep track of what is ruled in and ruled out in the manner of a position. You cannot do this by restricting yourself to one half (the formulas you take to be true), because the fact that you take $p \wedge q$ to be false cannot be encoded as a fact about something else being true. The language does not allow for that. To adequately represent the construction of valuations for our language means keeping track of what we take to be true and what we take to be false even handedly. Only then do we treat each connective on its own merits.

3.3 | BACK TO THE CUT RULE

Our proofs of the soundness and completeness theorem used the *Cut* rule. We appeal to *Cut* throughout the verification that any available position $[X : Y]$ is extended by some limit position: we walked through the entire language, appealing to *Cut* to convince ourselves for each formulas A that we could assign A to the left or to the right of the position, without rendering it unavailable. This argument *works*, but it is like cracking a nut with a steamroller. Given a finite position $[X : Y]$, there is no need to pass through the entire language in order to find a limit position extending it. If $[X : Y]$ is available, the goal is to find valuations for the atoms occurring X and Y that ensure that the formulas in X turn out to be true and the formulas in Y, false. The route from the formulas in X and the formulas in Y to their atoms is through the *subformulas* of those formulas. Going further is overkill. We can search for extensions of a position *locally* by processing formulas into their subformulas. In what follows, we are going to do without the *Cut* rule. This means that we can restrict our attention to a stricter criterion of availability, and say that a position $[X : Y]$ is *locally* available if there is no *Cut*-free derivation for $X \rhd Y$. Now, you know and I know that the locally available positions just *are* the available positions, since any sequent that can be derived with *Cut* can be derived without it. However, let's forget that for the moment, and we will see a totally different way to show that sequents derivable *with Cut* can also be derived without it.

A graphical way of representing this is in a tableaux or *tree* proof. It is a particular kind of upside down cut-free sequent derivation.

DEFINITION 3.20 [LOCAL EXTENSIONS OF AVAILABLE POSITIONS] The LOCAL EXTENSIONS of locally available positions are defined as follows:

- A local extension of $[X, \neg A : Y]$ is $[X, \neg A : A, Y]$.
- A local extension of $[X : \neg A, Y]$ is $[X, A : \neg A : Y]$.
- A local extension of $[X, A \wedge B : Y]$ is $[X, A \wedge B, A, B : Y]$.
- A local extension of $[X : A \wedge B, Y]$ is either $[X : A \wedge B, A, Y]$ or $[X : A \wedge B, B, Y]$, whichever is locally available.
- A local extension of $[X, A \vee B : Y]$ is either $[X, A \vee B, A : Y]$ or $[X, A \vee B, B : Y]$, whichever is locally available.
- A local extension of $[X : A \vee B, Y]$ is $[X : A \vee B, A, B, Y]$.
- A local extension of $[X, A \rightarrow B : Y]$ is $[X, A \rightarrow B : A, Y]$ or $[X, A \rightarrow B, B : Y]$, whichever is locally available.
- A local extension of $[X : A \rightarrow B, Y]$ is $[X, A : B, A \rightarrow B, Y]$.

LEMMA 3.21 [A LOCAL EXTENSION IS ALWAYS LOCALLY AVAILABLE] *If* $[X : Y]$ *is locally available and* $[X' : Y']$ *is one of its local extensions, it is also locally available.*

Proof: To show this, we verify that if $[X, \neg A : Y]$ is locally available, so is $[X, \neg A : A, Y]$, and so on, for the non-disjunctive extension clauses. For the disjunctive clauses, it suffices to show, for example, that if $[X : A \wedge B, Y]$ is locally available, then at least one of $[X : A \wedge B, A, Y]$ and $[X : A \wedge B, B, Y]$ is also locally available. The verifications of these facts are straightforward. If, for example, $X, \neg A : A, Y$ were *not* locally available, we would have a (*Cut*-free) derivation for $X, \neg A \rhd A, Y$. It would follow that we would also have a (*Cut*-free) derivation for $X, \neg A \rhd Y$, too:

$$\dfrac{\dfrac{X, \neg A \rhd A, Y}{X, \neg A, \neg A \rhd Y}\,{\scriptstyle \neg L}}{X, \neg A \rhd Y}\,{\scriptstyle W}$$

In a similar way, if neither $[X : A \wedge B, A, Y]$ nor $[X : A \wedge B, B, Y]$ were locally available, $[X : A \wedge B, Y]$ would also not be locally available:

$$\dfrac{\dfrac{X \rhd A \wedge B, A, Y \quad X \rhd A \wedge B, B, Y}{X \rhd A \wedge B, A \wedge B, Y}\,{\scriptstyle \wedge R}}{X \rhd A \wedge B, Y}\,{\scriptstyle W}$$

The verifications for each of the other rules are similar. ■

Let's call a local extension of a locally available position *proper* if it contains properly more formulas than the starting position. From our lemma, the following fact follows:

LEMMA 3.22 [LOCAL EXTENSION ENDS] *Each finite locally available position* $[X : Y]$ *is extended by some locally available position* $[X' : Y']$ *that itself has no proper local extensions.*

Proof: Each local extension of a position contains only subformulas of the formulas in the original position. Since the position contains only finitely many formulas, and these have only finitely many subformulas, any path down the tree of local extensions of $[X : Y]$ (and *their* local extensions, etc.) is finite. The leaves of this tree have no proper local extensions. ∎

Suppose $[X : Y]$ is locally available and it has no proper local extensions. That means it is downwardly closed in the following sense:

DEFINITION 3.23 [DOWNWARD CLOSURE FOR POSITIONS] $[X : Y]$ is DOWNWARDLY CLOSED if and only if

- If $\neg A \in X$ then $A \in Y$.
- If $\neg A \in Y$ then $A \in X$.
- If $A \wedge B \in X$ then $A \in X$ and $B \in X$.
- If $A \wedge B \in Y$ then $A \in Y$ or $B \in Y$.
- If $A \vee B \in X$ then $A \in X$ or $B \in X$.
- If $A \vee B \in Y$ then $A \in Y$ and $B \in Y$.
- If $A \to B \in X$ then $A \in Y$ or $B \in X$.
- If $A \to B \in Y$ then $A \in X$ and $B \in Y$.

It follows that if $[X : Y]$ is downwardly closed and locally available, then it determines a valuation v where $v \vDash [X : Y]$.

LEMMA 3.24 [VALUATIONS FROM DOWNWARDLY CLOSED POSITIONS] *If* $[X : Y]$ *is downwardly closed and locally available, any valuation* v *where* $v(p) = \text{true}$ *for* $p \in X$ *and* $v(p) = \text{false}$ *for* $p \in Y$ *models* $[X : Y]$*—and there are such valuations.*

Proof: Since $[X : Y]$ is downwardly closed and locally available, for no atom p do we have $p \in X$ and $p \in Y$. So, there are valuations v where $v(p) = \text{true}$ for $p \in X$ and $v(p) = \text{false}$ for $p \in Y$. We show by induction on the complexity of formulas A in X (in Y) that $v(A) = \text{true}$ (false). If $\neg A \in X$ (Y) then by the downward closure condition $A \in Y$ (in X), and by the induction hypothesis $v(A) = \text{false}$ (true), and hence $v(\neg A) = \text{true}$ (false).

If $A \wedge B \in X$ (Y), then by the downward closure condition $A \in X$ and $B \in X$ ($A \in Y$ or $B \in Y$), and so, by induction, $v(A) = \text{true}$ and $v(B) = \text{true}$ ($v(A) = \text{false}$ or $v(B) = \text{false}$). So, $v(A \wedge B) = \text{true}$ ($v(A \wedge B) = \text{false}$), as desired.

The reasoning for the other clauses is similar, and we declare the lemma proved. ∎

So, we have a stronger version of the completeness theorem. We have shown that if a position $[X : Y]$ is *locally* available, then it is extended by some downwardly closed and locally available position $[X' : Y']$. From this we can construct a valuation that models $[X' : Y']$, and hence $[X : Y]$. We have proved the following result:

THEOREM 3.25 [COMPLETENESS FOR CUT-FREE DERIVATIONS] *If* $X \succ Y$ *has no Cut-free derivation, there is some valuation* v *refuting* $X \succ Y$.

Contraposing this, we see that if there is no valuation v refuting $X \succ Y$, then it has some *Cut*-free derivation. If we combine this with the soundness theorem—to the effect that if a sequent $X \succ Y$ has some derivation (possibly involving *Cut*) then it has no valuation refuting it—we have shown that any sequent with a derivation has a *Cut*-free derivation.

COROLLARY 3.26 [THE ADMISSIBILITY OF CUT] *If* $X \succ Y$ *has a derivation with Cut, it also has some derivation without Cut.*

In the previous chapter, we had a constructive proof of this result, which showed how to *eliminate Cut* from a derivation. This result does no such thing. It only assures us that, since we can find a counterexample to any sequent without a *Cut*-free derivation, there are no sequents without *Cut*-free derivations which are not without derivations involving *Cut*. This does not give us an technique for removing *Cut*s from derivations.

However, if you look closely, you can see a technique for finding a *Cut*-free derivation for any sequent which has a derivation. Take the position corresponding to the sequent, and develop the tree of its local extensions. (This corresponds to doing a tableaux for the sequent.) The search must fail to locate an available sequent. That means, each branch in the search *closes* with sequents that are derivable. Turn this tree upside down to construct a sequent derivation.

3.4 | THE SIGNIFICANCE OF VALUATIONS

So, we have a connection between the proof theory of classical propositional logic, given in the sequent calculus, and the traditional two-valued 'semantics' given by valuation functions. We have proved soundness and completeness for classical logic, showing that each sequent either has a derivation δ or a counterexample v. In this section, we will examine a three the consequences of this result. (1) We will examine an argument, originally due to Georg Kreisel, to the effect that the formal soundness and completeness theorem can give us an intuitive justification that the concept of logical consequence formalised in classical logic captures a pre-theoretic understanding of validity. Next, (2) we will show how the connection between full classical sequents (of the form $X \succ Y$) and valuations is tighter and more natural than any connection that can be maintained between Set-Formula sequents, of the form $X \succ A$, which are more loosely connected to the behaviour of valuations. We can see that classical sequents provide richer expressive resources than the intuitionist sequents of the form $X \succ A$. (3) Following on from this, we will explore the connections between the *relative* notion of truth-according-to-a-valuation, which plays a role in the arguments in (1) and (2) and the *categorical* notion of *truth*, so important to many philosophers. But first, let's start with Kreisel's argument.

KREISEL'S SQUEEZING ARGUMENT

Here are questions one might ask concerning the classical sequent calculus. First, are the inference rules *correct*? (If we can deduce the sequent $X \succ Y$, does Y really follow from X in any appropriate sense?) Second, are the inference rules *comprehensive*? (If we cannot deduce the sequent $X \succ Y$, does Y *fail* to follow from X in that sense?) In other words, given

I say 'any' appropriate sense because I leave open the possibility that there is more than one notion of validity to which a formal theory could correspond. I am a logical pluralist [12, 13].

the language of propositional logic, does the classical sequent calculus precisely match the notion of validity for arguments expressed in that vocabulary which we were trying to model in the first place?

Stewart Shapiro gives an excellent accessible presentation of Kreisel's reasoning [223, 224], and it plays a significant role in Chapter 2 in Hartry Field's *Saving Truth from Paradox* [67].

Kreisel noted that that the soundness and completeness results give us the resources to address this question [120]. I will present a version of the argument, formulated in a manner suited to the resources of the classical sequent calculus, and the valuations we have defined here. The argument is, essentially, a *squeezing* argument, to the effect that our *target* notion of validity—the notion we want to model—is squeezed between formal validity, on the one hand (as defined by derivations) and formal invalidity, on the other (as defined by valuations). Because the target notion of validity is not formally defined, let's call it *informal* validity. If we can show, first, that any derivable sequent is informally valid, and second, that any sequent with a counterexample is informally invalid, then the soundness and completeness theorem (which tells us that valuation-validity and derivation-validity coincide) reassures us that formal validity and informal validity match.

So, to understand this argument, we need to say a little more about informal validity, in order to understand the connection between that notion and the formal validity defined by the sequent calculus. We will be taking two perspectives on our language $\mathcal{L}$ of formulas. The usual perspective of a formal logic is that formulas are objects to be manipulated. They occur in derivations, they are assigned values by valuation functions. Taking the perspective of the informal notion of validity, they are claims that can be *made*. They are judgements, with which we can agree or disagree—or withhold our own judgement. For the informal notion of validity, the language $\mathcal{L}$ is a language in *use*.

There are many different things we could mean when we call an argument *valid*, as we have seen. Should validity encode a notion of relevance or not? I am a pluralist about logical consequence, and I do not think that there is one well-defined intuitive notion of logical validity, if we fail to do more to fix that notion [12, 13, 194]. The target notion here is the idea that underwrites classical consequence, that there is no way to take the premise(s) to be true and the conclusion(s) false.

Given sets of sentences X and Y from a language in use, we will say that the sequent $X \succ Y$ is *informally valid* if and only if there is a *clash* involved in asserting each sentence in X and denying each sentence in Y. Informal validity goes beyond contingent truth. Take a sentence A that happens to be true, so it would be a mistake to deny it. Nonetheless, we would not want to be forced into saying that the sequent $\succ A$ is informally valid. Although it would be a mistake to deny A, denying A need not involve a *clash* in the salient sense. Validity judgements are used as lynchpin in discussions when interlocutors disagree. We can agree that the inference from A to B is valid or invalid, even when we disagree on whether A or B is true. So the kind of clash involved in asserting the premises and denying the conclusions of a valid argument involves more than just the truth of the premises or the falsity of the conclusions. We will get more of a sense of what is involved in this kind of clash when we examine the positive and negative sides of Kreisel's argument.

Suppose we have a derivation δ of some sequent $X \succ Y$ in our language $\mathcal{L}$. We will show, by induction on the structure of the derivation that there is a clash—in a very strong sense of the term—involved in asserting each member of X and denying each member of Y. In the case of axioms, There *is* a very strong clash involved in asserting p and denying p, for whatever p you choose. It is the clash grounded in the opposition between asserting and denying. The very acts of asserting p and denying p are opposed. We will say more in in Chapter 5 when we attend to the norms of assertion and denial in more detail. For now, let's settle with the idea that asserting p and denying p clash, for part of the role of denial of p is to oppose the assertion of p, and vice versa. To use the vocabulary of positions: there is no position involving both asserting p and denying p. To deny p is to attempt to take up a different position on the issue of whether p than any taken by someone who asserts p.

Now consider for the structural rules. If there is already a clash involved in asserting X and denying Y, in the sense that there is no available position involving those commitments, then adding *more* commitments by adding extra assertions or extra denials does not improve the situation. So if $X \succ Y$ is valid in this intuitive sense, so is $X, X' \succ Y, Y'$, any result of *weakening* the sequent. And as for contraction, repeating an assertion does not change the position taken, and neither does repeating a denial. The rule of contraction preserves intuitive validity in this sense. We do not *need* to consider *Cut* here, since as we proved in Theorem 2.25 (page 93) and also in Corollary 3.26 (page 135), anything derivable *with Cut* can be derived without it. (The intuitive admissibility of the *Cut* rule follows here as a consequence of the result we will prove.)

Of course, using the same words to make an assertion as you made before might have a new effect (suppose I point to one person, saying "You're on my team", and then I point to another, saying the same words. Here, two *different* things are said even though the same words were used. In the same way, *asserting* "You're on my team" (gesturing at one person) does not clash with *denying* "You're on my team" (gesturing at another).

As for the connectives, we can understand these rules as *definitional* of the meanings of the connectives, we can understand them to preserve intuitive validity, too. There are many things we can understand by the word "not", but if we agree to interpret $\neg A$ in such a way that asserting "$\neg A$" has the same upshot as denying "A", and denying "$\neg A$" has the same upshot as asserting "A", then the negation rules preserve intuitive validity. And the same goes for the other connective rules. (For more detail on how this could work, and why such definitions are always possible, see Chapter 6. For now, you will have to take it on faith that such definitions are acceptable.) With that understanding, we can see how to *read* a sequent derivation as certifying the validity of the underlying sequent in the language we use. Granting the connective rules as expressing the senses of the connectives, we see, then, that any derivable sequent is valid in the pre-theoretic sense.

Let's turn to the converse. We wish to show that any underivable sequent is *invalid* in the intuitive sense. This is, literally speaking, untrue, unless we refine our understanding of the informal sense of validity further. After all, arguments such as:

> All footballers are bipeds.
> Socrates is a footballer.
> Therefore, Socrates is a biped.

are valid in the intuitive sense, but are invalid when it comes to classical propositional logic. As far as my language $\mathcal{L}$ is concerned, the shape of this argument is

> p
> q
> Therefore, r

and that is certainly not valid. The structure responsible for the validity of this argument involves quantification, not just the propositional connectives. So, our sense of validity must be, at least partly, *formal*, in that it appeals to only parts of the structure in the language under discussion: the propositional connectives. Given sentences in $\mathcal{L}$ without any structure involving propositional connectives: that is, sentences that are not *conjunctions, disjunctions, conditionals, negations* or $\top$ or $\bot$, we must understand validity in such a way as to not connect different sentences. Each atomic sentence (that is, each sentence that is not a conjunction, disjunction, etc.) must be logically independent (in this salient sense) from any other sentence, for they have no structure that can be appealed to in *propositional* logic. We are concerned with validity in virtue of (propositional) logical form. (Notice that this sense of validity suffices for the verification that every derivable sequent $X \succ Y$ is intuitively valid, since the derivations we used only introduced propositional vocabulary. They did not impose any logical connection between distinct atomic sentences.)

Don't worry, we'll get to quantification in Chapter 8. It's not as if proof theory can tell us nothing about the quantifiers.

Now, we wish to show that any sequent $X \succ Y$ involving sentences from $\mathcal{L}$ which has no derivation is, in fact, intuitively invalid. We wish to show that there *is* a position (that is, a coherent position) involving asserting each member of X and denying each member of Y. We want to show that, in the relevant sense, asserting each member of X and denying each member of Y does not involve a clash. This is where we can appeal to the completeness theorem, connecting derivations and valuations. We know that since there is no derivation δ of $X \succ Y$, it follows that there is a valuation v according to which each member of X is true and each member of Y is false. We will use this v to show that there *is* a position involving asserting each member of X and denying each member of Y. The basic idea is that if we take the atoms v assigns the value true to be *true*, and the atoms v assigns the value false to be *false*, then this position takes each member of X to be true and each member of Y to be false. In other words, if things are as v takes them to be, we can agree with each sentence in X while denying each sentence in Y. *Why* is such a position available to us, no matter what valuation v we choose? It is because we have granted that there is no clash involved in asserting or denying any of the atoms in our language. These atoms have no structure on which propositional logic can take hold. The atoms are logically independent, in the sense that there is no clash involved (as far as propositional logic is concerned, anyway) in asserting any family of atoms and denying any disjoint family of atoms. Once that position is available, the rest follows immediately, given the meanings of the connectives as given in the sequent rules, and their corresponding truth conditions. In this position, each member of X is true (we follow the usual reasoning to verify this—if A is true in a position, $\neg A$ is false in that position, etc.) and each member of Y is false. It follows that any argument from X to Y which has no derivation can be given a counterexample, and this counterexample can be understood as explaining *how* we can grant the premises but deny the conclusion.

This is the intuitive sense corresponding to the formally defined notion of 'available position.'

This completes Kreisel's squeezing argument. We have used the completeness result to show that the formal notion of validity for this language coincides with a given intuitive notion of validity.

» «

This result can be used to diagnose issues concerning the potential mismatch between the seeming intuitive invalidity of certain arguments, and the actual classical derivability of those arguments. Take, as an example, one of the paradoxes of material implication:

$$p \succ q \to p$$

Sequents such as these have been taken to be invalid on the grounds of *relevance*. The truth of p should not be enough to vouch for the truth of the conditional that says that p follows in any way *from* q. That response is reasonable, but beside the point here, in two distinct but related ways. Kreisel's result tells us that this sequent is intuitively valid in the following sense: there is a clash involved in asserting p and denying $q \to p$. (This need not, in general, involve a connection between the premise and the conclusion. Perhaps the clash is due just to the premise, or just to the conclusion.) Here, there is some connection, but that connection does not go so far as to mean that the conclusion asserts that there is some connection between q and p. No, the clash brought to light by this sequent is the clash between asserting p and denying $q \to p$, because the conditional '$\to$', defined by the $\to L/\to R$ rules is such that denying $q \to p$ amounts to asserting p and denying q, and doing *that* certainly clashes with asserting p. It does it in a way that doesn't involve q in any way, of course, but it clashes nonetheless. Kreisel's result, which shows that classical derivability amounts here to intuitive validity helps direct us to the proper understanding of the connectives.

A good exercise for the reader is to do the same sort of diagnosis for other failures of relevance in classically valid sequents: $p \succ q \lor \neg q$ and $p \land \neg p \succ q$.

ON THE POWER OF SEQUENTS

Panu Raatikainen has recently revived [182] an argument due to Rudolf Carnap, to the effect that traditional proof-theoretical accounts of logic do not constrain the meaning of logical constants enough. In particular, there is nothing in the standard proof theories for classical logic that ensures that the standard valuations are the only possible valuations for the language. In this section, I'll explain this argument, and show that while it *is* an argument with significant consequence for some proof-theoretical accounts of classical logic, it does not have that consequence for the classical multiple-premise, multiple-conclusion sequent calculus.

A generalised valuation function maps formulas from $\mathcal{L}$ to the values true and false. Let's suppose that a consequence relation determines a family of sequents of the from $X \succ A$ (with some finite collection of premises X and a single conclusion A). Carnap and Raatikainen's target is either classical or intuitionistic logic, and for concreteness, we'll focus on classical logic here, so let's consider the sequents of the form $X \succ A$ where $X \succ A$ has a derivation.

DEFINITION 3.27 [SINGLE CONCLUSION COMPATIBILITY] We'll say that a valuation v is COMPATIBLE with our consequence relation if and only if whenever $X \succ A$ is in the consequence relation, if $v(B) = \text{true}$ for each $B \in X$, then $v(A) = \text{true}$ too.

The soundness theorem tells us that every standard valuation is single conclusion compatible. However, Carnap showed that there are non-standard valuations that are *also* compatible with single conclusion consequence. If the only resources that we can use to constrain valuations are those given by single conclusion consequence, these valuations should count as acceptable:

DEFINITION 3.28 [TWO NON-STANDARD VALUATIONS] The TRIVIAL (POSITIVE) valuation v^{true} assigns the value true to each formula. The TAUTOLOGOUS valuation $v^{\top}$ assigns the value true to every tautology, and false to every other formula.

LEMMA 3.29 [CARNAP'S LEMMA] *v^{true} and $v^{\top}$ are single conclusion compatible, but neither valuation satisfy the standard classical truth conditions.*

Proof: Clearly v^{true} is compatible with every single conclusion consequence relation. It never provides a counterexample to *any* argument of the shape $X \succ A$, as it never assigns any formula A false, so it never provedes a counterexample to any classically valid argument.

In addition, $v^{\top}$ is compatible with every classically valid sequent $X \succ A$, since if $v^{\top}$ assigns true to every member of X, then every member of X is a tautology, and we can only deduce other tautologies from members of X, so $v^{\top}$ must assign true to A if $X \succ A$ is valid. ∎

Notice that the trivial *negative* valuation v^{false}, which assigns false to every formula, is not single conclusion compatible. The sequent $\succ p \to p$ is valid, but v^{false} assigns false to $p \to p$, not true.

However, v^{true} and $v^{\top}$ are single conclusion compatible, and they do not satisfy the standard truth conditions. In particular, v^{true} fails to respect the negation condition, since $v^{\mathsf{true}}(p) = v^{\mathsf{true}}(\neg p) = \text{true}$. While $v^{\top}$ fails to respect the disjunction condition—$v^{\top}(p \vee \neg p) = \text{true}$ while $v^{\top}(p) = v^{\top}(\neg p) = \text{false}$—and it fails to respect the negation condition—$v^{\top}(p) = v^{\top}(\neg p) = \text{false}$. So, if what it takes for a valuation to be acceptable is given by a single conclusion consequence relation in the way given above, for a valuation to be acceptable, the consequence relation does not give us enough to pin down the traditional two-valued valuations.

That argument seems to me to be decisive. Single conclusion consequence is not enough to pin down standard valuations. We can do a little better if we expand the constraint to *single-or-zero* conclusion consequence, and admit sequents of the form $X \succ$, with empty right hand sides. Compatibility with *these* sequents rules out v^{true}. Here is why: consider what it is for a valuation to be compatible with $X \succ$? The natural proposal is to keep the same reading: that there is no valuation that makes everything on the left true while making something on the right false. Since in this case there is nothing in the right to make false, the constraint simplifies to the criterion that not everything in X is true. Now we can rule out ν^{true} because, for example, $A, \neg A \succ$ is valid, but ν^{true} assigns true to A and to $\neg A$. However, $\nu^{\top}$ has no problems with classically valid zero-conclusion sequents. It never assigns true to incompatible sentences, since it assigns true only to tautologies.

To rule out $\nu\top$, we need the full power of multiple conclusion sequents. If we require valuations to be compatible with multiple conclusion sequents, these deviant valuations are eliminated.

Julien Murzi and Ole Hjortland, in a reply to Raatikainen's paper [150], show that *bilateralist* conceptions of consequence give the resources to resist Carnap's conclusion, but they do not use multiple conclusion sequents to explain this. See Lloyd Humberstone's "Revival of Rejective Negation" [113] for a discussion of the relationship between bilateralist systems of natural deduction (in which formulas are signed positively or negatively for assertion and denial) and multiple conclusion sequents.

DEFINITION 3.30 [MULTIPLE CONCLUSION COMPATIBILITY] We'll say that a valuation ν is COMPATIBLE with our consequence relation if and only if whenever $X \succ Y$ is in the consequence relation, if $\nu(A) = \text{true}$ for each $A \in X$, then $\nu(B) = \text{true}$ for some $B \in Y$ too.

This still rules out ν^{true}, since $A, \neg A \succ$ is valid, as we have seen. So, any valuation compatible with classical multiple conclusion consequence must assign true to at most one of A and $\neg A$. But now, we have the space to reverse the reasoning: since $\succ A, \neg A$ is classically valid, any compatible valuation assigns true to one of A and $\neg A$. But if A is not a tautology, $\nu^{\top}$ assigns true to neither, so $\nu^{\top}$ is ruled out as incompatible with multiple conclusion consequence. This result can be strengthened in the natural way: the only valuations compatible with multiple conclusion consequence are the traditional boolean evaluations — and all such valuations are compatible.

THEOREM 3.31 [CLASSICAL COMPATIBILITY] *A valuation function ν is compatible with classical multiple conclusion consequence if and only if it satisfies the usual classical truth conditions.*

Proof: For the negation conditions, the sequents

$$\succ A, \neg A \qquad A, \neg A \succ$$

ensure that any compatible valuation assigns true to exactly one of A and $\neg A$. For conjunction, the sequents

$$A, B \succ A \wedge B \qquad A \wedge B \succ A \qquad A \wedge B \succ B$$

ensure that any compatible valuation assigns true to $A \wedge B$ if and only if it assigns true to both A and B. For disjunction,

$$A \vee B \succ A, B \qquad A \succ A \vee B \qquad B \succ A \vee B$$

ensure that any compatible valuation assigns true to $A \vee B$ if and only if it assigns true to at least one of A and B. For the conditional,

$$A \to B, A \succ B \qquad B \succ A \to B \qquad \succ A, A \to B$$

ensure that any compatible valuation assigns true to $A \to B$ if and only if it either assigns true to B or doesn't assign true to A. This shows that any compatible valuation satisfies the standard truth conditions. To show the converse, it suffices to show that any standard valuation is multiple conclusion compatible. That follows from soundness: if v is a standard consequence relation, it corresponds to a limit position $[X : Y]$ where every member of X is assigned true by v and every member of Y is assigned false. By soundness, this position is *available*, and hence, v is compatible with all classically valid multiple conclusion sequents. ■

So, there is a very intimate connection between standard valuations and classical multiple conclusion sequents. This connection is tighter than is available for single conclusion sequents.

SEQUENT RULES AND TRUTH CONDITIONS

What connection is there between the relative notion of truth in a position and the categorical notion of *truth* as such? We have had a chapter full of talk of truth in a position or truth according to a valuation. What has this to do with truth in the unrelativised sense?

Wilfrid Hodges gives a good treatment of this question in the context of Tarski-style model theory for classical first order logic [107].

Consider again the difference between *mentioning* the language, when we talk about sentences, and consider whether they hold in positions, and whether they are assigned the values true or false in some valuation, and what it is to *use* those sentences to assert and to deny. That practice is what grounds the categorical notion of truth. The disquotational truth-scheme tells us that "it is Tuesday" is true if and only if *it is Tuesday*. On the left hand side of the biconditional, the sentence is mentioned. On the right, it is used.

Consider, now, the traditional truth conditions for the connectives:

» $A \land B$ is true iff A is true and B is true.

» $A \lor B$ is true iff A is true or B is true.

» $A \rightarrow B$ is true iff A is not true or B is true.

» $\neg A$ is true iff A is not true.

How are these categorical truth conditions, which make no reference to valuations or to positions, to be related to the relativised truth conditions giving us truth and falsity relative to a position? A straightforward and swift answer is to identify truth with truth in some particular limit position—or equivalently, with truth in some valuation. As we saw in the previous section, the use of valuations (or limit positions) can be justified in terms of the proof theory of classical logic. These are the only valuations compatible with the classically valid multiple conclusion sequents.

The truth conditions for limit positions (for valuations) indeed *agree* with the categorical truth conditions for the connectives. If we had some reassurance that truth can be identified with truth in some particular limit position, we can be sure that the traditional truth conditions are satisfied, because all limit positions satisfy those truth conditions. But how can we get this reassurance? Limit positions are idealisations. How are we to pick out a particular limit position to play the role of categorical truth, if that involves answering every yes/no question? Limit positions are maximally opinionated, so finding the position which corresponds to the Truth seems like a task beyond our pay grade.

Thankfully, we don't need to independently specify a given limit position. Our task is much more modest. We need explain why what is true satisfies the conditions for a limit position, even if we cannot find (and even if we could *never* find) what position that actually is. But we have very nearly done this already when we defined the notion of multiple conclusion compatibility. Recall, a valuation v is multiple conclusion compatible (with classically valid sequents) if and only if whenever $X \succ Y$ is derivable, if $v(A) = \text{true}$ for each $A \in X$, then $v(B) = \text{true}$ for some $B \in Y$ too. We have proved that if v is multiple conclusion compatible, then v is a standard valuation. To show that *truth* corresponds to some standard valuation, we need just show that whenever $X \succ Y$ is derivable, if A is *true* for each $A \in X$, then B is *true* for some $B \in Y$. In other words, we need to show that all classically derivable sequents preserve *truth*. Once we have done that, all standard truth conditions are satisfied, as we have seen.

How can we do this? Again, the answer is found in the derivation δ of a derivable sequent $X \succ Y$. We can use the derivation to explain *how* (some member of) Y follows from (all members of) X. As before, the axioms are straightforward. If there is some sentence in both X and Y, then clearly if all members of X are true, so is some member of Y. The structural rules of contraction and weakening also obviously preserve validity. The work is done by the connective rules. Consider, for example

$$\frac{X, A \succ Y \quad X, B \succ Y}{X, A \vee B \succ Y}\ \vee L$$

If we have already reassured ourselves that whenever X and A are true, then (some member of) Y is true, and whenever X and B are true, then (some member of) Y is true. Does it follow that whenever X and $A \vee B$ is true, some member of Y is true? Well, yes it does, once we have reassured ourselves that if $A \vee B$ is true, then one of A and B is also true. (This is, of course, one half of the standard truth condition for disjunction. If this holds for $\vee$, depends of course on whether we can take $\vee$ to be *defined* by the rules of the sequent calculus. If we do take those as defining the meaning of the connectives, we can proceed. In Chapter 6, I will present an argument to the effect that we can do just that: I will explain why it is that we can safely take the connective rules to define the connectives. For now, we must take this as given, just as we did in the case of Kreisel's squeezing argument.) So, proceeding with this understanding of the meaning of $\vee$, we continue. *Which* member of Y depends on which of A and B are true, of course, but both premises tell us that a candidate can be found, either way.

Consider, another rule, this time, for negation:

$$\frac{X, A \succ Y}{X \succ \neg A, Y}\ \neg R$$

If we know that when X and A are true, then some member of Y is true, does it follow that when X is true, either $\neg A$ is true or some member of Y is true? It seems that endorsing this means we must be committed to the law of bivalence: that either A is true or $\neg A$ is true, and indeed this is part of what is involved in taking truth to be governed by the standard valuation clauses. Exactly one of A and $\neg A$ is *true*. How can this be justified? Must it be taken as an extra assumption, to be given independent justification? Again, we might appeal to the connective rules as *definitions* of the connectives themselves. If we can do so (and I will justify that debt in Chapter 6), our job will be done. In the absence of such an argument, we can at least answer the question concerning truth conditions *conditionally*: if we accept the rules of the classical sequent calculus, we have an explanation of why *truth*—categorically—satisfies the standard truth conditions.

3.5 | HISTORY

[[To be added.]]

3.6 | EXERCISES

[[To be added.]]

PART II

The Core Argument

TONK

4

In this central part of the book, we will use the tools of proof theory to address an important question in the philosophy of logic. How is it that rules confer meaning? The inference rules for logical concepts such as conjunction, negation, disjunction or the conditional seem to—in some sense—*define* those concepts. How does this work, if indeed, it *does* work. And what can this phenomenon tell us about the nature of logical concepts? To address these questions, we will grapple with what—if anything—makes logic distinctive. What do the tools of proof theory *do*? What is their scope, their range, their field of application?

We saw this question arise in a sharp manner at the end of Chapter 3, where we saw at certain points in Kreisel's argument, and in our examination of the connection between sequent rules and truth conditions, that we were prompted to take the sequent rules for the connectives as definitions. It is certainly tempting to do this, but perhaps this is a lacuna in our thinking. Perhaps these rules, which we took to be mere definitions, come with a price to pay. Perhaps we should resist them, rather than take them to be harmless clarifications of concepts. To see what is involved in taking rules to define concepts, we must address the issue head on.

To make the question even more concrete, to sharpen up this challenge, we will focus, in this chapter, on Arthur Prior's infamous challenge to explain how and why inference rules might define a concept. The literature on this challenge is extensive. In this chapter, I will give an account of Prior's challenge, and sketch the broad outlines of possible responses to that challenge, spending most of the chapter on the kind of response we can make using the tools of the first three chapters. Then, in the next chapter, on *positions* we will turn from these general considerations to the particular way that rules of inference interact with norms governing speech acts such as assertion, denial, supposition, and the like. With those connections at hand, we will then have the resources to give, in Chapter 6, a fresh account of the way particular kinds of rules might *define* a concept, and so meet Prior's challenge. So, let's begin to face Prior's challenge head on.

4.1 | PRIOR'S CHALLENGE

Suppose you notice that we use the word "and" in English in various ways. Your friend, who is a logician, wants you to focus on a particular sense in which you could mean "and", by drawing your attention to the following rules of inference:

$$\frac{A \quad B}{A \text{ } \mathit{and} \text{ } B}\,\mathit{andI} \qquad \frac{A \text{ } \mathit{and} \text{ } B}{A}\,\mathit{andE}_1 \qquad \frac{A \text{ } \mathit{and} \text{ } B}{B}\,\mathit{andE}_2$$

I use "*and*" here, not "$\wedge$" or "$\otimes$" since these natural deduction rules do not quite line up with those for either additive or multiplicative conjunction in the absence of either *Contraction* or *Weakening*.

If you had grasped some kind of practice of inference, then you have the means to grasp something about *this* concept of conjunction. In some sense, it seems to count as a kind of definition. Using these rules, for example, you could infer from the premise A *and* B to the conclusion B *and* A, so whatever this "*and*" means, the order of its conjuncts does not make a material difference, unlike in "I went out and I had dinner" if this does not entail "I had dinner and I went out". So, it seems that we can use the rules of inference of a proof system to define what we mean.

Provided, of course, that the order of premises in a proof does not matter, and that you can chain proofs together.

Arthur Prior, in an influential paper from 1960, challenged us to find the limits of such an approach. He asks us to consider the following definition of a new kind of connective [176]:

$$\frac{A}{A \textit{ tonk } B}\,{\scriptstyle tonkI} \qquad \frac{A \textit{ tonk } B}{B}\,{\scriptstyle tonkE}$$

As you can quickly see, if you use these two rules, it allows for rather many proofs. Here is a simple example:

$$\dfrac{\dfrac{2+2=4}{2+2=4 \textit{ tonk } 2+2=5}\,{\scriptstyle tonkI}}{2+2=5}\,{\scriptstyle tonkE}$$

From the (true) premise, to the effect that $2+2=4$, we have shown (somewhat surprisingly) that $2+2=5$. This is a powerful proof technique. Perhaps *too* powerful.

As Prior states it, if we take the rules for *tonk* to be *definitional* for the concept, then the proof is *analytically valid*. Its validity turns only on the meaning of the concept *tonk*. Since its inference rules are given by definition, the validity is given in virtue of the meanings of the terms involved.

Excursus on analytic validity: If you have been influenced by analytic philosophy from the 1960s, you will probably be familiar with Quine's critique of the traditional notion of analyticity [179, 180, 181]. The everyday notion of analytic truth involves a mixture of *metaphysical* and *epistemic* features. Following Boghossian, it is useful to distinguished these two elements in the following way:

- A sentence S is EPISTEMICALLY ANALYTIC if and only if "mere grasp of S's meaning by T suffice[s] for T's being justified in holding S true." [25, p. 334]
- A sentence S is METAPHYSICALLY ANALYTIC if "in some appropriate sense, it owes its truth-value completely to its meaning, and not at all to 'the facts'." [25, p. 334]

In our case, we are concerned not with analytic *truth* but analytic *validity*, and for Prior's purposes, the in question notion is more metaphysical than epistemic. We are not concerned so much with what we might know to be valid or be justified in holding to be valid, but simply in what *is* valid, so the relevant sense of analyticity in our notion of analytic validity is the metaphysical one. While we need not tarry over Boghossian's use of 'the facts' in opposition to 'meaning', the core notion in analytic validity is that an argument is analytically valid, if its validity is given completely by the meanings of the terms involved—and in particular, in the case of the argument involving *tonk*, its validity due to the meaning—the *definition*—of the term *tonk*. *End of excursus*

So we have Prior's complaint: if positing rules of inference is enough to define connectives, then we can find analytically valid proofs for any argument from a premise P to a conclusion C, by way of the intermediate step P *tonk* C. It follows, then, that the criterion of analytic validity is too weak to be useful in drawing distinctions between the valid and the invalid. If all that it takes to introduce a concept is to specify rules—without any other constraint on what those rules might be—and if all it takes for an argument from a premises X to conclusions Y to be analytically valid is for there to be some language in which there is some proof following the rules of that language which takes us from the premises X to the conclusion Y, then *any argument at all, from a given premise to a given conclusion* counts as analytically valid. This is absurd, so something must have gone wrong somewhere. But where?

» «

This is Prior's challenge for those who take inference rules to count as definitions, and for these rules—these definitions—to underwrite the validity of inferences. Having heard the challenge, let's sketch out the kinds of responses that might be available to us.

4.2 | WHAT COULD COUNT AS A RESPONSE?

Of the making of responses to Prior's paper, there is no end. There is no way I could give even cursory attention to the whole field of responses to Prior. We will start, instead, by charting the field of possible responses to the challenge of *tonk*.

According to Google Scholar, Prior's paper has been cited 535 times as of July 2018. The critical literature engaging with Prior's two-page paper is vast.

RESPONSE 1: NIHILISM One response to Prior's challenge is to concede its conclusion. You could agree with Prior that analytic validity, so understood, is hopelessly broad. Concepts such as *tonk* can indeed be defined by their characteristic inference rules, and so, definability is no guide to a boundary between the logical and the non-logical, or any boundary worth exploring. There is no notion of analytic validity worth saving.

This isn't to say that Prior agreed with the conclusion of the paper. Regardless, we will use the name "Prior" for a proponent of the view argued for.

RESPONSE 2: LOCALISM Another response is to grant that analytic validity draws a boundary, but that boundary is always relative to a given language. So, you can concede the conclusion this far: that for any argument, there's some language $\mathcal{L}$ in which that argument is analytically valid, but deny that it's analytically valid in *this* language. Whether a deduction is to be accepted is, then, not just a matter of its validity, but also whether the concepts employed in the deduction are available in the language at hand—and if not, then we do not take it as given that if something can be defined, then there is any sense in which it should be defined. So, if the question of the validity of the argument from A to B comes up, and an interlocutor says: let's define *tonk* as follows … and then proceeds to reason from A to A *tonk* B and thence to B, we might respond to the offer of defining *tonk* in that way with "I'd rather not." We grant that it is, in some sense, a definition, but it is one which leads us to a language we'd rather not speak. There is a cost associated with making a definition like that for *tonk*.

While this sort of response seems to be possible in some sense, it does seem to come with its own costs. When an interlocutor asks us to define a notion in a given way, there is some cost involved in rejecting that offer if the definition is in order. There is no problem in rejecting a proffered definition if that definition is, in some sense, out of bounds. If someone asks you to introduce the number term n as the number satisfying $n \times 0 = 1$, and then proceeds to prove that $1 = 2$, since $1 = n \times 0 = n \times (0 \times 2) = (n \times 0) \times 2 = 1 \times 2 = 2$, then it is a completely satisfying response to say that there is no such number as n since there is no number x at all where $x \times 0 = 1$, since $x \times 0 = 0$ for every number x. The purported definition for n is out of bounds. But the localist cannot make such a response, because according to her, the definition for *tonk* can be used to define a concept. It is not out of bounds in the sense that the purported definition of n is. Some other account must be given if we are to understand when a definition is acceptable and when it is not.

RESPONSE 3: REJECTION Perhaps we should not have granted Prior's conclusion that accepting the rules for *tonk* as definitions leads us inexorably to triviality. After all, the inference from A to B through A *tonk* B chained together two steps in natural deduction. Viewed as a derivation in the sequent calculus, it involves a *Cut* step:

$$\dfrac{\dfrac{A \succ A}{A \succ A \text{ } \textit{tonk} \text{ } B}\,\textit{tonkR} \qquad \dfrac{B \succ B}{A \text{ } \textit{tonk} \text{ } B \succ B}\,\textit{tonkL}}{A \succ B}\,\textit{Cut}$$

Perhaps we could block the argument in the following way. The argument from A to A *tonk* B is analytically valid (that's just *tonk* introduction). The argument from A *tonk* B to B is analytically valid (that's just *tonk* elimination). However, the argument from A to B is not analytically valid. Sometimes *Cut* fails. In particular, *Cut* can fail if the inferences use concepts such as *tonk* [42, 74, 203, 204, 243]. To be sure, rejecting the transitivity of consequence seems like a significant step. We have taken the validity of *Cut* for granted. In the first Part of this book, we only ever eliminated *Cut* to the extent that its removal as a primitive rule made no difference to what sequents were derivable. We never countenanced eliminating *Cut* in the stronger sense of countenancing having one proof from A to B, another, from B to C, and no way of chaining them together to have a proof from A to C. That would be a real *revision of*—and perhaps, a significant *abberation from*—our target notion of proof. If you make this step, however, you do have another kind of response to Prior's challenge.

» «

These three responses to Prior's challenge are, so far at least, the minority responses. The dominant theme of responses to Prior is an attempt to provide criteria distinguishing *good* rules, like those for *and*, which suffice to define a concept, and those for *tonk*, which fail to do so. Attempts to find such criteria themselves split into two distinct species. These species are well represented by the two initial responses made by two articles in *Analysis* immediately after the publication of Prior's original paper. The first is the approach taken by J. T. Stevenson in his 1961 paper "Roundabout the Runabout Inference-Ticket" [234], which gave a criterion for rules in terms of *model theory*. The second species of response is well represented by Nuel Belnap, in his 1962 paper "Tonk, Plonk and Plink" [15], which gave a criterion for rules in the vocabulary of *proof theory*.

response 4: model theoretic criteria Stevenson's reply argued that there was a significant difference between acceptable rules, such as the rules for conjunction, disjunction or negation, and unacceptable rules, like those for *tonk*. The difference was that the rules for the traditional connectives can be satisfied by a connective given an interpretation as a truth function, while there is no truth function available that can certify the rules for *tonk*. I will discuss this approach—and other model theoretic approaches like it—in more detail in the next section. Save to say that Stevenson has given a principled reason for allowing some inference rules to qualify as definitions while rejecting others.

response 5: proof theoretic criteria The boundary between acceptable and unacceptable rules can be drawn in a different way, using the intrinsic features of the rules themselves. Perhaps the most influential initial response to Prior's challenge was Nuel Belnap's reply, which argued that rules like those for *tonk* fail to be either *conservatively extending* or *uniquely defining*, and hence, cannot be taken to be definitions of the same kind as those for conjunction, disjunction, and other logical concepts. These criteria are stated in purely proof theoretic terms, and need make no recourse to truth functions or other model-theoretic structures.

After discussing such model theoretic approaches in the next section, the rest of this chapter will discuss Belnap's two criteria, that rules should be conservatively extending and uniquely defining, explain how they apply to different families of inference rules, and then relate these to other proof theoretic criteria, such as *Harmony*.

» «

Before we get to that, let's pause to reflect on the shared assumptions behind responses 4 and 5. These are the responses that take there to be differences to be found between one class of putative definitions and another, so let's call these responses *splitting* responses to Prior's challenge, distinguishing them from the *lumping* responses (1–3). This crucial point is that some putative definitions succeed in defining concepts of a particular kind, while others (such as those for *tonk*) do not. This could be understood in a number of different ways. One way is to take the rules for *tonk* to fail to define *anything at all*. This approach takes the criteria offered in splitting responses to be criteria to be satisfied by any definition. Since the rules for *tonk* do not satisfy the criteria, they do not define. Another way to understand these responses is conclude that the rules for *tonk* may count as a definition, but not for a logical connective. (The rules for conjunction, disjunction and the like, satisfying the relevant criteria, are taken to define logical connectives.) According to this approach, the criteria in the splitting responses are criteria for a concept's *logicality*. Yet another response is to take it that there is *some* distinction between the rules for *tonk* and those for *and*, and the criteria in either splitting response are ways for giving an account of just what such a difference might consist in. In the rest of this chapter, we will allow for each of these readings of the splitting responses. They all offer a distinction among putative definitions. This distinction might be understood as the distinction between a genuine definition and a failed definition; or a definition of a logical concept and the definition of something else; or between definitions with some other relevant difference, salient to Prior's challenge.

4.3 | ANSWERING WITH MODEL THEORY

So, let's return to Stevenson's response to Prior's challenge [234], and explain the distinction it draws. Recall the putative definition (or clarification) of the concept of conjunction, given by the rules *andI* and *andE*.

$$\frac{A \quad B}{A \text{ } \mathit{and} \text{ } B}\,{\scriptstyle \mathit{andI}} \qquad \frac{A \text{ } \mathit{and} \text{ } B}{A}\,{\scriptstyle \mathit{andE}_1} \qquad \frac{A \text{ } \mathit{and} \text{ } B}{B}\,{\scriptstyle \mathit{andE}_1}$$

If we consider this an introduction of an item of new vocabulary ("*and*") to our language, we then ask: is there something it could mean that satisfies these rules? It has the syntax of a binary connective. If we understand the validity of an argument as requiring that when the premises are true, so is the conclusion, then *andI* tells us that if A and B are both true, then A *and* B must also be *true*. If, for the moment, we call a sentence false whenever it is not true, this means that *and* must behave in the following way:

A	B	A *and* B
false	false	?
true	false	?
false	true	?
true	true	true

where the sentence A *and* B is unconstrained with respect to truth in the first three rows, but it must be assigned true when A and B are both assigned true.

To satisfy the *andE* rules, we require that if A *and* B is assigned true, then so must A and B. So, since in the first three rows, we do not have A and B both assigned true, we cannot assign true to the conjunction. To satisfy the *andE* rules then, we need the truth value of the conjunction to behave like this:

A	B	A *and* B
false	false	false
true	false	false
false	true	false
true	true	?

Putting these two partial tables together, we see that the *andI* and *andE* rules impose compatible constraints. They can be satisfied by a connective for which truth is assigned as follows:

A	B	A *and* B
false	false	false
true	false	false
false	true	false
true	true	true

The process works in much the same way for the other connectives and their usual natural deduction rules.

Excursus: I say the process works in *much* the same way, but it is not *quite* as straightforward as the conjunction case in general, as we saw when we encountered Carnap's Lemma (Lemma 3.29, page 140). The single conclusion natural deduction rules do not quite suffice to pin down the standard truth table interpretations for all connectives. They do not rule out non-standard valuations like ν^{true} (according to which, all formulas are true) or $\nu^{\top}$ (according to which, all tautologies are true and every other formula is false), so single conclusion rules for disjunction cannot rule out evaluations which make disjunctions true despite not making either disjunct true. However, these subtleties do not blunt Stevenson's point. The fact that the rules are not strong enough to eliminate non-standard interpretations is irrelevant to the point being made here, which is that at the concept like *and* can be given an interpretation in standard valuations which underwrites its inference rules. The fact that the rules are compatible with some non-standard valuations is beside *that* point. *End of Excursus*

However, things go very differently when we consider the rules for *tonk*. Consider Prior's rules:

$$\frac{A}{A\ \textit{tonk}\ B}\,\textit{tonkI} \qquad \frac{A\ \textit{tonk}\ B}{B}\,\textit{tonkE}$$

For *tonk* to satisfy the introduction rule, it needs to satisfy these constraints:

A	B	A *tonk* B
false	false	?
true	false	true
false	true	?
true	true	true

whenever A is true, so is A *tonk* B. To satisfy the elimination rule, the constraints are:

A	B	A *tonk* B
false	false	false
true	false	false
false	true	?
true	true	?

whenever B is *false* so is A *tonk* B. Combining these two tables, we see the problem:

A	B	A *tonk* B
false	false	false
true	false	true, false
false	true	?
true	true	true

The connective has been totally unconstrained in the third row (where A is false and B is true) and it is overconstrained in the second (where A is true and B is false). According to *tonkI*, A *tonk* B is true when A is true. According to *tonkE*, A *tonk* B is false when B is false. If A is true and B is false, we have a problem. Using *tonk* in this way places constraints on the concept that cannot be met.

This is Stevenson's criterion in action. There is no possible truth evaluation for sentences involving *tonk* that allow for the rules to be satisfied in their full generality. For connectives like conjunction, there *is* such an evaluation. This is the difference between rules for conjunction and Prior's *tonk*.

» «

What can be said about Stevenson's approach? For one thing, it does draw a genuine distinction among rules. When it comes to propositional connectives, the criterion seems well motivated. It does seem, at first glance, to rely on a commitment to truth functional semantics. One way of understanding Stevenson's constraint is to take it that the truth functions are the *genuine* semantic values of the propositional connectives, and the inference rules for a connective succeed in picking out the meaning of a concept if there is already a meaning (one of these semantic values) that satisfy the constraints expressed in the inference rule. The meaning is (and could independently be given by) the truth function. Inference rules do the job only as ways of describing those truth functions in a different vocabulary. This is *one* way of understanding Stevenson's criterion, but it is not the only way. It is an open question whether Stevenson himself understood the criterion in this way. He does write:

> The important difference between the theory of analytic validity as it should be stated and as Prior stated it lies in the fact that he gives the meanings of connectives in terms of permissive rules, whereas they should be stated in terms of truth-function statements in a meta-language. [234, p. 127]

For Stevenson, a rule is 'permissive' when it has the form "from these premises you are permitted to deduce this conclusion."

which seems to indicate that the meanings of connectives should be given by traditional truth functions. However, it may well be that non-truth functional connectives (such as modal operators, temporal operators, etc) can be given meaning in terms of truth functions in concert with other apparatus, such as worlds, times, etc. We can leave this aside, for the moment, because Stevenson's simple criterion can be seen to apply even if the meaning of a connective isn't fixed by a straightforward truth function.

To see why such an alternative understanding of this approach is possible, consider someone who seeks to clarify the way in which the modal operator $\Box$ is to be understood, by giving the following two inference rules:

$$\frac{\begin{array}{c}\vdots\\ A\end{array}}{\Box A}\,\Box I \qquad \frac{\Box A}{A}\,\Box E$$

The rule $\Box E$ is straightforward. It tells us that whenever $\Box A$ is true, so is A. The introduction rule is a little more complex. It is to be understood with the following side condition: if we have proved A *from assumptions that each start with a* $\Box$ then we can infer $\Box A$. The understanding is that if we prove something from premises that are (taken to be) necessary, then that thing is also (under the assumption of those premises) necessary. We can argue amongst ourselves whether this pair of rules should be understood as giving a definition of $\Box$, but before we do that, let's see how they fare under Stevenson's gaze.

This is intended to bring to mind a natural deduction rule for the universal quantifier $\forall$, for which we can infer $\forall x Fx$ if we have proved Fa from premises *which do not involve the term* a.

As we have already seen, the $\Box E$ rule tells us that that whenever $\Box A$ is true, so is A. The $\Box I$ rule doesn't tell us anything concrete about the truth value of $\Box A$ when A is true, for it tells us something when A is true and it has been proved in a particular way. (If you know some modal logic, you will be able to use some simple models to verify that these rules are satisfied in some circumstances where A is true while $\Box A$ is false.) But we need not stop here: one thing we *can* say is that if $\Box A$ were to act in the following manner:

A	$\Box A$
false	false
true	true

then the $\Box$ rules would be satisfied. The rules $\Box I$ and $\Box E$ satisfy Stevenson's constraints. There is *some* truth function that vindicates the rules, though the rules do not require that $\Box$ be interpreted in terms of that particular truth function. Consider an interpretation where sentences

In fact, these rules hold in more models than these. For those who know modal logic, these rules are satisfied in any model for the modal logic S4, which is strictly weaker than the logic S5, given by these particularly simple modal models with a universal accessibility relation.

are true or false at some number of *worlds*. A formula $\Box A$ is true at a world w if and only if A is true at all of the worlds. Under an interpretation of this shape, the rules can be seen to be vindicated, in the sense that for our rules, if the premises are true (at a given world) then the conclusion is also true (at that world). This clearly applies for $\Box E$. If $\Box A$ is true at a given world, then A is true at all worlds, so A is true at our given world. For $\Box I$ we reason as follows: We have a proof of A where all the premises have the form $\Box B$. So, in any world where each premise $\Box B$ is true, A is true. This means for *any* world, A is true, since if $\Box B$ is true in some world, it is true in all of them. So, $\Box A$ is also true at our given world.

So, there is an unending supply of different non-truth functional interpretations of $\Box$ that can also vindicate the rules. The rules $\Box I$ and $\Box E$ allow for a range of different interpretations, given different families of worlds.

We have seen, then, that the rules $\Box I$ and $\Box E$ have *an* interpretation in truth functions. This interpretation distinguishes this pair of rules from Prior's rules for *tonk*, which have no such interpretation. However, the rules for $\Box$ allow for multiple interpretations, while it seems that the rules for *and* allow nowhere near so much freedom. This raises the question of whether the rules $\Box I$ and $\Box E$ could be said to be *definitions*, if they allow for so much freedom. After all, I could learn to reason with a $\Box$ operator using these rules, and so could you, with no guarantee that we are talking "about" the "same thing." This can be seen in the following precise sense. Take a model in which formulas are evaluated as true or false at the numbers $0, 1, 2, \ldots$, and let $\Box^r A$ be true at the (odd) number $2n + 1$ iff A is true at every odd number; while $\Box^r A$ is true at the (even) number $2n$ iff A is true at every even number. We interpret $\Box^c A$ in a similar way: $\Box^c A$ is true at $2n$, and at $2n + 1$ iff A is true at both $2n$ and $2n + 1$. In other words, looking at the following diagram:

$$\begin{array}{cccccc} 0 & 2 & 4 & \cdots & 2n & \cdots \\ 1 & 3 & 5 & \cdots & 2n+1 & \cdots \end{array}$$

the sentence $\Box^r A$ counts as true at a number iff it is true everywhere along its *row*, while $\Box^c A$ counts as true at a number iff it is true at both numbers in that *column*. It is easy to see that the rules $\Box I$ and $\Box E$ preserve truth, when the $\Box$ in question is taken to be $\Box^r$ or $\Box^c$. Suppose I already have operators like $\Box^r$ and $\Box^c$ in my vocabulary, and you attempt to introduce $\Box$ to me, using the rules $\Box I$ and $\Box E$. I am well within my rights to ask you: which operator do you mean—$\Box^c$, $\Box^r$ or something else? There is nothing in the rules themselves to tell us which concept you intend. Perhaps it is better to conceive of these rules not as providing a *definition* of a single concept, but as giving us a *constraint* which may be satisfied by a number of different of concepts.

Instead of spending more time on this in the context of model theoretic criteria for rules, we will leave the rest of this discussion for Section 4.5 (see page 163), where we will consider uniqueness, understood proof theoretically, as a constraint on rules.

» «

Before turning to proof theoretic approaches to understanding the behaviour of rules as definitions, let's turn to the role of truth functions in Stevenson's response to Prior's challenge. For it is clear that the Stevenson's argument could apply more generally than simply in the realm of truth functions. If there is a different range of models (say, possible worlds models from modal logic, models constructed out of algebras of more than two semantic values, or some other such thing), and there is some way to evaluate validity of arguments with respect to those models, then we could use these models, instead of two valued Boolean evaluations, as certification for our rules. The form of the argument remains the same, but the tool for evaluation is different.

The proponent of Stevenson's approach is going to be *forced* to widen her arsenal beyond Boolean evaluations, if her attention turns to rules for the quantifiers. Let me propose using $\forall$ in the following fashion:

$$\dfrac{\begin{matrix}\vdots\\ A\end{matrix}}{\forall x A|^{n}_{x}}\,\forall I \qquad \dfrac{\forall x A}{A|^{x}_{t}}\,\forall E$$

where the formula $A|^{s}_{t}$ is found by taking the formula A and replacing each instance of the term s in A by the term t. In the rule $\forall I$, the side condition is that the term n does not occur in any of the premises active in the subproof of A. The intended interpretation is that whatever we have proved about n when we prove A holds *generally* (since the proof makes no assumptions concerning n, it is *arbitrary*), so it holds for *all* objects. This rule does not identify the meaning of $\forall x A$ with every instance of the formula A applied to every object for which we happen to have a name (that would be a so-called "substitution" interpretation of the quantifiers [63]). Rather, the interpretation involves a notion of generality. To derive $\forall x A$, we wish to derive $A|^{x}_{n}$ *generally*, in such a way that we do not assume anything in particular about n. So, it follows that what we have proved "about n" we have proved about *anything at all*.

Note: there are some subtleties hiding in the notion of substitution to do with free and bound variables in A and in t. See Chapter 8 for the details. Those details are not important here.

See Chapter 8 for more on the motivation and applications of rules of this kind.

The rules $\forall I/E$ are rules for a universal quantifier. Is there a truth functional verification for this pair of rules? If by "truth functional" we mean the traditional sense in which the truth value of a complex formula is determined as a function of the truth value of its constituent formulas, the answer is—not *really*. The constituent of $\forall x A|^{n}_{x}$ is the formula $A|^{n}_{x}$, in which the variable x is free. This is not traditionally thought of as a *sentence*, since it has a dangling unbound variable x. (Is "x is even" true? It depends on what we're doing with the variable x. Universally bound, it is false. Existentially bound, it is true.) However, we *could* assign it a truth value if we were forced to do so. Looking at the rule $\forall E$, we see that if $\forall x A$ is true, A must be true too. This would force the following constraint on the "truth functional interpretation" of $\forall$:

Or should that be '*mis*-interpretation'?

A	$\forall x A$
false	false
true	?

If we wish $\forall xA$ to ever be true at all (and we stick with our crazy plan to interpret $\forall$ as truth functional in this sense), we see that $\forall xA$ must be true when—and only when—A is.

A	$\forall xA$
false	false
true	true

But now consider a language where we have assigned Fa the value true and Fb the value false for two different names a and b. In this case, $\forall xFx$ must be false (otherwise, we could infer the false Fb from the true $\forall xFx$, using $\forall E$). So, using our truth functional interpretation for $\forall$, we must have Fx false. If the language contains *negation*, and if negation is given the usual truth functional interpretation, then $\neg Fx$ is true. It then follows that $\forall x\neg Fx$ is true, and by $\forall E$, so is $\neg Fa$. This contradicts the assignment of true to Fa (at least, if $\neg$ is interpreted using its usual truth function). It follows that if the universal quantifier is interpreted using this (terrible!) truth function, we cannot countenance interpretations where atomic predications give different values for different names. If Fa is true, so is Fb. This seems like too much of a price to pay to interpret the quantifier as truth functional.

Before you object too strenuously that no-one would think that the quantifier had that truth functional interpretation, notice that we were prepared to make the concession in the case of the modal operator $\Box$. No-one would think that a genuine notion of necessity is *intended* to be interpreted by the trivial truth function (which sends true to true and false to false), but the fact that it *could* be so interpreted was taken to be enough to assure us that the rules $\Box I$ and $\Box E$ are in some sense, safe. We cannot do the same for the universal quantifier, unless we restrict our attention to interpretations where names are doing no work—where atomic sentences Fa and Fb are coordinated to always have the same value, independently of which names feature. This is predicate logic in name only. The sleight of hand involved in interpreting a non-truth functional notion as truth functional just for the purpose of assuring us of its safety does not work in the case of the universal quantifier.

Of course, we might respond by saying that the reading of truth functionality in terms of *constituents* is altogether too tight. An interpretation may be found in which the truth value of the formula $\forall xA$ is given in terms of the truth values of the formulas $A|^x_a, A|^x_b, \ldots, A|^x_t, \ldots$ for each and every term which can feature in A. This is certainly stretching the notion of "truth functional" further than the simple case, given that the relevant function has infinitely many inputs but let's stick with this for a moment or two. The only appropriate truth function (in this sense) to interpret the universal quantifier would assign $\forall xA$ the value true if and only if *every* instance $A|^x_t$ is assigned true. (The *only if* part is required by $\forall E$. The *if* part is given by the desire that $\forall$ formulas turn out true at least some of the time—as is given by the fact that if $\top$ is derivable from no premises, so is $\forall x\top$.) This is certainly closer to something that a universal quantifier could *mean* than our previous attempts. However, this reading of the quantifier is still unsatisfactory. The new truth "function" rules out interpretations where $\forall x A$ is false, but where $A|^x_t$ is true for each term t in our language. However, this is completely coherent if we allow the possibility of our language not having a name for everything. If we take it that something is A (that not everything is not A), but nothing we have a term for is itself A. This interpretation is ruled out by taking the quantifier to be truth functional in this extended sense.

So, the defender of Stevenson's approach, when faced with quantifiers, finds herself pushed in the direction of tools and techniques beyond truth functions. The suitable tools are available: Tarski's models for first-order predicate logic—in which formulas are evaluated relative to assignments of values to the variables, where values are assigned from some given domain—are eminently suitable. But once we make this move, the appropriate interpretation for $\forall$ which will vouchsafe the rules $\forall I$ and $\forall E$ are no longer truth functions in a straightforward sense. They are at most, truth functions relative to a complex semantic machinery. If I propose a new set of rules, for a higher order quantifier, or an operator like a modal operator, or an intensional (non-truth functional) conditional, or something else entirely, it is very hard to state in advance *what* kind of place I should look to find appropriate interpretations for which I am to test my rules. The task of underwriting these rules is much more complex. Given that the model theory of an expression of arbitrary syntactic category is itself a difficult task, Stevenson's approach, when shifted out of the bounds of straightforwardly truth functional propositional connectives, becomes significantly more complex.

This is not, in itself, an argument against Stevenson's approach. If suitable model theoretic machinery is available to make distinctions between rules, then let us put it to use! However, the tools for the task may be difficult to find, and they may also prove difficult to wield, if we are fortunate enough to find them. Perhaps another approach is possible. Given that the rules have been presented to us as *inference rules*, it may well turn out that the criteria we might use to judge them could be found by assessing them on their own terms, without having to go through the detour of relating them to models. After all, as we have seen in Chapter 3, we can understand models as arising out of more fundamental proof theoretical concerns. In the light of these results, it should be no surprise that the kinds of constraints Stevenson attempts to mark out using models can be found using the tools from proof theory instead. This was Nuel Belnap's insight in his reply to Prior [15], so it is back to proof theory that we turn in the remaining sections.

4.4 | CONSERVATIVE EXTENSION

The first step to understand a putative definition—like Prior's rules for *tonk*—in its own terms is to understand the context. When we are offered a definition of a new item of vocabulary, we are offered a new *language*. Given our initial language (call it $\mathcal{L}_1$), with the addition of the new item of vocabulary, we define a new language (call it $\mathcal{L}_2$). Where the lan-

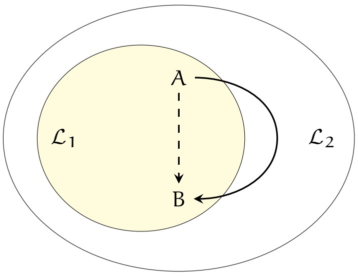
If $\mathcal{L}_1$ contains other connectives, such as $\wedge$, then presumably $\mathcal{L}_2$ contains sentences like $(A\ \textit{tonk}\ B) \wedge C$ too, though the details of exactly how $\mathcal{L}_2$ is formed from $\mathcal{L}_1$ do not matter here. The details will matter in Chapter 6, however, so we will be more explicit there.

guage $\mathcal{L}_1$ had formulas like A, B, etc., the new language has all of the same sentences as are in $\mathcal{L}_1$, but also new sentences, formed using the new item of vocabulary. So, $\mathcal{L}_2$ contains sentences such as $A\ \textit{tonk}\ B$, $A\ \textit{tonk}\ (B\ \textit{tonk}\ C)$, and so on.

However, the syntax of the new language is only a small part of the picture. $\mathcal{L}_2$ also has a notion of *proof*, given by adding *tonkI* and *tonkE* to the stock of inference rules inherited from $\mathcal{L}_1$. So, we have two languages, $\mathcal{L}_1$ and $\mathcal{L}_2$ and two families of proofs.

$$\mathcal{L}_1 \quad \rightsquigarrow \quad \mathcal{L}_2$$

Proofs in $\mathcal{L}_2$ can make use of vocabulary distinctive to $\mathcal{L}_2$ (like *tonk*), but they can also make use of vocabulary that was present in $\mathcal{L}_1$. We call the extension from $\mathcal{L}_1$ to $\mathcal{L}_2$ *conservative* over provability when the following condition applies:



DEFINITION 4.1 [CONSERVATIVE EXTENSION] $\mathcal{L}_2$ is a CONSERVATIVE EXTENSION of $\mathcal{L}_1$ if and only if whenever there is a proof *in* $\mathcal{L}_2$ from premises in the language $\mathcal{L}_1$ to conclusions also in that language, then there is also a proof from those premises to those conclusions *in* $\mathcal{L}_1$.

That is, $\mathcal{L}_2$ is a conservative extension of $\mathcal{L}_1$ if and only if there is no way go through $\mathcal{L}_2$ vocabulary to get a new $\mathcal{L}_1$ conclusion that we couldn't already arrive at, using the tools of $\mathcal{L}_1$ alone. Provided that $\mathcal{L}_1$ is a language where there is *some* pair of formulas A and B where there is no $\mathcal{L}_1$ proof from A to B, then Prior's addition of *tonk* to form the new language $\mathcal{L}_2$ is not conservative. For $\mathcal{L}_2$ contains the proof from A, through $A\ \textit{tonk}\ B$ (by *tonkI*) to B (by *tonkE*). There is no proof in $\mathcal{L}_1$ from A to B, but there is such a proof in $\mathcal{L}_2$. Here, $\mathcal{L}_2$ is a non-conservative extension of $\mathcal{L}_1$. The facts about what follows from what are not preserved. We have *revised* our account of consequence in the $\mathcal{L}_1$ vocabulary, not *extended* it. If we were satisfied that the account of consequence we had in $\mathcal{L}_1$ was correct, as far as it went, then extending our vocabulary with *tonk* to move to $\mathcal{L}_2$ would not be seen as an advance. We would take it to be a mistake.

This is not to say that Prior's *tonk* is always and from every starting point a bad move. Belnap's criterion of conservativeness is always a relative notion. Though we have seen that the addition of *tonk* is non-conservative over some languages, it is not that way *always*. If our initial language was $\mathcal{L}_0$, in which there is a single atomic formulas p_0, for which the single identity proof, from p_0 to itself, counted as the only proof, then from *here*, the addition of *tonk* counts as conservative. Now we have new vocabulary: $p_0\ \textit{tonk}\ p_0$ is in the new language and it wasn't present in $\mathcal{L}_0$. We can indeed make the inference from p_0, though $p_0\ \textit{tonk}\ p_0$, back to p_0, but this is not telling us any fact about consequence that we did not have before. The direct route from p_0 to itself was always available. So, *tonk* is conservative over some languages, but not others. To be sure, *tonk* fails to be conservative over a vast range of languages—any in which there are A and B where there is no proof from A to B, and where chaining proofs together yields another proof—but the point remains, this isn't quite *everything*. Belnap's criterion of conservative extension does not, in itself, yield a categorical judgement over putative definitions. It gives us a way to evaluate putative definitions as extensions from different possible starting points.

» «

We have seen that Prior's *tonk* fails Belnap's criterion of conservative extension, at least when it is offered as an extension to a language which has a pair of sentences with no proof from one to the other. Is there a sense for which the common logical vocabulary (say, the rules for *and*, which have been the other running example in this chapter) *are* conservative? How would we show that the addition of the rules for *and* do manage to be conservative over an initial language $\mathcal{L}_1$?

The reasoning would go as follows. Let's now take $\mathcal{L}_2$ to be the language extending $\mathcal{L}_1$ with the binary connective *and*, introduced with the following two rules:

$$\frac{A \quad B}{A \textit{ and } B}\,{\scriptstyle andI} \qquad \frac{A \textit{ and } B}{A}\,{\scriptstyle andE_1} \qquad \frac{A \textit{ and } B}{B}\,{\scriptstyle andE_2}$$

We wish to show that in the new language $\mathcal{L}_2$, if we have a proof from premises to conclusions in the $\mathcal{L}_1$ vocabulary, there is some proof doing the same job in the old vocabulary. As was the case for *tonk*, this is *not* always the case. Let's take our starting language $\mathcal{L}_1$ to consist of the language with a stock of atomic formulas and the binary conditional connective $\rightarrow$, governed by the linear natural deduction system of Chapter 1. Extending this language with the rules for *and* is non-conservative. We can see this, by attending to the following short proof:

Here, we return to Exercise 21 from Chapter 1.

$$\cfrac{\cfrac{\cfrac{p \quad [q]^{(1)}}{p \textit{ and } q}\,{\scriptstyle andI}}{p}\,{\scriptstyle andE_1}}{q \rightarrow p}\,{\scriptstyle \rightarrow I,1}$$

Now we have a proof in $\mathcal{L}_2$ from p to $q \rightarrow p$—both from the language $\mathcal{L}_1$—going through the new connective *and*, where there was no $\mathcal{L}_1$ proof from p to $q \rightarrow p$. (Remember, linear logic does not allow for vacuous discharge.) The rules are non-conservative—at least from this starting point.

It is not difficult to see why. The rules for *and*, when chained together, allow us to do something that we could not do in $\mathcal{L}_1$.

$$\cfrac{\cfrac{A \quad B}{A \textit{ and } B}\,{\scriptstyle andI}}{A}\,{\scriptstyle andE_1}$$

We can bring A and B together in a proof, so that *both* are present in a proof of the one formula A. In other words, we can bring an irrelevant side formula B alongside a formula A, adding it as a premise, even though there is no need for B to occur as a premise for A. These two conjunction rules allow us to make this manoeuvre, even if it was not something we could do before. In our starting language, $\mathcal{L}_1$, the rules for proof were restrictive enough to ensure that we could not always bring irrelevant side formulas in as extra premises in proofs. So, if we begin with $\mathcal{L}_1$, adding the rules *andI* and *andE* is non-conservative. Adding them constitutes a revision of $\mathcal{L}_1$'s account of proofs.

The discussion here is, of necessity, very brief. For more on this topic, consult a recent paper by Hjortland and Standefer [104], which addresses these issues more systematically.

So, when are the rules for *and* conservative? Are they ever safe, in Belnap's sense? Here, proof theory can also provide us with some answers. If the notion of proof is not that of linear implication, but allows for weakening, in the sense that we already had a proof in the $\mathcal{L}_1$ vocabulary from A and B to A (and another proof, to B)—as there is if we allowed for vacuous discharge, moving beyond linear logic to at least affine implication, then we could reason as follows: Suppose we have an $\mathcal{L}_2$ proof which uses the rules for *and*, but whose premises and conclusion do not feature the new concept, but are in the $\mathcal{L}_1$ vocabulary. If the language $\mathcal{L}_2$ allows for the normalisation theorem (Theorem 1.13, see page 22), and normal proofs have the subformula property (See Theorem 1.11 on page 19) then we transform the proof in $\mathcal{L}_2$ into a normal proof. Since the premises and conclusion in this proof are all $\mathcal{L}_1$ formulas, and since the normal proof has the subformula property, now all of the formulas in the proof are also $\mathcal{L}_1$ formulas. (We suppose that the constituents of sentences in $\mathcal{L}_1$ remain in $\mathcal{L}_1$.) So, there is no way for this normal proof to feature the rules for *and*, since these rules always feature the new vocabulary, restricted to $\mathcal{L}_2$.

For this reasoning to work, we need to check that normalisation works for $\mathcal{L}_2$ proofs. The crucial step for this is where we take a detour through an introduction and then an elimination of an *and*.

$$\cfrac{\cfrac{A \quad B}{A \text{ and } B}\ \textit{andI}}{A}\ \textit{andE}_1$$

How are we to transform this proof into a normal proof? There are two choices. One is require that $\mathcal{L}_1$ allow for explicit steps of weakening. We ask for this

$$\cfrac{A \quad B}{A}\ K$$

to be a *proof* in $\mathcal{L}_1$. This would do no violence to what is provable in $\mathcal{L}_1$, since there is already a proof from A and B to A, namely:

$$\cfrac{\cfrac{A}{B \to A}\ {\to}I \qquad B}{A}\ {\to}E$$

where the introduced B in the $\to I$ step is vacuously discharged. This, too, is not normal. Normalisation for *this* proof, using the standard techniques of normalisation, from Theorem 1.13, would transform the proof into the simple identity proof A, for which the extra premise B disappears. However, if we countenance explicit weakening steps, we don't need to do this: we could keep B around as an extra (irrelevant) premise. Or, if we like, we can drop the B, at no cost to the resulting logic, because there is nothing that requires that premises be *present* for them to be discharged. In either case, normalisation proceeds as normal, and we can transform any proof of a sequent from $\mathcal{L}_1$ vocabulary using *and* into a proof that restricts its resources to the $\mathcal{L}_1$ vocabulary alone.

» «

This is the first insight into how the tools of the first part of this book will play a role in answering the questions raised by Prior's *tonk*. Normalisation (and Cut elimination) and the resulting subformula property, will show us how and when proofs using a larger vocabulary can be transformed into proofs utilising a restricted vocabulary. This gives us a way to address the criterion of conservative extension, and thereby to ensure that a proposed extension to our language is indeed an extension, and not a revision.

There is much more to say about conservative extension, both why the constraint is an important one, and more importantly, the techniques we can use to assure ourselves that a proposed extension of the language is conservative. For an extended discussion of this constraint, in a wider setting, we will wait for Chapter 6.

4.5 | UNIQUENESS

Belnap's response to Prior is only partly on the criterion of conservative extension. Belnap's chief insight—and here he goes beyond Stevenson—is that putative definitions can fail in two ways. Not only can they be too strong (by being non-conservative). They can also fail by being too weak (by failing to be uniquely defining). We have already been given a hint of this phenomenon, when considering the fate of the proposed account of modal operators in the $\Box I$ and $\Box E$ rules on page 155. Here, we saw that the rules are weak, in the sense that they allow for multiple distinct realisations. We saw that we could have interpretations where there are distinct $\Box$-like connectives, $\Box^r$ and $\Box^c$, both satisfying the $\Box I$ and $\Box E$ rules, but with completely different interpretations. There is no sense of *the* modal operator $\Box$ satisfying the $\Box$ rules. There is not one, there are many. Belnap's second constraint on a putative definition is that it be *uniquely defining*. If we want to define a term, and not merely give a partial description of some of its features, then uniqueness is a constraint worth having. In the rest of this section, we will see what this might mean, and how we might find it when it is there to be had.

As with the case of conservative extension, it will be helpful to see how unique definability can be understood from a proof theoretical perspective, in a case like *and*, before moving on to other cases like the modal operators.

Let's return to Prior's original presentation of the rules for *and*, and suppose that instead of adding this item of vocabulary to a language that has no notion of conjunction, I *already* had a connective—which I'll write &—which is governed by these rules:

$$\frac{A \quad B}{A \,\&\, B}\,{\scriptstyle \&I} \qquad \frac{A \,\&\, B}{A}\,{\scriptstyle \&E_1} \qquad \frac{A \,\&\, B}{B}\,{\scriptstyle \&E_2}$$

What can we say about the relationship between *and* and &? Do we have one concept under two different guises, or do we have two concepts? One answer to the question is found by looking at what we can do in the new vocabulary. We can form the following two proofs:

$$\frac{\dfrac{A \textit{ and } B}{A}\,{\scriptstyle andE_1} \qquad \dfrac{A \textit{ and } B}{B}\,{\scriptstyle andE_2}}{A \,\&\, B}\,{\scriptstyle \&I} \qquad\qquad \frac{\dfrac{A \,\&\, B}{A}\,{\scriptstyle \&E_1} \qquad \dfrac{A \,\&\, B}{B}\,{\scriptstyle \&E_2}}{A \textit{ and } B}\,{\scriptstyle andI}$$

So, we have a proof from A *and* B to $A \,\&\, B$, and back. As far as proofs are concerned, any proof we have *from* one kind of conjunction can be transformed into a proof from the other, and any proof *to* one kind of conjunction can be transformed into a proof to the other. They are indistinguishable as far as provability is concerned. Well, except at the cost of inflating of the number of premises involved. If we are back in the land of linear logic, and do not—this time—allow multiple discharge, then this can make a material difference. If, on the other hand, we allow for contraction—for multiple discharge—then strictly speaking, anything we can prove with one conjunction we can prove with the other.

So, in this sense, the rules for conjunction are uniquely defining. The rules for the two conjunctions allow us to uniformly transform proofs involving one into proofs involving the other. If I have both conjunctions in my vocabulary, there is no way to split them, at least at the level of proof.

Excursus on meaning: This is not to say that all aspects of meaning or significance are shared between *and* and &. On some accounts of the meaning of the word '*but*', as a sentence connective this also satisfies the rules for *and*. From A *but* B I can infer A, and I can infer B. From A and from B I can infer A *but* B. Does it follow that '*and*' and '*but*' mean the same thing? I don't think it does. At the very least, one implicates something that the other does not. So, perhaps the kind of uniqueness delivered by proof rules like these can go so far toward uniqueness but not further. *End of excursus*

Unique definability is bound up with the provability of identity sequents for the introduced connectives. The derivation $Id_{A \textit{ and } B}$ (see the proof of Theorem 2.3 on page 53) which gives us A *and* $B \succ A$ *and* B, at least in the presence of contraction, can be transformed into derivations for A *and* $B \succ A \,\&\, B$ and $A \,\&\, B \succ A$ *and* B. The tools used for understanding the structural rules of *Identity* and *Cut* play a role in understanding unique definability and conservative extension.

» «

Let us now consider what might be said about unique definability and the rules $\Box I$ and $\Box E$. If we consider the case of two connective satisfying these rules—$\Box^r$ and $\Box^c$—let's see how we might attempt to prove $\Box^r A$ from $\Box^c A$. The 'proof' could go like this:

$$\dfrac{\dfrac{\Box^c A}{A}\,\Box^c E}{\Box^r A}\,\Box^r I$$

There is only one problem. This is not a proof. The rule $\Box^r I$ dictates that a $\Box^r$ can be introduced when the formula proved immediately above has been proved from premises prefixed by $\Box^r$. Here, the premise is not prefixed by $\Box^r$, but by $\Box^c$. As it stands, this is not a proof. If we had some prior notion of *modal formula* and we agreed that $\Box^r A$ and $\Box^c A$ both counted as modal formulas, and if we reinterpreted the $\Box I$ rule as saying that we could introduce a $\Box$ if the derived formula was proved on the basis of *modal formulas*, then all would be well, and this would now count as a proof. $\Box^c$ and $\Box^r$ would collapse into the one notion, and the $\Box$ rules would be uniquely defining—relative to the background notion of what counted as a modal formula.

Varying what counts as a 'modal formula' in this sense, is a well-known feature of natural deduction systems for modal logics, going back to Prawitz's groundbreaking work in the 1960s [169, Chapter 6].

Here, again, we see the subtle variations that are possible in the application of Belnap's conditions. Unique definability is also not an all-or-nothing thing. It is sensitive to what Belnap called the "antecedently given context of deducibility" [15, page 131]. This context might the structural rules (identity, *Cut*, contraction, weakening), and also the choice of syntax—such as some independent characterisation of what counts as a 'modal formula'. We will see more such choice points in the remaining chapters, where we consider further, and more refined applications of Belnap's constraints of conservative extension and unique definability. For now, though, let us end this chapter with some other characterisations of proof theoretical criteria for rules, used to reply to Prior's challenge.

4.6 | HARMONY

Conservativeness, as we have seen, is not intrinsic to the system of rules, but is a relational property, between rules and the antecedently given context of deducibility. It has seemed to many that there is some intrinsic property of rules which could be used to account for the difference between the rules for the standard connectives, and defective rules like those for *tonk*. Is there some intrinsic property which could do the work? Some cryptic remarks from Gentzen point to a possibility:

> … the introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. This fact may be expressed as follows: In eliminating

> a symbol, we may use the formula whose terminal symbol we are dealing only 'in the sense afforded it by the introduction of that symbol'. An example may clarify what is meant: We were able to introduce the formula $A \rightarrow B$ when there existed a derivation of B from the assumption formula A. If we then wished to use that formula by eliminating the $\rightarrow$ symbol . . . we could do this precisely by inferring B directly, once A has been proved, for what $A \rightarrow B$ attests is just the existence of a derivation of B from A. [80, page 80, 81 of the Szabo translation]

I have changed Gentzen's notation. His $\mathfrak{A} \supset \mathfrak{B}$ is our $A \rightarrow B$.

The motivation is clear. An introduction rule, like $\rightarrow I$ or *andI* is taken to be the *definition* of the introduced connective, and the elimination rules for the connective simply reverse the introduction rules in the sense that they allow us to return to the conditions from which we introduced the formula. In the case of $\rightarrow I$, since I can introduce the formula $A \rightarrow B$ when I am in possession of a proof from A to B, reversing this procedure means that I can infer *from* the assumption $A \rightarrow B$ that I can derive B *from* A:

$$\dfrac{\begin{matrix}[A]^{(i)}\\ \vdots \\ B\end{matrix}}{A \rightarrow B}{\scriptstyle \rightarrow I^{i}} \qquad \rightsquigarrow \qquad \dfrac{A \rightarrow B \quad A}{B}{\scriptstyle \rightarrow E}$$

Similarly, given that I can introduce A *and* B on the basis of A and of B, then, so the explanation goes, from A *and* B I can recover A, and recover B.

$$\dfrac{A \quad B}{A \textit{ and } B}{\scriptstyle andI} \qquad \rightsquigarrow \qquad \dfrac{A \textit{ and } B}{A}{\scriptstyle andE_1} \quad \dfrac{A \textit{ and } B}{B}{\scriptstyle andE_2}$$

This works, up to a point. In the absence of vacuous discharge, these *andE* rules do genuine damage to the proof. If our proofs encode a notion of relevance, then decomposing A *and* B into A alone (and leaving out the B) is not returning us to where we were before. A more subtle inference rule is required to do that. Suppose we have a proof which takes us from the *two* premises A and B to some conclusion C. Then instead of appealing to those two premises A and B, we could discharge them, and appeal to the one premise A*and*B instead. The new, more discriminating elimination rule for *and* has this shape:

$$\dfrac{A \textit{ and } B \quad \begin{matrix}[A, B]^{(i)}\\ \vdots \\ C\end{matrix}}{C}{\scriptstyle andE^{i}}$$

This rule discharges the two assumptions A and B, and it allows us to reason *from* A *and* B in exactly the same way as you reason from the premises used in its introduction rule—and it does this regardless of whether vacuous or multiple discharge are present in the underlying natural deduction system. A process like this can be used to generate elimination rules from introduction rules for other connectives, too.

You can recover the old elimination rules for *and* using vacuous discharge:

$$\dfrac{A \textit{ and } B \quad [A]^{1}}{A}{\scriptstyle andE^{1}}$$

We vacuously discharge a B.

» «

When we attempt this process on Prior's introduction rule for *tonk*, we see immediately that we don't generate his proposed elimination rule.

$$\frac{A}{A \textit{ tonk } B}\,{\scriptstyle tonkI} \quad\rightsquigarrow\quad \frac{A \textit{ tonk } B}{A}\,{\scriptstyle tonkE^*}$$

Instead, we get a straightforward elimination rule which returns us from A *tonk* B back to A. A *tonk* B has the same effect in proofs as A alone. There is no overgeneration here. The mismatch between *tonk*'s introduction and elimination rules is found, in that they fail to be related in the way that Gentzen took *all* his connective's rules to be related.

» «

But *why* should introduction and elimination rules be related in this way? For Michael Dummett—who developed Gentzen's insight in a systematic way from the early 1970s [58, 59]—the answer is given in a comprehensive theory of meaning. It is worth quoting a passage from Dummett's *Frege: Philosophy of Language* at length, to see how he addresses the question of constraints on a definition.

> No one now supposes that established linguistic practice is sacrosanct. The supposition that it is, which was the fundamental tenet of 'ordinary language' philosophy, rested on the idea that, since meaning is use, we may adopt whatever linguistic practice we choose, and our choice will simply determine the use, and hence the meanings, of our expressions: any proposal to alter established practice is therefore merely a proposal to attach different senses to our expressions, whereas we have the right to confer on them whatever senses we please. The error underlying this attitude lies, in part, in the failure to appreciate the interplay between the different aspects of 'use', and the requirement of harmony between them. Crudely expressed, there are always two aspects of the use of a given form of sentence: the conditions under which an utterance of that sentence is appropriate, which include, in the case of an assertoric sentence, what counts as an acceptable ground for asserting it; and the consequences of an utterance of it, which comprise both what the speaker commits himself to by the utterance and the appropriate response on the part of the hearer, including, in the case of assertion, what he is entitled to infer from it if he accepts it. [58, page 396]

For Dummett, meaning, understood as norms for appropriate use of an expression, has two aspects: (1) the conditions under which we can assert the sentence, and (2) the consequences of that assertion. These can be seen reflected in an account of proof as the *introduction* and *elimination* rules for an expression. Dummett states here that these two aspects of an expression's meaning should be in *harmony*. The proposal that *tonk* should have the introduction and elimination rules that Prior gives it is then a violation of this condition concerning harmony, and it can be rejected on these grounds.

These evocative terms are used by Robert Brandom for the introduction and elimination conditions for a concept [31, pages 193, 194]

Dummett has more to say on exactly *why* harmony is required between the *upstream* and *downstream* connections of an assertion — and more needs to be said about how these aspects of meaning arise out of our linguistic and conceptual practices. More, too, should be said about whether this constraint is one that allows us to *evaluate* concepts or whether it is somehow constitutive of being a concept. On the former approach, concepts can be disharmonious — and the disharmony incurred is a part of the "price of admission", the commitments you must be prepared to undertake, were you to use the concept. On the latter approach, no concepts are or can be disharmonious. Both options are open for the friend of harmony.

» «

The notion of harmony, as Dummett states it, involves a balance between the introduction and elimination rules for a concept. As far as that is concerned, there is no requirement that the introduction rule be primary. We could, after all, define something by means of its elimination rule, and then demand that the introduction rule be in harmony with it. For some proponents of Harmony—such as Dummett himself—the introduction rule nonetheless plays a primary role, motivated by a verificationist or a constructivist theory of meaning [59]. The meaning of an assertion is given, in the first instance, by the conditions for verification of that judgement. This is best expressed by its introduction rule. However, in the subsequent literature on harmony and the specification of "correct" inference rules, the priority of introduction rules has faded into the background, and the focus has turned to a more balanced reading of the fit between *I* and *E* rules. So, for example, in a clear recent discussion of harmony conditions, Florian Steinberger distinguishes four different ways to understand how rules for an expression can fail to be in harmony.

- *E-strong disharmony*: the E-principles are unduly permissive (relative to the corresponding I-principles);
- *E-weak disharmony*: the E-principles are unduly prohibitive (relative to the corresponding I-principles).
- *I-strong disharmony*: the I-principles are unduly permissive (relative to the corresponding E-principles);
- *I-weak disharmony*: the I-principles are unduly prohibitive (relative to the corresponding E-principles). [232, pages 621, 622]

Here, there is no set priority for introduction or elimination rules. They have been treated on a par: either the introduction or the elimination rules can be thought of as the standard by which the other rules are evaluated.

The failure of the *tonk* rules are seen to be quite drastic, on Steinberger's account. The fact that *tonkE* allows us to infer B from A *tonk* B where *tonkI* did not give B as one of the premises from which we could infer A *tonk* B means that the *tonk*'s *E*-principles are unduly permissive (relative to its *I*-principles), so it is *E*-strongly disharmonious. Taking the *I* rule as given, the *E*-rule produces something it shouldn't. But this can be seen also as a case of *I*-weak disharmony, as it could be seen as the 'fault' of the *I*-rule, in that it prohibited the case of the inference to A *tonk* B from B, which is required to match the *E*-rule. So, it is both a case of *E*-strong and *I*-weak disharmony.

But the *tonk* rules can be seen to fail on the *other* grounds, too. At least at first glance, the *tonkE* rule does not allow us to immediately recover A from A *tonk* B. Since it does not license that inference, the *tonkE* rule is unduly prohibitive, relative to its *I*-rule, so the rules are in *E*-weak disharmony. Reading this in reverse, it is also a case of *I*-strong disharmony. The rules are badly balanced in all four senses.

I said, in that last paragraph that this holds 'at first glance'. If you look for longer, you can see that even though the *tonkE* rule doesn't immediately license the recovery of A from A *tonk* B (since it only directly licenses the inference to B), this doesn't mean that we cannot infer from A *tonk* B to A by other means. After all, we have the *tonk* rules in their full power. Here is one way to do it: infer from A *tonk* B to (A *tonk* B) *tonk* A (using *tonk I*) and then to A (using *tonk E*). So, we can recover the conditions from which A *tonk* B was inferred, in a more roundabout way. The *E*-rule didn't *prohibit* the recovery of the premise A. It just didn't supply it all on its own. So, is this a case of *E*-weak disharmony after all?

Here we see that there is a tension in the *locality* or *intrinsicness* of harmony conditions such as these. Harmony conditions are motivated by their locality. It seems that there is something suspicious in the rules for *tonk* in and of themselves, without recourse to all the details of the wider system in which they are placed. Harmony considerations are intended as a way to give account of this intuition. There seems to be no need to look at the *entire* antecedently given context of deducibility in order to check whether the harmony conditions for a suite of rules are satisfied. However, some non-local features remain. In the verification that a system of rules satisfies the conditions for harmony, we need to at the very least compose inference rules. This relies on general features of the shape of inferences. Whether the rules for *and* are harmonious or not depends on the specifics of the discharge policy in the underlying natural deduction system. Recall: we can mimic vacuous discharge by weakening in an irrelevant B, by inferring A *and* B from A together with B and then eliminating that conjunction to leave A, now deduced "from" A and B. Are these rules in harmony? The answer is always and only relative to the structural rules of the underlying proof system.

There is no way to escape all of the non-local features of the system of inference rules, when reasoning about harmony. This should not be a surprise, since inference rules are designed to be composed: they are the building blocks for proofs, and they make sense only in the context of a wider space of possible combinations. But these are *exactly* what Nuel Belnap called the antecedently given context of deducibility in his initial response to Prior. The features Belnap singled out as salient are these [15, page 132]:

AXIOM	$A \vdash A$	
RULES	*Weakening*:	from $A_1, \ldots, A_n \vdash C$ to infer $A_1, \ldots, A_n, B \vdash C$.
	Permutation:	from $A_1, \ldots, A_i, A_{i+1}, \ldots, A_n \vdash B$ to infer $A_1, \ldots, A_{i+1}, A_i, \ldots, A_n \vdash B$.
	Contraction:	from $A_1, \ldots, A_n, A_n \vdash C$ to infer $A_1, \ldots, A_n \vdash C$.
	Transitivity:	from $A_1, \ldots, A_m \vdash B$ and $C_1, \ldots, C_n, B \vdash D$ to infer $A_1, \ldots, A_m, C_1, \ldots, C_n \vdash D$.

And you can see that these are a characterisation of the structural rules of some system of proof. Here, a decision is made in favour of multiple discharge and vacuous discharge (*Weakening* and *Contraction*). Here, the decision concerning atomic proofs is given (*Axiom*). It is also made explicit that the order of premises makes no difference, concerning proofs (*Permutation*). And finally, here, the principles of composition of proofs is laid down (*Transitivity*). These are the most important features concerning the ambient structure of proofs. A view about the structure of proof is implicit in all of these principles: we are reasoning with a consequence relation between a list of premises and a single conclusion. Only given a background context like this can *any* discussion of harmony, or of conservative extension and uniqueness, take place. Despite the motivation to find 'intrinsic features' concerning specifications of rules, in order to single out rules as harmonious or disharmonious, this ambient background context plays an important and unavoidable role.

So, it is to this background context that we will turn in the next chapter.

POSITIONS 5

In this chapter, we step back from the discussion of inference rules and definitions to focus on what Nuel Belnap called the "antecedently given background context of deducibility" [15]. We saw, in the last chapter, that the space of possible proofs is central to any account of inference rules and definitions of concepts. We need to understand the space in which our definitions are to be given, if we wish to have an overview of the scope and power—and limits—of our ability to define concepts by inference rules. Our background material in Part 1 has given us ample evidence that the shape of proofs, discharge policies, positions in sequent systems and structural rules play a central role in proof theory. This comes to a focus in the role that *positions* play, in proof theory, in the connections between proof theory and model theoretic semantics (see Chapter 3), and also to the rules governing the practice of our talk and thought. Positions are where the rubber of the formal calculus of proof theory hits the road of our practice. So, positions are our topic in this chapter.

5.1 | LANGUAGE

In the first part of this book, we examined proofs, derivations and positions, and the basic constituents of each of these structures were formulas — sentences in a formal language. Now that we are going examine the fundamentals of how proofs are grounded in our practice, we need to spend just a little time spelling out the nature of the language in question. In Chapter 1, as I introduced natural deduction for conditional formulas, I talked of conditional *judgements*. Notice that 'judgement' is an ambiguous expression. It is ambiguous between the *act* of judging, and the content of what is judged. The same is true of 'proposition'. A proposition might be the act of proposing, or the content of what is proposed. This double meaning is not a coincidence. The sentences in our languages do double duty. I can declare that *the Route 58 tram is leaving soon*, using the sentence "*The Route 58 tram is leaving soon*", and you can agree with me, and form another judgement with the same content—perhaps using the same words if you were to voice them aloud, or if some time passes, with different words (such as "*The Route 58 tram* was *leaving soon*"). There are individual acts of judgements (or considerations or wonderings or supposings or questionings whether …) which are always the acts of agents, and then there are the contents of those judgements, which are shared between agents. You can agree with something I say. We can jointly wonder whether something is the case. You can *prove* to me that something you have said follows from premises I believe. The contents of judgements are, in some sense, public, communicable and

"Now the term 'thought' is has a wide range of application, including such items as assumptions, the solving of problems, wishes, intentions and perceptions. It is also ambiguous, sometimes referring to *what* is thought, sometimes to the *thinking* of it." Wilfrid Sellars, "Language as Thought and as Communication" [222, page 65].

shared. None of this hangs on a detailed and worked out account of what a content *is*. I do not here presuppose some settled view on *what* it is that is shared when you and I disagree on whether *the Route 58 tram is leaving soon*.

There are two ways of understanding how this sharing of content arises. One way is to focus on our internal mental states and concepts, to attempt to account for shared concepts when different agents are related to the world in the same way (an externalist account of shared concepts) or are encoded in the same way or serve the same role (an internalist account of the same issue) and to think of our public languages are selected to express the antecedently existing mental representations. Another way to understand the relationship between language and mental states is that our acquisition and competence with language helps constitute those mental states—that our engagement with our public language is part of how those mental states are identified and individuated. We do not need to settle the details of the relationship between thought and talk for the project before us. However we prefer to settle this issue, it will do us well to remember that language, as a human and social practice, is not just a *game*, even though it is not only rule governed, but normatively constituted, and it involves scorekeeping and other gamelike aspects. No, although languages are gamelike, they are also embodied. The claim that *the Route 58 tram is leaving soon* is not merely an abstract marker like a bishop or a pawn or the square e4 which can be equally represented in wood or steel or paper or bits and bytes, and be totally abstract positions in a structure. The judgement involves the *world*, and an agent's making their way in that world. Language, as used, involves our beliefs, our commitments (even when we are engaged in make believe or deception or supposition), and our position in the world. It is not so much a *game* as a *sport*, in which embodiment plays a central and ineliminable role [128].

» «

What is important, in what follows, is that we have identified a notion of *samesaying* in our judgements. It is at least possible for me to assert something that you can deny — or agree with. If I suppose that p then that is (at least in principle) something you might do, too, and there is a fact of the matter about whether or not you are also supposing that p. In everyday reasoning, this is completely straightforward and unobjectionable. If I say "*the Route 58 tram is leaving soon*" and you nod and say "*a ha*," then I take you to agree with me, and also to take it that the tram is leaving soon. Sometimes, identifying what is said and whether two judgements succeed in samesaying is more difficult: I point in the general direction of some marsupials to the left in the zoo enclosure and say to you, "that's a wallaby over there" then if you interpret my pointing to be gesturing not at the wallabys but at the kangaroos that were also nearby, you might take me to have said that *that* (kangaroo) was a wallaby. The use of demonstratives and indexical expressions makes it a subtle matter to identify samesaying. Regardless of the difficulties, we will presuppose that we can account for straightforward cases of samesaying. Although this is a difficult matter to identify in general, in many of the languages we will consider, we will err on the side of simplicity, an in many cases, absent indexicals and demonstratives, the surface syntax of the sentences will suffice for specifying the samesaying relation.

Regardless of any difficulty or complexity in how different languages are implemented, we are not going to tarry with any of these details. This is still an account of *formal* logic, intended to apply equally wherever a practice of judgement and reasoning is in place. The argument of this chapter and the rest of this Part is unashamedly *formal*, not in the sense that it is not able to be *applied*, but to the contrary, because it can be applied in a number of ways. It is intended to be applied generally, so generally that in *this* chapter we abstract away from any of the detail concerning *any* of the contents of the judgements expressible in the languages under discussion. In the remaining chapters that follow, we focus on the patterns and features shared by many different languages and practices. The atomic formulas in our analysis are places where we allow for further analysis but give none of our own here and now. If it helps, you can think of a formulas A as a sentence in some public language. This will not lead you very far astray, as long as that sentence does not contain any indexicals or contextually shifting devices. The crucial matter is that each occurrence of A in a proof or a derivation or a position is *samesaying* with every other occurrence. Given that condition, the analysis can proceed.

That's what it is for this to be formal. The items of our formal language are contents that can be asserted and denied, but we will, in the next two chapters, play little heed to the nature of the content of those assertions and denials. The *structure* of norms governing the very acts of assertion and denial will give us enough to be going on with, especially when we attend to the individual and to the social norms governing those these speech acts.

5.2 | ASSERTION AND ITS NORMS

Assertion and denial are, fundamentally, communicative speech acts. An act of assertion has a source and a target—a speaker and a hearer. The assertion may be spoken, signed, written, or mediated in some other way. The hearer may be present with the speaker, or a long way away in space or in time. The audience may be a single person, or many. The hearer may be the speaker themselves, as assertion may be internalised. Assertions may be overheard and received by those other than the intended audience (you may overhear a conversation on the tram, or read the letters exchanged between two people [130]). In any case, assertion—and denial—is a directed act, involving a speaker and a hearer.

We have already seen that assertion and denial involve *content*. I can assert *that the Route 58 tram is leaving soon*, thereby informing you of the situation. Someone else could disagree with me, and *deny* that claim. The contents of assertions and denials are the kinds of things that we can agree with or disagree with, suppose or wonder about, and take as premises or conclusions in our reasoning. I could wonder whether the tram is leaving soon, or suppose that it is, or conclude that it is. These contents are, at first blush, the components of our proofs as attempt to reason and to resolve or clarify our disagreements.

Assertions and denials are also governed, in some sense, by norms. They can be correct or incorrect. The acts of asserting and denying can be judged appropriate or inappropriate, and can succeed or fail. Given the nature of the communicative act, these norms can be understood in in terms of either side of the transaction, focussing on the site of *production* (the one asserting) and *consumption* (the audience). We can consider norms governing of assertion and denial at a range of levels of analysis, focussing on the individual—the producer of the speech act—or on the communicative context into which the speech act is performed.

TRUTH, KNOWLEDGE AND MORE

Much of the literature on norms governing assertion focus on the norms by which we can judge an assertion to be *correct*, or which can help the speaker evaluate which assertions should (or should not) be made. One obvious candidate norm is a *truth* norm:

[TRUTH NORM] Assert that p only if is is the case that p.

If I say that the Route 58 tram is coming soon, and it turns out that it *doesn't* come soon, then my assertion was incorrect. I erred. I may have been sincere in my error. I may have believed—and had good reason to believe—that the tram was coming soon, and it may have been delayed without my knowledge. Although I have violated the truth norm, I did so sincerely. I have not violated the following *belief* norm:

[BELIEF NORM] Assert that p only if you believe that p.

Another norm like these two, which is very significant in the recent literature is much stronger than the belief norm and the truth norm (it entails both of them, but is entailed by neither—nor by both together). It is the *knowledge* norm.

[KNOWLEDGE NORM] Assert that p only if you know that p.

If I asserted that the Route 58 tram is coming soon—without having any information to that effect, merely on the basis of a whim—and it turns out that it does arrive, then I have not violated the first two norms, but I have violated the knowledge norm. I have been lucky and vindicated, through no effort of my own. I have asserted something beyond my knowledge.

The production and assessment of norms for assertion is a thriving industry in epistemology and the philosophy of language. I will not enter into the debate concerning the relationship between these norms. There is a sense in which incorrectness (violating the truth norm), insincerity (violating the belief norm), and unjustifiedness (violating the knowledge norm) are all possible grounds for criticising an assertion, so

For a good introduction to the literature, start with Brown and Cappellen's edited collection *Assertion: New Philosophical Essays* [33], and McKinnon's *The Norms of Assertion* [140].

any comprehensive account of norms of assertion must have something to say about all these norms, and the relationships between them. (This is not to say that the norms governing assertion are encompassed by these three. There are many other things one could say concerning these norms and other close relatives. For example, Jennifer Lackey presents a sustained criticism of the *Knowledge Norm* and proposes replacing it with what she calles the '*Reasonable to Believe Norm of Assertion*' [124].)

» «

There is something correct in accounts of norms of assertion like the truth, belief and knowledge norms. Nonetheless, something is missing. Each of these norms evaluate assertion insofar as it is as a production, but they say nothing at all about its significance for the hearer. To some degree, this independence from the hearer is to be expected in virtue of the normative structure of assertion. If I say to you that the tram is coming soon, and someone else overhears that, they also learn that the tram is coming soon. (This holds in general, for assertion, provided that the hearer understands the vocabulary and the contextually determined features of the utterance.) The same can not be said for a request. If I ask *you* to tell me when the tram is coming, and someone else overhears the request, I have not thereby asked *them*. The making of a request is directed to the addressee in a way that an assertion need not be. The *upshot* of the assertion for someone who overhears it is, in general, no different than its upshot for the direct addressee. Given this, it is not surprising that we can make significant headway concerning norms for assertion when we focus on 'quality control' norms for the speaker. However, this is not all that there is to say about assertion.

Rebecca Kukla and Mark Lance give an insightful analysis of norms governing speech acts in terms of the four-way distinction between particularity and generality for the *hearer*, and for the *speaker* in their book book *'Yo!' and 'Lo!': The Pragmatic Topography of the Space of Reasons* [123].

Focussing on the act of assertion without paying attention to the dialectical or conversational context in which an assertion is made threatens to obscure certain important features from view. Consider a conversation where we are disagreeing on something (say, whether or not the tram indicator board at our tram stop is receiving current information from the tram network), and I say "Suppose you're right and this sign is displaying current information. It says that the tram is coming in 3 minutes. So, it will be here soon." In this context, I seem to have asserted that the tram will be here soon, but I have done this under the scope of a supposition which I have explicitly made, and to which I have directly appealed. Suppositions are not constrained by any norm of belief or truth or knowledge. Has my assertion — under the scope of that supposition — failed in any way? It needn't express my *belief*, but it need not be insincere for all that. It has the grammatical form of an assertion. It can be used as a premise for further reasoning, it can be agreed with or disagreed with. It *is* an assertion, but not one which is used to express my current state of mind on the issue. It is an assertion that I am using for other effects — to help elaborate *your* state of mind and to draw out its consequences. In this dialogue, the source and the target, the speaker and the audience, both help constitute the field of play to which the norms apply.

Norms with the form of the truth, belief and knowledge norms —

which say that it is correct to assert p only when [some condition concerning the speaker and not the audience] — have nothing to say about the context of the audience, or the shape of any dialogue between speakers. Other accounts of norms must be attended to if we are to address these aspects of the acts of assertion and denial.

COMMITMENT

Brandom [29, 30, 31] and MacFarlane [135], in the neo-pragmatist tradition, give an account of the speech act of assertion in terms of commitment. They follow Charles Sanders Peirce, who articulated this approach as follows:

> Let us distinguish between the *proposition* and the *assertion* of that proposition. We will grant, if you please, that the proposition itself merely represents and image with a label or a pointer attached to it. But to *assert* that proposition is to make oneself responsible for it, without any definite forfeit, it is true, but with a forfeit no smaller for being unnamed. [160, p. 383–384]

On Peirce's approach (and here he is followed by Brandom and by MacFarlane) to assert that p is to commit yourself in some way concerning p, particularly, concerning p's truth. If I assert that p, then I risk some kind of forfeit or penalty if p turns out to not be true. The nature of the commitment that I undertake can be spelled out in a number of different ways, in terms of its upstream and its downstream connections. If I assert that p, I commit myself to discharge requests for justification, when asked. So, if I tell you that the tram is coming, it is appropriate for me to respond in some way if you ask "how do you know?" or "why is that?" If I can give no suitable answer, I should not be surprised if you don't agree with me that the tram is coming. In making the assertion also license you to refer back to me for a justification if you pass on the assertion. This is why testimony is possible. If I tell you that the tram is coming, since you rely on me for the justification of the claim, you can pass the claim on, allowing for the justification to pass on to me, if questions are raised. My link in the process of answering the question can be discharged by passing it on to you. Similar claims can be made concerning the nature of the sanction or forfeit that is applied if it is discovered that my assertion is false. If by asserting that p, I am committing myself to discharge requests to justify p, to answer questions such as "why is that?" or "how do you know?" concerning p. If it's discovered that p is not in fact the case, then there is no way I can discharge that debt. I am seen to be bankrupt—when it comes to p, at least. I have written a cheque that is seen to bounce.

So, with this understanding of the norms governing assertion, we see that an assertion is the undertaking of a commitment concerning the content of that assertion in a social practice of the pooling of information. We make information available for use, and play our part in answering for it.

» «

As MacFarlane points out in his “What is Assertion?” [135], this commitment analysis of the nature of assertion can help account for the applicability of the truth, belief and knowledge norms, where they in fact apply. If, when I assert that p, I make myself available to discharge a request to justify p, and I leave myself open to some kind of forfeit if p fails to be true, then you can see why I ought not assert falsehoods. In everyday cases, I run the risk of being exposed as a bankrupt if I make claims that are untrue, so it stands to reason that I ought not assert falsehoods. Of course, it is *possible* to lie or deceive, or otherwise tell untruths. For example, in teaching contexts, it is often desirable to make oversimplified (and literally *false*) claims as the student is beginning to learn, and only to refine those claims as the learner gains the conceptual capacity to understand more complex matters. Think, for example, of the teacher who introduces Newtonian mechanics, despite knowing that literally speaking, these claims are untrue. This can be understood as straightforward assertion, where the teacher has a reasonable expectation of how the claim is going to be used and understood, what questions are going to be asked (and, in particular, what questions are *not* going to be asked). The teacher also knows that if the students go around asserting claims about Newtonian motion to others, and refer back to the teacher to vouch for them, anyone to whom the distinction between Newtonian and relativistic mechanics matters will understand to take the students’ claims as simplifications. These simplifications do no harm: as far as this discourse goes, the claims being made are enough to be going on with. The same goes for assertion under the scope of a supposition. If, in some context, I we decide to *assume* p, and then we go on to reason to conclude q, we may well *assert* it at this point in the discussion, despite not taking it to be *true*. In this case, our obligation to answer questions concerning q can be discharged *relatively*. We can point to the assumption of p, and make out the case relative to that. Similarly, our license for others to assert q still obtains, *relative* to that supposition.

What goes for the truth norm can also be seen to apply to the belief norm, and the knowledge norm. If I assert that p and so, I indicate that am happy to vouch for p, and for others to do so to, referring back to me to vouch for them, then it is reasonable for you to take me to believe it, or at least to have further questions if you discover that I don’t. For example, I might conclude something quite surprising and initially unbelievable from premises that I do believe. Suppose I spell out the consequences of a theory I believe, and discover that by way of some proof, that this apparently overturns something else I have believed. In this case, I well assert that conclusion in the course of the proof, while also indicating that I don’t really know what to make of it, or whether I believe it. In this case, the assertion seems well in order. I can vouch for the claim (by way of its proof from a theory I’m committed to), you can also use this justification if you care to pass the claim on, or defer back to me, despite the fact that my own cognitive state has not necessarily caught up with what I have discovered. In this case, the assertion seems completely in order, despite the violation of the belief norm, and the account given here seems to give us some way of showing why it is so.

Attending finally to the knowledge norm, we can see, again, that it can be explained in terms of the commitments undertaken in giving an assertion — if I am going to vouch for the truth of p, what better way to discharge the responsibility than to know that p? — and we can also see that violations of the knowledge norm can also be explained in cases of exceptions. (The instruction case, as well as the surprising conclusion case are both violations of the knowledge norm, given that knowledge involves truth, and belief.) Other cases in which the knowledge norm is sometimes taken to be too demanding (when the asserter has reason to believe which falls short of *knowledge*) can be accommodated in some way by attending more carefully to the details of the commitment to vouch for the truth of the asserted content. Instead of paying close attention to these debates, we will instead turn to one other significant account given of the nature of assertion, also characterising the dialogical and social consequences of assertion, on how it changes the *common ground*.

COMMON GROUND

Another approach to norms governing assertion is found in Robert Stalnaker's analysis of assertion as a proposal to update the *common ground* in a discourse [229, 230, 231]. This, along with Brandom's commitment model of assertion, is a way of keeping track of the *score* in a conversational context [132]. The common ground in a dialogue is the set of propositions that are treated as presupposed or taken for granted by the participants, or equivalently (for Stalnaker), the set of possible worlds that are left open as genuine possibilities. As assertions are made, more propositions are added to the common ground, and the corresponding set of open possibilities shrink. This model of conversation and information update has proved incredibly fertile in semantics, in giving an account of the dynamics of discourse and the behaviour of many different linguistic phenomena, such as presupposition, illocutionary mood, parentheticals, and not-at-issue content [71, 149, 225]. A great deal can be learned in the shift to a dynamic perspective, understanding speech acts as moves that update the scoreboard, rather than as static representation [242, 247]. The crucial feature of common ground is that it is *common* to participants in a dialogue. For me to take something to be in the common ground is (in part) to take you to take it to be so, as well. So goes the public feature of dialogue. If you grant something that I assert — or at least, if you do not deny it — it is entered into the common ground, and can be built on in what follows. This is a part of what makes conversation and dialogue a shared activity, what it means for information to be pooled.

It is important to recognise that allowing something to enter the common ground does not mean that all participants *believe* it. It is possible to agree (or concede) "for the sake of the discussion". Nonetheless, as far as the discussion goes, the point is agreed, and in the absence of moving to have the item struck from the record, future moves in the discussion can point back to that entry or build upon it. The common ground is a way to give an account of the shared part of the scoreboard of the conversation, as well as the individual scores taking into account each different perspective of each participant.

» «

In "What is Assertion?" [135], John MacFarlane complains that common ground approaches make little sense of assertions conducted outside the canonical setting of a conversation between a sharply delimited set of speakers. For example, it is harder to make use of the conception of a common ground in circumstances where one is speaking to oneself, or to no-one in particular, and public assertions, such as in the context of a filmed interview, where there is a conversation taking place, but the interview is being conducted with an eye to it being widely distributed. We could easily add many more cases, like *written* assertions, where the reception is distributed widely through time and space, and all forms that assertion might take across social media, where claims are sent out into the world, quoted, retweeted, commented upon, etc. What sense could be made of *common ground* in cases like these, that stretch far beyond everyday face-to-face conversation?

MacFarlane writes: "I am sure that a Stalnakerian can give some account of what is going on in these cases, but they do put pressure on the idea that the "essential effect" of assertion is to add information to a common ground" [135].

There is no doubt that these cases stretch the notion of the common ground, but the notion is clear enough to see how it might apply in cases, both of distributed communication, as well as solitary musing. The case of solitary "speaking to oneself" is particularly straightforward. Here, there is a conversation with very few participants, and the common ground is what is taken for granted *in that conversation*. Notice that even in this case, common ground does not collapse into the speaker's belief. I can try on an idea, and develop its consequences, speaking to myself, perhaps playing more than one role, as I internally develop a debate between two positions, without committing myself to agreeing with any conclusion along the way. Here, the internal mono(dia)logue has its own common ground, as claims are made, filtered, added and thrown out.

MacFarlane has a second criticism of common ground analyses of assertion, to the effect that assertion, on this account, cannot cut back on the common ground. We will come to this matter on page 184.

The case of taking solitary assertion and reasoning as internalised speech is considered in very different ways by Wilfrid Sellars [219] and Catarina Dutilh Novaes [64, Chapter 4].

More interesting, and more challenging, are cases of distributed assertions, such as the assertions you are reading in this book. I, the author, am making many assertions (as well as engaging in many other speech acts) and you read them, at some time and space removed from when I wrote. With any luck, there will be more readers than you, as the manuscript is read in different times and places. Where is the "common ground" in cases like these? The author is not present in each and every act of reading in the same way that the participants in a conversation are present and active, and updating the common ground together. Nonetheless, reading (or listening) contains many of the aspects of a dialogue. Active readers argue back, accept some things, and reject others. The author, in this case, has prearranged their contribution to the conversation, and cannot change it in response to its reception with this or that reader. A book is not much different, except in scope and scale from a long rehearsed monologue from someone in the middle of an otherwise back-and-forth conversation. In cases like these, the common ground is that body of information that is built up in the "dialogue", where one partner has done most of the speaking — and is not there to respond to the reactions of the other participants. The norms governing dialogue still apply, in attenuated senses in these cases. The "conversational" score is updated. Presuppositions, parentheticals, not-at-issue content, anaphora, and other issues well analysed in terms of the dynamics of conversation apply equally well in written text as they do in face-to-face conversation. Retraction works, in part, too. If you aren't convinced by some claim, or you don't want to go along with it, nothing can force you to accept it. You can argue with the author, who isn't there to argue back. Or, you can go along with what is said, "for the sake of the argument." The author, on the other hand, having prepared their contribution to the conversation beforehand, cannot adapt it in response, if you don't decide to go along with what is said. The common ground is the common ground of the varying participants in the communicative act, however those participants are to be delineated. In the case of the author and reader, it amounts to the commitments of the author, as filtered through the reading actions of the reader. In the case of an online conversation, it is determined by the producers and the readers of the texts in question in any communicative activity. Identifying the borders of such "conversations" becomes more complex the more distributed the participants, but we seem to apply the same norms of interpretation, update and scorekeeping as we do in face-to-face communication. So, the norms governing assertion apply, as do the same kinds of dynamics, even though attenuated by time and space, and even if some participants are no longer *there* to be able to respond.

» «

To conclude this section on norms governing assertion, we are left with a synthesis, according to which making an assertion that p is understood as a commitment to p, involving being prepared to vouch that p if it is called into question, and licensing others to do the same. Whenever this assertion is *received* (whether in a face-to-face conversation or in a more mediated fashion), it is received as a bid to update the common ground for that context of reception.

I am not alone in taking it to be appropriate to synthesise different analyses of assertion. In a recent paper, Paal Antonsen gives an account intended to explicitly harmonise Brandom's and Stalnaker's accounts of the normative upshot of assertion [5].

Such a synthesis of the commitment and common ground approaches to assertion allows both for the social pooling of information and dynamics of update which has proved so useful in linguistics on the one hand, and the individual normative fine texture of individual responsibility central to Brandom and MacFarlane on the other. So it is with this dual perspective on assertion that that we will now attend to the specifics of the relationship between *assertion* and *denial*.

5.3 | ASSERTION, DENIAL, AND OTHER SPEECH ACTS

In earlier work on the normative upshot of proof theory [195] — and the sequent calculus, in particular — I have argued that assertion and denial work hand in hand: the derivation of a sequent $X \succ Y$ shows us how it is that asserting each member of X and denying each member of Y is out bounds. On this perspective, assertion and denial are treated as equal partners. Instead of defining the denial that p in terms of the assertion that $\neg p$, we use denial (and assertion, and the relationship between them) to *define* negation. This approach belongs to the tradition of *bilateralist* approaches to logic, among whose proponents we find Huw Price [171], Timothy Smiley [227] and Ian Rumfitt [208, 209]. Bilateralist approaches to logic analyse proof and consequence in terms of assertion and denial, in contrast to unilateralist approaches, which attend to *one* speech act, typically assertion. For a bilateralist approach to have any chance at success, it needs to clearly specify what counts as assertion and what counts as denial, and to explain the relationship between these speech acts. The core motivating factor in bilateralist approaches to logic is the *clash* between assertion and denial. Whatever a denial of p is, it aims to rule out the assertion of p, in some sense. As we have seen in the previous section, assertion is a complex action. An assertion may be *ruled out* or *rejected* on a number of grounds or in a number of different ways. This point has been forcefully and clearly made in the literature in a number of different places [52, 115]. To begin thinking about what is involved in different kinds of denial, think of the opposites of each of the norms for assertion that we encountered ath beginning of the previous section. Assertions can be defective on account of being untrue, or unjustified, or insincere. Given the structure of the norms governing assertion, *failure* can occur in various ways. Is there any hope sketching a picture where assertion and denial are on a par with one another, or is the project of a bilateralist understanding of the sequent calculus for classical logic doomed to failure?

ASSERTION & DENIAL, BOTH STRONG & WEAK

To address the structure of assertion and denial, and to chart the significance of the different ways in which a claim might be denied, it will be helpful to consider a very particular way we might assert — or deny — by answering a question. Suppose Abelard asks a yes/no question of Eloise:

(1) ABELARD: Is Astralabe in the study?

ELOISE: Yes.

Here, Eloise's "Yes" is an assertion, to the effect that Astralabe *is* in the study. The claim that Astralabe is in the study enters the common ground. Eloise is vouching for this, and if a question is raised again about where Astralabe is (say, if Abelard goes to the study and can't see Astralabe there because he is hiding), it is appropriate for him to respond with something like "Are you sure? I can't see him here", and for her to reply "he was there five minutes ago," as the conversational negotiation continues. If, on the other hand, she had answered "No"

(2) ABELARD: Is Astralabe in the study?

Abelard and Eloise (often simply abbreviated $\forall$ and $\exists$) are characters who often appear in modern treatments of dialogues and games in logic [106, e.g., pp. 23, 24]. Peter Abelard (1079–1142) was a medieval French theologian and logician. Héloïse d'Argenteuil (?–1164) was a French theologian and abbess. They had a troubled relationship, which could variously be described as a forbidden love affair between peers or as sexual abuse of a younger student by an older teacher. Before being forcibly separated, they had a child, whom Héloïse named Astralabe. For more on the life, thought and correspondence of Peter Abelard and Héloïse d'Argenteuil, consult Constant Mews' *Abelard and Heloise* [144].

ELOISE: No.

then her answer is as an assertion to the effect that Astralabe is *not* in the study — or better, to the effect that it's not the case that Astralabe is in the study. Now, *that* claim is placed in the common ground, and Eloise is vouching for it, etc. However, there is another way we can put things: we can say that she *denies* that Astralabe is in the study. There is reason to prefer calling Eloise's response a denial rather than an assertion, for we can think of the content of the claim in question as being fixed by the question, and Eloise's response, positive or negative, is an act whose significance is given by the content of the question. One advantage of taking this perspective is that it makes sense of the distinction between being able to deny claims presented in some language, on the one hand, and express claims involving negation, on the other. It is possible to countenance agents who are able to respond positively or negatively to claims presented in a simple vocabulary, where that vocabulary does not yet feature some way to express negation. To allow for this possibility, we can think of Eloise's response here as a *denial* of the claim that Astralabe is in the study. If we allow for both denial and assertion as speech acts, a common ground, then, is not just a set of claims that have been ruled in (taken for granted to *hold*), but a set of claims ruled *in* and set of claims ruled *out*. If we think of yes/no (or *polar*) questions, of the form 'p?' as asking to *settle* issues, then the two ways to settle an issue are to rule it in, and to rule it out. Such an understanding of how it is that we increase our stock of information makes sense of the enterprise, whether or not the language in which those issues are expressed itself contains ways to express the notion of exclusion. So, we will continue with this, dual perspective on common ground as involving claims that are ruled *in* and claims that are ruled *out*, and as understanding the salient sense of *denial* that p as the act that settles the question "p?" in the negative.

What is the point of being able to *deny* as well as to assert? What is the nature of the clash between assertion and denial, or between the two ways of settling a polar question? One way to begin to address this question is given by Huw Price. He asks us to consider a dialogue conducted by two people who are *Ideological Positivists*, the fanatical disciples of Norman Vincent Peale, who have removed from their language any way to express negative thinking. In this dialogue, the two speakers want to find their friend Fred.

> Think how it might go for us as Positivists:
>
> > ME: 'Fred is in the kitchen.' (Sets off for kitchen.)
> >
> > YOU: 'Wait! Fred is in the garden.'
> >
> > ME: 'I see. But he is in the kitchen, so I'll go there.' (Sets off.)
> >
> > YOU: 'You lack understanding. The kitchen is Fred-free.'

> ME: 'Is it really? But Fred's in it, and that's the important thing.' (Leaves for kitchen.)
>
> Your problem is to get me to appreciate that your claims are incompatible with mine. Even in such a trivial case, we can see that it would be useful to have a device whose function was precisely to indicate that an incompatible claim was being made: precisely to *deny* an assertion or suggestion by somebody else. [171, p. 224]

Why is a notion of incompatibility important? How could it have helped in this dialogue? What is missing is that there is no shared position between taking Fred to be in the garden and taking Fred to be in the kitchen. (Let's presuppose that Fred is not so large as to be in both places at once.) When the 'ME' of the dialogue appears to grant that Fred is in garden and that Fred is in the kitchen (by saying 'I see' in response to the claim that Fred is in the garden, while reiterating that Fred is in the kitchen) there is no way for 'YOU' object. The language, when free of denial and free of negation, offers no block to 'ME' simply granting everything that 'YOU' say, with no pressure on 'ME' to retract the claim that Fred is in the kitchen. Denial, in this strong sense, offers us the expressive power to make claims just like this. If I answer 'yes' and you answer 'no' to the same polar question, then we disagree, and there is no shared position incorporating both of our responses.

» «

So, at the very least, we express a denial when giving a 'no' response to a polar question. What about denial in response to an *assertion*? Suppose the Abelard and Eloise's dialogue had gone like this, instead:

(3) ABELARD: Astralabe is in the study.

ELOISE: No. He's in the kitchen.

Here, it is also natural to read Eloise's "no" as the denial of the claim that Astralabe is in the study, with exactly the same effect as saying "no" would have in response to the question. Eloise elaborates on her "no" by asserting an alternative: Astralabe is in the kitchen. In this case, too, it is natural to think of Eloise's "no" as objecting to Abelard's claim, and bidding to update the common ground with the information that Astralabe is not in the study, and then elaborating this claim with further detail concerning Astralabe's whereabouts.

However all uses of "no" perform that function. Consider this different way that the dialogue could have gone instead:

(4) ABELARD: Astralabe is in the study.

ELOISE: No. He's in *either* in the study or in the kitchen.

Here, Eloise never commits herself to the claim that Astralabe is *not* in the study, or equivalently, she does not *rule out* that possibility. If she thought that he weren't — or wanted Abelard to think that — then the elaborating claim is much less informative than it could have been. She could have just said that Astralabe is in the kitchen. The more natural and immediate interpretation of Eloise's response is that she is denying Abelard's claim, not in the sense of taking it to be *untrue*, but in the sense of being *too far*. Suppose, for example, that they are searching the house for Astralabe, and have checked every room other than the study and the kitchen. Abelard concludes that he is in the study, while Eloise warns him to not conclude this so quickly. So, Eloise's denial is not intended to settle the question of whether Astralabe is in the study in the negative, but rather, to *unsettle* it. Abelard has attempted to settle it in the positive, and Eloise is rejecting that update. Here, instead of attempting to update the common ground by adding to it the information that Astralabe isn't in the study, she bids to update by *retracting* the claim that Astralabe is in the study from the common ground — or resisting its being placed there in the first place. Here, we have a denial, to be sure, but not a denial that updates the common ground with further information. It is a weak denial, a bid to retract information from the common ground.

To give more evidence for this understanding of the difference in significance between "no" as expressing strong denial (adding negative information to the common ground) and weak denial (retracting positive information from the common ground), consider the following three reformulations of the dialogue, where Eloise gives three different answers to Abelard's question:

(5) ABELARD: Is Astralabe in the study?

ELOISE: *No. He's in either in the study or in the kitchen.

ELOISE: Perhaps. He's in either in the study or in the kitchen.

ELOISE: No. He's in the kitchen.

Here, the "no" in the first response is infelicitous. The response is better formulated with the "perhaps" given in the second. On the analysis of denial given here, these responses are easily explained. The "no" is inappropriate here, because Eloise is neither updating the common ground with the negative information excluding Astralabe from the study, nor is she attempting to retract any positive information from the common ground. There is nothing appropriate *in* the common ground for Eloise to deny. No, if all the information Eloise has is that he is either in the study or the kitchen, then a better response is given by saying "perhaps". The "perhaps", in response to the question, at least attempts to ensure that the common ground doesn't rule out the claim that Astralabe is in the study (by attempting to retract its denial, if it is there), and the assertion of the disjunction is a bid to put the weaker claim, that disjunction, into the positive part of the common ground. The "no" is inappropriate as a weak denial here, because the claim was not in the common ground to be retracted. However, the "no" in response to the question is appropriate if we settle it in the negative, as in the last answer, where Eloise settles the issue, with the claim that Abelard is in the kitchen.

This example shows that not only are there two forms of denial, a weak one, which aims to retract positive information from the common ground, as well as a strong one, which aims to add negative information: There are also two notions of assertion. Strong assertion is a bid to update the common ground with positive information, while weak assertion (expressed in our dialogue by the "perhaps") is a bid to update the common ground by retracting negative information. If I say "perhaps p", this is a bid to update the common ground to remove p's denial.

For more on "perhaps" as expressing the speech act of weak assertion, and its relationship to epistemic modals like "maybe" and "might", see J. J. Schödler's 2018 thesis *Assertion and rejection* [214].

Excursus: The difference between weak denial and strong denial is not marked in English. (Eloise says 'no' both in response to a polar question and in response to a bare assertion.) Other languages differ. Roelofsen and Farkas give the example of Romanian, where the particle combination *ba nu* can be used as a negative response to a positive assertion (requiring a retraction) but it *cannot* be used as a negative response to a question (where a retraction is not required), where only *nu* is felicitous [206, pp. 364, 365]. This is some further evidence that there is a significant distinction between weak and strong denial. *End of Excursus*

» «

There are many further issues to be settled concerning weak denial and weak assertion, in comparison to their strong counterparts. Consider, for example, cases often described as 'metalinguistic negation'.

(6) ELOISE: I am not Abelard's wife. He is my husband.

According to Graham Priest [174, p. 77], in a case like this, Eloise does not here deny or negate the claim that she is Abelard's wife. Given our account of the significance of assertion and denial, there is no need to take Eloise's apparent denial as anything other than the same kind of weak denial we have already seen. This case causes no further difficulty. She is bidding to remove the claim that she is Abelard's wife from the common ground, to replacing it with the claim that Abelard is her husband. It is true, of course, that — under the usual understandings of the terms 'husband' and 'wife' — the claim that she is Abelard's wife *follows* from what is in the common ground after her update has taken effect.

If the common ground is understood as a set of possible worlds (as is Stalnaker's approach, and the approach of all who follow him), this retraction and addition presumably leaves us with exactly the same set of worlds with which we started. If, on the other hand, a common ground features not just a set of worlds but a set of ways to specify those worlds — for example, the assertions and denials made to shape the common ground — then the common ground bears the marks of Eloise's declaration. She (implicitly) grants that she is Abelard's wife (given the usual understanding of the terms, which she need not contest), but she does not want the fact put *that* way. There is much more to say about these and other issues concerning forms of denial [35, 84, 109, 110, 202] but we have learned enough about assertion and denial for us to continue down our path to defining positions.

I say "presumably" because retraction is a difficult thing to specify. There are many ways to expand the space of possibilities so as to remove a proposition from the common ground. There is, in general, no canonical choice, and so, it may be that when we add worlds and then subtract, we don't necessarily get back to where we started. However, it is a reasonable sanity constraint on retraction that removing p and then adding p (or something analytically equivalent to p) to the common ground returns us to the same set of worlds with which we started.

OTHER SPEECH ACTS

Positions will be used to 'keep score' in dialogues and other contexts where claims are being made. Assertion and denial — both weak and strong — will play an important role in such contexts, but they are not the only speech acts of importance [14]. We have already used polar questions — an in particular, *answers* to polar questions — as a means to help distinguish strong denial from weak denial. But not all questions are polar. For example, suppose Abelard has been away from Astralabe for some time, and has not kept up with how his language learning has been going. If Abelard asks whether Astralabe has learned Latin or Greek, this question can be interpreted in two ways. One as a *closed* question, which *presupposes* that one of the two options is true, and is asking for which. In English, such a question is often marked by rising intonation on the first option and falling intonation on the second.

(7) ABELARD: Has Astralabe learned ↑Latin or ↓Greek?

ELOISE: He's learned Latin.

ELOISE: He's learned Greek.

ELOISE: *Yes.

Here, the "yes" answer is infelicitous, since *asking* the question put the disjunction into the common ground (if it wasn't there already) and so, answering with a bare "yes" is uninformative.

The same words could, on the other hand, be asked as an *open* question, which makes no such presupposition. In English, this is marked by keeping the rising intonation is kept on the second option.

(8) ABELARD: Has Astralabe learned ↑Latin or ↑Greek?

ELOISE: Yes, he's learned Latin.

ELOISE: Yes, he's learned Greek.

ELOISE: ?Yes.

ELOISE: Yes, but I'm not sure which.

ELOISE: No.

Here, a "yes" answer is acceptable, especially if followed up by an explanation that neither option is known to be true. The semantics and pragmatics of open and closed questions, and other non-polar questions, together with their interaction with polarity items like "yes" and "no" is a complex and rich field, with many interesting results [206].

» «

We have already seen, on page 178, that in a dialogue it is possible for a partner to *concede* a point, for the sake of the argument. When we concede an assertion (or a denial), it is entered into the common ground, but at this point (if not before) the common ground of the discussion may diverge from the commitments of at least one of the participants of the discussion. In dialogue (or when one reads, or listens to recorded speech, etc.) it is possible to let things pass into the common ground, without objection, in order to see where the discussion leads. If this occurs, the common ground is no longer an expression of the shared beliefs of those engaged in the discussion, but it remains an elaboration of a position nonetheless. There is no need, of course, for the discussion to elaborate the *beliefs* of any of the participants. Perhaps one dialogue partner will concede one point, another concedes a different point, etc., so the focus of the discussion can be a position that is ultimately believed by none of the participants, but is nonetheless explored for the sake of the argument. What can be done piecemeal in this way (each participant conceding something), can be done wholesale in one go. Everyone in a dialogue can concede something that no-one believes if the claim is merely *supposed*. If, in a dialogue I ask you to suppose that p, I ask for p to be entered into the common ground, to allow for the norms of dialogue to operate as before, without the thought that the results of such claims express our beliefs. We may assert (or deny) p merely for the sake of the argument, in order to see where that argument leads.

One reason for admitting supposing into our repertoire of dialogue is well understood. Supposing underlies the practice of inference rules that discharge, such as $\to E$ (in natural deduction) or $\to R$ (in the sequent calculus). In proof systems with $\to R$, for example, to prove a conditional $A \to B$ from the position $[X : Y]$, we prove the conclusion B from the position $[X, A : Y]$. In other words, we grant A for the sake of the argument — using it as we would use any other assertion, but without any thought to check whether it's true, or needing to justify it in any way — in order to conclude $A \to B$, from a position where A is no longer so granted. In proof systems like these, the norms governing this kind of 'supposing' are straightforward. To suppose p, simply add it to the positive part of the common ground for the purposes of the argument — with an eye to removing it later. While it is there, is is treated as any other assertion, with the exception that if it is called into question, instead of attempting to justify it in the usual manner, we simply point out that we granted it for the sake of the argument. When we *suppose* in this sense, you do not need to take anything else out of the common ground in order to make room for p. The paradoxes of material implication follow immediately. If p is in the common ground, and I suppose q — simply by adding it to the common ground, retaining p — then discharging q (using that $\to R$ step from before) we have $q \to p$. In other words, we have shown that p brings along with it $q \to p$, because the supposition of q did nothing to 'push out' the p that we had already granted.

Perhaps not all forms of supposition work in that straightforward way where assumptions are simply temporarily added to our common ground. If I grant that p is the case, but then ask you to suppose what would have happened if p *hadn't* been the case instead, it looks like I am asking you to consider things going differently. I am not asking you to consider that p both holds and fails to hold. I am asking you instead to countenance an *alternative*. Here, the assumption that p fails (perhaps adding p as a *denial*) is not intended to be simply added to the position we have already staked out, but rather, asks you to consider an alternative.

Exactly how we might make sense of this will be the topic for Chapter 9.

» «

Finally, for this section, I must mention one other speech act: the act of expressing *inference*. If I say something of the form

p, so q

then I have not not only asserted p, and asserted q, but I have, in some sense, *inferred* q from p. This does not mean that believe q on the basis of believing p, or that I have concluded q only because I was committed to p first. In presenting p and q in this way, I am offering p as some kind of reason for q. Inference, in this sense, is another kind of speech act, along with assertion and denial, supposition and many more. It is related to entailment and logical consequence, but not to be identified with them, for I may recognise that A entails B and not infer B from A. Perhaps I have granted A and I discover that A entails B and as a result of this, I decide to retract my commitment to A, instead of granting B. Entailment is a relationship between propositions (or sentences that express them) irrespective of whether they have been asserted or not. Inference, in the sense discussed here, is a speech act, connecting one assertion (or perhaps a denial) to other assertions (or perhaps a position, consisting of assertions and denials) [100]. Inference, in this sense, is involved in the practice of justification, or answering 'why' questions. If I am asked 'why B?' I could respond 'B, since A', giving A as a reason for B. We have already seen the importance of this practice when considering the connection between assertion and commitment on page 176ff, where we sketched an account of assertion which involves the asserter in being open to being called on to vouch for what is asserted. One way to attempt to discharge that debt is to offer a reason: another assertion (or perhaps more: a package of assertions and denials) from which the claim can be *inferred*. Of course, this is not the only way we could attempt to discharge that debt if called upon. Instead of inferring the claim from another claim, we could do something else, such as *showing* our inquirer that the claim is true (by putting them in a situation where they can see it for themselves), or pass on the request to someone else (in the case of testimony), and so on. Justification requests may be discharged in a number of ways — many of those ways involve inference.

As you can see here, inference plays a *role* in our account of the norms governing assertion, denial, and logical consequence, but it is not the central role. As a result, the approach being marked out here is, in MacFarlane's terms, a *normative pragmatism*, without being an *inferentialist* in any deep sense [134].

We will return to matters concerning inference later in this chapter when we consider the connection between positions, bounds and inference in §5.5 on page 202, and on page **??** in Chapter 7, we will address an important puzzle concerning the normative force of inference, the paradox of Achilles and the Tortoise, due to Lewis Carroll.

5.4 | POSITIONS AND STRUCTURAL RULES

Now we have all the materials we need to not only introduce positions, but explain how they determine a particular choice of the antecedently given context of deducibility, and hence, provide the setting where Belnap's constraints concerning conservative extension and unique definability play out.

DEFINITION 5.1 [POSITIONS] Given a language $\mathcal{L}$, a POSITION $[X : Y]$ IN $\mathcal{L}$ is a pair of sets of sentences from $\mathcal{L}$.

This simple definition is meant quite literally. Positions are individuated by the sets of assertions and denials made in them. Positions take no account of the order in which assertions or denials are made, nor the number of times they are made. If a sentence A is asserted twice, it is recorded in the position merely once. This presupposes that the language does not allow that the sentence A could say two different things. If I say "that is a cat" gesticulating at one animal, and you say "that is a cat", pointing at another, these two utterances be different items in the language.

This could be regimented with different demonstratives: "$\textit{that}_1$ is a cat" and "$\textit{that}_2$ is a cat" are distinct sentences, for the purpose of an appropriately regimented language, as would be "it's 6pm now" and "it's 6pm now" said at different times. Considerations like these are one aspect of the significance of the discussion of samesaying on page 172.

A position is a kind of *score*. It records the state of play in some practice. Rules concerning the practice refer back to the score. In a game of cricket, the *score* could be narrowly construed as the number of runs scored by each team and the number of wickets taken, or more broadly, as all of the information recorded on the most detailed scoreboard — the runs scored, and balls faced by each batter, the number of balls bowled by each bowler, the wickets taken and their mode of dismissal, etc. A position $[X : Y]$ is a rather thin scoreboard, which records only what has been explicitly ruled *in* (the elements of X) and what has been explicitly ruled *out* (the elements of Y). A position can be used to represent the common ground in a discussion, in which some things have been ruled in and others have been ruled out. There are many more features of a discourse or a text that are semantically salient, of course, and richer notions of score will be available to play a role in giving account of those phenomena. For our concerns this notion of score will be rich enough to address very many issues concerning proof, meaning and defining rules.

Richer structures for positions will be introduced in Chapter 9, in order to give an account of different kinds of modal operators.

BOUNDS

Not all positions have equal status. Some positions are, in an important sense, out of bounds. If assertion and denial are connected in the manner that was sketched in the previous section, then assertion and denial are tied together in an intimate way. If I (strongly) assert A and you (strongly) deny A, then I am putting in a bid for A to be in the common ground *positively* (and for it to be removed from the negative part of the common ground if it were already there), and you are putting in a bid for A to be in the common ground *negatively* (and for it to be removed from the positive part of the common ground if it were already there). We are attempting to modify the common ground in incompatible ways. While it is possible for a position to have the form $[X, A : A, Y]$ where A is both ruled in and ruled out — after all, that is a pair of sets of sentences, as per the definition — no common ground takes this position. In this way, such a position is *out of bounds*. The answers "yes" and "no" to the polar question "p?" are opposed to each other. If you and I take up those answers, we to take up distinct and incompatible positions on the issue of p.

This leaves an important question: What about those times when you are asked such a question, and you want to answer "yes and no"? Does this not show that assertion and denial are indeed *not* always opposed attitudes?

I address the semantic paradoxes as a motivation for asserting contradictions in Chapter 7, starting at page 264.

Here, the answer is a clear *no*. Putting aside the issue of semantic paradoxes as a motivation for dialetheism, straightforward cases where are are motivated to give "yes and no" answers are circumstances which call out for disambiguation. If Eloise asks Abelard "Is Astralabe happy?" and Abelard answers "yes and no", this asks for further elaboration, of the form "he is happy about x but unhappy about y" or "he is generally happy, but he's unhappy at the moment" or some other way of more finely individuating the question, in such a way that it is possible to answer one question affirmitively and the other negatively. Here, rather than thinking of these cases as pressing against the requirement that assertion and denial clash with one another, this norm plays a central part in the process of refining and articulating our concepts to push us toward greater expressive power. Our constraint, to nail the colours to the mast *pro* or *con* (and to not be happy to acquiesce in both) on each issue drives us forward to clarify and develop more discriminating concepts, to discover that we were not, after all, asking one question, but more (in distinguishing mass and weight, for example) or to discover that what we thought was an absolute matter (for example, simultaneity) is in fact, relative to a frame of reference.

» «

This constraint, recording the clash between assertion and denial, introduces the notion of *bounds* on the positions in our language $\mathcal{L}$. Positions of the form $[X, A : A, Y]$, in which A is both asserted and denied involve a clash, and are out of bounds. These bounds provide a field of play, inside which any position is in some sense, possible, and outside of which play cannot, in some sense, proceed. This constraint between assertion and denial is not the only possible bound on positions. One way to increase our expressive power is to include in our language vocabulary items that impose further bounds on positions. For example, if a language contains conjunction, then asserting the conjunction $A \wedge B$ may be no different from asserting A and asserting B, while *denying* a conjunction allows us to do something new: to rule out A and B obtaining together, without ruling out either on their own. To do this, we impose other bounds beyond the initial constraint that asserting A and denying A is out of bounds. For example, we may have the constraint that the position

$$[A, B : A \wedge B]$$

is also out of bounds — that there is a clash involved in asserting A, asserting B and denying $A \wedge B$. This will then have the effect that denying $A \wedge B$ will rule out A and B holding together, without thereby ruling out either of A or B by themselves. We are able to do something that we could not do in a language without conjunction. Exactly *how* these bounds might be imposed is a matter we will address in the next chapter. For now, it is enough to say that positions in a language come equipped with relation determining their bounds, and this need not be just be the simple matter of X and Y in $[X : Y]$ having no intersection. Whether a position $[X : Y]$ is out of bounds, or not, depends on the meanings, in some senses, of the expressions in X and Y.

DEFINITION 5.2 [BOUNDS FOR POSITIONS] Given a language $\mathcal{L}$, a set $\mathfrak{B}$ of positions in $\mathcal{L}$ is a set of BOUNDS for $\mathcal{L}$ if and only if it satisfies the following two conditions.

- For each $A \in \mathcal{L}$, $[A : A] \in \mathfrak{B}$.
- For each $A \in \mathcal{L}$, $[X : Y] \in \mathfrak{B}$ iff $[X : A, Y] \in \mathfrak{B}$ and $[X, A : Y] \in \mathfrak{B}$.

Given a set $\mathfrak{B}$ of bounds, we will also say, for positions $[X : Y]$ that are *not* out of bounds that they are *available*, or more colourfully, they are *on the field of play*.

These two conditions are simple, but they are rich with consequences. The first represents the constraint that if a position consists only of the assertion of A and the denial of A, it is out of bounds.

The second condition relates bounds concerning an arbitrary position $[X : Y]$ and its extension with the formula A. Reading the biconditional from left to right, we see that if a position is out of bounds, then *adding* an assertion — or adding a denial — is not the way to return to the field of play. The kind of defect suffered by a position that is out of bounds is preserved by any extension of that position. In other words, this is the weakening rule we saw in Chapter 2, phrased in the vocabulary of bounds.

The right to left half of the second condition is another structural rule we have already seen. It says that if the positions $[X : A, Y]$ and $[X, A : Y]$ are both out of bounds, then $[X : Y]$ is out of bounds too. This is a reformulation of the *Cut* rule. Contraposing this, we have that if $[X : Y]$ is in the field of play, then so is one of $[X, A : Y]$ and $[X : A, Y]$. In other words, wherever we are in the field of play if the question "A?" arises, at least one of the answers *yes* and *no* is also available. No question places us in a *severe quandry* where no option is available. One way to further illustrate this is to note the connection between undeniability and assertion. If $[X : Y]$ is available, but $[X : A, Y]$ is not, then there is a strong sense in which A is, relative to the position $[X : Y]$, *undeniable* — adding its denial would take us out of bounds. So, if the issue question concerning A arises, the only answer remaining is 'yes'. Denying it is not an option. Or to put the same point in another way, if, given the available position $[X : Y]$, the addition of A to the assertion side is out of bounds, then when the question "A?" arises, the answer must be *no*, because it is already implicitly ruled out by the position already taken. To explicitly deny it is merely to make explicit what is already implicit in the position $[X : Y]$.

You could be in a minor quandry, where both options concerning A seem equally problematic. However, according to *Cut* they cannot both be *unavailable*.

"SOCRATES: And no one can deny that all percipient beings desire and hunt after good, and are eager to catch and have the good about them, and care not for the attainment of anything which its not accompanied by good. PROTAGORAS. That is undeniable." [165]

This reasoning, in defence of the *Cut* rule under this interpretation of bounds and sequents, has been called into question (chiefly by Dave Ripley [204]). There is more to be said in defence of *Cut*, and I will address these criticisms below. However, before we do this, we will finish the presentation of bounds.

» «

The only structural features of bounds in Definition 5.2 are *reflexivity* (the first condition), *weakening* (the left to right part of the second condition) and *Cut* (the right to left part). In Chapter 2, we considered other structural rules, in particular, *contraction*. Contraction is not imposed as a separate condition on positions, because it is implicit in the choice of a position as a pair of *sets* of formulas. Here, there is no question of a difference between the status of $[X, A, A : Y]$ and $[X, A : Y]$, since these are literally the same position. On this account of the common ground, as given by the items ruled in and ruled out, there is no record of the number of times an item has been ruled in or out, and so, contraction, in this form, is a given, for this notion of positions and their bounds.

Another significant condition on bounds, implicit in their connection with the sequent calculus or any other system of proof where proofs are essentially finite, is the condition of *compactness*. Notice that positions allow for finite sets of formulas and also for infinite sets of formulas. They do not discriminate. In a proof system in which proofs are finite, the only constraints imposed on positions is by way of their constraints on what we might call *finitary* positions.

We say 'finitary' rather than 'finite', since *every* position is just a *pair* of sets of formulas, so in one sense, every position is finite: it contains *two* components.

DEFINITION 5.3 [FINITARY POSITIONS] $[X : Y]$ is a *finitary* position if X and Y are both finite sets of formulas. A position $[X : Y]$ that fails to be finitary is said to be *infinitary*.

If a derivation of the sequent $X \succ Y$ shows that $[X : Y]$ is out of bounds, then derivations initially constrain finitary positions (since any derivation is of a sequent, which is finitary), and only by proxy constrain infinitary positions. The resulting bounds satisfy the following further constraint: *compactness*.

DEFINITION 5.4 [COMPACTNESS] The bounds $\mathfrak{B}$ are COMPACT if and only for every position $[X : Y]$, $[X : Y] \in \mathfrak{B}$ if and only if there are finite $X' \subseteq X$ and $Y' \subseteq Y$ where $[X' : Y'] \in \mathfrak{B}$.

In other words, if bounds are compact, then whenever a position is out of bounds, that fact is witnessed by a finitary part of that position. There is no way for an position to be out of bounds essentially involving infinitely many sentences. Compact bounds are in some sense, well behaved or tractable. If the bounds are compact, this means that if any position at all is out of bounds, then there can be some finite certification of this fact (recorded by the finitary part of the position that is also out of bounds, and, for example, its derivation). If the bounds are imposed by way of some system of proofs, or some family of rules which can be used to coordinate among speakers in some dialogue, then compact bounds allow for any violation of the rules to be somehow independently certified. If I recognise that a given position is out of bounds, then I have some way to communicate that fact to you in some way that you can independently verify: by giving you, say, the derivation for the finitary position that is at fault. I can (at least in principle, if not in practice) literally spell out the entire derivation and show my working in a way that you can independently check. If, however, the bounds fail to be compact, then this need not always be possible. Non-compact bounds allow for the possibility position breach the bounds of play without that fact being communicable in any finite way.

Nonetheless, in what follows, we will not assume that bounds *must* be compact. After all, it seems conceivable to take the following position

$$[Nx : x = 0, x = 1, x = 2, \ldots]$$

to be out of bounds, if we fix on the interpretation of "Nx" as x *is a natural number*, "$=$" as the identity predicate, and the numerals 0, 1, etc., with their usual interpretations. Given that constraint (which is to admit the ω-rule as valid), this position is out of bounds, but no finitary part of that position is out of bounds. The bounds, specified in this way, fail to be compact, with all the costs that this can impose on our ability to check those bounds. In Chapter 8, when we consider the connections between generality, quantifiers, and bounds, we will consider the costs and benefits of remaining within the constraint of compact bounds, or allowing for bounds to be non-compact. Until we get there, it will be enough to note that our definition of bounds allows for failures of compactness, but regardless, most of the systems of bounds that we will study throughout the remainder of book will be compact.

AN EXAMPLE: COMPARATIVES

Let's consider a concrete example, to show how the bounds may be applied in a specific vocabulary, and how much work can be done with the simple structure of bounds, without any recourse to the connectives or other complex logical vocabulary.

It is well known that natural languages feature categories for singular terms and predicates. We can not only refer to *Eloise* and *Abelard*, but we can *describe* them, saying that *Eloise is tall* or *Abelard is happy*. Predicates can be combined with more than one singular term (*Eloise is tall, Abelard is tall*), just as terms can be combined with more than one predicate (*Eloise is tall, Eloise is happy*). Many predicates not only feature in these individual speech acts of predication, but can also be *compared*. Tallness (or height) and happiness are not simply an all or nothing thing. They are also a matter of *degree*. It makes sense to ask not only whether Eloise is tall, but also whether *Eloise is taller than Abelard*, and similarly, whether *Abelard is happier than Eloise*. There is a surprising amount of structure to be found in the interaction between predication and comparatives, and this structure can be clarified by paying attention to the bounds for such judgements.

» «

To make this precise, let's consider a language with some stock $a, b, c, \ldots$ of SINGULAR TERMS, and a single unary (one-place) PREDICATE F, which we will take to be *gradable* (for which comparisons are salient). So, we add to our language two binary (two-place) PREDICATES $>_F$, and $\geqslant_F$.

Whenever s and t are any two singular terms from our vocabulary, the formulas in our little formal language are one of the following three shapes:

$$Fs \qquad s >_F t \qquad s \geqslant_F t$$

where Fs is understood as saying that s *is* F; $s >_F t$ says that s *is more* F *than* t *is*, and $s \geqslant_F t$ says that s *is at least as* F *as* t *is*. We will call $>_F$ the *strong* (F-)*comparative*, and $\geqslant_F$ the *weak* (F-)*comparative*.

The first thing to notice about comparatives $>_F$ and $\geqslant_F$ is that they *compare*. It is very plausible that it is part of our competence with such comparatives that we take the following positions to be out of bounds:

$$s >_F t,\ t >_F u \ \succ\ s >_F u$$

$$s \geqslant_F t,\ t \geqslant_F u \ \succ\ s \geqslant_F u$$

whatever singular terms s, t, u we choose. That is, to grant that s is more F than t is, and that t is more F than u is, but to deny that s is more F than u is to misunderstand the notion of comparison in play. Similarly for the weak comparative in place of the strong.

Excursus: This is not to rule out failure of transitivity in all kinds of comparison. If, for example, we are ranking three things (a, b and c) on three criteria (F, G, H) where on the scale for F we have $a >_F b >_F c$, for G, we have $b >_G c >_G a$ and for H we have $c >_H a >_H b$, where each of these orderings is understood transitively, then in some sense, $a >_{most} b$ since a is ranked over b according to F and H (but not G), and $b >_{most} c$ since b is superior to c according to F and G (but not H) and similarly, $c >_{most} a$, (by G and H, but not F). Here, the aggregating $>_{most}$ is *not* a comparative in the sense discussed here, while $>_F$, $>_G$ and $>_H$ are. *End of Excursus*

The strong and the weak comparatives vary with regard to how they police *reflexive* comparisons.

$$t >_F t \ \succ \qquad\qquad \succ\ t \geqslant_F t$$

Here, it is out of bounds to assert that t is more F than itself—to do that would be to misunderstand the strong comparative—and similarly, out of bounds to deny that t is at least as F as itself—that would be to misunderstand the weak comparative.

Furthermore, the comparatives are connected, at least in the following way:

$$s >_F t \ \succ\ s \geqslant_F t$$

To deny that s is *at least as* F as t while maintaining that s is *more* F than t is to misunderstand the relationship between comparatives. Can we say more than this about the relationship between them? After all, everything we have said about $>_F$ would also be satisfied by the comparative $\gg_F$, of being *much more* F *than*. Is there anything that ties $>_F$ and $\geqslant_F$ more intimately than this? Plausible candidates are the following pair of conditions on the bounds:

$$s >_F t,\ t \geqslant_F s \succ \qquad\qquad \succ s >_F t,\ t \geqslant_F s$$

This has the effect of taking $s >_F t$ and $t \geqslant_F s$ to be *contradictories*. It is out of bounds to *assert* both (they are at least *contraries*), and it is out bounds to deny both (they are also *subcontraries*). For the former, it seems clear that to take s to be strictly *more* F than t clashes with the judgement that t is at least as F as s is. That is straightforward enough—but again, note that this would be satisfied were we to replace $s >_F t$ by $s \gg_F t$, so the first condition does nothing to pin $>_F$ down to some notion of being *merely more* F *than*. For that, we need something like the subcontrary condition. Consider what it would be to deny that t is at least as F as s is, and *also* to deny that s is more F than t. If the degrees of F-ness are totally ordered, this ruled out. To deny that t is at least as F as s is to rule out for b all of the places on the F-scale from s's position and beyond. The only places remaining are for t to be somewhere *below* s on that scale, so denying those positions for t is to find nowhere to place t at all.

These are bounds on the comparatives $>_F$ and $\geqslant_F$, but we have said nothing yet about the relationship between the comparatives and *predication*. This is straightforward. An obvious condition is expressed by this sequent:

$$Fs,\ t \geqslant_F s \succ Ft$$

which says that it is out of bounds to assert Fs and to deny Ft, while also asserting that t is at least as F as s.

Yes, this reasoning assumes that the only positions we can take for comparative F judgements are *pro* and *con*. If you have a richer language which might take there to be singular terms for which it is a *category mistake* to judge with respect to F—for example, if the object s doesn't exist (if the term 's' does not denote) then perhaps one can deny both $s >_F t$ and $t \geqslant_F s$ at no cost—then another technique will have to be used. For now, we stick with our simple little language to illustrate the bounds. See Chapter 8 for more detail on how to model non-denoting terms in this framework.

Let's collect these conditions into a single definition:

DEFINITION 5.5 [COMPARATIVES] The language $\mathcal{L}$ has STRONG AND WEAK COMPARATIVES $>_F$ and $\geqslant_F$ for the predicate F if and only if its bounds include the following, for each singular term s, t and u.

$$\begin{aligned}
\textit{strong transitivity}&:\ s >_F t,\ t >_F u \succ s >_F u\\
\textit{weak transitivity}&:\ s \geqslant_F t,\ t \geqslant_F u \succ s \geqslant_F u\\
\textit{strong irreflexivity}&:\ s >_F s \succ\\
\textit{weak reflexivity}&:\ \succ s \geqslant_F s\\
\textit{contraries}&:\ s >_F t,\ t \geqslant_F s \succ\\
\textit{subcontraries}&:\ \succ s >_F t,\ t \geqslant_F s\\
\textit{strength}&:\ s >_F t \succ s \geqslant_F t\\
\textit{preservation}&:\ Fs,\ t \geqslant_F s \succ Ft
\end{aligned}$$

There are other conditions on the bounds that are as plausible as these, but many of these follow from the conditions we have mentioned here, by way of *Cut*. In fact, there is already a little redundancy among the list of conditions. We need only one reflexivity condition (*strong irreflexivity* or *weak reflexivity*). *Subcontraries* and *strong irreflexivity* give us *weak reflexivity*, and *contraries* and *weak reflexivity* give us strong *irreflexivity*.

$$\dfrac{\dfrac{\textit{subcontraries}}{\succ s >_F s,\ s \geqslant_F s} \qquad \dfrac{\textit{strong irrefl.}}{s >_F s \succ}}{\succ s \geqslant_F s}\,\textit{Cut} \qquad\qquad \dfrac{\dfrac{\textit{weak refl.}}{\succ s \geqslant_F s} \qquad \dfrac{\textit{contraries}}{s >_F s,\ s \geqslant_F s \succ}}{s >_F s \succ}\,\textit{Cut}$$

More is present in the bounds than what we have explicitly listed. For example, the *antisymmetry* of the strong comparative ($s >_F t,\ t >_F s \succ$) follows from strong transitivity and strong irreflexivity:

$$\dfrac{\dfrac{\textit{strong transitivity}}{s >_F t,\ t >_F s \succ s >_F s} \qquad \dfrac{\textit{strong irrefl.}}{s >_F s \succ}}{s >_F t,\ t >_F s \succ}\,\textit{Cut}$$

Similarly, we can show that the strong comparative interacts with predication just as much as the weak comparative does. We have *strong preservation*: $Fs, t >_F s \succ Ft$.

$$\dfrac{\dfrac{\textit{strength}}{t >_F s \succ t \geqslant_F s} \qquad \dfrac{\textit{preservation}}{Fs,\ t \geqslant_F s \succ Ft}}{Fs,\ t >_F s \succ Ft}\,\textit{Cut}$$

We also have this condition, which is like preservation, except now the strong comparative is an alternate conclusion: $Fs \succ s >_F t,\ Ft$.

$$\dfrac{\dfrac{\textit{subcontraries}}{\succ s >_F t,\ t \geqslant_F s} \qquad \dfrac{\textit{preservation}}{Fs,\ t \geqslant_F s \succ Ft}}{Fs \succ s >_F t,\ Ft}\,\textit{Cut}$$

In other words, it would be a mistake to assert Fs and to deny both Ft and the claim that s is more F than t is. To put this more positively, if we find ourself agreeing that Fs, and we are curious about Ft, we are at least assured that we are in one of the two cases: (1) Ft, (2) $s >_F t$.

The contrary and subcontrary connections between $s >_F t$ and $t \geqslant_F s$ also ensure that the transitivity properties for $>_F$ and $\geqslant_F$ interact appropriately. For example, we have $s \geqslant_F t,\ t >_F u \succ s >_F u$.

$$\dfrac{\dfrac{\textit{subcontraries}}{\succ u \geqslant_F s,\ s >_F u} \qquad \dfrac{\dfrac{\textit{weak transitivity}}{u \geqslant_F s,\ s \geqslant_F t \succ u \geqslant_F t} \qquad \dfrac{\textit{contraries}}{u \geqslant_F t,\ t >_F u \succ}}{u \geqslant_F s,\ s \geqslant_F t,\ t >_F u \succ}\,\textit{Cut}}{s \geqslant_F t,\ t >_F u \succ s >_F u}\,\textit{Cut}$$

(A similar demonstration justifies $s >_F t,\ t \geqslant_F u \succ s >_F u$.)

» «

Before we leave the example of bounds for comparatives, it will be worthwhile to consider the behaviour of limit positions in this little language.

Recall Definition 3.12 (page 126), which introduced the notion of a limit position, a partition of the language $[X : Y]$ such that there is never a derivation for $X \rhd Y$ — that is, a derivation of a sequent $X' \succ Y'$ where $X' \subseteq X$ and $Y' \subseteq Y$. We showed (in Lemma 3.13, on page 126) that if the language contains propositional connectives satisfying the standard rules, and if the bounds satisfy *Contraction* and *Weakening*, then limit positions behave like two-valued boolean evaluations. If the bounds satisfy the conditions for comparatives, we have the following:

THEOREM 5.6 [LIMIT POSITIONS FOR COMPARATIVES ARE LINEAR ORDERS] *Each* $[X : Y]$ *is a limit position for a language* $\mathcal{L}$ *with comparatives induces a linear order on the singular terms of the language* $\mathcal{L}$ *in the following sense: Define the relation* $\succsim$ *on the set* T *of singular terms from* $\mathcal{L}$ *by setting* $s \succsim t$ *iff the sentence* $s \geqslant_F t$ *is in* X*. Then,* $\succsim$ *is*

- TRANSITIVE: *if* $s \succsim t$ *and* $t \succsim u$ *then* $s \succsim u$,
- REFLEXIVE: $s \succsim s$ *for each* s, *and*
- TOTAL: *for each* s *and* t, *either* $s \succsim t$ *or* $t \succsim s$.

Furthermore, the following two conditions are satisfied:

- *The formula* $s >_F t$ *is in* X *if and only if* $t \not\succsim s$.
- *The set* $\mathcal{F}$ *consisting of all terms* s *where* Fs *is in* X, *is an* $\succsim$*-upward closed subset of* T: *If* $s \in \mathcal{F}$ *and* $t \succsim s$, *then* $t \in \mathcal{F}$, *too.*

Proof: Let's first show that $\succsim$ as defined is a total order on T.

For *transitivity*, if $s \succsim t$ and $t \succsim u$ (so $s \geqslant_F t$, and $t \geqslant_F u$ are in X), then since the bounds contain $s \geqslant_F t,\ t \geqslant_F u \succ s \geqslant_F u$, we cannot have the formula $s \geqslant_F u$ in Y, since $[X : Y]$ is available. Since $[X : Y]$ is a partition of the language, we must have $s \geqslant_F u$ in X, and hence $s \succsim u$ as desired.

For *reflexivity*, since $\succ s \geqslant_F s$ is in the bounds, we cannot have $s \geqslant_F s$ in Y, so it must be in X (since $[X : Y]$ is a limit position), and hence $s \succsim s$.

For *totality*, the *subcontrary* condition tells us that the bounds contain $\succ s >_F t,\ t \geqslant_F s$, and the strength condition gives $s >_F t \succ s \geqslant_F t$, so *Cut* ensures that the bounds contain $\succ s \geqslant_F t,\ t \geqslant_F s$, which means that we cannot have both $s \geqslant_F t$ and $t \geqslant_F s$ in Y, which, since $[X : Y]$ is a partition, ensures that one of $s \geqslant_F t$ and $t \geqslant_F s$ is in X, and hence either $s \succsim t$ or $t \succsim s$.

For the remaining two conditions, connecting $>_F$ and F to the underlying partial order, we reason as follows: First, the *contrary* condition on the bounds ($s >_F t,\ t \geqslant_F s \succ$) tells us that we do not have $s >_F t$ and $t \geqslant_F s$ both in X, which means that if $s >_F t$ is in X, then $t \not\succsim s$. The *subcontrary* condition on the bounds ($\succ s >_F t,\ t \geqslant_F s$) ensures that one of $s >_F t$ and $t \geqslant_F s$ is in X, and so, if $t \not\succsim s$, then $s >_F t$ is in X. Second, if $s \in \mathcal{F}$ (that is, Fs is in X), and $t \succsim s$ (so $t \geqslant_F s$ is also in X), since the bounds contain $Fs,\ t \geqslant_F s \succ Ft$, we cannot have Ft in Y, and hence, Ft is also in X, and hence $t \in \mathcal{F}$. ∎

So, the bounds for comparatives give rise to linear orders, when taken to the limit. We can be sure that these bounds make no *more* claim on $\succsim$ than this by way of another result, which shows the converse: that linear orders can give rise to bounds. To state this result, it will help to have the following definition:

DEFINITION 5.7 [LANGUAGE FOR A POSET] If $\langle \mathsf{T}, \succsim, \mathcal{F} \rangle$ is a set T equipped with a binary relation $\succsim$ (not necessarily a partial order) and a subset $\mathcal{F}$ of T, then a LANGUAGE FIT FOR $\langle \mathsf{T}, \succsim, \mathcal{F} \rangle$ has a distinct singular term $\underline{t}$ for each object $t \in \mathsf{T}$ (and no other singular terms) a predicate F and the formulas have the form $\mathsf{F}\underline{t}$, $\underline{s} \geqslant_{\mathsf{F}} \underline{t}$, and $\underline{s} >_{\mathsf{F}} \underline{t}$ for each $s, t \in \mathsf{T}$.

In such a language, a pair $[X : Y]$ of sets of formulas is a TOTAL DESCRIPTION of $\langle \mathsf{T}, \succsim, \mathcal{F} \rangle$ is a partition of the language, such that (1) $\mathsf{F}\underline{t} \in X$ iff $t \in \mathcal{F}$, (2) $\underline{s} \geqslant_{\mathsf{F}} \underline{t} \in X$ iff $s \succsim t$, and (3) $\underline{s} >_{\mathsf{F}} \underline{t}$ iff $t \not\succsim s$.

THEOREM 5.8 [LINEAR ORDERS GIVE COMPARATIVE BOUNDS] *If* $\langle \mathsf{T}, \succsim, \mathcal{F} \rangle$ *is a set* T *equipped with a linear order* $\succsim$*, and an upwardly closed subset* $\mathcal{F}$ *of* T*, then a total description* $[X : Y]$ *of* $\langle \mathsf{T}, \succsim, \mathcal{F} \rangle$ *is a limit position in its language, that respects the bounds for comparatives.*

Proof: Clearly $[X : Y]$ is a partition of the language. To consider *which* bounds it respects, let's look at the coarsest bounds it respects, its *extensional bounds*, specified by setting $U \succ_{[X:Y]} V$ iff either $U \cap Y \neq \emptyset$ or $V \cap X \neq \emptyset$. That is, $U \succ_{[X:Y]} V$ holds iff either some member of U is "false" (according to $[X : Y]$) or some member of V is "true." Any such extensional bounds clearly satisfy *Identity*, *Contraction*, *Weakening* and *Cut*, immediately.

To show that the conditions for comparatives are satisfied, it suffices to show that they hold extensionally. For strong transitivity, we show that the sequent

$$\underline{s} >_{\mathsf{F}} \underline{t}, \underline{t} >_{\mathsf{F}} \underline{u} \succ \underline{s} >_{\mathsf{F}} \underline{u}$$

holds extensionally by showing that if $\underline{s} >_{\mathsf{F}} \underline{t}$, $\underline{t} >_{\mathsf{F}} \underline{u}$ are both in X (i.e., not in Y) then $\underline{s} >_{\mathsf{F}} \underline{u}$ is in X. But this holds if and only if whenever $t \not\succsim s$ and $u \not\succsim t$ then $u \not\succsim s$, but if $\succsim$ is a total order, this is immediate. $t \not\succsim s$ holds iff $s \succ t$, and $\succ$ is also transitive. Justification for each of the other comparative conditions is similarly immediate. ∎

» «

There is much more that could be said about this simple logic of comparatives, but this is enough to illustrate what can be done with bounds. The vocabulary of the bounds provides a rich and powerful context in which to give an account of logical interconnections, and in a way that sticks closely to the norms governing our practices of assertion and denial. The logic here is orthodox and standard as far as it goes, but notice one special feature of this analysis: we have not had to express our conditions using the vocabulary of propositional logic. It is possible, of course, to add propositional connectives, and, so, to derive formulas such as $(s \geqslant_{\mathsf{F}} t \land t \geqslant_{\mathsf{F}} u) \to s \geqslant_{\mathsf{F}} u$, or $s >_{\mathsf{F}} t \to \neg(t \geqslant_{\mathsf{F}} s)$. However, there is no need to proceed in this way. We could, instead, proceed as we have done here, examining the logical relationships between judgements of the form Fs, $s >_F t$ and $s \geqslant_F t$ in their own rights, without concerning ourselves with other vocabulary. The bounds give us the means to do this in such a way that quite a lot of the behaviour of comparatives can be articulated.

This simple language is also a good illustration of the naturalness of attending to sequents with multiple conclusions. Here, the duality manifest between the *contrary* and *subcontrary* conditions between $s >_F t$ and $t \geqslant_F s$ is completely at home in this sequent setting. Adapting this to sequents of the form $X \succ A$ with one and only one formula in the consequent position would be constraining here. We would not see the natural duality present in these judgements.

5.5 | BOUNDS, CUT AND INFERENCE

There is more to say about positions and their bounds, but this is enough for us to proceed. In Chapter 9 we will extend our account of positions to allow for a more complex kind of scorekeeping, to keep track of the kind of reasoning and dialogue important for modal concepts. (There, positions will take more structure, allowing for certain different kinds of hypothetical assertion and denial, and as a result we move from sequents to hypersequents.) For now, though, this account of positions and their bounds will sustain us throughout the discussion of the next few chapters, as we address defining rules, propositional and predicate logic, and the various consequences of this understanding of the scope and power of proof. In the next rest of this chapter, we address the relationship between bounds (constraining assertion and denial) and the speech act of *inference*, and then address some other challenges and concerns that have been raised for this account of the connection between logic, assertion and denial. Before we get there, however, it is time to return to the *Cut* rule and its critics.

CUT, AND ITS CRITICS

As I indicated in the previous section, the *Cut* condition on bounds is not without its critics. It is true that one can accept the other conditions on bounds (reflexivity and weakening) and not accept the *Cut* condition. That is formally consistent, there is nothing that compels us to take 'bounds' to satisfy *Cut*.

Nothing, that is, except the nature of assertion and denial.

Here, I will elaborate on the connection between the norms governing assertion and denial that we have already seen, and show how they give us further reason to endorse *Cut* as a constraint on bounds. To do this, I will respond to some interesting arguments by Dave Ripley, in his paper, "Anything Goes" [204]. Ripley writes:

> Cut says that if asserting the Xs and denying the Ys blocks

I have changed Ripley's notation from Γ and Δ to X and Y to match ours. Ripley discussing *multiplicative Cut*, which proceeds from $X \succ A, Y$ and $X', A \succ Y'$ to $X, X' \succ Y, Y'$, instead of the *additive Cut* we use here. This difference is immaterial to the discussion. In the presence of weakening and contraction, they stand or fall together.

> you from denying A without going out of bounds, and asserting the X′s and denying the Y′s blocks you from asserting A without going out of bounds, then if you do all those things together, asserting the Xs and the X′s and denying the Ys and the Y′s, then you're already out of bounds. You don't have to say anything about A (A doesn't appear in the conclusion-sequent of the cut); you're already clashing before A even comes up.
>
> That is, cut says that if you've done some things that rule out your denying A, and you've done some other things that rule out your asserting A, then it's already too late. What you've already done doesn't fit together. You've got to leave yourself a path open to take on A, either leave open the option of asserting it or the option of denying it, *for any* A *always*.
>
> I don't see any particular reason why we should expect that. It certainly seems very straightforward, very coherent, to reject that kind of constraint, and say no, there are some sorts of things that the problem is with *them*. They can't be coherently asserted, and they can't be coherently denied, but that's not my fault. I'm perfectly well in bounds. It's the thing itself that's the source of the trouble with either asserting or denying it. [204, page 30]

Ripley is right to say that there is nothing in the very idea of bounds that force us to recognise *Cut* as governing the bounds. There is an interesting and rich theory of systems of proof according to which *Cut* is not only not accepted as a primitive rule, but according to which *Cut* is invalid. Ripley's own work is significant testament to that [41, 205, 203, 204]. However, the rejection of *Cut* comes with significant cost. Restricting the chaining of proofs — deciding that a proof from A to B and another proof from B to C give you no guarantee of a proof from A to C — is a not insignificant amputation of our logical toolkit. The restriction here allows for a great flexibility elsewhere: Ripley shows that in the absence of *Cut* a simple answer is available to Prior's puzzle we explored in the last chapter. Anything goes! Just about any proof rule, including those for *tonk*, can define a concept without disturbing the previous consequence relation. This does mean that proofs involving *tonk* cannot compose, so logic for *tonk*-afficionados looks rather different than it does for those of us who like to compose proofs, use lemmas, and appeal to *Cut*. Life without *Cut* seems to be a formal possibility. There is nothing in the structure of bounds that *demands* it.

However, there is something in the norms governing assertion and denial that speaks to the validity of the *Cut* rule. The norms governing assertion and denial interact in such a way as to give us a way to explain how *Cut* serves as a justifiable constraint on bounds.

Recall the analysis of assertion in terms of commitment given from page 176. To assert something (strongly) is to undertake a commitment concerning the content of that assertion, ruling it *in*, making that information available for use, and allowing others to call on us to vouch for it. To deny something (strongly) is to undertake a commitment concerning the content of that denial, ruling it *out*, making information to that effect available for use, and allowing others to call on us to vouch for it. Suppose, then, that we are in a dialogue with $[X : Y]$ as the common ground. Each member of X has been ruled in, each member of Y has been ruled out, and the position is in bounds. Suppose the question concerning A arises — someone asks: A? After a little work we demonstrate that $[X, A : Y]$ is out of bounds. Ruling it *in* has been decisively ruled *out*. (That is, we have not simply ruled out its assertion as *unjustified* or even *untrue* — A is *inconsistent* with our other commitments. There is no available position extending $[X : Y]$ at which A is ruled in.) Consider what it would take, then, to deny A. Given the understanding of denial we have already seen, we have what we need to make the denial. The answer concerning the question "A? *yes* or *no*?" is settled by the commitments already undertaken. The answer concerning A has to be *no*. *Yes* is ruled out, so we are in no way unsettled. The information ruling A out is already present in the position $[X : Y]$ and no further commitments need to be undertaken. If someone asks us to justify our denial of A, we can do so in terms of the demonstration of the sequent $X, A \succ Y$, and if questions concerning any other components of $[X : Y]$ arise, we use our preexisting answers for those questions, too. The material for answering the questions are all there. We have the resources to make the denial from the commitments we have already undertaken, so there is no possibility of a clash if we make that denial explicit, as opposed to implicit. *Cut* follows, in this way, from these principles governing assertion and denial.

» «

Now, the principles appealed to in this justification are, in an important sense, simply a restatement of the *Cut* rule. With such basic logical principles such as these, it is hard to go *deeper*. Explanations may well be circular, but sometimes the scenic circular route shows how the concepts are interconnected, and along that circular route, you see what else you might need to *avoid* if you do not wish to arrive at the place where you accept the *Cut* rule. Here, the *Cut* rule is implicit in the principle that if *one* answer to a yes/no question is ruled out, the other remains. To avoid *Cut*, you need to say that for a problematic sentence A (for which *Cut* is thought to fail at the position $[X : Y]$), when in the position $[X : Y]$ the question "A?" is raised the answer must be *no*, but not the kind of "*no*" that would be a *denial*. On this view, there must be two kinds of "no" answers to polar questions: the kind of "no" that you say when you truly *deny* A, and a second kind of "no" that rules out asserting A as out of bounds, but somehow does not count as a denial.

Ripley takes me to task for circular reasoning, when he addresses an earlier argument of mine addressing this issue [204, pp. 32, 33].

This two-level account of denial might seem familiar. After all, we have noticed that "no" features in more than one way, to *strongly* deny — to rule out the content in such a way that this is registered in the common ground — and to *weakly* deny. Isn't weak denial enough to be the kind of denial the opponent of *Cut* needs? Unfortunately, it is not. Weak denial is simply the bid to retract the content from the common ground. It is true that if the answer "*yes*" is ruled out for the issue A, in the position $[X : Y]$ then one response might be to bid to retract A, to weakly deny it. However, the position $[X : Y]$ involves much more about A than the information to the effect that A is unwarranted or goes beyond what we have reason to assert. The position $[X : Y]$ has already ruled out the *yes* answer for A. To reject *Cut* is to say that there are two ways that a position $[X : Y]$ can rule out a content such as A. One is *explicit* (where A is in Y), and the other, implicit (where $[X, A : Y]$ is out of bounds) and these can never, in general, be made to agree. To accept *Cut* is to say that this implicit ruling out can be made explicit, without further transgressing the bounds.

To phrase this in the equivalent way, swapping assertion and denial, we see that if a position $[X : Y]$ in the field of play is such that $[X : A, Y]$ is out of bounds, then $[X, A : Y]$ remains in the field of play. What does this mean? It is that if we are in the field of play and we consider A and recognise that it is *undeniable*, in the strong sense that denying it would be inconsistent with the commitments we have already undertaken, then the move familiar in Socratic dialogue would be a non-sequitur. If you reject *Cut* you have a move available to you if you encounter the Eleatic Stranger. If in dialogue he manoeuvres you into a position where you recognise that A is undeniable. No matter. If you reject *Cut*, it does not follow that you have any reason to accept it. If you reject *Cut*, perhaps you *cannot* accept A, especially if it is formed using a defining rule like that for *tonk*.

Recall Protagoras from the *Philebus*, among many of Socrates' dialogue partners, quoted on page 191.

BOUNDS AND INFERENCE

This criticism I have laid at the foot of those positions which would have us reject *Cut* has been laid at *my* feet, too. After all, the analysis of positions and bounds given here tells us that when we have shown that $[X : A, Y]$ is out of bounds (that is, when we have somehow derived $X \succ A, Y$), all we have shown is that denying A is ruled out. If we accept *Cut* this also shows us that there is no *extra* cost incurred in asserting A — the position $[X, A : Y]$ is available, just as $[X : Y]$ is. But according to some critics this is not enough to explain the force of *inference*. A derivation which tells us that A is undeniable is not enough. Apparently, what is wanted is not only that we *can* assert A, but that in some sense, in some context, we *must*. Is there any sense in which we *must* assert A, given that we are in the position $[X : Y]$, and the position $[X : A, Y]$ is out of bounds?

Which critics? Two examples are Hartry Field [68, 69], and Graham Priest [175, page 268].

There is no such force in the straightforward sense. I have discussed this at some length in "Multiple Conclusions" [195], but it is worth reiterating the major points here. Suppose we are in a position $[X : Y]$, and we recognise that $[X : A, Y]$ is out of bounds (say by constructing a derivation for $X \succ A, Y$). Does this mean that we ought to go on an assert A, or to accept it in some sense? In some cases, clearly not. There may be independent reason to reject A. In this case, perhaps the thing to do is to retract our commitment to something in the position $[X : Y]$, before going on to accept A. There is meaning in the old saw that one person's *modus ponens* is another's *modus tollens*. There is more than one thing that one could in an argument. Except for the case where $[X : Y]$ is *empty* (so we have a derivation for $\succ A$, that A in and of itself is *undeniable*), we could always cut away the derivation at some point in its foundation. This must be acknowledged before we go on to discover where there *is* a push to conclude the inference.

Another point that we must concede is that we may have reason *not* to consider or concede A even if it does follow from the commitments we have already made. Suppose the starting position involves the axioms of Peano Arithmetic, and suppose that A is some unbelievably complex expression which, if we were *able* to spell it out and we had the time to confirm, we 'would' (in some sense) recognise as following from the axioms we have granted. (Suppose it is some large expression too complex for any human to retain, of the form "$1+1+1+1+1+1+1+\cdots+1$ is even" for an inconceivably large, but finite, string of "1"s.) Clearly there is no sense in which we *must* assert that claim. We couldn't assert it.

We could assert something else that entails that claim. (We could assert the axioms of PA, for example.) We could perhaps describe the claim, too, in some roundabout way. But even to describe it and to say that it is true is not the same as asserting it.

But suppose we are not in a case like that. Suppose we are in the position $[X : Y]$, committed positively to everything in X and negatively to everything in Y. And suppose we have a derivation for $X \succ A, Y$, which assures us that A is undeniable. Suppose the question concerning A arises. Is there *any* pressure to accept A or to assert A?

We have already seen the answer in our discussion of the *Cut* rule. Suppose the question concerning A arises and we ask: "A?" The derivation for $X \succ A, Y$ shows us that (relative to the commitments we have undertaken) that A is undeniable. Given the polar question prompts for a yes/no answer, *no* has been ruled out, and so, the issue is settled. It is no longer unsettled between "yes" and "no", no uncertainty remains. Once the question has been asked, the position $[X : Y]$ and the derivation for $X \succ A, Y$ provide the answer. It must be 'yes'. If we go on to assert A, then we can meet all the commitments required for making that assertion, if we could do so for the starting position $[X : Y]$. We use the derivation $X \succ A, Y$ to show how the "no" pole of the polar question is ruled out. A is undeniable and has already been ruled in by the commitments we have already made. If questions are asked concerning the other assumptions in this derivation, we refer to whatever backed up *those* assertions. In this way, we use the bounds and the commitments we have already undertaken together, in concert, to discharge our commitments concerning A. This is how we can make explicit what was already implicit in the commitments we have undertaken.

What we have said for assertion can also be said for *inference*. The materials we have seen — whatever it is that shows us that $X \succ A, Y$ holds, that $[X : A, Y]$ is out of bounds — can be understood as underwriting the *inference* to A from $[X : Y]$. If we have already made the commitments in the position $[X : Y]$, if that is in the common ground, then the justification for $X \succ A, Y$ can be seen as grounding the inference to A, since it answers the question, *why* A. It shows how A *follows* from $[X : Y]$. In this sense, we not only make the assertion A, but we flag in advance where we can prompt a questioner to look if ever a question arises concerning A. This is one reason to make inference *explicit*, to show, in advance, how the commitment to vouch for A can be discharged.

» «

In this way, the sequent calculus, the structural rules, derivations, positions, bounds, common ground, assertion, denial, and inference are all interconnected aspects of our practice. Together, they constitute the antecedently given context of deducibility in which we can address questions of rules and definitions, and the status of logical vocabulary. Before we consider those questions and directly face Prior's challenge of *tonk* in the next chapter, we will address some of the challenges and concerns some people have to this approach.

5.6 | CHALLENGES AND CONCERNS

Thanks to Hannes Leitgeb, Lloyd Humberstone and Nico Silins for independently raising this issue with me.

Here is a puzzle, which might remind you of Moore's paradox. On this view of the bounds given here, is this position not automatically out of bounds?

$$[p : \text{I assert } p]$$

Surely there is a clash between *asserting* p on the one hand, and *denying* that you assert p on the other. But if we admit this, it seems that the proponent of a bounds account of positions is committed to take the argument from p to the claim that I assert p to be valid. Surely, since p could be true without being asserted, we have many reasons to resist the inference, and as a result, we have reason to deny our analysis of the connection between logical consequence and norms governing assertion and denial. Or do we?

Consider carefully what is going on in the case just described. A position in which I assert p and deny that I assert p is certainly *mistaken*. I cannot correctly assert p and also correctly deny that I assert p. However, it is perfectly possible for *you* to be committed to both parts of this position. Furthermore, there is no problem in you explicitly taking it up, by asserting p and denying that I assert p. (In fact, this happens regularly, when people argue with me, and take issue with my commitments, saying that there is some truth that I fail to recognise as true, and so, won't assert.) So, there is certainly no problem with this position, and it need not be out of bounds, even though there is a sense that it can never be something I can explicitly avow. Yes, it is Moore paradoxical in that limited sense, but that is not enough for it to be out of bounds.

Furthermore, I can recognise that the case is coherent and within bounds even if I cannot truly assert p and deny that I have asserted p. Recall that things can enter into the common ground without being recognised as *true*. I can happily proceed as follows:

> Suppose that p, and that I have not asserted that p.

Here, I am asking us to consider the position [p : I assert p], without asserting its content. This is a totally coherent supposition, and there need be no clash in the bounds.

There is no need to make the supposition so explicit to make this point. Suppose that we are not, in fact, committed to p, but are considering what would follow from p, were it to be the case. Given that dialectical scenario, we could happily reason as follows: suppose that p is the case. Does it follow, under the context of that supposition that I *assert* p? No, because I may well be ignorant of p, or think it too impolite to assert, or I could lie. We are very happy to grant scenarios in which p is the case but I do not assert p. (After all, I am not omniscient!) The strong norms of deductive logic are norms that govern deduction, and deductive consequence applies not only to reasoning from premises taken to be true, but also, reasoning from premises which are merely *supposed*, or *granted for the sake of argument*, as we have already conceded in our discussion of the common ground. This is enough to defeat the worry that our norms cut too coarsely, and license as valid inferences about *assertion* or *denial* that are not themselves valid.

We consider different norms governing kinds of supposition in Chapter 9 when we consider modal logics.

» «

Another concern for this approach to the antecedently given context of deducibility, is that it does not avail itself of enough resources available to us when we consider meaning and semantics. After all, the tension between *tonk* and other proposed definitions of concepts is concerning the ways *define* terms. Defining terms involves articulating or making precise what they *mean*. If we are going to address questions of meaning, it is important to use the resources that are available to us. Positions and their bounds might be a way to articulate *consistency* and the field of possible positions that we might take up in some language. But does it offer all that we want in a *semantics*? It is one thing to have a theory of logical consistency given by these bounds on positions, but does this say anything about the semantics of the expressions in our language? This question is stark, because it arises even at this stage, when we have not considered any logically complex expressions, and have an otherwise unanalysed language. Consider the atoms in such a language. Is there any sense in which we have given them a semantics by giving some account of the bounds for the language? Doesn't the power of model theory give us more to work with when it comes to issues in semantics? Doesn't starting here, with proof theory, limit our options, like depriving a boxer the use of their fists?

After all people call model theory "semantics" and proof theory is left out in the cold as "mere syntax."

In short, the answer is negative. Model theory for a language $\mathcal{L}$ is simply another way to determine the bounds. There is nothing in Tarski-style model theory which distinguishes one model from another. Another principle must be given to select out those models which are *good* models of how things are, and those that aren't. There are different choices to be made here, than in the case of proof theory, but the options of the same shape. We have seen how proof theory can connect with model theory in Chapter 3 when we discussed Kreisel's squeezing argument in

§3.4 on page 135. Here we showed that at least in the case of propositional valuations, we have a correspondence between model theory and the bounds determined by the classical connective rules. The same will hold in the case of predicate logic, though once we go beyond compact bounds, matters will get more delicate. However, once we consider non-compact bounds, model theory and the class of bounds defined in this way transcend the level of the tractable. Throughout the rest of this book, we will see how the bounds determined by concepts and the rules used to defined them, give rise to the standard structures of classical model theory. However, as we saw in Chapter 3 this will show how we can give model theory a distinctly normative pragmatic flavour, connecting it to our practices of assertion and denial, explaining how models are grounded in our practices and our positions. We need not be deceived into reversing the picture and thinking that our practices have the significance they do just because they happen to have models of a particular kind. Rather, the practices we have determine the possible field of play. Chess is the game it is because of the rules governing the practice. It would be very difficult to reverse engineer the rules of the game of chess simply by looking at the class of possible games.

DEFINING RULES

6

The first part of the book leads up to a single argument. This chapter presents that argument. Here, we will apply the techniques of proof theory to address Prior's question concerning whether rules can define concepts, and if they do, what kinds of concepts they define.

The argument of this chapter shows that given an appropriate antecedently given context of deducibility, with a language given by a set of bounds satisfying *Identity* and *Cut*, and for which any connectives are governed by *defining rules* (a notion which will be made precise in this chapter) then the addition of any new concept also governed by a defining rule is both *conservative* and *uniquely defining* from that starting point. In particular, given the antecedently given context arising out of norms governing assertion and denial discussed and defended in the previous chapter, then the rules defining the classical concepts of conjunction, disjunction, negation and the material conditional are, indeed, conservative and uniquely defining. Furthermore, given the syntactic distinction between singular terms and predicates, then the universal and existential quantifiers, and the identity predicate are also conservative and uniquely defining. More concepts than these can also be shown to be conservative and uniquely defining given a richer structure in the background context of deducibility, as we will see in the chapters ahead. However, the argument as it stands in this chapter applies more generally than this. If you prefer a background context in which contraction or weakening (or both) are rejected, then this argument will still apply. The crucial structural rules for the bounds are *Cut* and *Identity*, as far as this argument goes.

As I will argue in this chapter, concepts that are both uniquely defining and conservatively extending can play a particular important and distinctive role in our conceptual practice in a way that concepts that fail either criterion do not. The distinction between concepts that are conservative and uniquely defining in this sense, and those that aren't, is natural, it is worth drawing, and it provides a way to answer Arthur Prior's challenge to distinguish *tonk* from our everyday logical concepts like *and*, *or*, *not*, *some*, *all* and *identity*.

The result is of this chapter is simple to state in these general terms, but making it precise and proving it requires two different kinds of elaboration. *First*, the crucial concepts (the antecedently given context of deducibility, and defining rules) must be carefully made out, in a way that is general enough for these ideas to be applied to a wide range of concepts, but not so general so as to be hopelessly abstract and unintelligible. We have covered some of this ground in Chapter 2, and we will cover the remaining issues in Section 6.2. *Second*, there is the elaboration of the proof itself. This will take up Sections 6.3 and 6.4. The proof falls into two parts. First, we show that any system given an appropriate background context of deducibility and a system of defining rules can be converted into system consisting not of defining rules but of *Left* and *Right* rules in Gentzen's sense, using the *Cut* rule. Then, in the second part, we show that derivations in this system using the *Cut* rule can be reformulated without the use of that rule, and that the resulting derivations have some version of the subformula property — which then shows that the addition of new defining rules is *conservative*, because no sequent not involving new vocabulary must make use of that vocabulary in its derivation.

'Hauptsatz' is German, and the literal meaning is 'main clause.' In mathematics, it has come to mean '*main theorem*'. Gentzen's Hauptsatz is his cut elimination theorem.

This is our *Hauptsatz*, the central result of the book. Once we have completed that result, there is much to be done to draw out its significance. We will begin our descent from the mountain in this chapter in the final two sections, in Section 6.5 with an explanation of exactly how the result of this Chapter answers Prior's original question concerning *tonk*, and then look forward to the different ways this result might apply, to the kinds of concepts (connectives, modal operators, quantifiers, identity) which could count as properly *logical*, given this result.

The main theorem of this chapter, and its proof, are formulated in a very general way, because it is intended to apply to a very wide range of concepts and defining rules. This means that at first glance it may be hard to see *how* it is meant to apply in any particular case, or why it is formulated in the way that it is. So, in the next section we will start with a particular concrete instance of the result. and then, with that in hand, we will proceed to the argument in its generality.

6.1 | DEFINING A BICONDITIONAL

Let us start with a language, $\mathcal{L}_1$, and for the moment, suppose that it features no concepts used to create complex statements out of other, simpler statements. However, let us also suppose that it is used for assertion and denial, and is equipped with a family of positions and a bounds relation on those positions, satisfying the rules of *Identity* and *Cut*. The rules of *Weakening* and *Contraction* will play no central role in the argument of this chapter. We are free to assume contraction and weakening, and free to refrain from them. For the purposes of this section, we will assume that the bounds for $\mathcal{L}_1$ are closed under contraction and weakening, but we will note, as the argument proceeds, that they play no essential role.

The items in $\mathcal{L}_1$ need not be logically independent from one another. The bounds may feature significant relations in the language itself, provided that they are closed under *Identity* and *Cut*. To make matters concrete, and to ensure that the language has *some* distinctive features, we will use the example from Section 5.4 on page 193, a language for comparatives. We will take $\mathcal{L}_1$ to contain a single predicate F, two comparative relations $\geqslant_F$ and $>_F$ for "... at least as F as ..." and "... more F than ...", together with a stock of singular terms. The bounds for $\mathcal{L}_1$ will be given by the sequents in Definition 5.5, repeated here for convenience:

$$
\begin{array}{rl}
\textit{strong transitivity}: & s >_F t,\ t >_F u \succ s >_F u \\
\textit{weak transitivity}: & s \geqslant_F t,\ t \geqslant_F u \succ s \geqslant_F u \\
\textit{strong irreflexivity}: & s >_F s \succ \\
\textit{weak reflexivity}: & \succ s \geqslant_F s \\
\textit{contraries}: & s >_F t,\ t \geqslant_F s \succ \\
\textit{subcontraries}: & \succ s >_F t,\ t \geqslant_F s \\
\textit{strength}: & s >_F t \succ s \geqslant_F t \\
\textit{preservation}: & Fs,\ t \geqslant_F s \succ Ft
\end{array}
$$

These specify the bounds for $\mathcal{L}_1$, given the structural rules of identity, cut, contraction and weakening. $\mathcal{L}_1$ is our starting point for extending our language with a connective.

Instead of extending $\mathcal{L}_1$ with a connective that we have already defined in previous chapters, try extending the language with a *new* connective, one we could have defined in terms of the more familiar connectives, but which we can also add on its own: a *biconditional*. We aim to have in our new language $\mathcal{L}_2$ (extending $\mathcal{L}_1$) for each formula A and B, the formula $A \leftrightarrow B$. So, not only will statements like

$$Fa \leftrightarrow Fb \qquad Fa \leftrightarrow a >_F b$$

be formulas in $\mathcal{L}_2$ (which combine two formulas from $\mathcal{L}_1$) but so are

$$Fa \leftrightarrow (Fb \leftrightarrow a >_F b)$$

and other more complex statements of arbitrary depth. Notice that we have a well-founded *subformula* relation on the formulas in $\mathcal{L}_2$: the subformulas of a formula are simpler, smaller formulas, and there is no infinite descending chain of subformulas.

The bounds for $\mathcal{L}_2$ are fixed by requiring that $\mathcal{L}_2$ be closed under *Cut* and *Identity*, as well as *Weakening* and *Contraction*. The only extra constraint on the bounds is distinctive to the new vocabulary of the biconditional. We will impose the following *defining rule*.

$$\frac{X, A \succ B, Y \qquad X, B \succ A, Y}{X \succ A \leftrightarrow B, Y}\ {\scriptstyle \leftrightarrow Df}$$

This tells us that a position in which $A \leftrightarrow B$ is denied (together with asserting each element of X and denying each element of Y) is out of bounds if and only if the position in which A is asserted and B is denied (together with asserting each element of X and denying each element of Y) is out of bounds, *and* the position in which B is asserted and A is denied (together with asserting each element of X and denying each element of Y) is also out of bounds. Here is something that follows from this particular rule, given the bounds on $\mathcal{L}_1$ — at least in the presence of *Weakening* (*K*) — $s \geqslant_F t, t \geqslant_F s \succ Fs \leftrightarrow Ft$.

$$\dfrac{\dfrac{\dfrac{\scriptstyle preservation}{s \geqslant_F t,\ Fs \succ Ft}}{s \geqslant_F t,\ t \geqslant_F s,\ Fs \succ Ft}{\scriptstyle K} \qquad \dfrac{\dfrac{\scriptstyle preservation}{t \geqslant_F s,\ Ft \succ Fs}}{s \geqslant_F t,\ t \geqslant_F s,\ Ft \succ Fs}{\scriptstyle K}}{s \geqslant_F t,\ t \geqslant_F s \succ Fs \leftrightarrow Ft}{\scriptstyle \leftrightarrow Df}$$

Notice what is significant about this shape of the defining rule for the biconditional. It tells us what we need to know about the bounds for positions in which the newly introduced formula $A \leftrightarrow B$ is asserted in terms of positions involving simpler formulas, A and B. These formulas may not be formulas in $\mathcal{L}_1$, but they are at the very least, formulas which involve formulas with fewer instances of the newly introduced biconditional than $A \leftrightarrow B$. The rule is two-way, or *invertible*, which means that the bounds concerning $X \succ A \leftrightarrow B, Y$ are totally determined in terms of the bounds for $X, A \succ B, Y$ and $X, B \succ A, Y$. The facts it tells us about positions like these (in which $A \leftrightarrow B$ is denied) apply to *every* different position in which a biconditional is denied, since the rule applies to every position of the form $X \succ A \leftrightarrow B, Y$, and here, X and Y are unconstrained. The rule applies to any position in which $A \leftrightarrow B$ is denied.

DERIVING THE LEFT RULES

We are told less, however, about norms governing *asserting* a biconditional. That is because we do not *need* to specify such a rule independently. The constraints governing assertion of a biconditional are given by the structural rules which connect assertion and denial—*Identity* and *Cut*. For example, given *Identity*, we see that at least *these* two norms govern the assertion of formulas of the form $A \leftrightarrow B$.

$$\dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, A \succ B}\,{\scriptstyle \leftrightarrow Df} \qquad \dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, B \succ A}\,{\scriptstyle \leftrightarrow Df}$$

Asserting $A \leftrightarrow B$ clashes with asserting A and denying B, and also with asserting B and denying A. These two constraints follow from the defining rule for the biconditional, and the constraint that whatever biconditional judgements are, they are also constrained by the rule *Identity*. Given this, and the defining rule for biconditionals, we have at least *some* constraints on denying a biconditional. (If we did not impose *Identity* in the new language, this constraint would not follow from the defining rule.) Denying biconditionals are constrained, because the new language is taken to be closed under *Identity* and *Cut*.

But we can do more than this. This pair of rules tells us something about constraints governing asserting a biconditional $A \leftrightarrow B$ and its connections to positions involving its constituents A and B. We can generalise from this by using *Cut* to replace the A and B in these sequents by arbitrary formulas. For example, we can *Cut* with $X \succ A, Y$ and with $X', B \succ Y'$ to derive something more general:

$$\dfrac{\dfrac{X \succ A, Y \qquad \dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, A \succ B}\,{\scriptstyle \leftrightarrow Df}}{X, A \leftrightarrow B \succ B, Y}\,{\scriptstyle Cut} \qquad X', B \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\,{\scriptstyle Cut}$$

Given that the premise sequent $A \leftrightarrow B \succ A \leftrightarrow B$ in this derivation is justified by identity, we have shown that the following rule:

$$\frac{X \succ A, Y \qquad X', B \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L_1}$$

is a *derived* rule, using *Identity*, *↔Df* and *Cut*. We can do the same with the other sequent $A \leftrightarrow B, B \succ A$ which is also justified using *↔Df*.

$$\frac{\dfrac{X \succ B, Y \qquad \dfrac{A \leftrightarrow B \succ B \leftrightarrow A}{A \leftrightarrow B, B \succ A}\ {\scriptstyle \leftrightarrow Df}}{X, A \leftrightarrow B \succ A, Y}\ {\scriptstyle Cut} \qquad X', A \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle Cut}$$

and so, we can derive the second *↔L* rule,

$$\frac{X \succ B, Y \qquad X', A \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L_2}$$

which is also given by *Identity*, *↔Df* and *Cut*. The two rules, $\leftrightarrow L_1$ and $\leftrightarrow L_2$, give an account of when it is that asserting a biconditional is out of bounds, in an arbitrary position. Notice, however, that unlike the defining rule, these rules are not invertible. For one thing, there are *two* rules:

$$\frac{X \succ A, Y \qquad X', B \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L_1} \qquad \frac{X \succ B, Y \qquad X', A \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L_2}$$

and if we are inquiring as to whether or not the position $X, X', A \leftrightarrow B \succ Y, Y'$ is out of bounds, it may be due to the constraint given in the first rule, or that given in the second. Furthermore, the rules invoke a *splitting* of the position. If I am inquiring as to whether $U, A \leftrightarrow B \succ V$ is derivable, then to apply these rules, I would consider different ways to split U and V between the two premises (of both rules). In the presence of *Weakening* and *Contraction* we can simplify matters a little, and do away with need to split the position between the premises. Given *Contraction*, we can *justify* non-splitting versions of the rules, like this:

$$\frac{\dfrac{X \succ A, Y \qquad X, B \succ Y}{X, X, A \leftrightarrow B \succ Y, Y}\ {\scriptstyle \leftrightarrow L_1}}{X, A \leftrightarrow B \succ Y}\ {\scriptstyle W} \qquad \frac{\dfrac{X \succ B, Y \qquad X, A \succ Y}{X, X, A \leftrightarrow B \succ Y, Y}\ {\scriptstyle \leftrightarrow L_2}}{X, A \leftrightarrow B \succ Y}\ {\scriptstyle W}$$

and given *Weakening* and the non-splitting rules ($\leftrightarrow L'_{1,2}$) we can retrieve the splitting versions:

$$\frac{\dfrac{X \succ A, Y}{X, X' \succ A, Y, Y'}\ {\scriptstyle K} \qquad \dfrac{X', B \succ Y'}{X, X', B \succ Y, Y'}\ {\scriptstyle K}}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L'_1}$$

$$\frac{\dfrac{X \succ B, Y}{X, X' \succ B, Y, Y'}\ {\scriptstyle K} \qquad \dfrac{X, A \succ Y}{X, X', A \succ Y, Y'}\ {\scriptstyle K}}{X, X', A \leftrightarrow B \succ Y, Y'}\ {\scriptstyle \leftrightarrow L'_2}$$

Regardless, even in the presence of *Weakening* and *Contraction*, these rules are not invertible. Nonetheless, they give us all the information we need concerning the bounds for positions in which a biconditional is asserted, in terms of the bounds for simpler positions, and they are derivable rules, given *Identity* and *Cut*, and the defining rule $\leftrightarrow Df$. Although our starting point was a sequent rule that is quite different from Gentzen's left and right rules for the connective — an invertible defining rule which pins down the behaviour of the connective on one side of a position — we have shown that in the presence of *Identity* and *Cut* we can reconstruct left and right rules more familiar to us. The *right* rule for the biconditional is just the top-to-bottom half of the defining rule, and the *left* rules are derived in the way we have seen. However we have motivated the defining rule as a kind of *definition*, showing how to extend the language $\mathcal{L}_1$ and its bounds to a new, wider language, $\mathcal{L}_2$ in which more can be expressed. In the rest of these sections we will see how such a definition is guaranteed to meet Belnap's criteria of uniqueness and conservative extension, and why it is worth calling such definitions *purely logical*.

UNIQUENESS

We will start with Belnap's second criterion, to the effect that the introduced concept should be *uniquely defined*. There are many different ways we could understand this constraint, but the one we will focus on here is simple.

DEFINITION 6.1 [INDISTINGUISHABILITY WITH RESPECT TO THE BOUNDS] The formulas A and A' are indistinguishable with respect to the bounds in language $\mathcal{L}$ when

- $X, A \succ Y$ is derivable iff $X, A' \succ Y$ is derivable — that is, A and A' are *left*, or *assertion* indistinguishable.
- $X \succ A, Y$ is derivable iff $X \succ A', Y$ is derivable — that is, A and A' are *right*, or *denial* indistinguishable.

That is, the bounds constrain A and A' in exactly the same way.

LEMMA 6.2 [INDISTINGUISHABILITY ON ONE SIDE IS ENOUGH] *If the bounds for $\mathcal{L}$ are closed under Identity and Cut, then when A and A' are either* assertion *indistinguishable or* denial *indistinguishable, they are* both.

Proof: Suppose A and A' are assertion indistinguishable, and consider these derivations:

$$\dfrac{X \succ A, Y \qquad \dfrac{A' \succ A'}{A \succ A'}\,{\scriptstyle AI}}{X \succ A', Y}\,{\scriptstyle Cut} \qquad\qquad \dfrac{X \succ A', Y \qquad \dfrac{A \succ A}{A' \succ A}\,{\scriptstyle AI}}{X \succ A, Y}\,{\scriptstyle Cut}$$

In these derivations, we have uses *Identity*, assertion indistinguishability (labelled *AI*) and *Cut* to move between $X \succ A, Y$ and $X \succ A', Y$. Dual derivations (swapping left and right) show us how to use denial indistinguishability to move between $X, A \succ Y$ and $X, A' \succ Y$. ∎

If two formulas are indistinguishable, they function in the same way with respect to the bounds. There is no *give* or *play* between them as far as the bounds are concerned. If we add A and A' to some language in such a way that they are indistinguishable, there is no way to spot any difference, up to the bounds. A more familiar way to show that A and A' are indistinguishable is to prove the following two sequents:

$$A \succ A' \qquad A' \succ A$$

That is, if A and A' *entail each other*. Indeed, if A and A' entail each other then (in the presence of *Identity* and *Cut*) they are indistinguishable.

lemma 6.3 [co-entailment is indistinguishability] *A and A' entail each other in $\mathcal{L}$ (in the presence of Identity and Cut) if and only if they are indistinguishable in $\mathcal{L}$.*

Proof: Clearly, since $A \succ A$ is derivable (that's *Identity*), if A and A' are indistinguishable in $\mathcal{L}$ then they entail each other in $\mathcal{L}$. Conversely, using co-entailment and *Cut*, we have

$$\frac{X \succ A, Y \quad A \succ A'}{X \succ A', Y}\,\mathit{Cut} \qquad \frac{X \succ A', Y \quad A' \succ A}{X \succ A, Y}\,\mathit{Cut}$$

which gives us denial indistinguishability, and by the previous lemma, this suffices for indistinguishability. ∎

Now we can return to the question of uniqueness for $\leftrightarrow$, as given by its defining rule $\leftrightarrow$*Df*. The crucial question concerns the case where we we use the defining rule *twice*, once to define one connective (let's use '$\Leftrightarrow$' for this case) and again, to define another (let's use '$\leftrightharpoons$' for this case). In other words, we take our starting language $\mathcal{L}_1$ and we end up with a new language $\mathcal{L}_2^*$ constrained by the following two new defining rules:

$$\frac{X, A \succ B, Y \quad X, B \succ A, Y}{X \succ A \Leftrightarrow B, Y}\,{\Leftrightarrow}\mathit{Df} \qquad \frac{X, A \succ B, Y \quad X, B \succ A, Y}{X \succ A \leftrightharpoons B, Y}\,{\leftrightharpoons}\mathit{Df}$$

Is there a difference between $\Leftrightarrow$ and $\leftrightharpoons$? The way we will answer this question is by asking a related one: are $A \Leftrightarrow B$ and $A \leftrightharpoons B$ distinguishable in the language $\mathcal{L}_2^*$? The answer is simple: $A \Leftrightarrow B$ and $A \leftrightharpoons B$ are *indistinguishable* for all formulas A and B in $\mathcal{L}_2^*$. This follows straightforwardly, given these derivations, and Lemma 6.3, which assures us that co-entailment suffices for indistinguishability.

$$\frac{\dfrac{A \Leftrightarrow B \succ A \Leftrightarrow B}{A \Leftrightarrow B, A \succ B}\,{\Leftrightarrow}\mathit{Df} \quad \dfrac{A \Leftrightarrow B \succ A \Leftrightarrow B}{A \Leftrightarrow B, B \succ A}\,{\Leftrightarrow}\mathit{Df}}{A \Leftrightarrow B \succ A \leftrightharpoons B}\,{\leftrightharpoons}\mathit{Df}$$

$$\frac{\dfrac{A \leftrightharpoons B \succ A \leftrightharpoons B}{A \leftrightharpoons B, A \succ B}\,{\leftrightharpoons}\mathit{Df} \quad \dfrac{A \leftrightharpoons B \succ A \leftrightharpoons B}{A \leftrightharpoons B, B \succ A}\,{\leftrightharpoons}\mathit{Df}}{A \leftrightharpoons B \succ A \Leftrightarrow B}\,{\Leftrightarrow}\mathit{Df}$$

The reasoning here is general. (We have appealed to no features of $\mathcal{L}_1$ other than that it is closed under *Identity* and *Cut*.) If I have used the defining rule for the biconditional to govern my language (and I use $\Leftrightarrow$ in my language to represent the biconditional I introduced), and you have used the same rule (and I use $\leftrightharpoons$ in your language), then once we recognise this, and form a shared language in which we use $\Leftrightarrow$ and $\leftrightharpoons$ as connectives both governed by rules of the same form, then they are *indistinguishgable*. There is no way to split them as far as the bounds are concerned.

Notice, too, that nothing in our reasoning concerning uniqueness has appealed to *Weakening* or *Contraction*. The argument for uniqueness works whether either of those structural rules are imposed, or if they are left out.

» «

It is worth pausing to explain that this is not always the case, when a concept is governed by bounds of a particular shape. We have seen another example in the bounds for comparatives. Clearly a language can have two predicates F and G, paired with comparatives $>_F$, $\geqslant_F$ and $>_G$ and $\geqslant_G$ both satisfying the constraints of Definition 5.5, but there is no need at all for Fa and Ga, or for $a >_F b$ and $a >_G b$ to be indistinguishable. (Take a simple interpretation in which F is understood as '*fast*' and G, '*green*'.) There is no need to move from the premise that two concepts are governed (in some sense) by the same 'logic' to the conclusion that we really have one concept rather than two. In these cases, the bounds for those concepts do not give us a *definition* of what the concept *is* in any sense at all, but merely a *description* singling out some features the concept may have. In the case of a concept given by a defining rule, like the biconditional, enough is given to ensure that the concept is singled out, up to indistinguishability.

In fact, we could have a language in which F and G *are* indistinguishable, but $>_F$ and $>_G$ aren't. It is instructive to develop a class of models in which F and G always agree on what objects are in or out, but disagree concerning their *scales*.

CONSERVATIVE EXTENSION

Uniqueness is one of Belnap's two criteria. The other is conservative extension. Is the move from $\mathcal{L}_1$ to $\mathcal{L}_2$ conservative? That is, are the bounds for $\mathcal{L}_2$, when restricted to positions from $\mathcal{L}_1$ exactly the same as the bounds for $\mathcal{L}_1$. And if $\mathcal{L}_2$ is a conservative extension of $\mathcal{L}_1$, how could we show this? Thankfully, the tools are at hand. We have already shown conservative extension results using Gentzen's technique of Cut Elimination (see Theorem 2.27 on page 94 for details). We will do the same here, though for us, the starting point is not the same as Gentzen's. For us, the left and right rules for our biconditional are *derived*, and not primitive. Furthermore, the *Cut* rule is central in defining the bounds for $\mathcal{L}_1$ and $\mathcal{L}_2$. How does eliminating *Cut* help?

See Section 4.4 on page 159 for more discussion of the details of conservative extension, if you need a refresher.

Here is how the Cut Elimination argument will help us prove that $\mathcal{L}_2$ is a conservative extension of $\mathcal{L}_1$. The argument has three stages:

1. Show that the bounds for $\mathcal{L}_2$ are also defined by a sequent system where we replace $\leftrightarrow Df$ by the rules $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$.

2. Show that anything derivable in that system can also be derived without appeal to the *Cut* rule. (That is Gentzen's Cut Elimination result.)
3. Conclude that if a sequent $X \succ Y$ from the language $\mathcal{L}_1$ could have been derived in our *Cut* free system, it has a derivation using none of the rules featuring $\leftrightarrow$, that is, it is already out of bounds in $\mathcal{L}_1$.

Step 1 involves showing that in the presence of *Cut* and *Identity*, the rules $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$

$$\frac{X \succ A, Y \qquad X', B \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}{\scriptstyle \leftrightarrow L_1} \quad \frac{X \succ B, Y \qquad X', A \succ Y'}{X, X', A \leftrightarrow B \succ Y, Y'}{\scriptstyle \leftrightarrow L_2} \quad \frac{X, A \succ B, Y \quad X, B \succ A, Y}{X \succ A \leftrightarrow B, Y}{\scriptstyle \leftrightarrow R}$$

give us no more and no less than $\leftrightarrow Df$. We have already seen (page 211) that the left rules are derivable from $\leftrightarrow Df$, so the system in which the defining rule is replaced by the left and right rules will not *overgenerate* (allow us to derive too much). To show that it will not undergenerate, we need to show that from the left and right rules we can recover the defining rule, using *Identity* and *Cut*. The right rule gives us the top to bottom reading of the defining rule immediately. For the bottom to top reading, we can reason as follows:

$$\frac{X \succ A \leftrightarrow B, Y \quad \dfrac{A \succ A \quad B \succ B}{A \leftrightarrow B, A \succ B}{\scriptstyle \leftrightarrow L_1}}{X, A \succ B, Y}{\scriptstyle Cut} \quad \frac{X \succ A \leftrightarrow B, Y \quad \dfrac{B \succ B \quad A \succ A}{A \leftrightarrow B, B \succ A}{\scriptstyle \leftrightarrow L_2}}{X, B \succ A, Y}{\scriptstyle Cut}$$

So, Step 1 is complete. The bounds as given by extending those of $\mathcal{L}_1$ with $\leftrightarrow Df$ (in the presence of *Identity* and *Cut*) are exactly the same as the bounds given by extending those of $\mathcal{L}_1$ with $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$ (in the presence of *Identity* and *Cut*). Notice that at no stage of this argument we have appealed to *Weakening* or *Contraction*. This step of our reasoning works whether either of those structural rules are imposed, or if they are left out.

For Step 2, we show how derivations using $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$ and *Cut* (and the bounds of $\mathcal{L}_1$, which include *Identity* for $\mathcal{L}_1$ sequents, and perhaps *Weakening* and *Contraction*). The chief step of this argument is the eliminability of matching principal constituents. We need to show that if in a derivation we appealed to a *Cut* on a biconditional formula $A \leftrightarrow B$ which is immediately introduced in both premises of the *Cut*, by way of the left and right rules, then this *Cut* could have been replaced by *Cut*s on the simpler formulas, A and B. One case is this:

$$\frac{\dfrac{\overset{\vdots\, \delta_1}{X, A \succ B, Y} \quad \overset{\vdots\, \delta_2}{X, B \succ A, Y}}{X \succ A \leftrightarrow B, Y}{\scriptstyle \leftrightarrow R} \quad \dfrac{\overset{\vdots\, \delta_3}{X' \succ A, Y'} \quad \overset{\vdots\, \delta_4}{X'', B \succ Y''}}{X', X'', A \leftrightarrow B \succ Y', Y''}{\scriptstyle \leftrightarrow L_1}}{X, X', X'' \succ Y, Y', Y''}{\scriptstyle Cut}$$

and here, the *Cut* can indeed be simplified to

$$\dfrac{\dfrac{\overset{\vdots\,\delta_3}{X' \succ A, Y'} \quad \overset{\vdots\,\delta_1}{X, A \succ B, Y}}{X, X' \succ B, Y, Y'}\,{\scriptstyle Cut} \qquad \overset{\vdots\,\delta_4}{X'', B \succ Y''}}{X, X', X'' \succ Y, Y', Y''}\,{\scriptstyle Cut}$$

and the same holds for a *Cut* where the biconditional was introduced with $\leftrightarrow L_2$ instead of $\leftrightarrow L_1$, but in this case it will be the other premise of the $\leftrightarrow R$ step that is chosen to *Cut* with the premises of the $\leftrightarrow L_2$ step.

Is it merely a coincidence that this *Cut* is eliminable? In this case, the answer is a clear *no*. For the balance between the left and the right rules are due to the fact that both arise out of the single defining rule. Consider the particular $\leftrightarrow L_1$ inference used before the *Cut* step, and let's see how it was composed out of the more primitive $\leftrightarrow Df$ rule.

$$\dfrac{\dfrac{\overset{\vdots\,\delta_3}{X' \succ A, Y'} \quad \dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, A \succ B}\,{\scriptstyle \leftrightarrow Df}}{X', A \leftrightarrow B \succ B, Y'}\,{\scriptstyle Cut} \qquad \overset{\vdots\,\delta_4}{X'', B \succ Y''}}{X', X'', A \leftrightarrow B \succ Y', Y''}\,{\scriptstyle Cut}$$

Notice that *Cut*s on A and on B are present here inside this decomposition of $\leftrightarrow L_1$. If we replace that instance of $\leftrightarrow L_1$ inside the derivation with the *Cut* on $A \leftrightarrow B$ with its decomposition in terms of the defining rule and the simpler *Cut*s, you see something quite distinctive:

$$\dfrac{\dfrac{\overset{\vdots\,\delta_1}{X, A \succ B, Y} \quad \overset{\vdots\,\delta_2}{X, B \succ A, Y}}{X \succ A \leftrightarrow B, Y}\,{\scriptstyle \leftrightarrow R} \qquad \dfrac{\dfrac{\overset{\vdots\,\delta_3}{X' \succ A, Y'} \quad \dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, A \succ B}\,{\scriptstyle \leftrightarrow Df}}{X', A \leftrightarrow B \succ B, Y'}\,{\scriptstyle Cut} \qquad \overset{\vdots\,\delta_4}{X'', B \succ Y''}}{X', X'', A \leftrightarrow B \succ Y', Y''}\,{\scriptstyle Cut}}{X, X', X'' \succ Y, Y', Y''}\,{\scriptstyle Cut}$$

The *Cut* formula at the conclusion of the $\leftrightarrow L_1$ rule, now expanded into its derivation from the defining rule, is *passive* along the two *Cut*s we used to define it. If we commute the *Cut* up that shaded path, to occur at the top rather than at the bottom, we get the following:

$$\dfrac{\dfrac{\overset{\vdots\,\delta_3}{X' \succ A, Y'} \quad \dfrac{\dfrac{\overset{\vdots\,\delta_1}{X, A \succ B, Y} \quad \overset{\vdots\,\delta_2}{X, B \succ A, Y}}{X \succ A \leftrightarrow B, Y}\,{\scriptstyle \leftrightarrow R} \quad \dfrac{A \leftrightarrow B \succ A \leftrightarrow B}{A \leftrightarrow B, A \succ B}\,{\scriptstyle \leftrightarrow Df}}{X, A \succ B, Y}\,{\scriptstyle Cut}}{X, X' \succ B, Y, Y'}\,{\scriptstyle Cut} \qquad \overset{\vdots\,\delta_4}{X'', B \succ Y''}}{X, X', X'' \succ Y, Y', Y''}\,{\scriptstyle Cut}$$

If you look at this remaining *Cut* on $A \leftrightarrow B$, you can see that both premises of the *Cut* are instances of $\leftrightarrow Df$. The left premise is the top-to-bottom version, from the premises $X, A \succ B, Y$ and $X, B \succ A, Y$, and the right premise is the bottom-to-top version, applied to the identity sequent $A \leftrightarrow B \succ A \leftrightarrow B$, and selecting the *first* component of the defining rule. The result is that we *undo* the defining rule which produced $X \succ A \leftrightarrow B, Y$, trading it in for its first premise, $X, A \succ B, Y$. Deleting that roundabout piece of reasoning and staying with that first premise instead of taking the journey through $A \leftrightarrow B$, we get

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_3\\ X' \succ A, Y'\end{array} \quad \begin{array}{c}\vdots\,\delta_1\\ X, A \succ B, Y\end{array}}{X, X' \succ B, Y, Y'}\,{\scriptstyle Cut} \quad \begin{array}{c}\vdots\,\delta_4\\ X'', B \succ Y''\end{array}}{X, X', X'' \succ Y, Y', Y''}\,{\scriptstyle Cut}$$

which is *exactly* the derivation we wanted. The elimination of the complex *Cut* on $A \leftrightarrow B$ can thus be explained by the way in which the left and right rules arise out of the single, invertible *defining* rule. That justifies the elimination of matching active $A \leftrightarrow B$ formulas in a derivation. The *rest* of the argument for the elimination of *Cut* is exactly as before, from Chapter 2. The only wrinkle is to be found in the bounds for $\mathcal{L}_1$. We have said that the bounds for $\mathcal{L}_1$ are given by the conditions for comparatives, and these are closed under *Cut*. Our argument for eliminating *Cut* does not deal with *these* sequents. For the purposes of our new presentation of the bounds, we take all of the bounds from $\mathcal{L}_1$ as given. (If you like, take every sequent in the $\mathcal{L}_1$ vocabulary that is out of bounds as an *axiom*.) We extend these bounds simply by using $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$, and *Contraction* and *Weakening* if they are present. That is how we can define the bounds for $\mathcal{L}_2$. Anything we could show to be out of bounds using *Cut*, in addition to these rules is out of bounds already. Instead of going through the remaining details for that proof here — they are standard, and exactly as given in Chapter 2 — I will save the verification for the next section, where we show that this phenomenon for the rules for $\leftrightarrow$ is totally general, and applies to all defining rules. For now, though, we declare Step 2 of our argument complete.

Before continuing on to Step 3, I will again pause to note again that nothing in this argument appealed to *Contraction* or to *Weakening*, so again, we can take or leave these structural rules, as desired.

This leaves Step 3. Suppose we have a derivation in $\mathcal{L}_2$ of the sequent $X \succ Y$, where this sequent is in the vocabulary of $\mathcal{L}_1$. If we have such a derivation in our presentation without *Cut* but with $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$, it follows that the derivation *cannot* have used the rules for $\leftrightarrow$, for these rules introduce $\leftrightarrow$ formulas, and none of the rules in the calculus (*Contraction*, *Weakening* if present, and the $\leftrightarrow L_{1,2}$ and $\leftrightarrow R$ rules) discard materials from the premises. Whenever a formula is present in the premises, it is present in the conclusion, at least as a subformula, so in this sense, the rules are *compositional*, when read from top to bottom. The rules used to derive the sequent $X \succ Y$ in the vocabulary of $\mathcal{L}_1$ are restricted to rules featuring in $\mathcal{L}_1$. So, its extension to $\mathcal{L}_2$ is conservative. Step 3 is complete.

» «

So, our extension from $\mathcal{L}_1$ to $\mathcal{L}_2$ is conservative. This means that the transition from $\mathcal{L}_1$ to $\mathcal{L}_2$ leaves the bounds for $\mathcal{L}_1$ positions *totally unchanged*. In other words, any open position in $\mathcal{L}_1$ remains open in $\mathcal{L}_2$. Any closed position in $\mathcal{L}_1$ remains closed in $\mathcal{L}_2$. The transition to $\mathcal{L}_2$ allows for more things to be said, things that could not be said in $\mathcal{L}_1$ alone. The extension is through giving a defining rule, not simply stating that some formula in $\mathcal{L}_1$ could be rewritten in a new way. If the speaker of $\mathcal{L}_1$ does not want to use the concept from $\mathcal{L}_2$, it makes no difference to her, as far as the bounds in $\mathcal{L}_2$ are concerned. The extension to $\mathcal{L}_2$ is pure gain. Nothing one could express in $\mathcal{L}_1$ is altered in any way, as far as the space of possible positions is concerned. In this way, the extension is *free*. Since the extension is uniquely defined (as per the previous subsection), the extension is *defined* rather than merely *described*, in that the bounds for $\mathcal{L}_2$ are determined up to indistinguishability.

Before spending the next section generalising this result to *arbitrary* defining rules of this shape, it is worth sketching how a concept like *tonk* fares, in comparison to the biconditional. It will come as no surprise to learn that *tonk* has *no* defining rule. Recall, in the sequent setting, Prior's rules for *tonk* have the following forms:

$$\frac{X, B \succ Y}{X, A \mathbin{\textit{tonk}} B \succ Y}\,\textit{tonkL} \qquad \frac{X \succ A, Y}{X \succ A \mathbin{\textit{tonk}} B, Y}\,\textit{tonkR}$$

If we take the *left* rule for *tonk* as a possible defining rule, then we get this:

$$\frac{X, B \succ Y}{\overline{X, A \mathbin{\textit{tonk}_L} B \succ Y}}\,\textit{tonk}_L\textit{Df}$$

This certainly defines a connective, but A $\textit{tonk}_L$ B so defined is indistinguishable from B, and this in no way supports the *tonkR* rule. We could do the same for the *right* rule:

$$\frac{X \succ A, Y}{\overline{X \succ A \mathbin{\textit{tonk}_R} B, Y}}\,\textit{tonk}_R\textit{Df}$$

but A $\textit{tonk}_R$ B, so defined, is indistinguishable from A, and this in no way supports the *tonkL* rule. Neither the left *tonk* rule nor the right is enough to deliver the other rule when used as a defining rule.

Is there any *other* way we could introduce *tonk* by way of a defining rule? To answer that question, we need to discuss defining rules in their full generality. That will be our topic in the next section. The sections after that will then spell out the three steps of our argument in the full generality of those defining rules. Section 6.3 shows how to define left and right rules from defining rules, and shows how these rules have exactly the same power as the defining rules. Section 6.4 then shows that *Cut* can be eliminated from the resulting system, and then in Section 6.5, we show how this results in conservative extension, and so, an answer to Prior's original question concerning rules and definitions.

6.2 | defining rules defined

In an aim to be general, we will consider the case of a defining rule for an n-ary connective, $\sharp$, where n may be any finite number. The defining rule for $\sharp$ will specify the bounds for positions, in which a $\sharp$ formula is either asserted (on the left) or denied (on the right). That is, it will either define bounds for positions of the form

$$X, \sharp(A_1, \ldots, A_n) \succ Y$$

where the formula is asserted, or it will define bounds for positions of the form

$$X \succ \sharp(A_1, \ldots, A_n), Y$$

where the formula is denied. Instead of always having to deal with two different cases for a defining rule, one for formulas on the left, and the other for formulas on the right, we will use a single unified notation for these two different cases. We will represent the sequent $X \succ Y$ as $\mathfrak{S}$, and we will represent the addition of the formula A into the sequent $X \succ Y$ on one side (either left or right) as $\mathfrak{S}\langle A\rangle$. Here, the angle brackets represent a choice of *one* side of the sequent, and we will use the square brackets to represent the *other* side. So, this notation

In all of the reasoning in the next three sections, we will explicitly consider only multiple conclusion sequents. However, every definition can be understood, too, in a single conclusion setting, though you must take care to respect the constraints of single conclusion sequents. For example, in a single conclusion setting, $\mathfrak{S}\langle A\rangle$ is defined where $\langle\,\rangle$ selects for the right, only when $\mathfrak{S}$ has an empty right hand side. Showing exactly what remains in a single conclusion setting will be left for an *excursus*, starting on page 231.

$$\frac{\mathfrak{S}[A]}{\mathfrak{S}\langle\neg A\rangle}$$

— where we interpret $\langle\,\rangle$ as selecting the *left* of the sequent — represents the usual $\neg L$ rule. If we interpret $\langle\,\rangle$ as selecting the *right* of the sequent, then this notation represents the $\neg R$ rule, as given below.

$$\frac{X \succ A, Y}{X, \neg A \succ Y}\,{\scriptstyle \neg L} \qquad \frac{X, A \succ Y}{X \succ \neg A, Y}\,{\scriptstyle \neg R}$$

We can combine the use of the two brackets in the one sequent. If we understand $\langle\,\rangle$ as selecting the right of the sequent, then the $\to R$ rule has the following shape:

$$\frac{\mathfrak{S}[A]\langle B\rangle}{\mathfrak{S}\langle A \to B\rangle}\,{\scriptstyle \to R} \qquad \frac{X, A \succ B, Y}{X \succ A \to B, Y}\,{\scriptstyle \to R}$$

where $\mathfrak{S}[A]\langle B\rangle$ is the sequent $X \succ Y$ with A added to the left and B added to the right. This is how we will use the two sets of brackets, to represent inserting formulas into different sides of a sequent.

A defining rule, then, for the n-ary connective $\sharp$ will give the bounds for the sequents $\mathfrak{S}\langle\sharp(A_1, \ldots, A_n)\rangle$, where $\langle\,\rangle$ selects either for the *left* or for the *right*. Let's see the rule $\leftrightarrow Df$ from this perspective. It is

$$\frac{\mathfrak{S}[A_1]\langle A_2\rangle \quad \mathfrak{S}[A_2]\langle A_1\rangle}{\mathfrak{S}\langle A_1 \leftrightarrow A_2\rangle}\,{\scriptstyle \leftrightarrow Df} \qquad \frac{X, A_1 \succ A_2, Y \quad X, A_2 \succ A_1, Y}{X \succ A_1 \leftrightarrow A_2, Y}\,{\scriptstyle \leftrightarrow Df}$$

where $\langle\,\rangle$ selects for the right. It places the subformulas of $A_1 \leftrightarrow A_2$ in different positions in each of the premise sequents. Not all defining rules select each of the subformulas in each premise. For example, the defining rule for $\vee$ (in the absence of *Contraction* or *Weakening*, at least) has the following shape:

$$\frac{\mathfrak{S}\langle A_1\rangle \quad \mathfrak{S}\langle A_2\rangle}{\mathfrak{S}\langle A_1 \vee A_2\rangle}\,{\vee Df} \qquad \frac{X, A_1 \succ Y \quad X, A_2 \succ Y}{X, A_1 \vee A_2 \succ Y}\,{\vee Df}$$

where $\langle\,\rangle$ now selects for the left of the sequent. Here, the subformulas A_1 and A_2 are distributed between the premise sequents. The general form of a defining rule, then, can be rather involved. If we define an n-ary connective $\sharp$, we may have m different premise sequents, each of which select some of the subformulas $A_1, \ldots, A_n$ for the *same* side as the subformula occurred, and some of the subformulas for the *other* side. I don't say some *other* of the formulas, because it is quite permissible, in general, to have the same formula occurring on both sides of the sequent, though, in practice, this makes sense only in the absence of the *Weakening* rule. (In the presence of *Weakening*, the sequent $X, A \succ A, Y$ is always out of bounds. In *Weakening*'s absence, it may not be.) Furthermore, in the absence of *Contraction*, it may be that a subformula occurs multiple times in the same premise of a defining rule. For example, we could have a defining rule for A^2, which is also definable as $A \otimes A$, with the following shape:

$$\frac{\mathfrak{S}\langle A, A\rangle}{\mathfrak{S}\langle A^2\rangle}\,{{}^2Df}$$

where $\langle\,\rangle$ selects for the left. In the absence of *Contraction*, A^2 is distinguishable from A. In fact, in the absence of *Weakening*, A^2 is also distinguishable from A. (To derive $A \succ A^2$, you need *Contraction* or something like it. To derive $A^2 \succ A$, you need *Weakening*, or something like it. In linear logic, neither is derivable.)

So, the general shape of a defining rule is rather involved. The n-ary connective $\sharp$ may be defined by an m-premise defining rule, given a choice of a side for $\langle\,\rangle$, and given a selection, for each $1 \leqslant i \leqslant m$, the selection of numbers j_i and k_i, and formulas $B^i_1, \ldots, B^i_{j_i}$ and $C^i_1, \ldots, C^i_{k_i}$ from the distinct subformulas $A_1, \ldots, A_n$ of $\sharp(A_1, \ldots, A_n)$. The defining rule then has the form:

$$\frac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}] \text{ (for each } 1 \leqslant i \leqslant m)}{\mathfrak{S}\langle \sharp(A_1, \ldots, A_n)\rangle}\,{\sharp Df}$$

We have already seen how some familiar connectives fit in this scheme. Figure 61 contains the family of each defining rule for a connective given on the *left*. It is worth explaining some of the distinctive features of the rules given in Figure 61.

On the top line, notice that we have two rules for 0-ary connectives, that is, *constants*. Being 0-ary, they have no subformulas, and so, there are no formulas to insert into the sequent $\mathfrak{S}$. The two formulas so defined here take the two possible responses to this constraint. For $\bot$, we select *zero* premise sequents, and hence, nothing is required to derive

$$\frac{}{\mathfrak{S}\langle\bot\rangle}\ \bot Df \qquad \frac{\mathfrak{S}}{\mathfrak{S}\langle t\rangle}\ tDf \qquad \frac{\mathfrak{S}[A_1]}{\mathfrak{S}\langle\neg A_1\rangle}\ \neg Df$$

$$\frac{\mathfrak{S}\langle A_1, A_2\rangle}{\mathfrak{S}\langle A_1 \otimes A_2\rangle}\ \otimes Df \qquad \frac{\mathfrak{S}\langle A_1\rangle \quad \mathfrak{S}\langle A_2\rangle}{\mathfrak{S}\langle A_1 \vee A_2\rangle}\ \vee Df$$

$$\frac{\mathfrak{S}\langle A_1\rangle[A_2]}{\mathfrak{S}\langle A_1 - A_2\rangle}\ -Df \qquad \frac{\mathfrak{S}\langle A_1\rangle[A_2] \quad \mathfrak{S}\langle A_2\rangle[A_1]}{\mathfrak{S}\langle A_1 \bowtie A_2\rangle}\ \bowtie Df$$

Here, $\langle\,\rangle$ is the *left* of a sequent, and $[\,]$ is the *right*.

Figure 61: DEFINING RULES FOR SOME FORMULAS ON THE LEFT

the concluding sequent $\mathfrak{S}\langle\bot\rangle$. It is *always* derivable. In other words, $X, \bot \succ Y$ is axiomatic. Asserting $\bot$ is always out of bounds. This rule is invertible, in the sense that from its concluding sequent $X, \bot \succ Y$, the premise sequents follow — but in this case, there is no problem, because there are *no* premise sequents to derive. The other constant we define on the left is t. It takes the other response to the constraint that there are no subformulas, in that it requires nothing to be inserted into the sequent $\mathfrak{S}$. In other words, $X, t \succ Y$ is out of bounds if and only if $X \succ Y$ is out of bounds. (In this case, t is a formula that makes no difference to the bounds when asserted.)

The rules for $\neg$, $\otimes$ and $\vee$ are familiar, so we do not need to discuss them. The connectives $-$ and $\bowtie$ are new. Asserting $A - B$ has the same effect as asserting A and denying B. It is the *difference* connective, standing as the dual to the conditional $A \to B$ (where *denying* $A \to B$ has the same effect on the bounds as asserting A and denying B). In classical logic, $A - B$ is indistinguishable from $A \wedge \neg B$, but in weaker logics, these come apart. The last connective $A \bowtie B$ is symmetric difference, which is the dual to the biconditional $A \leftrightarrow B$. A quick inspection of the defining rules shows that $A \bowtie B$ is equivalent to $(A - B) \vee (B - A)$.

What we can do on the left of the sequent we can also do on the right. The connectives and operators $\top$, f, $\neg$, $\oplus$, $\wedge$, $\to$ and $\leftrightarrow$ can be defined on the right, as is shown in Figure 62.

These are by no means the only connectives definable by way of defining rules. We can define more complex concepts by choosing options however we wish. However, each of the connectives so definable could have been made out of the connectives we have seen here. For example, consider the following defining rule (where $\langle\,\rangle$ selects for the *right*), for a 4-place connective $\flat$.

$$\frac{\mathfrak{S}\langle A_1, A_2, A_2\rangle[A_3] \quad \mathfrak{S}[A_1, A_4]}{\mathfrak{S}\langle\flat(A_1, A_2, A_3, A_4)\rangle}\ \flat Df$$

$$\frac{}{\mathfrak{S}\langle \top \rangle}\,\top Df \qquad \frac{\mathfrak{S}}{\mathfrak{S}\langle f \rangle}\,f Df \qquad \frac{\mathfrak{S}[A_1]}{\mathfrak{S}\langle \neg A_1 \rangle}\,\neg Df$$

$$\frac{\mathfrak{S}\langle A_1, A_2 \rangle}{\mathfrak{S}\langle A_1 \oplus A_2 \rangle}\,\oplus Df \qquad \frac{\mathfrak{S}\langle A_1 \rangle \quad \mathfrak{S}\langle A_2 \rangle}{\mathfrak{S}\langle A_1 \wedge A_2 \rangle}\,\wedge Df$$

$$\frac{\mathfrak{S}[A_1]\langle A_2 \rangle}{\mathfrak{S}\langle A_1 \to A_2 \rangle}\,\to Df \qquad \frac{\mathfrak{S}[A_1]\langle A_2 \rangle \quad \mathfrak{S}[A_2]\langle A_1 \rangle}{\mathfrak{S}\langle A_1 \leftrightarrow A_2 \rangle}\,\leftrightarrow Df$$

Here, $\langle\,\rangle$ is the *right* of a sequent, and $[\,]$ is the *left*.

Figure 62: DEFINING RULES FOR SOME FORMULAS ON THE RIGHT

Decomposing the steps of the rule, you can see that $\flat(A_1, A_2, A_3, A_4)$ is indistinguishable from $(A_3 \to (A_1 \oplus (A_2 \oplus A_2))) \vee \neg(A_1 \otimes A_4)$, since applying the defining rules for $\to$, $\oplus$, $\neg$ and $\otimes$ we get

$$\frac{\dfrac{\dfrac{\dfrac{\mathfrak{S}\langle A_1, A_2, A_2 \rangle [A_3]}{\mathfrak{S}\langle A_1, A_2 \oplus A_2 \rangle [A_3]}\,\oplus Df}{\mathfrak{S}\langle A_1 \oplus (A_2 \oplus A_2) \rangle [A_3]}\,\oplus Df}{\mathfrak{S}\langle A_3 \to (A_1 \oplus (A_2 \oplus A_2)) \rangle}\,\to Df \qquad \dfrac{\dfrac{\mathfrak{S}[A_1, A_4]}{\mathfrak{S}[A_1 \otimes A_4]}\,\otimes Df}{\mathfrak{S}\langle \neg(A_1 \otimes A_4) \rangle}\,\neg Df}{\mathfrak{S}\langle (A_3 \to (A_1 \oplus (A_2 \oplus A_2))) \vee \neg(A_1 \otimes A_4) \rangle}\,\vee Df$$

which shows that our formula $(A_3 \to (A_1 \oplus (A_2 \oplus A_2))) \vee \neg(A_1 \otimes A_4)$ is indistinguishable from $\flat(A_1, A_2, A_3, A_4)$, being derivable from exactly the same starting sequents, by way of invertible rules.

THEOREM 6.4 [WHAT DEFINING RULES DEFINE] *Any n-ary connective definable by way of a defining rule of this shape is indistinguishable from some concept composable from $\bot$, t, $\otimes$, $\vee$ and $\neg$.*

Proof: Consider the connective $\sharp$, given by the defining rule

$$\frac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i} \rangle [C^i_1, \ldots, C^i_{k_i}] \text{ (for each } 1 \leqslant i \leqslant m)}{\mathfrak{S}\langle \sharp(A_1, \ldots, A_n) \rangle}\,\sharp Df$$

where each $B^i_{j'}$ and $C^i_{k'}$ is chosen from $A_1, \ldots, A_n$, and where $\langle\,\rangle$ selects for the left of the sequent. It is immediate that the formula $\sharp(A_1, \ldots, A_n)$ is indistinguishable from the m-fold disjunction

$$\sharp_1(A_1, \ldots, A_n) \vee \cdots \vee \sharp_m(A_1, \ldots, A_n)$$

where the connective $\sharp_i$ is given by the single premise defining rule

$$\frac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i} \rangle [C^i_1, \ldots, C^i_{k_i}]}{\mathfrak{S}\langle \sharp_i(A_1, \ldots, A_n) \rangle}\,\sharp_i Df$$

which singles out the i-th premise of the defining rule for $\sharp$. (In the case where m is *zero*, it is a zero-ary additive disjunction, namely $\bot$.) Now consider the $j_i + k_i$-fold multiplicative conjunction

$$B^i_1 \otimes \cdots \otimes B^i_{j_i} \otimes \neg C^i_1 \otimes \cdots \otimes \neg C^i_{k_i}$$

Unpacking the defining rules for $\neg$ and for $\otimes$, we see that we have

$$\dfrac{\dfrac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}]}{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}, \neg C^i_1, \ldots, \neg C^i_{k_i}\rangle}\,{\neg Df}}{\mathfrak{S}\langle B^i_1 \otimes \cdots \otimes B^i_{j_i} \otimes \neg C^i_1 \otimes \cdots \otimes \neg C^i_{k_i}\rangle}\,{\otimes Df}$$

In other words, $\sharp_i(A_1, \ldots, A_n)$ is indistinguishable from $B^i_1 \otimes \cdots \otimes B^i_{j_i} \otimes \neg C^i_1 \otimes \cdots \otimes \neg C^i_{k_i}$. (In the case where $j_i + k_i$ is *zero*, the appropriate formula is the zero-ary multiplicative disjunction, namely t.) Putting this together, the formula $\sharp(A_1, \ldots, A_n)$ is indistinguishable from an additive disjunction of a multiplicative conjunction formulas either chosen from $A_1, \ldots, A_n$ or their negations.

This leaves the case of defining rules where $\langle\,\rangle$ selects for the *right* of the sequent:

$$\dfrac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}] \text{ (for each } 1 \leqslant i \leqslant m)}{\mathfrak{S}\langle \flat(A_1, \ldots, A_n)\rangle}\,{\flat Df}$$

Here, $\flat(A_1, \ldots, A_n)$ is indistinguishable from $\neg\sharp(A_1, \ldots, A_n)$, where $\sharp$ is given by a defining rule on the *left*, and hence, is definable in terms of $\bot$, t, $\otimes$, $\vee$ and $\neg$, and as a result, so, too, is $\flat$. ■

Despite this result, it is important to keep the option of defining concepts directly, using defining rules, rather than going through intermediates. For example, a concept like $\leftrightarrow$ can be defined in terms of $\rightarrow$ and $\wedge$, but the reverse is not necessarily the case. There is no way, in general, to recover the *asymmetric* formula $A \rightarrow B$ from the thoroughgoingly symmetric $\leftrightarrow$. (In some cases, it can be done in the presence of other connectives, of course. $A \leftrightarrow (A \wedge B)$ is indistinguishable from $A \rightarrow B$ in the presence of *Weakening*. In linear logic, however, it is indistinguishable only from $(A \rightarrow B) \wedge (A \rightarrow A)$, which is, in general, *distinguishable* from $A \rightarrow B$.) A language in which we only have the biconditional is worth studying on its own, to find the limits of its expressive power, without adding extraneous concepts. So, we will continue to consider defining rules in their generality, for individual complex concepts may be worth treating in their own right, by way of their own defining rules.

» «

Before we leave this section, there is one more simple result concerning the nature of defining rules that is worth making explicit.

THEOREM 6.5 [DEFINING RULES DEFINE UNIQUELY] *If $\sharp$ and $\sharp'$ are introduced with defining rules of the same form, then $\sharp(A_1, \ldots, A_n)$ is indistinguishable from $\sharp'(A_1, \ldots, A_n)$.*

The result is an easy consequence of Lemma 6.2 (see page 212), and the shape of defining rules.

Proof: If $\sharp$ and $\sharp'$ are governed by $\sharp Df$ and $\sharp' Df$, which have the same shape, then we can reason as follows:

$$\dfrac{\dfrac{\mathfrak{S}\langle \sharp(A_1, \ldots, A_n)\rangle}{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_1}\rangle[C^i_1, \ldots, C^i_{k_i}]}\,\sharp Df \qquad \text{(each } i\text{)}}{\mathfrak{S}\langle \sharp'(A_1, \ldots, A_n)\rangle}\,\sharp' Df$$

It follows that $\sharp(A_1, \ldots, A_n)$ is indistinguishable from $\sharp'(A_1, \ldots, A_n)$ on the side of sequents selected by $\langle\,\rangle$. Lemma 6.2 tells us that indistinguishability on one side suffices for indistinguishability, so as far as the bounds are concerned, $\sharp$ and $\sharp'$ are indeed indistinguishable. ■

So, defining rules satisfy one half of Belnap's constraints. To show that they also satisfy the other will require more work. This is the topic of the next two sections.

6.3 | DEFINING RULES AND LEFT/RIGHT RULES

Given a defining rule for an n-ary connective $\sharp$, our technique for showing that $\sharp$ can be conservatively added to a language will be to first show how we can define the pair of left and right rules that are equal in power to $\sharp Df$. So, consider the defining rule:

$$\frac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}] \text{ (for each } 1 \leqslant i \leqslant m)}{\mathfrak{S}\langle \sharp(A_1, \ldots, A_n)\rangle}\,\sharp Df$$

where $\langle\,\rangle$ selects a side of the sequent, and for each $1 \leqslant i \leqslant m$, the we have chosen numbers j_i and k_i, and formulas $B^i_1, \ldots, B^i_{j_i}$ and $C^i_1, \ldots, C^i_{k_i}$ from the distinct subformulas $A_1, \ldots, A_n$ of $\sharp(A_1, \ldots, A_n)$. One of the left and right rules for $\sharp$ will be the top-to-bottom reading of this defining rule. So, if $\langle\,\rangle$ selects for the left, we have the left rule, and we must define right rules to match it. (There will be m of them, one for each premise of the defining rule.) If, on the other hand, $\langle\,\rangle$ selects for the right, we have the right rule, and we must define m corresponding left rules.

Here is how we do it. We apply the defining rule to the identity sequent for $\sharp(A_1, \ldots, A_n)$ in the following way:

$$\frac{\sharp(A_1, \ldots, A_n) \succ \sharp(A_1, \ldots, A_n)}{\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}, \sharp(A_1, \ldots, A_n)]}\,\sharp Df$$

The conclusion of this rule is the *empty* sequent, to which we have added the formulas B^i_1 to $B^i_{j_i}$ on the $\langle\,\rangle$ side, and added the formulas C^i_1 to $C^i_{k_i}$ and $\sharp(A_1, \ldots, A_n)$ on the *other* side. To convert this into an appropriate conclusion for a left or right rule, we must replace each of the formulas

$B^i_{j'}$ and $C^i_{k'}$ using *Cuts*. To replace the first such formula, B^i_1 (if present), we *Cut*, like this.

$$\dfrac{\mathfrak{S}_1[B^i_1] \quad \dfrac{\sharp(A_1,\ldots,A_n) \succ \sharp(A_1,\ldots,A_n)}{\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle \sharp Df}}{\mathfrak{S}_1\langle B^i_2,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle Cut}$$

To replace the first formula C^i_1 on the other side (if present), we *Cut*, like this.

$$\dfrac{\dfrac{\mathfrak{S}_1[B^i_1] \quad \dfrac{\sharp(A_1,\ldots,A_n) \succ \sharp(A_1,\ldots,A_n)}{\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle \sharp Df}}{\mathfrak{S}_1\langle B^i_2,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle Cut} \quad \mathfrak{S}'_1\langle C^i_1\rangle}{\mathfrak{S}_1\mathfrak{S}'_1\langle B^i_2,\ldots,B^i_{j_i}\rangle[C^i_2,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle Cut}$$

Doing the same to each $B^i_{j'}$ and $C^i_{k'}$, we have the following rule:

$$\dfrac{\mathfrak{S}_1[B^i_1]\cdots\mathfrak{S}_{j_i}[B^i_{j_i}] \quad \mathfrak{S}'_1\langle C^i_1\rangle\cdots\mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}[\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}$$

with $j_i + k_i$ premise sequents. It is the ith of m rules on the $[\,]$ side for $\sharp$, to pair with the single $\langle\,\rangle$ rule with m premises.

LEMMA 6.6 [LEFT AND RIGHT RULES EQUIVALENT TO THE DEFINING RULE] *The rules $\langle\sharp\rangle$ and $[\sharp]_i$ (for each $1 \leqslant i \leqslant m$) are interderivable with the defining rule $\sharp$Df, in the presence of Cut and Identity.*

Proof: We have shown one half of this result already, defining $\langle\sharp\rangle$ and $[\sharp]_i$ (for each $1 \leqslant i \leqslant m$) from the defining rule $\sharp$*Df*, using *Cut* and *Identity*. To complete the lemma, we need to do the reverse. We need to show that the each bottom-to-top reading of the defining rule can be recovered from the rules $[\sharp]_i$, using *Cut* and *Identity*. Choosing the ith instance of the rule, our target is to recover this inference

$$\dfrac{\mathfrak{S}\langle\sharp(A_1,\ldots,A_n)\rangle}{\mathfrak{S}\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i}]}$$

using $[\sharp]_i$, which has this form:

$$\dfrac{\mathfrak{S}_1[B^i_1]\cdots\mathfrak{S}_{j_i}[B^i_{j_i}] \quad \mathfrak{S}'_1\langle C^i_1\rangle\cdots\mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}[\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}$$

The structure of the derivation is simple, if you recall the case of the biconditional, from page 215. First, apply $[\sharp]_i$ to the identity sequents for each $B^i_{j'}$ and $C^i_{k'}$. as follows:

$$\dfrac{\langle B^i_1\rangle[B^i_1]\cdots\langle B^i_{j_i}\rangle[B^i_{j_i}] \quad \langle C^i_1\rangle[C^i_1]\cdots\langle C^i_{k_i}\rangle[C^i_{k_i}]}{\langle B^i_1,\cdots,B^i_{j_i}\rangle[C^i_1,\cdots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}$$

Then from this sequent, we can use *Cut*, to recover our defining rule from identity sequents and the $[\sharp]_i$ rule

$$\dfrac{\mathfrak{S}\langle\sharp(A_1,\ldots,A_n)\rangle \qquad \dfrac{\langle B^i_1\rangle[B^i_1]\cdots\langle B^i_{j_i}\rangle[B^i_{j_i}] \quad \langle C^i_1\rangle[C^i_1]\cdots\langle C^i_{k_i}\rangle[C^i_{k_i}]}{\langle B^i_1,\cdots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}{\scriptstyle [\sharp]_i}}{\mathfrak{S}\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i}]}{\scriptstyle Cut}$$

We have used a single *Cut* to trade in the formula $\sharp(A_1,\ldots,A_n)$ with the subformulas $B^i_1,\ldots,B^i_{j_i}$ and $C^i_1,\ldots,C^i_{k_i}$, inserted into the appropriate parts of the sequent. In this way, the left and right rules match the defining rule precisely. ∎

This argument is perfectly general, applying to any defining rule of whatever number of premises, for whatever n-ary connective you can extract from this scheme. We can pass freely between the invertible defining rule and a pair of left and right rules, whether *Contraction* and *Weakening* are present, or not.

» «

We have seen that once we trade in in our defining rules for left and right rules, we lose nothing when it comes to the bounds. However, the left and right rules have features not shared by defining rules. In particular, they are, in an important sense, *compositional* (when read from premises to conclusion) or *decompositional* (when read in reverse). As we know, this has important consequences for *Cut* and conservative extension, which we will discuss in the next section. It also has important consequences for *Identity*. It turns out that when we move to the left and right rules for a connective like $\sharp$, we no longer need to separately impose *Identity* for $\sharp$ formulas. If we add $\sharp$ to our language by way of its left and right rules (as motivated by the defining rule for $\sharp$) then *Identity* for $\sharp$ formulas is derivable from identity for the prior vocabulary. This can be proved in a way that naturally generalises Theorem 2.3 (page 53).

THEOREM 6.7 [LEFT AND RIGHT RULES GIVE IDENTITY] *The left and right rules $\langle\sharp\rangle$ and $[\sharp]_i$ for a connective $\sharp$ given by a defining rule are sufficient to derive the identity sequent $\sharp(A_1,\ldots,A_n) \succ \sharp(A_1,\ldots,A_n)$ from the identity sequents $A_i \succ A_i$.*

Proof: This derivation does the work:

$$\dfrac{\dfrac{\langle B^i_1\rangle[B^i_1]\ \cdots\ \langle B^i_{j_i}\rangle[B^i_{j_i}] \quad \langle C^i_1\rangle[C^i_1]\ \cdots\ \langle C^i_{k_i}\rangle[C^i_{k_i}]}{\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i},\sharp(A_1,\ldots,A_n)]}{\scriptstyle [\sharp]_i} \quad \text{(each } i\text{)}}{\sharp(A_1,\ldots,A_n) \succ \sharp(A_1,\ldots,A_n)}{\scriptstyle \langle\sharp\rangle}$$

Each premise sequent is an instance of $A_l \succ A_l$ for some l from 1 to n, since each B^i_j, and C^i_k, is chosen from A_1 to A_n. ∎

6.4 | ELIMINATING CUT

Now we will proceed to the second step of our argument, eliminating *Cut*. To make clear the scope of the setting, we need to pay a little more attention to the transition from our starting language $\mathcal{L}_1$ to our new language $\mathcal{L}_2$ with the addition of a newly definable connective $\sharp$, given by a defining rule. To be more general, we will no longer suppose that our starting language has no concepts which allow for the formation of complex formulas out of simpler formulas. We will allow this, with one constraint. These concepts allowing for the formulation of complex formulas will *themselves* be given by defining rules. So, the bounds in $\mathcal{L}_1$ are given by a set of *axiomatic* sequents, themselves closed under *Identity* and *Cut* (and perhaps *Contraction* and *Weakening*, according to taste), and we have already added some number (perhaps zero, perhaps more) of connectives by way of defining rules.

One can see ... main technical limitations in current proof-theory: The lack in *modularity*: in general, neighbouring problems can be attacked by neighbouring methods; but it is only exceptionally that one of the problems will be a corollary of the other ... Most of the time, a completely new proof will be necessary (but without any new idea). This renders work in the domain quite long and tedious. For instance, if we prove a cut-elimination theorem for a certain system of rules, and then consider a new system including just a new pair of rules, then it is necessary to make a complete new proof from the beginning. Of course 90% of the two proofs will be identical, but it is rather shocking not to have a reasonable «modular» approach to such a question: a main theorem, to which one could add various «modules» corresponding to various directions. Maybe this is inherent in the subject; one may hope that this only reflects the rather low level of our conceptualization!

— Jean-Yves Girard, *Proof Theory and Logical Complexity*, Vol. 1, pp. 16–17 [87]

The argument of this section will show that the addition of a new connective $\sharp$ by way of its defining rule $\sharp Df$, to form the language $\mathcal{L}_2$, is conservative, by showing how to eliminate *Cut* from derivations of the bounds in the new language, using the left and right rules for each of the connectives in $\mathcal{L}_2$. We show that any sequent derivable in $\mathcal{L}_2$, using the bounds of $\mathcal{L}_1$, *Identity*, *Cut*, and the left and right rules for the connectives in $\mathcal{L}_2$ can be derived without appeal to the *Cut* rule. Then, inspecting the rules which remain — the original bounds from $\mathcal{L}_1$, including *Identity*, the left and right rules (arising out of defining rules), and perhaps *Contraction* and *Weakening*— we see that all of the rules *introduce* concepts but never *eliminate* them, so the only *Cut*-free derivations of $\mathcal{L}_1$ sequents were already present in $\mathcal{L}_1$. It follows that the addition of $\sharp$ must be conservative over the original bounds. The new concept, $\sharp$, is a free addition to the language. Adding it does not close off any positions that were available in $\mathcal{L}_1$. It merely gives us *new* positions, that were not expressible before.

» «

The argument follows the general structure of the *Cut Elimination Theorem* from Chapter 2, specifically, Theorem 2.25 (see page 93). To show that the argument applies in the case of the bounds for $\mathcal{L}_2$, we need to show that the proof system we have defined for $\mathcal{L}_2$ satisfies the conditions laid out in Theorem 2.25. For this, we need to (1) provide an analysis of the rules in the proof system in such a way as to define *active* and *passive* formulas, and their *ancestry* in each of the rules (see Definition 2.16, page 84); (2) verify that the rules are closed under sequent substitution (see Lemma 2.20, page 86); (3) verify that our rules satisfy the sole active orphan condition (see Lemma 2.22, page 88); and so, (4) verify that all *Cut*s can be commuted up into orphan active *Cut*s (see Lemma 2.23, page 89), and finally (5) check that active orphan *Cut*s can be reduced into *Cut*s on the subformulas of the *Cut* formula (see Lemma 2.24, page 90). With those pieces in place, Theorem 2.25 applies, and *Cut*s can be eliminated from derivations.

These five verifications are straightforward. They take exactly the same shape as the versions we have already proved in Chapter 2, except we prove them in the generality of our broader setting. The breadth of our setting takes two forms. The more obvious form is that we allow for an arbitrary collection of formulas given by way of defining rules. The other, less obvious form is the breadth we allow for the bounds in our initial language $\mathcal{L}_1$. Here, we make no concession concerning the bounds for $\mathcal{L}_1$ except that they are closed under *Identity* and *Cut*, and that any connectives in $\mathcal{L}_1$ are given by defining rules of their own. $\mathcal{L}_1$ may, for example, contain the bounds for comparatives for a number of unary predicates. Or many other substantial claims governing the non-logical vocabulary, far beyond the *thin* assumptions of Chapter 2, where the only axioms governing the basic vocabulary were *Identity* axioms. Here, we may have much more. As I have mentioned, the only assumption we make concerning $\mathcal{L}_1$ bounds is that they are closed under *Identity* and *Cut*. In particular, we will single out the *basic* bounds for $\mathcal{L}_1$ to be the sequents of the form $X \succ Y$ where X and Y contain no connectives at all. These bounds themselves contain all sequents $A \succ A$ where A is a basic formula from $\mathcal{L}_1$, and they are also closed under *Cut*. The *basic* bounds will play the role in our *Cut* elimination argument in the same way that the *Identity* axioms played in Chapter 2.

(1) The *analysis* of the rules in our proof system proceeds in much the same way as in Chapter 2, in Definitions 2.15 and 2.16 (from page 83). We shall say that the formulas in a basic sequent $X \succ Y$ from the bounds of $\mathcal{L}_1$ (featuring only atomic formulas) are *active*, and not *passive*. The formulas occurring in each instance of *Weakening* or *Contraction* are passive, and not active (as in our previous analysis), and the formulas occurring in a *Cut* step are passive, except for the *Cut* formula, which is active. In the case of the left and right rules for any connective arising out of a defining rule,

$$\frac{\mathfrak{S}\langle B^i_1, \ldots, B^i_{j_i}\rangle[C^i_1, \ldots, C^i_{k_i}] \text{ (each } i)}{\mathfrak{S}\langle \sharp(A_1, \ldots, A_n)\rangle}\,\langle\sharp\rangle$$

$$\frac{\mathfrak{S}_1[B^i_1] \cdots \mathfrak{S}_{j_i}[B^i_{j_i}] \quad \mathfrak{S}'_1\langle C^i_1\rangle \cdots \mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}[\sharp(A_1, \ldots, A_n)]}\,[\sharp]_i$$

each formula occurrence in the sequent structures $\mathfrak{S}$, $\mathfrak{S}_i$ or $\mathfrak{S}'_i$ are passive, while each displayed instance of $B^i_{j'}$, $C^i_{k'}$, and the introduced formula $\sharp(A_1, \ldots, A_n)$ is *active*.

As for parenthood and ancestry, the analysis agrees completely with that of Definition 2.16. The parents of a formula occurring passively in a conclusion of a rule are the formulas occurring in the corresponding position in the structure falling under the same schematic letter in the presentation of the rule. So, for example, a formula occurring passively in the sequent $\mathfrak{S}$ in which the formula $\sharp(A_1, \ldots, A_n)$ is introduced in a $\langle\sharp\rangle$ inference has one parent in each premise of that inference, an occurrence of the same formula in the same position in each occurrence of $\mathfrak{S}$ as a premise. Ancestry for the structural rules is exactly as given in Definition 2.16. Since ancestry is not an issue in the basic axiomatic sequents from $\mathcal{L}_1$ (they are not *rules* but *axioms*), this is the only other generalisation needed.

(2) With this definition, we need to check that the rules are closed under substitution of a sequent for a formula occurring in passive position. (This is needed for when we commute an instance *Cut* above an occurrence of a rule.) In the case of $\mathcal{L}_2$ this means that we need to check that our left and right rules are closed under sequent substitution, in the sense that if a formula D occurs passively in the conclusion of one of these rules, then the result of substituting a new sequent $X' \succ Y'$ for D (for *that* instance of D in the conclusion, and all of its parents—if any—in the premises) in the rule gives us another instance of the rule.

Verifying this fact is immediate, because in the case of both rules $\langle\sharp\rangle$ and $[\sharp_i]$, all passive occurrences of formulas occur as formulas in a sequent variable $\mathfrak{S}$ (in the case of $\langle\sharp\rangle$) or some $\mathfrak{S}_j$ or $\mathfrak{S}'_j$ (in the case of $[\sharp]_i$). In either case, replacing the formula D by a whole sequent simply involves replacing the original sequent with another. In the case of $\langle\sharp\rangle$ this means replacing the sequent in all of the premises of the rule, as $\mathfrak{S}$ occurs in all premises, and in the case of $[\sharp]_i$ this means replacing the sequent in one of the premises, since each passive occurrence will be an occurrence in one of the sequents $\mathfrak{S}_j$ or $\mathfrak{S}'_j$. Regardless, the result is another instance of the rule, so the sequent substitution condition is verified.

(3) We need to show that active formulas are the sole orphan in any occurrence of a rule with premises. In the case of zero premise rules (like those for $\bot$ or $\top$), this does not apply, but in the case of rules $\langle\sharp\rangle$ (with one or more premises) and $[\sharp]_i$, the result is verified by inspection. In the case of these rules, provided that a premise is *present*, the only orphan in the conclusion is the introduced formula $\sharp(A_1, \dots, A_n)$, so this condition is verified.

(4) Given the analysis from (1) and the verification of (2) sequent substitution and (3) the sole orphan condition for active formulas in conclusions in rules, we can see that the proof of Lemma 2.23 (page 89) applies directly. We can commute a *Cut* on C occurring in a derivation upward, until the only remaining *Cut*s on C are those in which C is an orphan and active in both premises of the *Cut*.

(5) The only remaining component to verify is the following: that a *Cut* on active orphan formulas may be simplified into *Cut*s on subformulas, or eliminated entirely. In the case where the formula has a connective $\sharp$ dominant, we focus on the following *Cut* step.

$$\dfrac{\dfrac{\mathfrak{S}_1[B^i_1]\cdots\mathfrak{S}_{j_i}[B^i_{j_i}]\quad \mathfrak{S}'_1\langle C^i_1\rangle\cdots\mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}[\sharp(A_1,\dots,A_n)]}{\scriptstyle[\sharp]_i}\qquad \dfrac{\mathfrak{S}\langle B^{i'}_1,\dots,B^{i'}_{j'_i}\rangle[C^{i'}_1,\dots,C^{i'}_{k_i}]\ \text{(each } i'\text{)}}{\mathfrak{S}\langle\sharp(A_1,\dots,A_n)\rangle}{\scriptstyle\langle\sharp\rangle}}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}\mathfrak{S}}{\scriptstyle \mathit{Cut}}$$

Here, the *Cut* on $\sharp(A_1, \dots, A_n)$ can be traded in for $j_i + k_i$ cuts on the formulas B^i_1 to $B^i_{j_i}$ and C^i_1 to $C^i_{k_i}$, appealing to the single i premise of the $\langle\sharp\rangle$ inference which matches the choice of the $[\sharp]_i$ rule featured in the *Cut*.

$$\frac{\mathfrak{S}_1[B^i_1]\cdots\mathfrak{S}_{j_i}[B^i_{j_i}] \quad \mathfrak{S}'_1\langle C^i_1\rangle\cdots\mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle \quad \mathfrak{S}\langle B^i_1,\ldots,B^i_{j_i}\rangle[C^i_1,\ldots,C^i_{k_i}]}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}\mathfrak{S}}\,\textit{Cuts}$$

This process is simply undoing the *Cut* that we used to define $[\sharp]_i$ from the defining rule in the first place, given on page 225. The connective $\sharp$ is completely arbitrary, so this works for all pairs of left and right rules given by defining rules. That is, this argument applies to all connectives from our vocabulary $\mathcal{L}_2$.

The only remaining case to verify is that the *Cut* formula is atomic. That is, it is connective-free. In this case, our *Cut* step has the shape:

$$\frac{X \succ A, Y \quad X', A \succ Y'}{X, X' \succ Y, Y'}\,\textit{Cut}$$

and we have the further information that A is *active* and *orphan* in both premises of the *Cut*. The only rules in which a basic formula is active is an axiom, the basic axioms from $\mathcal{L}_1$. In all other rules, these formulas are passive. By hypothesis, the bounds from $\mathcal{L}_1$ were closed under *Cut*, so if $X \succ A, Y$ and $X', A \succ Y'$ are basic bounds from $\mathcal{L}_1$, so is $X, X' \succ Y, Y'$, so we can eliminate the appeal to the premises of this *Cut* step and appeal instead to the conclusion directly.

In checking conditions (1) to (5), we have completed all of the work we need to verify the following result.

THEOREM 6.8 [CUT ELIMINATION FOR DEFINING RULES] *Given a language $\mathcal{L}$ with a system of bounds closed under Identity and Cut, extended by a family of connectives $\sharp$ given by defining rules $\sharp Df$, then any derivation appealing to the bounds of $\mathcal{L}$, Cut, (with Weakening and Contraction, if present) and the left and right rules $\langle\sharp\rangle$ and $[\sharp]_i$ for each defined connective, can be systematically transformed into a derivation that does not appeal to the Cut rule.*

This Theorem gives us all we need to complete the second step of our argument. Whenever we have a language $\mathcal{L}_1$ in which the bounds are closed under *Identity* and *Cut*, and in which any connectives are already given by defining rules, then if we introduce *new* connectives by means of further defining rules, we can transform derivations of the bounds in $\mathcal{L}_2$ into a form in which the *Cut* rule is no longer used.

LEMMA 6.9 [SUBFORMULA PROPERTY] *If δ is a Cut-free derivation of a sequent $X \succ Y$ from basic bounds, using left and right rules $\langle\sharp\rangle$ and $[\sharp]_i$ for each defined connective $\sharp$, and perhaps the structural rules of Contraction and Weakening, then all formulas occurring in δ are subformulas of formulas in the conclusion sequent $X \succ Y$.*

Proof: This result is immediate, given the structure of the rules $\langle\sharp\rangle$ and $[\sharp]_i$. All formulas occurring in the premises of the rules must occur in the conclusion, either as whole formulas or as subformulas of the introduced formula $\sharp(A_1, \ldots, A_n)$. The same holds for the structural rules of *Contraction* and *Weakening*. (The basic bounds are axioms, and not rules, so they do not have the opportunity to introduce any material when passing from conclusion to premises.) ∎

Now we have all of the pieces required for our Hauptsatz.

theorem 6.10 [defining rules are conservative] *For any language $\mathcal{L}_1$ whose bounds are closed under Identity and Cut, and whose connectives (if any) are governed by defining rules, the extension to a new language $\mathcal{L}_2$ by the addition of new connectives* also *governed by defining rules, is conservative.*

Proof: Consider the language $\mathcal{L}_2$, and suppose that $X \succ Y$ is a sequent from the vocabulary of $\mathcal{L}_1$ that is derivable from the rules for $\mathcal{L}_2$. By Lemma 6.6, $X \succ Y$ can be derived using the left and right rules, rather than the defining rules. By Theorem 6.8, we can eliminate all instances of *Cut* from the resulting derivation. By Lemma 6.9, any such derivation only involves subformulas of the formulas in $X \succ Y$, that is, formulas from $\mathcal{L}_1$. So, this derivation cannot appeal to any of the rules for new concepts added in $\mathcal{L}_2$, since these introduce concepts not in $\mathcal{L}_1$. So, this sequent $X \succ Y$ was already out of bounds in $\mathcal{L}_1$, and hence, the extension of $\mathcal{L}_1$ to $\mathcal{L}_2$ is conservative. ∎

Excursus: We mentioned when introducing defining rules that the reasoning of these sections also applies in the case of single conclusion sequents, but that care must be taken to ensure that we stay within the bounds of sequents with no more than one conclusion. In this excursus, we will take a look at exactly what this involves. There are four points at which we must be careful. (1) Understanding the restriction in scope for defining rules. (2) Showing that all connectives definable by single conclusion defining rules generate single conclusion left/right rules, and verifying that the argument to the effect that the left/right rules and defining rules are equal in power carries over unchanged. (3) Understanding the restriction to sequent substitution in this setting, and verifying that sequent substitution suffices to allow for commuting *Cut*s past passive occurrences of formulas in rules. (4) verifying that *Cut*s on active orphan formulas may still be simplified into *Cut*s on subformulas. With these conditions checked, the argument to *Cut* elimination will remain unchanged, except for the restriction of sequents to single conclusion.

If you do not care about restricting rules to single conclusion sequents, then skip this excursus. We return to our regular programming on page 234.

(1) In a single conclusion setting, if $\langle\,\rangle$ selects for the left hand side, then expression $\mathfrak{S}\langle A\rangle$ defines a single conclusion sequent, no matter what $\mathfrak{S}$ may be. But if $\langle\,\rangle$ selects for the right, $\mathfrak{S}\langle A\rangle$ is a sequent only when $\mathfrak{S}$ has an empty right hand side. This restricts the kinds of defining rules that may be constructed in a single conclusion setting. If $\langle\,\rangle$ selects for the left, then these defining rules unproblematically work:

$$\frac{}{\mathfrak{S}\langle\bot\rangle}\ \bot Df \qquad \frac{\mathfrak{S}}{\mathfrak{S}\langle t\rangle}\ t Df \qquad \frac{\mathfrak{S}\langle A_1, A_2\rangle}{\mathfrak{S}\langle A_1 \otimes A_2\rangle}\ \otimes Df \qquad \frac{\mathfrak{S}\langle A_1\rangle \quad \mathfrak{S}\langle A_2\rangle}{\mathfrak{S}\langle A_1 \vee A_2\rangle}\ \vee Df$$

While these have premises that are not well-formed if $\mathfrak{S}$ has a non-empty right hand side.

$$\frac{\mathfrak{S}[A_1]}{\mathfrak{S}\langle\neg A_1\rangle}\,\neg Df \qquad \frac{\mathfrak{S}\langle A_1\rangle[A_2]}{\mathfrak{S}\langle A_1 - A_2\rangle}\,{-Df} \qquad \frac{\mathfrak{S}\langle A_1\rangle[A_2] \quad \mathfrak{S}\langle A_2\rangle[A_1]}{\mathfrak{S}\langle A_1 \bowtie A_2\rangle}\,{\bowtie}Df$$

When $\langle\ \rangle$ selects for the *right* hand side of a sequent, these rules are unproblematic:

$$\frac{}{\mathfrak{S}\langle\top\rangle}\,\top Df \qquad \frac{\mathfrak{S}}{\mathfrak{S}\langle f\rangle}\,f Df \qquad \frac{\mathfrak{S}[A_1]}{\mathfrak{S}\langle\neg A_1\rangle}\,\neg Df \qquad \frac{\mathfrak{S}[A_1]\langle A_2\rangle}{\mathfrak{S}\langle A_1 \to A_2\rangle}\,{\to}Df$$

$$\frac{\mathfrak{S}[A_1]\langle A_2\rangle \quad \mathfrak{S}[A_2]\langle A_1\rangle}{\mathfrak{S}\langle A_1 \leftrightarrow A_2\rangle}\,{\leftrightarrow}Df \qquad \frac{\mathfrak{S}\langle A_1\rangle \quad \mathfrak{S}\langle A_2\rangle}{\mathfrak{S}\langle A_1 \wedge A_2\rangle}\,{\wedge}Df$$

While this rule generates a premise sequent with two formulas on the right hand side, and so, is out of bounds.

$$\frac{\mathfrak{S}\langle A_1, A_2\rangle}{\mathfrak{S}\langle A_1 \oplus A_2\rangle}\,{\oplus}Df$$

So, of our familiar connectives, subtraction ($-$) and symmetric difference ($\bowtie$) and multiplicative disjunction ($\oplus$) cannot be defined in this setting, and we have only one defining rule for negation (where $\neg A$ is introduced on the *right*, not the left). The other familiar defining rules remain. The general restrictions can be understood in this way. This is a defining rule

$$\frac{\mathfrak{S}\langle B^i_1, \dots, B^i_{j_i}\rangle[C^i_1, \dots, C^i_{k_i}] \text{ (for each } 1 \leqslant i \leqslant m)}{\mathfrak{S}\langle\sharp(A_1, \dots, A_n)\rangle}\,\sharp Df$$

when each $B^i_{j'}$ and $C^i_{k'}$ is a selection from $A_1, \dots, A_n$, and (a) if $\langle\ \rangle$ selects for the *left*, then $k_i = 0$ for each i (that is, there are no $C^i_{k'}$ formulas inserted into the right hand side of the sequent in any premise), and (b) if $\langle\ \rangle$ selects for the *right*, then $j_i \leq 1$ for each i (that is, there is at most one $B^i_{j'}$ formula inserted into the right hand side of the sequent in any premise, since that position has been vacated by $\sharp(A_1, \dots, A_n)$). You can see that condition (a) is violated by the defining rule for subtraction (it attempts to insert A_2 from $A_1 - A_2$ on the right, while that position may be occupied already), while condition (b) is violated by the defining rule for $\oplus$ (it attempts to insert two formulas on the right, the disjuncts of $A_1 \oplus A_2$).

(2) Once we have our restrictions in place in the defining rule for an arbitrary connective $\sharp$, we must check to see what corresponding single conclusion left/right rules $[\sharp]_i$ are generated. In the unrestricted setting, the defining rule for $\sharp$ (given above) generates the rules

$$\frac{\mathfrak{S}_1[B^i_1] \cdots \mathfrak{S}_{j_i}[B^i_{j_i}] \quad \mathfrak{S}'_1\langle C^i_1\rangle \cdots \mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}_{1-j_i}\mathfrak{S}'_{1-k_i}[\sharp(A_1, \dots, A_n)]}\,[\sharp]_i$$

for each i. In the case where $\langle\,\rangle$ selects for the *left*, our restriction tells us that there are no formulas $C^i_{k'}$ present, so the rule $[\sharp]_i$ has the form

$$\frac{\mathfrak{S}_1[B^i_1] \;\cdots\; \mathfrak{S}_{j_i}[B^i_{j_i}]}{\mathfrak{S}_{1-j_i}[\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}$$

which is indeed well formed: all substitutions occur on the right, and each sequent has a single formula on the right. In the case where $\langle\,\rangle$ selects for the *right*, our restriction tells us that in each premise i there is at most one formula $B^i_{j'}$. This means that our rule $[\sharp]_i$ has the form

$$\frac{\mathfrak{S}[B^i_1] \quad \mathfrak{S}'_1\langle C^i_1\rangle \;\cdots\; \mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}\mathfrak{S}'_{1-k_i}[\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}$$

which is also (perhaps surprisingly!) well formed. Since $\langle\,\rangle$ selects for the right, the sequents $\mathfrak{S}'_{k'}$ have empty right hand sides (into which $C^i_{k'}$ is inserted), and only $\mathfrak{S}$ may have a formula on the right. It is this formula (the RHS of $\mathfrak{S}$) that is the only formula present (if any) in the RHS of the concluding sequent $\mathfrak{S}\mathfrak{S}'_{1-k_i}[\sharp(A_1,\ldots,A_n)]$ of the rule.

Furthermore, the argument (Lemma 6.6) showing that from the rules $\langle\sharp\rangle$ and $[\sharp]_i$ we can recover the defining rule remain within the bounds of single conclusion sequents given that the restriction applies to the defining rule itself, as you can easily check. (The argument starts by applying the rule $[\sharp]_i$ to identity sequents, and then uses *Cut* to recover the generality of the defining rule.)

(3) Sequent substitution works as before in our setting, provided that the conditions on the sequents used in the *Cut* are themselves single conclusion. The details are unchanged from Lemma 2.19 (page 86) in Chapter 2, so we can declare this condition verified.

(4) *Cut*s on orphan active formulas can be eliminated in just the same way as in the multiple conclusion setting, too, provided that the single conclusion restrictions are satisfied by the left and right rules. In the case where $\langle\,\rangle$ selects for the left, an orphan active *Cut* has the shape:

$$\frac{\dfrac{\mathfrak{S}\langle B^i_1,\ldots,B^i_{j_i}\rangle \text{ (each } i)}{\mathfrak{S}\langle\sharp(A_1,\ldots,A_n)\rangle}\,{\scriptstyle\langle\sharp\rangle} \qquad \dfrac{\mathfrak{S}_1[B^i_1] \;\cdots\; \mathfrak{S}_{j_i}[B^i_{j_i}]}{\mathfrak{S}_{1-j_i}[\sharp(A_1,\ldots,A_n)]}\,{\scriptstyle [\sharp]_i}}{\mathfrak{S}\mathfrak{S}_{1-j_i}}\,{\scriptstyle Cut}$$

where $\mathfrak{S}$ is the only sequent frame possibly with a non-empty RHS. (The $B^i_{j'}$s and $\sharp(A_1,\ldots,A_n)$ occupy the RHS of the other sequents.) This *Cut* can be traded in by *Cut*s on the formulas $B^i_{j'}$ as follows:

$$\frac{\mathfrak{S}\langle B^i_1,\ldots,B^i_{j_i}\rangle \quad \mathfrak{S}_1[B^i_1] \;\cdots\; \mathfrak{S}_{j_i}[B^i_{j_i}]}{\mathfrak{S}\mathfrak{S}_{1-j_i}}\,{\scriptstyle Cuts}$$

where we *Cut* once with each sequent $\mathfrak{S}_{j'}$, and each intermediate step is a single conclusion sequent, whose RHS is the RHS of $\mathfrak{S}$.

When $\langle\,\rangle$ selects for the right, in each premise i of the rule $\langle\sharp\rangle$ there is at most one formula $B^i_{j'}$. A *Cut* then looks like this:

$$\cfrac{\cfrac{\mathfrak{S}\langle B^i_1\rangle[C^i_1,\ldots,C^i_{k_i}]\ (\text{each } i)}{\mathfrak{S}\langle\sharp(A_1,\ldots,A_n)\rangle}\langle\sharp\rangle \qquad \cfrac{\mathfrak{S}'[B^i_1]\quad \mathfrak{S}'_1\langle C^i_1\rangle\ \cdots\ \mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}'\mathfrak{S}'_{1-k_i}[\sharp(A_1,\ldots,A_n)]}[\sharp]_i}{\mathfrak{S}\mathfrak{S}'\mathfrak{S}'_{1-k_i}}\mathit{Cut}$$

Here, $\mathfrak{S}'$ is the only sequent frame which could have a non-empty RHS, since $\mathfrak{S}$ and each $\mathfrak{S}'_{k'}$ have formulas substituted into their RHS. Again, applying the *Cut*s to the subformulas B^i_1 and the $C^i_{k'}$s, we have

$$\cfrac{\mathfrak{S}\langle B^i_1\rangle[C^i_1,\ldots,C^i_{k_i}]\quad \mathfrak{S}'[B^i_1]\quad \mathfrak{S}'_1\langle C^i_1\rangle\ \cdots\ \mathfrak{S}'_{k_i}\langle C^i_{k_i}\rangle}{\mathfrak{S}\mathfrak{S}'\mathfrak{S}'_{1-k_i}}\mathit{Cuts}$$

which is a single conclusion derivation, as desired. So, the procedure for eliminating *Cut*s manages to stay within the bounds of single conclusion sequents if we apply it to rules which respect those constraints.

These conditions verified, we see that the argument works in the single conclusion setting, though the relative inflexibility of single conclusion rules means that certain connectives (like $\oplus$) can no longer be defined. *End of Excursus*

» «

This is where we return to our regular programming.

We have completed the third stage of the argument for conservative extension. Now we have all the tools we need to answer Prior's question. Before we do that, it will be worth examining the scope of this result. How far can it go? What are its limits?

The first thing to note is that the argument as given tells us nothing about quantifiers, or other putatively logical notions which may be given rules which appeal not to bounds and judgements themselves, but to something involving the inner structure of those judgements, such as the division into predicates and singular terms. The central features of the analysis can be extended in this direction — for the details of that extension, see Chapter 8 — the omission is merely temporary. It would have made this central argument even *more* abstract were we to attempt to generalise to quantifiers and other notions. The second thing to note is that we have said nothing, either, about modal operators, and concepts which may be defined by appealing to more generous notions of a position, generalisations of the two-sided sequent structure that has been our focus until now. Again, this omission is temporary. To see what is involved in meeting this challenge, wait for Chapter 9.

If we restrict our focus to languages with *connectives*, a question remains about the restriction that connectives in our initial language $\mathcal{L}_1$ are themselves given by defining rules. This is quite a restriction. To see what is involved, let's consider a putative case of non-conservative extension where the extending connective is given by way of a defining rule.

Here is the case: Suppose we have a language $\mathcal{L}_1$ with a stock of atomic formulas and an *intuitionistic* conditional $\to$. So, we at have $X \succ$

A in the bounds for $\mathcal{L}$ whenever there is a derivation of the sequent $X \succ A$ in the *single conclusion* sequent calculus, with *Identity*, *Cut*, *Weakening* and *Contraction*, and where $\rightarrow$ is governed by the rules $\rightarrow L$ and $\rightarrow R$ when restricted to sequents with single formulas on the right. We can think of the bounds as applying to pairs of sets of formulas as before, but $[X : Y]$ is out of bounds in $\mathcal{L}_1$ only when there is a single conclusion derivation for $X \rhd Y$, that is a derivation for $X \rhd A$ where A is a formula in Y. The bounds for $\mathcal{L}_1$ are closed under *Identity*, *Cut*, *Contraction* and *Weakening*, but it is not difficult to show that $\rightarrow$, so defined, is *not* given by the defining rule $\rightarrow Df$:

$$\frac{X, A \succ B, Y}{X \succ A \rightarrow B, Y}\ {\rightarrow}Df$$

It is well known that there is is no *single conclusion* derivation for Peirce's law $(((p \rightarrow q) \rightarrow p) \rightarrow p)$, but this can be given by the defining rule in a multiple conclusion setting, in the presence of *Contraction* and *Weakening*.

Exercise 7.1 on page 43 asks you to show that Peirce's Law has no normal natural deduction proof. A similar argument shows that it has no *Cut*-free single conclusion derivation.

$$\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{p \succ q, p}\ K}{\succ p \rightarrow q, p}\ {\rightarrow}Df \qquad \dfrac{(p \rightarrow q) \rightarrow p \succ (p \rightarrow q) \rightarrow p}{(p \rightarrow q) \rightarrow p, p \rightarrow q \succ p}\ {\rightarrow}Df}{\dfrac{(p \rightarrow q) \rightarrow p \succ p, p}{(p \rightarrow q) \rightarrow p \succ p}\ W}\ Cut}{\succ ((p \rightarrow q) \rightarrow p) \rightarrow p}\ {\rightarrow}Df$$

So, $\mathcal{L}_1$, as given, does not feature Peirce's Law. Suppose we add negation to $\mathcal{L}_1$, governed by its defining rule:

$$\frac{X, A \succ Y}{X \succ \neg A, Y}\ \neg Df$$

where we impose *no* restriction on Y, and so we allow for all the classical derivations concerning negation. In particular, it will justify the usual negation on the left rule in this way:

$$\dfrac{X \succ A, Y \qquad \dfrac{\neg A \succ \neg A}{A, \neg A \succ}\ \neg Df}{X, \neg A \succ Y}\ Cut$$

and Double Negation Elimination, like this:

$$\dfrac{\dfrac{A \succ A}{\succ A, \neg A}\ \neg Df \qquad \dfrac{X \succ \neg\neg A, Y}{X, \neg A \succ Y}\ \neg Df}{X \succ A, Y}\ Cut$$

With $\neg L$ and *DNE* in hand, we can repurpose the derivation of Peirce's Law using Double Negation Elimination, seen on page 71 in Chapter 2.

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \succ p}{\neg p, p \succ}\,\neg L}{\neg p, p \succ q}\,KR}{\neg p \succ p \to q}\,{\to}R \qquad p \succ p}{(p \to q) \to p, \neg p \succ p}\,{\to}L}{(p \to q) \to p, \neg p, \neg p \succ}\,\neg L}{(p \to q) \to p, \neg p \succ}\,W}{(p \to q) \to p \succ \neg\neg p}\,\neg R}{(p \to q) \to p \succ p}\,DNE}{\succ ((p \to q) \to p) \to p}\,{\to}R
$$

Notice that *this* derivation uses the conditional rules only with single conclusion sequents, so we have not used rules that were not present in $\mathcal{L}_1$ to govern the conditional. Yet now, using the added rules for negation (in particular, $\neg L$ and *DNE*, justified by the defining rule for negation), we have a non-conservative extension of $\mathcal{L}_1$. What gives?

The immediate response is that the conditional, as given by $\to L$ and $\to R$ in a single conclusion setting, is not given by a defining rule in the setting of multiple conclusion positions. If we restrict our sequents to be single conclusion, then a defining rule *will* generate the intuitionist conditional, but then we do not have the resources to generate our negation, which obeys *DNE*, by means of a defining rule that complies with the single conclusion constraints. As Nuel Belnap understood, the antecedently given context of deducibility plays a significant role.

That response explains a structural distinction between these two approaches to extending our initial language $\mathcal{L}_1$. If we take $\mathcal{L}_1$ to be specified by a multiple premise multiple conclusion family of bounds, then its conditional is not given by a defining rule, and it is not conservatively extended by the negation given by $\neg Df$. If we take $\mathcal{L}_1$ to be specified, instead, by a defining rule in which the sequents are given in multiple premise and single conclusion form, then the conditional *is* given by a defining rule, and can be conservatively extended with any other connective definable in those terms. But that leaves out the negation given by $\neg Df$ in its multiple conclusion generality. The restricted defining rule for a negation in that form does not justify *DNE*, and furthermore, it does no violence to the intuitionist conditional.

Before thinking we are left with no guidance as to whether to choose one option (single conclusion defining rules) over another (multiple conclusion defining rules), it is important to remember that we are given more to work with than these formal systems *qua* formal systems. If our linguistic practice involves (strong) assertion and (strong) denial, then we have the resources to define concepts with multiple premise and multiple conclusion sequents. If positions consisting of assertions and denials satisfy the constraints of *Contraction* and *Weakening*, (and I argued in Section 5.4 that they *do*), then we have all the resources of classical multiple premise and multiple conclusion sequents at our disposal. Restricting ourselves to single conclusion reasoning is interesting, useful and important for certain purposes, but we have the resources to avail ourselves of *more*, and to do so in a way that respects the connection between formal rules and our concrete practices.

You can be a *pluralist*, who finds space for more than one different logical consequence relation, as I am [13].

6.5 | ANSWERING PRIOR'S QUESTION

Now, let's return to Prior's challenge, which we discussed in Chapter 4. How can an inference rule confer meaning? Which inference rules *define* concepts? How can we distinguish inference rules like those for conjunction or negation (or the quantifiers, and identity) that define logical concepts, on the one hand, from those for *tonk*, which do not?

Given the tools at our disposal from this chapter, the answer is straightforward. *Defining rules* define logical concepts. Defining rules introduce concepts in ways that are both *conservative* (Theorem 6.10) and *uniquely defining* (Theorem 6.5), given an appropriate background context of deducibility. Gentzen's left and right rules for the logical connectives — or corresponding natural deduction introduction and elimination rules — can be seen to arise out of defining rules, in the presence of *Identity* and *Cut*. Prior's rules for *tonk* do not arise out of any defining rules. In the rest of this section I will elaborate and defend these claims, connect them to our earlier discussion concerning assertion and denial, and examine what these connections can tell us about what is distinctive about logical concepts.

» «

Suppose we have a language in which claims are made. We make assertions and denials, and we keep track of the bounds for positions. Perhaps we recognise that certain combinations of assertions and denials are out of bounds, beyond the basic constraint that asserting p together with and denying p is out of bounds. However, let's suppose that our language is, as yet, relatively *thin*. We do not, in this language, have any parts of speech (or we do not express any concepts) which allow for unbridled combination of assertions (or denials) to make new assertions (or denials). Just one example of this would be a language with unary predicates and comparatives. We can assert and deny claims like Ft or $t >_F s$, but there is, as of yet, no way to combine these (or other) claims into new claims. Nonetheless, this language may well be quite rich and logically textured. Certain combinations of assertions and denials are out of bounds, while others remain on the field of play. (The borderline between what is in bounds and what is out of bounds characterises linear orders and an extension upwardly closed under that order.) What are the costs incurred and the benefits to be gained in extending this language by way of a *connective* by way of inference rules?

Given that the bounds for our language satisfy the structural rules of *Identity*, *Cut*, *Weakening* and *Contraction*, were we to add to our language any connectives by way of multiple premise and multiple conclusion defining rules, then we would be extending our practice in a quite powerful way. We would not only be adding to our language new contents, which could be asserted and denied, but we would also be constraining further extensions of the language. Given a connective like the conditional, not only are we able to form conditionals $A \to B$ consisting of judgements from our earlier vocabulary, and not only are we able to form conditionals whose constituents contain judgements from our *new* vocabulary (conditionals with other conditionals as antecedents and consequents), we are also constraining further extensions of the language to the extent that if we add further *newer* vocabulary, the rules for conditionals constrain *those* formulas too. Connectives, in this sense, are open ended and can apply to any content, whether expressible now, or later. So, to add a connective to a vocabulary is to incur a debt to be paid off in all future extensions of that vocabulary. Such a debt may not always be easy to pay off.

So, should we add a conditional (for example) to our basic vocabulary? If we add one using the defining rule $\to Df$, then that extension would be conservative (by Theorem 6.10) and uniquely defining (by Theorem 6.5). That can reassure us to some extent, but notice that if we were to define a conditional using the weaker, single conclusion rules for the conditional, the result would *still* be conservative and uniquely defining. Do Belnap's criteria give us any reason to prefer one defining rule over the other? Here, we can consider what happens when we project into future language extension. Theorem 6.10 reassures us that the multiple conclusion rule $\to Df$ is not only conservative over our original vocabulary, but is also conservatively extended by any other multiple conclusion defining rule. The same cannot be said for the single conclusion defining rule. (This is only conservatively extended by other single conclusion rules.) So, if we have some principled reason to constrain our bounds to single conclusion bounds, there is some reason to prefer single conclusion rules. But if our bounds are more general than that — and in the case of comparatives we have good reason to pay attention to combinations of denials, such as in the *subcontrary* condition (which says that denying both $s >_F t$ and $t \geqslant_F s$ is out of bounds) — then extending our language by a weaker defining rule may prove insufficient to future language extension. This is the power of defining rules. They not only provide conservative extension (and unique definability) but they work well together. Given a sequent setting (single conclusion or multiple conclusion, contraction and weakening present, or not) then all defining rules conservatively extend each other, when added in any order.

So, there are three senses in which adding a connective by way of a defining rule is free. First, it is unconstrained by any details of what is present in the underlying vocabulary. It is free of any constraint concerning *content* or *form* of the judgements in the initial vocabulary. Provided that we have bounds satisfying *Identity* and *Cut*, then adding a concept by way of a defining rule is conservative and uniquely defining, no matter *what* the initial vocabulary contains. The connectives operate freely on any kinds of judgements, provided that the bounds are respected. They do not look for judgements of a particular form.

This level of absolute freedom is not quite met by *quantifiers*, which will need us to peek inside our judgements to the level of finding a particular class of terms and generalising into their position. The detail for how we can apply defining rules *for quantifiers*, and what this means for logicality and quantification is our topic in Chapter 8.

Second, adding connectives by way of defining rules is a pure extension of our prior practice because any position available before that connective was added remains available afterward. If a rule failed to be conservative in *that* sense, then someone who was occupying the position which would be ruled out were the connective added would have good reason to resist the addition of the connective! However, it is also a pure extension of our practice in a *third* sense. The defining rule not only imposes no new constraints on the bounds that we had not already made, but also in imposes no further constraints on extending our practice by way of *further* defining rules. It is pure addition, leaving the field as open for other defining rules of its kind as it was previously. It imposes no new constraints on positions that were not already present. In this sense, any concept given by a defining rule is cost free.

Furthermore, the addition of a concept by way of a defining rule is a *constrained* extension to the bounds because that extension is uniquely defined, up to indistinguishability. A defining rule is not merely a description of a *kind* of extension of the language and the bounds, which could be realised in more than one way. The extension to the bounds is fully defined to the extent that any two extensions with the one defining rule differ nowhere, with regard to the bounds. They are indistinguishable, as far as the bounds are concerned. Adding a concept by way of a defining rule gives us a way to decompose one act (assertion or denial) involving this new concept in terms of acts involving its components (for example, denying a conditional has the same impact on the bounds as asserting the antecedent and denying the consequent, no more, no less), but the dual act (in this case, asserting a conditional) is *new*, and does not consist of any combination of previous acts. The norms governing this act, however, are specified enough so as to constrain the concept up to indistinguishability.

This is how adding concepts to our language by way of a defining rules satisfies Belnap's constraints of conservative extension and unique definability in a very strong sense. In this way, defining rules allow us to define logical concepts.

» «

What, then, can we say about Prior's rules for *tonk*? At the face of it, they are not defining rules, but is there any way to massage them *into* defining rules? There is no such way, at least in defining rules for multiple premise, multiple conclusion sequents (or their restrictions to single conclusions). Any such defining rule conservatively extends a system of bounds satisfying *Identity* and *Cut*, into a larger system of bounds still satisfying *Identity* and *Cut*. Provided that *some* non-trivial position is in bounds (say, for our language for comparatives, the position $[Fs : Ft]$ for two different terms s and t), then there is no extension of this language with a defining rule for a concept *tonk* that results in a language that renders the *tonk* rules valid. This is because the new language is still closed under *Cut*, and were the *tonk* rules satisfied, the following two sequents would be validated:

$$\mathsf{Fs} \succ \mathsf{Fs}\ \textit{tonk}\ \mathsf{Ft} \qquad \mathsf{Fs}\ \textit{tonk}\ \mathsf{Ft} \succ \mathsf{Ft}$$

and by *Cut*, we could conclude $\mathsf{Fs} \succ \mathsf{Ft}$, which would show the addition of *tonk* to be non-conservative, since $[\mathsf{Fs} : \mathsf{Ft}]$ was available, before the addition of *tonk*. Given that the addition of any concept given by means of a defining rule *is* conservative, no defining rule gives us *tonk*.

The same could be said for other, creative attempts to define concepts that satisfy different proposals for harmony between introduction and elimination rules, but still produce deviant connectives. One such creative proposal is Stephen Read's *bullet*. It is a propositional constant, introduced by the following (strange, to be sure) introduction rule in a natural deduction context [188, p. 571]:

$$\dfrac{\begin{matrix}[\bullet] \\ \vdots \\ \bot\end{matrix}}{\bullet}\ {\bullet}I$$

which tells us that one way to introduce $\bullet$ is to derive a contradiction from the assumption of $\bullet$. In a sequent setting (whether multiple conclusion or single conclusion) the corresponding rule is:

$$\dfrac{X, \bullet \succ Y}{X \succ \bullet, Y}\ {\bullet}R$$

where in the case of a single conclusion system, we require Y to be empty. Read's bullet is, to be sure, a very strange beast. If the constraint on rules is *harmony* (generating the elimination rule *from* the introduction rule, or the *left* rule from the *right*), then Read shows that harmony can be satisfied. The corresponding elimination rule is also rather strange. It is

$$\dfrac{\bullet \quad \bullet}{\bot}\ {\bullet}E$$

and in the sequent system it has the form

$$\dfrac{X \succ \bullet, Y}{\bullet, X \succ Y}\ {\bullet}L$$

which is the converse of the right rule! (This is very close to being a defining rule.) Read shows that we can unwrap an introduction rule followed by an elimination rule in the usual sense, by making this transformation:

$$\dfrac{\begin{matrix}\vdots\ \pi_1 \\ \bullet\end{matrix} \qquad \dfrac{\begin{matrix}[\bullet] \\ \vdots\ \pi_2 \\ \bot\end{matrix}}{\bullet}\ {\bullet}I}{\bot}\ {\bullet}E \quad \rightsquigarrow \quad \begin{matrix}\vdots\ \pi_1 \\ \bullet \\ \vdots\ \pi_2 \\ \bot\end{matrix}$$

in which the introduction and elimination rules are deleted, but the proofs π_1 and π_2 remain. The same works in the sequent setting. A *Cut* on $\bullet$ introduced on the right and the left can be eliminated like this:

$$\dfrac{\dfrac{\begin{matrix}\delta_1\\ \vdots\\ X, \bullet \succ Y\end{matrix}}{X \succ \bullet, Y}{\scriptstyle \bullet R} \qquad \dfrac{\begin{matrix}\delta_2\\ \vdots\\ X' \succ \bullet, Y'\end{matrix}}{X', \bullet \succ Y'}{\scriptstyle \bullet L}}{X, X' \succ Y, Y'}{\scriptstyle Cut} \quad\rightsquigarrow\quad \dfrac{\begin{matrix}\delta_2\\ \vdots\\ X' \succ \bullet, Y'\end{matrix} \qquad \begin{matrix}\delta_1\\ \vdots\\ X, \bullet \succ Y\end{matrix}}{X, X' \succ Y, Y'}{\scriptstyle Cut}$$

Of course, there is something strange in both cases. In the elimination of the *Cut*, we trade in a *Cut* on a $\bullet$ in terms of a *Cut* on *another* $\bullet$. The same feature holds in the processing of the $\bullet I/\bullet E$ transition. The interface between the two proofs π_1 and π_2 is *another bullet*. In the case where the bounds contain *Contraction* (or the natural deduction system allows duplicate discharge) things can go to hell in a handbasket rather quickly.

$$\dfrac{\dfrac{\dfrac{\bullet \succ \bullet}{\succ \bullet, \bullet}{\scriptstyle \bullet R}}{\succ \bullet}{\scriptstyle W} \qquad \dfrac{\dfrac{\bullet \succ \bullet}{\bullet, \bullet \succ}{\scriptstyle \bullet L}}{\bullet \succ}{\scriptstyle W}}{\succ}{\scriptstyle Cut} \qquad\qquad \dfrac{\dfrac{\dfrac{[\bullet]^1 \quad [\bullet]^1}{\bot}{\scriptstyle \bullet E}}{\bullet}{\scriptstyle \bullet I^1} \qquad \dfrac{\dfrac{[\bullet]^2 \quad [\bullet]^2}{\bot}{\scriptstyle \bullet E}}{\bullet}{\scriptstyle \bullet I^2}}{\bot}{\scriptstyle \bullet E}$$

This derivation (on the left) is not *Cut* free. If you try eliminating the *Cut* in the derivation, you'll see that the *Cut* proliferates, and the result (eventually) is another derivation of exactly the same shape. The *Cut* cannot be eliminated. The proof (on the right) is not normal. If you try rewriting the $\bullet I/\bullet E$ transition in the proof, you will see that the non-normality proliferates, and the result (eventually) is another proof of exactly the same shape. The proof cannot be normalised.

In the context of *Contraction* (or of duplicate discharge) the addition of $\bullet$ using these rules need not be conservative. If the bounds did not already contain the empty sequent (which would be totally trivialising if we also had *Weakening*), then the addition of $\bullet$ is not conservative. It is a violent revision of the bounds, given *Identity* and *Cut*.

However, in the absence of *Contraction*, the addition of $\bullet$ *can* be conservative! In the case of *Cut* elimination, if *Contraction* is absent, it is straightforward to show that each step in eliminating *Cut* strictly reduces the size of the derivation. There is no need to appeal to a well-founded subformula relation, and no need to require that the *Cut* on a bullet reduce to a *Cut* on a smaller formula. In the absence of *Contraction*, the resulting derivation is always *shorter*, and so, the process of eliminating *Cut* will terminate, and we end up with a *Cut* free derivation. The same holds for normalisation. Without duplicate discharge, the process of normalisation terminates in a normal proof. With $\bullet$, the inversion principle (and Read's *General-Elimination Harmony*) comes apart from conservative extension in the presence of *Contraction*, while in its absence, bullet manages to be conservative, even though it is very strange.

How strange? Here's how: $\bullet$ is equivalent to $\neg\bullet$.

How does $\bullet$ look from the perspective of defining rules? Its sequent rules look very much like a defining rule, when superimposed.

$$\frac{X, \bullet \succ Y}{X \succ \bullet, Y}\ \bullet Df?$$

However, when we look at using such a rule to *add* $\bullet$ to a vocabulary in which it is not present, we have been given no advice as to what to do, at least when couched in terms of the bounds of the prior vocabulary. We are told that denying $\bullet$ is out of bounds (in some context) if and only if asserting it is out of bounds (in that context). Of course, that is *strange* (and in the presence of *Contraction*, this clashes with *Cut*, to be sure), but the problem, from the point of view of defining rules, is not just that is clashes with the combination of *Contraction* and *Cut*. The more significant problem is that it is too *unconstrained*. We are not given the bounds for either asserting or denying $\bullet$ in terms of the prior bounds. We are just told to extend the bounds in such a way as to satisfy $\bullet$*Df?*. If we don't have *Contraction*, the problem is that there are *too many* options for interpreting $\bullet$. We could have two strange propositions

$$\frac{X, \bullet \succ Y}{X \succ \bullet, Y}\ \bullet Df? \qquad \frac{X, \circ \succ Y}{X \succ \circ, Y}\ \circ Df?$$

both of which can freely move from left to right in a sequent, and nothing will force $\bullet$ to match $\circ$. All we end up with is *two* items of information with this strange property.

Excursus: Here is a quick and dirty verification of this fact. Define the bounds on a language in which $\bullet$ and $\circ$, and other atoms are the only formulas. A sequent $X \succ Y$ is out of bounds if and only if each formula in that sequent appears an even number of times — and it contains at least two formulas. So, for example, $p \succ p$ and $p, p \succ$ and $\circ, \bullet, \bullet \succ \circ$ are out of bounds, but $p \succ \bullet, p$ and $\circ, \bullet \succ \circ$ are not. Clearly *Identity* sequents $A \succ A$ are all out of bounds, and *Cut* is satisfied because if the premises of a *Cut* satisfy this condition for being out of bounds, so does its conclusion. (If the number of occurrences of each atom in $X \succ C, Y$ and $X', C \succ Y'$ are both even, this holds, too, for $X, X' \succ Y, Y'$, whatever C is.) *Contraction* fails, of course, as it erases the distinction between even and odd numbers of occurrences, and *Weakening* also fails for the same reason. In this setting, $\circ$ and $\bullet$ are distinguishable, since $\bullet \succ \bullet$ is out of bounds, but $\bullet \succ \circ$ is not. *End of Excursus*

The putative defining rule for $\bullet$ is not a defining rule at all, it is a description of a strange property that a proposition might have, not something that determines it up to indistinguishability. Read's rules, though allowing for reductions of *Cut* (and their elimination, in *Contraction*-free systems), do not meet the standards of *defining* rules, and nor should they. These rules merely describe a possible concept, they do not define one.

The moral of this story is simple: Freely mixing and matching introduction and elimination rules — or left and right rules — is at most asking for a concept that satisfies those conditions. The result of choosing such a pair of rules does not, in and of itself, *define* a concept for us. To be sure that we have done that, we should use a defining rule.

» «

We have said a great deal of how it is that defining rules manage to define concepts, and how other putative definitions, which do not meet the high standards of defining rules, can fail to do so. We have said less about what this distinction has to do with what is usually called *logic*. Is it a mere coincidence that the notions we can define by way of defining rules bear a suspicious resemblance to the traditional propositional connectives? To answer *that* question, we will consider Ian Hacking's analysis of what should be involved in answering Prior's challenge [95, p. 304]:

> Why should rules for "pure logic" have the subformula property and be conservative? There are several kinds of answers. They correspond to criteria of adequacy like the following. (A) The demarcation should give the "right" logicist class of logical constants and theorems. That is, it should include the traditional (and consistent) core of what logicists said was logic and should exclude what they denied to be logic. (B) Since the demarcation is couched in terms of how logical constants are characterized, it should provide the semantics for the constants called "logical." (C) It should explicate why logic is important to the analytic program. Although (A) and (B) are important, (C) is essential. A demarcation of logic that leaves the analytic program unintelligible is of little philosophical interest (unless the point is to show that the analytic program *is* unintelligible).

For Hacking, the *analytic program* involves showing how (for example) arithmetic can be obtained from logic and definitions of the arithmetic concepts.

To appropriately address Hacking's concerns, we need to (A) show that defining rules can define the concepts we have ordinarily taken to be logical — or to explain the divergence, if there is any; (B) explain how it is that defining rules can give us *semantics* for the concepts we define; and most importantly, (C) give some explanation of the *use* of the concepts given by defining rules in the analytic program.

For (A), we have already seen that defining rules give us the traditional logical constants, and if we have *Contraction* and *Weakening* and multiple conclusion sequents (as I have argued that the norms governing assertion and denial warrant), then these connectives behave classically. We have a justification of classical propositional logic in terms of the bounds for positions. We have not given a justification of quantifiers or identity — other concepts taken to be central to the logicist program, foundational at the start of the 20th Century and the development of modern logic. For that, we will need to wait for Chapter 8, but once we have shown how defining rules can give an account of quantifiers and identity, we will have addressed Hacking's first concern.

For (B), we need to give some kind of *semantics* for expressions given by defining rules. This is especially pressing if rules are mere formalism without any interpretation. But defining rules used to extend the bounds of positions are no such things. They give us norms for *asserting* or *denying* new judgements for our vocabulary. We have expanded our vocabulary with new sentences, and we have already given them an interpretation. However, it may be that you prefer *semantics* to be understood in terms of traditional forms of interpretation, or *models*, in which sentences are assigned *truth conditions*. In the case of bounds satisfying *Contraction* and *Weakening*, an answer is at hand. We have shown in Chapter 3 that limit positions give rise to traditional boolean evaluations which validate the standard truth conditions for the connectives. In this way, the propositional logical constants are given a semantics, and this semantics is justified in terms of the defining rules which introduce those constants. In this way, the account of logical constants given in terms of defining rules answers concern (B), at least for the propositional connectives. (See Chapter 8 to see how the answer can be extended beyond, to quantifiers and identity.)

As Hacking points out, issue (C) is the most interesting and challenging, and it goes some way further to explain a distinctive point for our having logical vocabulary. We will leave for Chapter 8 some of the details of how *arithmetic* may be treated in our context. For the connection between logic and the broader project of *analysis* and *definition*, however, some comments can be made. To see the connection, consider what it is for us to *introduce*, or to *explicate* a concept by means of an explicit definition. For example, when explaining what I mean by *square* by saying something like "a *square* is a *rectangle* with *equal sides*." To be more precise, I could say:

$$x \text{ is a } \textit{square} =_{\text{df}} x \text{ is a } \textit{rectangle} \land x \text{ has } \textit{equal sides}.$$

In other words I either show how the concept *square* could be expressed in my original vocabulary, using a compound expression, using concepts I already possess, and using the concept of additive conjunction to do the combining. Or, in case I didn't already possess the concept of additive conjunction, I can introduce *that* concept by way of a defining rule, and then use it to combine the other two concepts I already have. The traditional logical notions allow us to freely *combine* concepts we already possess to form complex concepts using simple concepts as constituents. Such combinations are not possible unless we have a mode of combination. There is no vocabulary for the formation of complex concepts, or the analysis of concepts into others in a language in which there are no complex combining expressions. Defining rules allow us to introduce ways of combining which allow for just this sort of composition of new concepts out of old ones, in ways which underwrite the behaviour of explicit definition.

One criterion for a good explicating definition is that the definition allows for no room to move: if we agree on the definition, we can use this as a fixed point for marshalling disagreement. If you and I disagree on whether something is *square* — and we agree on the analysis of *square* given above — then we should be able to shift this disagreement onto the question of whether it is a *rectangle* and whether it has *equal sides*. If we cannot do that, then the concept *square* is floating freely, independent of the component from which it is defined. If conjunction, used here, is given by its defining rule, we can see that this condition is satisfied. The behaviour of conjunction is fixed, and disagreements concerning $A \wedge B$ are impossible without some kind of difference concerning A and concerning B. Concepts given by defining rules provide fixed points which ensure that we can extend our conceptual repertoire by way of combining old concepts into new ones.

There is more to be said concerning the behaviour of definition and of analysis and the role of logical vocabulary in articulating meanings, but these comments should be enough to indicate the direction that the story will go in the remaining chapters. For now, we have the broad strokes of our answer to Prior's challenge. We have seen that defining rules can indeed safely and securely introduce concepts into a vocabulary, that not any just *any* set of inference rules satisfy these constraints, and we have some idea of the distinctive and constructive roles that such concepts can play. In the remaining chapters, we will spell out the consequences of that account for issues in the philosophy of logic, of semantics, epistemology and metaphysics, and also fill in more details concerning defining rules and quantifiers, and modal operators.

6.6 | HISTORY

The special features of invertible proof rules for the connectives have been discussed in the literature for nearly as long as the sequent calculus has been with us. The importance of invertible rules was (as far as I know) first discussed by Karl Popper in 1948 [168], and they were further extensively discussed by Dana Scott in a number of papers in the 1970s [216, 217].

Ian Hacking's 1979 paper "What is Logic?" [95], cited above, marks out the question concerning sequent rules for logical constants as defining, and does not isolate *invertible* rules as providing a key to the answer to the question. However, he does discuss the centrality of *Cut* elimination, and the derivability (eliminability) of *Identity* sequents in the new vocabulary as an important aspect of any criterion for logicality. Alberto Naibo and Mattia Petrolo's recent discussion of Hacking's connection between the derivability of *Identity*, and Belnap's criterion of uniqueness is a helpful characterisation of some of the issues here [152].

Kosta Došen, in the 1980s, proposed that the defining rules (in our sense, in a multiple conclusion sequent setting) gave a criterion for logicality [54, 55] which adequately answered Prior's concerns about inference rules. For Došen, logical constants reflect into the language the underlying "punctuation" of the sequent system. The present analysis extends Došen's in a number of ways, in particular, generalising it to different systems of structural rules, and to the single conclusion sequent setting, as well as showing systematically how left/right rules are equal in strength to defining rules, and how *Cut* can uniformly be eliminated from the system featuring left/right rules. In addition, the interpretation of the result in terms of positions and norms governing assertion and denial is original here. Došen's pioneering work is an important precursor to the approach I take here.

Giovanni Sambin's work with his colleagues on *Basic Logic* [10, 66, 212, 213] in the late 1990s and beyond, is also closely related to this approach to defining rules, in that invertible rules play a central role. In Basic Logic, connectives are introduced by invertible rules but there, the rules have a slightly less general shape than defining rules. Connectives are introduced by invertible rules in which the formula is either the entire LHS or the entire RHS. A multiplicative conjunction is defined by this invertible rule:

$$\frac{A, B \succ Y}{\overline{\overline{A \otimes B \succ Y}}}$$

and Sambin and colleagues show [213] that in the presence of *Identity* and *Cut*(in their setting, a *single formula Cut*, where we proceed either from $X \succ A$ and $X', A \succ Y'$ to $X, X' \succ Y'$, or from $X \succ A, Y$ and $A \succ Y'$ to $X \succ Y, Y'$) that this gives rise to a corresponding *right* rule:

$$\frac{X \succ A \quad X' \succ B}{X, X' \succ A \otimes B}$$

The setting is distinctive, and Sambin's work with his colleagues is the first that I'm aware of where invertible rules are used to generate pairs of left and right rules. I have generalised that technique to the broader setting of arbitrary sequents to recover more sequent calculi, and to better fit with the motivation in terms of norms governing *positions*. (For a defining rule to tell us norms governing the assertion of $A \otimes B$, for example, we would like to know how to understand this in the presence of any other assertions and denials, so applying the rule to a sequent of the form $X, A \otimes B \succ Y$ seems more appropriate, rather than restricting the rule to those instances where $A \otimes B$ is the sole formula on the left.)

The *generality* of the current approach has precursors in Nuel Belnap's Display Logic [16, 18], and in my treatment of Cut Elimination in substructural logics [193, Chapter 6], though it is further generalised here to arbitrary defining rules.

PART III

Insights

MEANING AND PROOF 7

Our topic in this chapter is the *philosophy of propositional logic*, in its many different guises. Having equipped ourselves with tools from proof theory in the first part of the book, and having navigated the central argument concerning defining rules in the second part, we will use the chapters that remain to elaborate the consequences of all of this for the philosophy of logic, the philosophy of language, for epistemology and for metaphysics. In this chapter, our focus is propositional logic, since we will address the questions that arise concerning logic, proof and meaning *as such*, without focussing our attention to specifics of singular terms, objects, and quantifiers, or of modal operators like necessity and possibility. (Those issues are left for remaining chapters.) The language of propositional logic, though it is by no means all there is to say about *logic* is rich and subtle enough to raise many different philosophical questions, and they are our topic for this chapter.

We will start the chapter by looking at the semantics of the propositional constants themselves—at issues concerning conjunction, disjunction, the material conditional and, especially, negation. Once we have addressed those questions, in the remaining sections we ascend to a higher vantage point, and address issues of logical consequence, proof and meaning.

7.1 | CONNECTIVES

What do defining rules for the connectives have to do with our everyday use of the words, "*and*", "*or*", "if" and "*not*"? On the surface of things, there is no immediate connection. The overwhelming majority of users of these words — or of the concepts expressed by those words — have never encountered a defining rule. You do not gain access to logical vocabulary by first learning the sequent calculus and committing defining rules to memory. You encounter logical vocabulary in the same way that you encounter *any* vocabulary, in the midst of learning to listen and speak, to read and think, to argue and persuade. The vocabulary of propositional connectives arises alongside many other components of our conceptual apparatus; we gain competence with logical vocabulary together as we acquire many other cognitive and discursive abilities. These practices bear some kind of relationship to the formal rules of a calculus, but that connection is not constituted by a learner explicitly or consciously following those formal rules in order to enter into the practice.

There is a great deal of empirical research on the acquisition of logical concepts [2, 43, 233].

Perhaps an analogy can help us think about the relationship between our everyday concepts and faculties and the formal rules that have been the object of study of this book. Consider the connection between our learning to count and our mastery of everyday numerical concepts on the one hand and the formal apparatus of axioms of arithmetic on the other. What is the relationship between learning to count, to add and to multiply — skills mastered by many in their early years — and the basic axioms of Peano Arithmetic? Here is one presentation of the axioms and rule of Peano Arithmetic:

SUCCESSOR AXIOMS:

PA1: $s(n) = s(m) \succ n = m$ PA2: $0 = s(n) \succ$

ADDITION AXIOMS:

PA3: $\succ n + 0 = n$ PA4: $\succ n + s(m) = s(n + m)$

MULTIPLICATION AXIOMS:

PA5: $\succ n \times 0 = 0$ PA6: $\succ n \times s(m) = (n \times m) + n$

INDUCTION RULE:

$$\text{PA7:} \quad \frac{X \succ F0, Y \quad X, Fn \succ Fs(n), Y}{X \succ Fn, Y} \quad (n \text{ not present in } X, Y, F)$$

These axioms and rules are presented in terms of bounds, and we have a language with a two-place identity predicate, arbitrary singular terms n and m for natural numbers, and a constant term 0 (for the number *zero*), and function symbols s (for successor, which picks out, for each number, the *next* one), $+$ and $\times$ for addition and multiplication. The first six axioms PA1 to PA6 give the behaviour of successor, addition and multiplication. PA1 and PA2 govern successor and zero, ensuring that we have an unending sequence of numbers, starting at zero. PA3 and PA4 fix the interpretation of addition for each number starting from zero, by recursion, while PA5 and PA6 do the same for multiplication. The *rule* PA7 is a rule for *induction*. It says that if $F0$ is undeniable in the position $[X : Y]$, and it is also out of bounds to assert Fn and deny $Fs(n)$ in $[X : Y]$ (where that position makes no commitment concerning n, so n is arbitrary), then Fn itself is undeniable in $[X : Y]$ — ensuring that if the first number, zero, is F, and if from a number being F it follows that the next one is F, then all the numbers are F. This is an elegant, simple characterisation of basic principles governing natural numbers.

Well, it is *relatively* simple. The axioms hide some complexity in the behaviour of the *identity* predicate. For details of the logic of the identity predicate, you can skip ahead to Chapter 8, but those details are not important for the point we need to make here.

No-one — *I hope* — learns arithmetic by *first* learning these axioms and rules. We learn terms for numerals by learning to count small collections of things; we learn to add by counting different collections of things, separately and then together (we count two oranges and three apples as five pieces of fruit and say that $2 + 3$ is 5); we learn to multiply by learning to count collections (two lots of three apples is six apples; 2×3 is 6), and so on. Different learners are trained in different ways, but the general shape and structure of arithmetic competence is identifiable across a wide range of different practices. Many of us have learned to count with the scaffolding of Arabic numerals. Others learn with a tally system. Machines keep track of numbers in their own very different ways. Almost all of us have learned to count in ways that are very different from any explicit application of Peano Arithmetic. Nonetheless,

The details of *how* competence in number vocabulary is gained is an active research question. See the papers in Sorin Bangu's 2018 collection *Naturalizing Logico-Mathematical Knowledge* for a good introduction to the state of the field [7].

those axioms bear an important relationship to the specific practices by which we learned to count, and the underlying architecture of our numerical concepts. In particular, we can see the formalisation arithmetic as one way of distinguishing a *kind* of practice. Given that we can learn arithmetic vocabulary in a range of different ways, one way to check for agreement on our vocabulary is to check that we agree on PA1 to PA7. If I discover that you are using vocabulary in such a way as to allow for zero to be the successor of some number — and when this is revealed to you, you don't correct yourself — then we can conclude that you are diverging from the *natural number practice* somehow, either by diverging on *zero*, or *successor* or *natural number*. It is through checking with agreement on these basic bounds that we can discern that we are engaging in different practices. (Your "*number*" might better match my "*integer*", where positive and negative numbers are allowed, for example.) You might not learn addition in such a way that you appeal to the recursive conditions PA3 and PA4 as *fundamental*, but insofar as you and I arrive at a place where we both grant PA1 to PA7, we have a great deal of agreement in our practices — enough to provide a way of translating between your vocabulary and mine, at least as far as arithmetic goes. Here, a formalisation provides one way of understanding how some bounds are established, and different concrete practices, established in different ways, can comply with the bounds as presented in some formalisation.

What goes for arithmetic can go for the connectives, too. You need not first learn "and" by learning this defining rule:

$$\frac{X, A, B \succ Y}{X, A \otimes B \succ Y} \otimes Df$$

You don't learn conjunction by first learning the defining rule for additive conjunction, either.

$$\frac{X \succ A, Y \quad X \succ B, Y}{X \succ A \wedge B, Y} \wedge Df$$

If the bounds are closed under *Weakening* and *Contraction*, additive and multiplicative conjunction are indistinguishable. I focus on multiplicative conjunction here, because its defining rule has a single premise, and is therefore more straightforward to apply.

but if you use "AND" in such a way as to treat the assertion of A *and* B as out of bounds if and only if the assertion A and the assertion of B are jointly out of bounds, then the formal system's "$\otimes$" is a good match for your "*and*." Again, not all uses of the word "*and*" match this rule, in just the same way that not all uses of "number" match the use of numerical vocabulary in Peano arithmetic. This rule can be used in order to see whether some vocabulary is treated in such a way as to match *these* norms. In the case of logical vocabulary, it is all the more obvious that we do not acquire the vocabulary by consciously reading and understanding the defining rule as explicit instructions to be followed. The instructions themselves *use* a concept like conjunction. They say that the assertion of the conjunction $A \otimes B$ is out of bounds iff the assertion of A *and* the assertion of B is out of bounds. To apply this, I need to be able to check whether A is asserted *and* B is asserted. It is hard to follow this instruction without understanding what the "*and*" in the instruction involves.

So, instead of treating the defining rule as a manual for the learner to follow, it can be understood as either a checklist for the *interpreter* to understand how to classify a practice, or for a *teacher* to explain what needs to be imparted in order for a student to learn a concept. So, how could we decide whether the defining rule for $\otimes$ is a good fit for the use of our everyday concept "*and*"? By asking questions such as these, about the connection between assertion and denial and conjunction — about their relationship to the bounds. What is the point of coordinating on these? It means we've fixed things up to indistinguishability. We can be sure that disagreement about conjunctions, then, can be transformed into disagreement elsewhere.

Defining rules, then, are not so much explicit instructions to follow to gain a concept for the first time. They are tools for showing what a practice should feature in order for something to count as a concept of a particular kind. If our practice satisfies the constraints of $\otimes Df$, then we have, indeed, defined a *conjunction*. If we already find ourselves with something very much like the concept of conjunction, then we can think of the defining rule $\otimes Df$ as an *explication* of a particular, precisely delimited, concept of conjunction. It is a way of isolating, within an informal and loosely structured practice, a particular way of forming assertions and denials with a well defined and precisely articulated shape. For example, explicating our conjunction practice by way of the defining rule $\otimes Df$ is a way to make clear that *this* sense of conjunction is one for which $A \otimes B$ is indistinguishable from $B \otimes A$ as far as the bounds are concerned — so any sense in which we can (strongly) assert A *and* B but (strongly) deny B *and* A is a sense for "*and*" which diverges from what we have explicated with $\otimes$, for any position in which we have asserted A and asserted B is a position in which we have asserted B and asserted A (that is, positions, as we have defined them, have no record of *order*).

A. .W. Carus' *Carnap and Twentieth-Century Thought* is a fruitful discussion and defence of Carnap's project of explication [36]. Carnap's most explicit and sustained discussion of explication is in the first chapter 'On Explication' in his *Logical Foundations of Probability* [34].

That is *strong* assertion and denial here, and not their weak counterparts. Recall: it is permissible to weakly deny C and then strongly assert D, indistingushable from C (temporarily removing C rom the common ground, only to enter D into the common ground instead), which has the effect of placing C back into the common ground by courtesy, as a consequence of D. (Consider the case where Eloise weakly denies "I am Abelard's wife" then to assert "Abelard is my husband.") It is unproblematic to deny $A \otimes B$ weakly and then to assert $B \otimes A$. When pressed on whether $A \otimes B$, you could explicitly answer "yes, but I wouldn't put it *that way*." Taking $A \otimes B$ to be indistinguishable from $B \otimes A$ simply means that they are the same as far as the bounds are concerned. No more need follow than that.

Let's move from considering conjunction to *disjunction*. If we focus on the following sense of disjunction, as explicated by the defining rule $\oplus Df$

$$\frac{X \succ A, B, Y}{X \succ A \oplus B, Y}\ \oplus Df$$

we see that denying a disjunction $A \oplus B$ has the same effect on the bounds as denying both A and B. This is exactly the kind of constraint we point to when we distinguish *inclusive* and *exclusive* notions of disjunction. Here, denying $A \oplus B$ has exactly the same effect on the bounds as denying A and denying B. In particular, asserting A and asserting B and denying $A \oplus B$ is out of bounds because asserting both A and B and denying both A and B are (jointly) out of bounds — in two different ways. So, any exclusive sense of "*or*" — for which asserting A *or* B but involves commitment to A and B not *both* holding — diverges from the notion of disjunction expressed by the rule $\oplus Df$.

In the rest of this section we will consider the defining rules $\to Df$ and $\neg Df$, and we will revisit some of the more substantial logical and philosophical disagreements which have been litigated over conditionals and negation in the light of these defining rules.

THE (MATERIAL) CONDITIONAL

Recall the defining rule $\to Df$:

$$\frac{X, A \succ B, Y}{X \succ A \to B, Y}\ \to Df$$

What does this defining rule say about our everyday uses of conditional expressions? Not a great deal, at first glance. The rule $\to Df$ fixes on a notion of conditionality for which denying $A \to B$ has the same upshot as asserting A and denying B. This is *somewhat* related to everyday notions of conditionality (denying a conditional has *something* to do with granting the antecedent and denying the consequent), but the match is nothing like a perfect fit. I deny the conditional

If I have a cup of coffee from House of Cards, I feel regret.

House of Cards is a coffee vendor and social enterprise at the Parkville campus of the University of Melbourne.

but I do *not* (now) assert the antecedent (*I have a cup of coffee from House of Cards*) because as I am writing this, I am at home, and I don't have a coffee from *House of Cards* — I have some coffee I made myself. Nonetheless, I deny the conditional. If I am to be consistent, then denying a conditional like the one I have denied here must involve something other than flatly asserting the antecedent and denying the consequent. In some sense, it seems to involve *hypothetically* committing oneself to the antecedent, which involves changing the context in such a way as to incorporate the antecedent (*remembering* when I last had coffee from *House of Cards*, *considering* future circumstances when I might, and the like). As far as $\to Df$ is concerned, denying a conditional (in the sense of $A \to B$) just involves asserting A and denying B. Coordinating on the sense of the *material* conditional involves explicating *"if"* by means of this defining rule.

Given this understanding of the concept given by $\to Df$, arguments about the so-called "paradoxes of implication", like

$$q \succ p \to q \qquad \neg p \succ p \to q \qquad \succ p \to q, q \to r$$

can be sidestepped, relatively quickly. Yes, to be sure, there may be *other* notions of conditionality where a conditional does not follow from its consequent, or from the negation of its antecedent. Yes, it would be surprising to think that for some *strong* notion of implication, that for any three statements, p, q and r, either the first implies the second or the second implies the third. However, the salient notion of conditionality given by $\to Df$ delivers these so-called 'paradoxes' immediately. After all, $A \to B$ is indistinguishable from the disjunction $\neg A \vee B$, and these so-called 'paradoxes' are much less paradoxical when rewritten in this form:

$A \to B$ is indistinguishable from $\neg A \vee B$ in the presence of *Contraction* and *Weakening*, and multiple conclusions. In all of our logics with multiple conclusion sequents $A \to B$ is indistinguishable from the multiplicative disjunction $\neg A \oplus B$.

$$q \succ \neg p \vee q \qquad \neg p \succ \neg p \vee q \qquad \succ \neg p \vee q, \neg q \vee r$$

The paradoxes of implication are problematic only for an everyday notion of conditionality, which is not what is delivered by the material conditional. What remains surprising, of course, is that something quite unlike a conditional ($\neg A \vee B$, or $\neg A \oplus B$) can nonetheless do some of the work *of* a conditional. That is underwritten by its defining rule. If we read $X, A \succ B, Y$ as telling us that in the presence of $[X : Y]$, when A is given (added to the assertion side), B is *undeniable*, then it follows that in the presence of $[X : Y]$, the conditional $A \to B$ is also undeniable. In other words, to answer the question of whether or not $A \to B$ holds, we can (temporarily) grant A, and ask the question concerning B. That is, of course, a re-presentation of the conditional introduction rule *→I*, familiar from natural deduction. We can prove a conditional by assuming the antecedent and deriving the consequent. For *everyday* reasoning it seems that assuming the antecedent involves more than adding it naïvely to our collection of assertions. Sometimes we (temporarily) update the common ground with the antecedent in such a way as to (temporarily) remove some things we have granted which *clash* with the antecedent. In this way, we could resist the inference to $\neg p \succ p \Rightarrow q$ (for a conditional '$\Rightarrow$' that hews more closely to our everyday practice), because when we assume p we may remove ¬p from the stock of information at hand. Or, perhaps we consider an *alternative possibility* where p holds. The virtue of the material conditional and the simple-minded way of treating antecedents — by piling them onto the stack of assertions — is that it does remarkably well when it comes to the behaviour of conditionals in mathematical reasoning. The formalisation of conditionals in natural mathematical reasoning by way of the standard classical (or intuitionistic) reasoning arising out of *→Df* (whether in the presence of single conclusions or multiple conclusions) is a sign that the material conditional, though bad at expressing modally robust connections, remains one useful explication of our notion of conditionality. While we still have reason to consider other ways to extend our notion of a position in order to explicate our notions of conditionality and alternative possibilities, the material conditional gives us something to work with in our standard sequent setting.

One way to understand "considering an alternative possibility" in our sequent setting will be a central theme in Chapter 9.

» «

It is worth mentioning, too, that the structural rule *Weakening* also plays a role in the paradoxes of implication. To derive the paradoxes, we use *Weakening*, in each case:

$$\dfrac{\dfrac{p \succ p}{q, p \succ p}\,K}{q \succ p \to q}\,{\to}Df \qquad \dfrac{\dfrac{\dfrac{p \succ p}{p \succ p, q}\,K}{\neg p, p \succ q}\,\neg Df}{\neg p \succ p \to q}\,{\to}Df \qquad \dfrac{\dfrac{\dfrac{p \succ p}{p, q \succ q, r}\,K}{p \succ q, q \to r}\,{\to}Df}{\succ p \to q, q \to r}\,{\to}Df$$

The notion of positions and bounds defended in Chapter 5 directly motivates *Weakening*. For example, since there is a clash involved in asserting p and denying p, there is a clash involved in asserting p and in denying p and denying q. Since asserting p and denying p is out of bounds, so is asserting p and q and denying q and r. You do not avoid your mistake by piling on more commitments. Nonetheless, there may be ways of understanding bounds in such a way as to undercut *Weakening*. For example, we may decide to keep track of *basic* bounds: $[p, q : q, r]$ may be out of bounds but not *basically* so, and if we were to keep track not only of what is out of bounds, but what is basic to those violations, then, a rule like *→Df* could be maintained without giving rise to the paradoxes of implication. However, if your focus is on accounts of positions which satisfy

Weakening (as mine is, here) the paradoxes are an inevitable upshot of the concept of the conditional given by $\rightarrow Df$.

» «

There is yet one more issue concerning the logic of conditionals that is worth considering. Vann McGee showed, in 1985, that there are some plausible counterexamples to the rule *modus ponens* ($\rightarrow E$), using the natural language notion of conditionality [139]. Consider the following example, due to McGee:

> I see what looks like a large fish writing in a fisherman's net a ways off. I believe
>
>> If that animal is a fish, then if it has lungs, it's a lungfish.
>
> That, after all, is what one means by "lungfish." Yet, even though I believe the antecedent of the conditional, I do not conclude
>
>> If that creature has lungs, it's a lungfish.
>
> Lungfishes are rare, oddly shaped, and, to my knowledge, appear only in fresh water. It is more likely that, even though it does not look like one, the animal in the net is a porpoise. [139, p. 462–463]

If we formalise the claims in the obvious way (F is that animal is a $\underline{\mathsf{F}}$ish; H is that animal $\underline{\mathsf{H}}$as lungs; L is that animal is a $\underline{\mathsf{L}}$ungfish; and $\Rightarrow$ is the conditional in use), then it looks like McGee wants to take up the following position:

$$[\mathsf{F} \Rightarrow (\mathsf{H} \Rightarrow \mathsf{L}), \mathsf{F} : \mathsf{H} \Rightarrow \mathsf{L}]$$

by *asserting* $\mathsf{F} \Rightarrow (\mathsf{H} \Rightarrow \mathsf{L})$ (that is a part of what it means to be a lungfish, let us grant), and F (the animal certainly looks like a fish), but *denying* $\mathsf{H} \Rightarrow \mathsf{L}$. The grounds for denying $\mathsf{H} \Rightarrow \mathsf{L}$ are stated at the end of the passage I quoted, and they seem to be very good reason for denying the conditional. Knowing what we *know*, and adding to this the hypothesis that the creature has lungs, it is much more likely that the creature is a porpoise, and *not* a lungfish. It seems that we not only have no reason to accept $\mathsf{H} \Rightarrow \mathsf{L}$, but we have grounds to deny it.

If this is an available position, we have a counterexample here to *modus ponens* for $\Rightarrow$. This position would not be available, were the conditional governed by a defining rule with the same shape as $\rightarrow Df$, or *any* conditional for which all instances of the sequent

$$\mathsf{A} \Rightarrow \mathsf{B}, \mathsf{A} \succ \mathsf{B}$$

are derivable, for our position has that form, in the specific case where the consequent B is itself another conditional. If the position McGee stakes out is available, then we are certain that the '*if*' in use is not the material conditional, and furthermore, it is a notion of '*if*' for which *modus ponens* indeed fails: it is (on occasion) an available position, to assert a conditional, and its antecedent, and to deny its consequent. This behaviour is, perhaps, troubling, because *modus ponens* is naturally understood to be a central feature of conditionals.

I won't tarry here long to explain what kind of understanding of conditionals might vindicate the judgements McGee wants to make concerning the conditionals expressed in this particular case, but we would do well to understand at least one formulation of the conditional which will deliver McGee's verdict, to reassure us that the rejection of *modus ponens* can be well motivated in a systematic way. To see how the sense of conditionality functions here, start with the reasons McGee gives for denying $H \Rightarrow L$: the *likelihood* that L holds, given the addition of the assumption H, is low. One way to regiment this thought is to take this vocabulary literally, and think of the threshold for assertion or denial in terms of probability, and to evaluate a conditional, we will use *conditional probability*. The probability of L *given* H (that is, $Pr(L \mid H)$, which can be formally defined as $Pr(L \wedge H)/Pr(H)$) is low. We can grant that H itself has a high probability. How can we define the conditional probability for the conditional $F \Rightarrow (H \Rightarrow L)$? Literally speaking, the conditional probability $Pr(B \mid A)$ cannot be identified with the probability of some conditional $A \Rightarrow B$ [131], so we cannot calculate $Pr((H \Rightarrow L) \wedge F)/Pr(F)$. However, if we *can* calculate a conditional probability corresponding to the conditional $F \wedge H \Rightarrow L$. For that, we calculate the conditional probability $Pr(L \mid F \wedge H)$, which is $Pr(L \wedge (F \wedge H))/Pr(F \wedge H)$. This conditional probability can be high — in fact, it can be 1. To make things specific, if $Pr(F) = 0.99$, $Pr(H) = 0.01$, and if we set L to be $H \wedge F$, with the probability 0.0001, then the probabilities for our setup are:

$$\begin{array}{rrcl}
F \Rightarrow (L \Rightarrow H): & Pr(L \wedge (F \wedge H))/Pr(F \wedge H) & = & 1 \\
F: & Pr(F) & = & 0.99 \\
L \Rightarrow H: & Pr(L \wedge H)/Pr(H) & = & 0.01
\end{array}$$

and we see that the premises of our *modus ponens* have very high probability (1 and 0.99), and the conclusion probability is very low (0.01). If we use those reasonable thresholds for assertion and denial, we can see the sense in which asserting the premises and denying the conclusion is appropriate.

McGee's counterexample is no argument against our analysis of connectives given by defining rules. We have already granted that there may be senses of '*if*' that diverge from the sense given by $\to$*Df*. If the conditional appearing in McGee's counterexample can be given a coherent interpretation, then this will be such a sense. We can see that the conditional of McGee's counterexample is not the conditional given by $\to$*Df*, because the reasoning of the counterexample does not apply to the material conditional. We have been given no grounds to deny the (weaker) claim $H \to L$, for *that* denial cannot be vouchsafed by the likelihood that the animal is a porpoise. To deny $H \to L$ involves implicitly asserting H and denying L, and this is much stronger indeed. When the conditional is read materially, we have no counterexample to *modus ponens*.

Here is the upshot of this discussion of the conditional: There really do seem to be stable meanings of *"if" other* than the sense expressed in $\rightarrow Df$. However, a conditional satisfying $\rightarrow Df$ marks out a stable and useful notion that can do *some* good work. The defining rule helps mark out what the material conditional *is*, and what it *isn't*.

NEGATION

The logic of negation provokes more discussion than any other propositional connective or operator. In this section, I will examine two different concerns that have led some people to reject the logic of negation given to us by the defining rule $\neg Df$. If we can address these arguments, we will better understand the scope and the power of defining rules. The concerns to address are *vagueness* and truth value gaps (questioning $\succ A \lor \neg A$), and *paradoxes* and truth value gluts (questioning $A \land \neg A \succ$). We will start with *vagueness*.

VAGUENESS

Consider a vocabulary for describing a continuously varying quantity, such as a strip shading from one colour (say, *red*) to another (say, *yellow*), and a language with a predicate R (say, the predicate *"is red"*), and comparatives $>_R$ and $\geqslant_R$ (for *"more red than"* and *"at least as red as"* respectively). Enrich the language with a collection of singular terms, and let's use these to picking out patches on the strip, from one end to the other, and ensure that there are enough patches picked out that the difference in shade between adjacent patches is very very small. Let's further suppose that we all agree that the predicate R is *vague*, in the sense that we take there to be patches on the strip that are red (so there are terms t where we would be happy to assert Rt) and patches on the strip that aren't red (so there are other terms t where we would be happy to deny Rt), but there are patches along the strip where we neither would be happy to rule *in* or rule *out* the claim that the strip is red at that patch. Let's suppose that the terms are numbered from a_1 to $a_{10\,000}$, and we'll grant that Ra_1 is clearly true (we are happy to assert it), $Ra_{10\,000}$ is clearly false (we are happy to deny it), and we are neither happy to assert nor to deny $Ra_{5\,000}$. Patch 5 000 is clearly in the borderline between the clearly red and the clearly not-red parts of the strip (wherever *those* parts of the strip begin and end). Let's grant, too, that $a_1 >_R a_{5\,000}$ and $a_{5\,000} >_R a_{10\,000}$ are both clearly true, and that at least $a_1 \geqslant_R a_2 \geqslant_R \cdots \geqslant_R a_{9\,999} \geqslant_R a_{10\,000}$, that is, each point along the strip is at least as red as points further along (toward the yellow direction). Such a description of the way we might use the predicate R and the comparatives $>_R$ and $\geqslant_R$ is, so far, at least, agnostic between different philosophical positions on the nature of vagueness, and also between different positions concerning the logic of negation and the other propositional connectives. So far, everything we have said is consistent with all of the standard views of vagueness, whether

Are we happy to grant $a_1 >_R a_2 >_R a_3 >_R \cdots$ too? It depends on how we understand "more *red* than". Perhaps we understand degrees of redness in a way that can transcend our power to perceive it. (In the same way that I can understand "*longer than*" as precisely defined while taking "*long*" to be vague.) And perhaps we take "more *red* than" to be vague, too. Nothing hangs on the difference here, provided that we *at least* take R to be vague.

For an introduction to the rather large literature on vagueness, Rosanna Keefe and Peter Smith's reader is a good place to start [118] to get you to the end of the 20th Century. For more recent work, Richard Dietz and Sebastian Moruzzi's *Cuts and Clouds* [?] is a good next step.

they be non-classical degree-of-truth accounts, or supervaluational, or epistemicist, or some other such account.

Before pressing on to consider the consequence of extending our little linguistic practice describing our strip using vague predicates with the defining rules for the connectives, we would do well to examine the behaviour of the bounds a little more closely, and in particular, the rules of *Identity* and *Cut*. According to *Identity*, the position $[Ra_n : Ra_n]$ is out of bounds. If I assert that patch n is red and you deny it, we disagree — there is no available position that incorporates our two claims. We are taking opposing stands on n (and on how to understand R). This seems straightforward enough. Even though there may be different ways we could make R sharper, some of which take Ra_n to be true, and others which take it to be false, there is no one position that makes *both* of those choices simultaneously. Given our background commitment, to the effect that $a_1 \geqslant_R a_2 \geqslant_R a_3 \geqslant_R \cdots \geqslant_R a_{10\,000}$ (which I will abbreviate as RO, the R *ordering* claims), the positions that are open to us are restricted. For example, the positions

$$[RO, Ra_1 : Ra_{10{,}000}] \qquad [RO, Ra_{2\,000} : Ra_{8{,}000}]$$

are equally available, but

$$[RO, Ra_{2\,000} : Ra_1, Ra_{8{,}000}]$$

is not, since $RO, Ra_{2\,000} \succ Ra_1$ is derivable. You cannot (consistently) assert the ordering facts RO, assert Ra_n and deny Ra_m, whenever $m \leqslant n$. In fact, for any position of the shape

$$[RO, Ra_n : Ra_m]$$

the claims $Ra_{n'}$ are *implicitly asserted* for each $n' \leqslant n$, and the claims $Ra_{m'}$ are *implicitly denied* for each $m' \geqslant m$, since we can derive

$$RO, Ra_n \succ Ra_{n'} \qquad RO, Ra_{m'} \succ Ra_m$$

whenever $n' \leqslant n$ and $m \leqslant m'$. In other words, given that we grant RO, once we rule patch n *in* as red, and rule patch m *out* as red, we implicitly rule all patches we have taken to be *as least as red as* n as also counting as red, and all those patches we take m to be *at least as red as*, to fail to be red. This allows us to be non-committal concerning patches on the strip between those we have ruled in as red and those we have ruled out. However, the structural rule of *Cut* tells us more about which positions must be available.

Given that the position $[RO, Ra_1 : Ra_{10\,000}]$ is available, by *Cut*, at least one of the positions

$$[RO, Ra_1, Ra_n : Ra_{10\,000}] \qquad [RO, Ra_1 : Ra_n, Ra_{10\,000}]$$

is available, for each n. We can do the same reasoning with every patch across the whole of the strip, appealing to *Cut* each time, to conclude that at least *some* positions of the form

$$[RO, Ra_1, \ldots, Ra_n : Ra_{n+1}, \ldots, Ra_{10\,000}]$$

are available. If we restrict ourselves to purely *structural* features of the bounds (saying that the only bounds at all are those given by the logic of comparatives), then *all* such positions are available, for there is nothing in logic, other than the comparatives in RO, that imposes *any* logical constraint among the formulas of the form Ra_m. In other words, each *maximally opinionated* position — in which every spot up to a_n counts as *red*, and the others further along (even if visually inseparable from spot a_n) do not — is available. There is, according to the bounds, no clash involved in treating R as if it were precise, as if there were a sharp cutoff between those parts of the strip that count as red and those that do not. If the bounds involve *more* than the logic of comparatives (if there are specific norms governing the interpretation of colour terms, say) then perhaps not *all* such positions are available, but regardless, by *Cut*, given that $[\mathrm{RO}, Ra_1 : Ra_{10\,000}]$ is available, then *some* position of the form $[\mathrm{RO}, Ra_1, \ldots, Ra_n : Ra_{n+1}, \ldots, Ra_{10\,000}]$ is also available, by picking some patch as yet undetermined concerning redness, and appealing to *Cut* to show that either it may be coherently added to the left, or to the right, and continuing the process. Provided a distinction *can* be coherently drawn that rules patch 1 in as red and rules patch 10 000 out, by *Cut* it may be *coherently* made sharply, even if to do so is to go far beyond our powers of observation or our ability to judge redness.

This is not to say that any such sharp way of delineating a border between the red and the not-red counts as "the truth of the matter" (whatever that might be), but the logic of the bounds at least says that taking up such a position is coherent, and it cannot violate any of the norms governing positions and their bounds. There is no clash involved in taking the positions up to n to be red, and the positions after that to *not* be red. Nothing forces us to take up such a position, or to say anything concerning what the "facts of the matter" dictate. Everything here is consistent with a range of different analyses of what is involved in matters of vagueness, whether this is to be understood in terms of degrees of truth, of semantic indecision, or the essential limits on our knowledge. All we have is an account that takes the bounds (satisfying *Cut* and *Identity*, as well as *Weakening* and *Contraction*) to apply to our language involving a vague predicate. It will go beyond our perceptual or cognitive capacities to make settled claims concerning redness for every patch along the strip, but to do so would not transgress the bounds.

The scare quotes here do not indicate any skepticism concerning truth on my part. They just serve to warn you that there are more things that would need to be settled if we are to adjudicate on the matter of truth in any different way of spelling out this specific case.

With the bounds for our language clarified, let us turn to consider what would be involved in extending our language with propositional connectives, in particular, *negation*, satisfying our defining rules. If we extend our vocabulary with $\neg Df$, then we can not only assert or deny claims of the form Ra_n, but we can also ask whether or not $\neg Ra_n$ holds too. Asserting $\neg Ra_n$ has the same force as denying Ra_n (as far as the bounds are concerned) while denying $\neg Ra_n$ has the same force as asserting Ra_n. So much seems straightforward. Furthermore, given the results of Chapter 6, we are assured that such an extension of our practice is conservative, and uniquely defined up to indistinguishability. I have argued that it is a free extension of our earlier practice. Does it come at any cost?

Let's see. One feature of our new practice, extending the vocabulary with negation is this: since $[Ra_n : Ra_n]$ is out of bounds, we also have

$$\succ Ra_n, \neg Ra_n$$

In other words, simultaneously denying Ra_n and denying $\neg Ra_n$ is out of bounds. You cannot coherently deny both Ra_n and $\neg Ra_n$, for to deny Ra_n is to make the same impact on the bounds as *asserting* $\neg Ra_n$. But what of the case where n is a position somewhere in the unsettled zone of the strip, which I don't want to rule *in* as red, and I don't want to rule *out*, either? Don't I then want to deny Ra_n and deny $\neg Ra_n$ too? Isn't this a clear case of wanting some kind of *gap* between truth and falsity a gap which classical negation, as defined by $\neg Df$ is too quick to close off?

No, it isn't. The work we did on assertion and denial in Chapter 5 will help clarify what is involved in the connection between negation, assertion and denial in cases like these. We can see that there is an important sense in which the denial of Ra_n and the denial of $\neg Ra_n$ are both appropriate if I am unsettled about whether patch n is red or not. If you were to assert Ra_n, I would resist such an assertion, because I take it to be inappropriate for Ra_n to be added to the common ground, since I am unsettled about it (and, let's say, you have done nothing to update my view and convince me that patch n *is* red). Similarly, if you were to assert $\neg Ra_n$, I would also resist, because I would also take it to be inappropriate for $\neg Ra_n$ to be added to the common ground. But these are not *strong* denials — they are *weak* denials, in the sense marked out in Chapter 5 (see page 184). They are bids to retract information from the common ground. They are not, primarily, bids to add to the common ground. This distinction is important in everyday cases of assertion and denial, and is not an *ad hoc* fix to deal with cases of vagueness. Here, the weak denial of Ra_n does not have the same significance as the assertion of $\neg Ra_n$, so we can deny both Ra_n and $\neg Ra_n$, *weakly*, without being concerned that the position $[\ \ : Ra_n, \neg Ra_n]$ should somehow be available. That would make no sense, for this position has the incoherent force of both answering *no* to the question "Ra_n?" (that is the force of strongly denying Ra_n) and *ruling out* exactly that answer (that is the force of strongly denying $\neg Ra_n$). That is not what the person who is committed to maintaining that position n is in the gap between what is claimed to be red and what is claimed to not be red wants to say.

You could extend the account of the vocabulary in some way to give positive content to weak denial in these cases. Perhaps when I weakly deny p in this context, I am, in effect, strongly denying Sp, the claim that p is *settled* (where this could be spelled out on pragmatic or epistemic or semantic or ontological grounds or some combination of these four, depending on your favoured analysis of vagueness). Different accounts of vagueness give different accounts of the semantics of an operator like this.

No, the weak denial of Ra_n and $\neg Ra_n$ has the upshot, not of attempting to stake out that incoherent position, but of resisting adding Ra_n to the assertion side, and adding Ra_n to the denial side. The weak denial expresses a modesty concerning taking a position, not a positive claim about when to answer 'yes' and when to answer 'no' to questions concerning whether a patch is red. Once we are careful to understand the connections between strong and weak denial, and assertion and negation, we see that the new bounds for negation (like $\succ A, \neg A$) need not pose problems.

Perhaps I have been too swift with this conclusion. Diagnosing *one* sequent thought to be problematic is no guarantee that all such sequents pose no problems. In particular, I have not considered what happens in the presence of other logical vocabulary, also given by defining rules. In particular, given multiplicative disjunction with its defining rule

$$\frac{X \succ A, B, Y}{X \succ A \oplus B, Y}\ \oplus Df$$

we quickly find that the bounds tell us $\succ A \oplus \neg A$ for all A, and hence, even when n is squarely in the borderline between what we take to be red and what we take to be not red, once we add $\oplus$ and $\neg$ to our vocabulary, we find ourselves committed to

$$\succ Ra_n \oplus \neg Ra_n$$

In other words, for each n, either patch n is red, or it isn't. Now *that* seems well and truly beyond what many who use vague vocabulary would be prepared to endorse. Yes, there are *some* (like those who endorse supervaluational [70, 117] or epistemicist [245, 246] accounts of vagueness) who are happy to endorse $Ra_n \oplus \neg Ra_n$ for borderline cases, but there are also others (like those who prefer a degree of truth semantics [228] or a semantics involving truth value *gaps* [241] for vagueness) who would prefer to leave scope for sentences like $Ra_n \oplus \neg Ra_n$ (*"patch n is either red or it isn't"*) to fail to be *true*. Here, at least, it seems that we must part ways with at least some of our fellow travellers along the road of understanding the semantics of vague terms, and that the fact that our negation and disjunction enter as conservative extensions of initial practice is not enough to reassure our *degree of truth* or *truth-value gap* colleagues to endorse these defining rules. However, nothing that we have said about assertion, denial and the bounds has committed us to any particular understanding of truth, and in particular, we have made no claims concerning whether truth values can admit of degrees. If this seems correct to you, feel free to skip the following *excursus*. If you would like a demonstration that the account of the bounds that we have given can cohere with degree theories of truth, then read on.

Excursus: Here is why our friends who endorse *degree of truth* or *truth-value gap* accounts of vagueness could go with us on the initial stage of our journey. Suppose we have a set D of *degrees of truth* selected from the numbers between 0 and 1 inclusive, which at least includes 0, and includes 1 and includes at least one other intermediate value. Let's say that a *degree model* m given our collection $a_1, \ldots, a_{10\,000}$ of patches assigns to each patch a_i some degree $m_R(a_i)$ (its degree of redness) such that $m_R(a_i) \geqslant m_R(a_j)$ whenever $i \leqslant j$ (that is, the degrees of redness never *rise* as you go from a_1 to $a_{10\,000}$). Let's assign the degree of truth — relative to the model m — to statements as follows:

- $m(Ra_i) = m_R(a_i)$. (So, the degree to which Ra_i is true is identified with the degree to which a_i is *red*.)
- $m(a_i \geqslant_R a_j) = 1$ iff $m_R(a_i) \geqslant m_R(a_j)$, and is 0 otherwise. (So, a_i is at least as red as a_j is *true* if the degree to which a_i is red is

at least as high as the degree to which a_j is red, and is false otherwise. In other words, $\geqslant_R$ is a sharp predicate tracking the degree of redness.)

- $m(a_i >_R a_j) = 1$ iff $m_R(a_i) > m_R(a_j)$, and is 0 otherwise. (So $>_R$ is also a sharp comparative, the strict complement of the converse of $\geqslant_R$.)

This is a very simple degree of truth model of the kind used by friends of degrees of truth. Let's say that a model m *verifies* a position $[X : Y]$ to threshold ϵ (where $\epsilon < \frac{1}{2}$) if and only if according to m, each member of X is assigned the value greater than or equal to $1 - \epsilon$, while each member of Y is assigned the value less than or equal to $0 + \epsilon$, and let's say a position $[X : Y]$ is *out of bounds* according to the degree semantics if and only if there is no model m that strongly verifies it. (That is, there is no way, according to the degree semantics, for each member of X to be very close to fully true and each member of Y to be very close to fully false.) It is not difficult to show that if $\epsilon > 0$, the bounds, so defined, satisfy *Identity*, (no model assigns a formula both close to 1 and close to 0, since $\epsilon > \frac{1}{2}$) *Contraction*, *Weakening*, (both trivially satisfied by the structure of the modelling conditions) and most importantly, *Cut*.

To verify *Cut*, suppose we have some model in m which verifies $[X : Y]$ and suppose A is some formula. If A is a comparative, its value in m is either 1 or 0, the model either verifies $[X, A : Y]$ or $[X : A, Y]$, depending on the value of $m(A)$. If A is of the form Ra_i, where $m_R(a_i)$ is neither close to 1 nor close to 0 (so $1 - \epsilon > m_R(a_i) > 0 + \epsilon$) then decide whether we wish to make Ra_i closer to *true*, or closer to *false*—either is possible. Let's choose *true*, without loss of generality. Then *renormalise* our degrees of redness, forming a new model m', where the the patches whose degrees are up to $m_R(a_i)$ have their degrees re-scaled to the interval 1 to $1 - \epsilon$ (preserving the strict ordering) and the other degrees are left unchanged. This preserves all ordering facts, and the new model m', verifies $[X, A : Y]$ as desired.

The bounds for comparatives are satisfied by this interpretation, as we can see if we apply the reasoning used in the proof of Theorem 5.8 (see page 198). It follows that the constraints of the bounds do not themselves impose anything too alien to the notion of degrees of truth, at least as far as primitive vocabulary is concerned. *End of Excursus*

So, suppose that we have a language with a vague predicate (and, perhaps, comparatives), and we are sympathetic to a degree of truth analysis of vagueness, or at least, we have *some* view of vagueness that leads us to initially balk at accepting each instance of $A \oplus \neg A$. What does the fact that the bounds for our language are conservatively extended by the defining rules for $\oplus$ and $\neg$ mean for our view? At the very least, we should pay attention to what the defining rules for $\oplus$ and $\neg$ involve, and what they don't, and what the facts concerning the bounds commit us to, were we to extend our language to include $\oplus$ and $\neg$. Here is a derivation of $\succ A \oplus \neg A$. Let's see where the friend of gaps or degrees of truth — or any other account of vagueness encouraging the rejection of some instances of $A \oplus \neg A$ — might call this derivation into question:

$$\dfrac{\dfrac{A \succ A}{\succ A, \neg A}\,{\scriptstyle \neg Df}}{\succ A \oplus \neg A}\,{\scriptstyle \oplus Df}$$

Reading from the top to bottom, the axiom tells us that it is out of bounds to assert A and deny A. So far, so good, even if A is a claim like Ra, where a is a borderline case of redness. The inference to the next step tells us that it is out of bounds to deny both A and $\neg A$. We have defended this already, given that we have interpreted asserting $\neg A$ as having the same force as denying A, and denying $\neg A$ has the same force as asserting A. (If we have a degree theory of truth, and there is some threshold for degrees of truth which marks acceptability of assertion, then presumably there is some way to understand the interaction of degrees and negation and the threshold for *denial* in such a way as to deliver an appropriate reading of this constraint. One easy way to manage this is to *identify* degree of falsity of A with the degree of truth of $\neg A$, and the degree of falsity of $\neg A$ with the degree of truth of A, and the threshold for denying A is the same threshold required to assert $\neg A$.) On this reading, there is absolutely no problem with $\succ A, \neg A$, which tells us that it is always out of bounds to deny both A and $\neg A$, for this tells us nothing more and nothing less than that it is out of bounds to simultaneously deny A and to assert it, which we had already granted.

To move from this to claim that $A \oplus \neg A$ is undeniable, we have to appeal to the definition of $\oplus$, to the effect that to deny $A \oplus B$ has the same effect as denying A and denying B. If $\oplus$ has its intended meaning, this is the meaning with which we have endowed it, and this is what the sequent tells us: $A \oplus \neg A$ is *undeniable*. Any position in which it is (strongly) denied is one which is out of bounds. Why? To deny $A \oplus \neg A$ is to deny A and to deny $\neg A$, but to deny $\neg A$ is to (implicitly) implicitly assert A, and that is ruled out if we have already denied A. So, $A \oplus \neg A$ is undeniable, or so the reasoning goes.

Notice that this reasoning has its appeal, even if A is a borderline case as far as our vague language is concerned. Even when A neither meets the threshold for assertion or the threshold for denial, nonetheless, since we cannot deny both, the disjunction $A \oplus \neg A$ remains undeniable in the strong sense that its denial is out of bounds. It seems that nothing here is problematic for the degree of truth (or truth value *gap*) analysis, provided that $\oplus$ is understood in terms of its defining rule. Of course, given that interpretation a disjunction $A \oplus B$ may reach the threshold for truth — at least in the sense of being undeniable — without either A or B meeting that threshold. This may seem to be problematic if we are committed to the notion of degrees of *truth*, but perhaps this could be made more palatable by noticing that $A \oplus B$ is indistinguishable here from $\neg A \rightarrow B$, and it is much less problematic to think that $\neg A \rightarrow \neg A$ could reach the threshold for assertion when neither A nor $\neg A$ reach that threshold — even when A is squarely in the unsettled border of some vague predicate. If $A \oplus \neg A$ says nothing more (and nothing less) than $\neg A \rightarrow \neg A$, then being committed to truth value gaps or degrees of truth do not, in themselves, give us a strong reason to resist using defining rules to provide the vocabulary of negation and disjunction.

So *this* is the sense that our practice is conservatively extended by negation, even in the presence of vagueness. It is true, some views of vagueness allow for the interpretation of the classical logical vocabulary more straightforwardly and immediately than others, but even views like those of degrees of truth or truth value gaps seem to be consistent with the meanings of the connectives as given by the defining rules. There is no insuperable bar to using defining rules in the presence of vague vocabulary — in fact, this vocabulary is totally appropriate in this context, since defining rules give us rules for extending the practice of assertion and denial in terms of bounds, and we can use vague language to assert and deny, just as we can use precisely delimited concepts. There is nothing in the bounds that means that vague vocabulary is ruled out. Vague language admits of enough structure to allow for the structure of bounds and defining rules, even when the result of those defining rules is classical propositional logic.

PARADOX

Another important motivation for a non-classical account of negation is found in semantic paradoxes. *Dialetheists* take the paradoxes, such as the liar paradox:

$$(\lambda) \qquad \lambda \text{ is not true.}$$

Here, 'λ' is a singular term, naming a sentence, so 'λ is not true' is a sentence, involving negation and a predicate expressing the property of being true in the same way that a sentence 'Kermit is not green' is a sentence, involving negation a predicate expressing the property of being green.

to give us good reason to think that some sentences are both *true* and *not true*. Since the dialetheist is prepared to assert the sentence λ and also to assert its negation, they at least implicitly reject the sequent

$$A, \neg A \succ$$

since they take it to be *in* bounds to assert some sentence and its negation [11, 172, 175]. (Of course, they may think it is *sometimes* out of bounds to assert a sentence and its negation, but in taking it to be acceptable to assert λ and its negation, at least *some* contradictory pairs are on the field of play.) Taking this approach to the paradoxes is not compatible with accepting the defining rule $\neg Df$ for negation, for this rule (given *Identity*) leads immediately to $A, \neg A \succ$. So, there is tension between a dialetheist position, for which there are true contradictions, the account of assertion and denial underlying our analysis of sequents (for which a derivation of a sequent of the form $X \succ Y$ tells us that asserting each member of X and denying each member of Y is out of bounds) and our account of defining rules. In this section, I will explore this tension, and use it to illuminate the role of defining rules in our account of the logical connectives.

To set the scene, and to give dialetheists their due, let's consider the motivation for this non-classical treatment of negation by way of the paradoxes of self reference. As indicated above, one such motivation is found in the liar paradox, a sentence that says of itself that it is not true. There are a number of different ways to represent the structure of reasoning involved in the paradox. Figure 71 presents one such derivation, reducing the claim that the sentence λ is the sentence that says of itself that it is not true (formalised $\lambda = \ulcorner\neg T\lambda\urcorner$) to absurdity, given the structural rules of *Identity*, *Cut* and *Contraction*, two rules for a truth predicate, the left and right rules for negation (arising out of its defining rule), and two pairs of rules TL and TR governing the truth predicate, and a rule $=L$ governing the binary *identity* predicate. The rule for the identity predicate takes the following forms

For more on rules for identity, and how these can indeed take the shape of defining rules, see Chapter 8, and specifically, Section 8.2.

$$\frac{X \succ Y}{s = t, X|^s_t \succ Y|^s_t}{=}L \qquad \frac{X \succ Y}{s = t, X|^t_s \succ Y|^t_s}{=}L$$

which states that at the cost of adding the identity statement $s = t$, we can replace instances of s by t (or instances of t by s) in a sequent. The rules for the truth predicate make use of *quotation*, which forms for each sentence A a singular term $\ulcorner A\urcorner$. The rules

$$\frac{X, A \succ Y}{X, T\ulcorner A\urcorner \succ Y}TL \qquad \frac{X \succ A, Y}{X \succ T\ulcorner A\urcorner, Y}TR$$

have the effect of making the predication $T\ulcorner A\urcorner$ equivalent to the sentence A. In other words, to claim that the sentence A is true is to say nothing more and nothing less than would be said by saying A itself. Then, a liar sentence, like λ above, is a sentence that says of itself that it is not true. In other words, the term 'λ' picks out the same sentence as the quotation term '$\ulcorner\neg T\lambda\urcorner$'. However the derivation in Figure 71 shows that assertion of the claim that $\lambda = \ulcorner\neg T\lambda\urcorner$ is itself out of bounds, using only the rules we have specified.

Could TL and TR be motivated as arising out of a *defining rule* for the truth predicate?

$$\frac{X, A \succ Y}{X, T\ulcorner A\urcorner \succ Y}TDf$$

TDf is certainly an invertible rule and it is equivalent to the left and right rules in the presence of *Cut* and *Identity*. However, this rule involves two separate notions, the *truth predicate* and *quotation*, and is not a defining rule within the restricted meaning of the term. In Chapter 8 we will explore whether this should count as a defining rule. To see why it isn't a defining rule in the sense we have considered, notice that it tells us nothing about norms governing the assertion of an arbitrary sentence of the form Tx.

The reasoning for the liar paradox is relatively simple. If the derivation is to be blocked, then we must reject one of the principles involved, whether a structural rule, like *Identity*, *Cut* or *Contraction*, or the rules for negation, truth or the identity predicate. Pointing the finger at the identity predicate is almost certainly non-starter, because the same sort of reasoning can be reconstructed in similar contexts without making use of identity at all. For example, *Russell's paradox* concerning set membership, can be formalised without the use of identity. To do this, let us enrich our vocabulary with a way to construct *terms* (for *sets*) from *predicates*. In other words, given the predicate $\phi(x)$, with the free variable x we will form the term $\{x : \phi(x)\}$, which we will understand to be picking out the collection of all and only those things x which have the property $\phi(x)$. Given a collection s and an object t we can inquire whether t is in the collection s, or whether $t \in s$. The criterion for whether $t \in \{x : \phi(x)\}$ is simple: it is $\phi(t)$. This motivates a pair of rules for membership:

Again, we could ask whether these rules should arise out of a *defining* rule for $\in$. Notice that these rules also feature a term forming operator as well as the concept supposedly being 'defined'.

$$\frac{X, \phi(t) \succ Y}{X, t \in \{x : \phi(x)\} \succ Y}{\in}L \qquad \frac{X \succ \phi(t), Y}{X \succ t \in \{x : \phi(x)\}, Y}{\in}R$$

Russell's Paradox arises when we consider $\{x : x \notin x\}$, the collection of

Here, '$s \notin t$' is an abbreviation for '$\neg s \in t$', the negation of '$s \in t$.'

$$
\begin{prooftree}
\AxiomC{$T\lambda \succ T\lambda$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg T\lambda, T\lambda \succ$}
\RightLabel{TL}
\UnaryInfC{$T\ulcorner\neg T\lambda\urcorner, T\lambda \succ$}
\RightLabel{$=L$}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, T\lambda, T\lambda \succ$}
\RightLabel{W}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, T\lambda \succ$}
\RightLabel{$\neg R$}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner \succ \neg T\lambda$}
\RightLabel{TR}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner \succ T\ulcorner\neg T\lambda\urcorner$}
\RightLabel{$=L$}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, \lambda = \ulcorner\neg T\lambda\urcorner \succ T\lambda$}
\RightLabel{W}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner \succ T\lambda$}
\AxiomC{$T\lambda \succ T\lambda$}
\RightLabel{$\neg L$}
\UnaryInfC{$\neg T\lambda, T\lambda \succ$}
\RightLabel{TL}
\UnaryInfC{$T\ulcorner\neg T\lambda\urcorner, T\lambda \succ$}
\RightLabel{$=L$}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, T\lambda, T\lambda \succ$}
\RightLabel{W}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, T\lambda \succ$}
\RightLabel{Cut}
\BinaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner, \lambda = \ulcorner\neg T\lambda\urcorner \succ$}
\RightLabel{W}
\UnaryInfC{$\lambda = \ulcorner\neg T\lambda\urcorner \succ$}
\end{prooftree}
$$

Figure 71: A DERIVATION FOR THE LIAR PARADOX

non-self-membered things. Is this collection a member of itself? One argument says that it isn't (since if it were self-membered, it satisfies the criterion for being in itself, which means that it isn't). But then, if it isn't a member of itself, it satisfies the condition for self-membership, so it is. We can formalise this reasoning straightforwardly: Letting r stand for $\{x : x \notin x\}$ the derivation in Figure 72 shows us that *Identity*, *Cut*, *Contraction*, $\neg L/R$ and $\in L/R$ are enough to derive the empty sequent. Notice that the identity predicate ('$=$') plays no role in this derivation.

What is the problem with deriving the empty sequent? Given *Weakening*, it would mean that every sequent is derivable, which would rather defeat the purpose of deriving things. Without *Weakening*, deriving the empty sequent it is not ***trivialising***, but it does amount to a derivation of $t \succ f$, and deriving something false from something true does seem problematic.

$$
\begin{prooftree}
\AxiomC{$r \in r \succ r \in r$}
\RightLabel{$\neg L$}
\UnaryInfC{$r \notin r, r \in r \succ$}
\RightLabel{$\in L$}
\UnaryInfC{$r \in r, r \in r \succ$}
\RightLabel{W}
\UnaryInfC{$r \in r \succ$}
\RightLabel{$\neg R$}
\UnaryInfC{$\succ r \notin r$}
\RightLabel{$\in R$}
\UnaryInfC{$\succ r \in r$}
\AxiomC{$r \in r \succ r \in r$}
\RightLabel{$\neg L$}
\UnaryInfC{$r \notin r, r \in r \succ$}
\RightLabel{$\in L$}
\UnaryInfC{$r \in r, r \in r \succ$}
\RightLabel{W}
\UnaryInfC{$r \in r \succ$}
\RightLabel{Cut}
\BinaryInfC{$\succ$}
\end{prooftree}
$$

Figure 72: A DERIVATION FOR RUSSELL'S PARADOX, WHERE r IS $\{x : x \notin x\}$

So, the standard structural rules (*Identity*, *Cut*, *Contraction*, *Weakening*) and the defining rules for negation are incompatible with these principles concerning truth (TL/R) or set membership ($\in L/R$). How are we to understand this tension? There are a number of options to be explored. One is to accept the standard structural rules, and the strong principles

We need not stop here. We could consider the property *heterologicality* had by properties which do not apply to themselves, and derive a similar paradox with *property* abstraction: $\lambda x.\phi(x)$ applies to s if and only if $\phi(s)$.

for the truth predicate or for set membership, and to restrict the rules for negation. A second is to accept the defining rules and the strong principles for truth or for sets, and to reject a structural rule. A third is to reject the strong principles for truth or set membership. We will consider all three in turn, to see what we can learn about defining rules and negation.

» «

For a dialetheist like Graham Priest [172, 174, 175], or Jc Beall [11], it is more natural to accept the standard structural rules and to reject the defining rule $\neg Df$. The favoured propositional logic for many dialetheists like Priest at least includes distributive lattice logic for conjunction ($\wedge$) and disjunction ($\vee$), and any natural proof for the distributive law

$$A \wedge (B \vee C) \succ (A \wedge B) \vee (A \wedge C)$$

involves *Weakening* and *Contraction* in some form. To see how the structural rules are involved, see the derivation on page 63. Given this setting, if we wish to retain the strong principles for the truth predicate or set membership, we are to reject $\neg Df$. Given that *adding* $\neg Df$ in the presence of the rules for truth (or set membership) leads to paradox, it seems to follow that defining rules are not as cost-free as we have been led to believe. Understanding this phenomenon will help us clarify what is involved — and what is *not* involved — in taking defining rules to be purely definitional.

The situation concerning contraction for these dialetheists is somewhat complex. They logics they prefer may involve a conditional $\rightarrow$ and a multiplicative conjunction $\oplus$, defined in terms of a "punctuation mark" in sequents for which *Contraction* is rejected, while at the very same time accepting *Contraction* for a different mode of combination, involved in the definition of conjunction. See work by Belnap, Gupta and Dunn [19], Read [184], Slaney [226], and my earlier work [192, 193] for more on this tradition.

Our first step will be to notice that despite appearances, our conservative extension result still holds. The $\neg Df$ rules *do* conservatively extend a language with the full complement of structural rules, and with a truth predicate satisfying truth rules ($\mathsf{T}L/R$). The fly in the ointment for the dialetheist is that the result of the conservative extension is a language in which the truth rules ($\mathsf{T}L/R$) *no longer hold*. To explain how this could be, we will first specify a language with bounds satisfying the structural rules (*Identity*, *Cut*, *Contraction*, *Weakening*) and the truth rules ($\mathsf{T}L/R$), and then we will explain how this language *can* be conservatively extended by a negation satisfying $\neg Df$.

To be specific, let's consider a language $\mathcal{L}$, which a one-place predicate F, with comparatives $>_F$ and $\geqslant_F$, the two-place identity predicate $=$, a family of constants $s_0, s_1, \ldots$, conjunction ($\wedge$) and disjunction ($\vee$). We will suppose that the bounds for $\mathcal{L}$ are closed under the standard structural rules, $\wedge Df$, and $\vee Df$, the bounds for comparatives for F, reflexivity for identity ($\succ s = s$, for each singular term s), the identity rule $=L$. This is our language $\mathcal{L}$. $\mathcal{L}_T$ extends $\mathcal{L}$ with a new predicate T, new singular terms $\lambda_0, \lambda_1, \ldots$ (to allow self-referential sentences) and quotation device $\ulcorner\cdot\urcorner$ such that for any sentence A in the language $\mathcal{L}_T$, $\ulcorner A\urcorner$ is a singular term. The bounds for the extended language include the bounds of $\mathcal{L}$, and the new identity bounds $\succ t = t$ for each new term in the language (whether a quotation term or a λ_i term) as well as $\ulcorner A\urcorner = \ulcorner B\urcorner \succ$ for each pair of *different* sentences A and B, as well as the truth predicate rules ($\mathsf{T}L/R$). This will be our extended language $\mathcal{L}_T$. The bounds for this language are non-trivial, in the sense that $\mathcal{L}_T$ is a conservative extension of $\mathcal{L}$. Adding quotation terms and the truth rules makes no change to the bounds of $\mathcal{L}$.

Showing that this conservative extension holds involves some hard work. Nothing in the rest of this section stands on the details of how it is shown, and the construction uses the tools of *model theory* and not proof theory, so I will leave the demonstration to an excursus below. If you are happy to take the result on faith, or if you would rather stick to proof theory, then skip ahead to page 271. If you can cope with the technical details of a model construction, read on.

And as a matter of fact, the result fails if you add *quantifiers* to $\mathcal{L}$, if the logic of $\mathcal{L}$ is no stronger than classical logic. Models of $\mathcal{L}$ could well be finite — we can have models in which $(\forall x)(\forall y)x = y$ holds — and there is no way to conservatively add the quotation terms of $\mathcal{L}_T$ to $\mathcal{L}$, let alone the truth rules, since we have $\ulcorner F\lambda_0 \urcorner = \ulcorner F\lambda_1 \urcorner \succ$ in $\mathcal{L}_T$, contradicting $(\forall x)x = x$. This is no problem for *us*, for we have no way to express the general judgement $(\forall x)(\forall y)x = y$ in our $\mathcal{L}$.

Excursus: We will show the conservative extension result by extending any model for $\mathcal{L}$ into a model for $\mathcal{L}_T$. A model for $\mathcal{L}$ can be — according to Theorems 5.6 and 5.8 — specified by $\langle T, \succsim, \mathcal{F} \rangle$, a set T equipped with a linear order $\succsim$, and some upwardly closed subset $\mathcal{F}$ of T. The structure $\langle T, \succsim, \mathcal{F} \rangle$ gives us enough information to interpret the predicate F and its comparatives $>_F$ and $\geqslant_F$, together with conjunction, disjunction and the identity predicate. The remaining items in $\mathcal{L}$ are the singular terms $s_0, s_1, \ldots$, and we can interpret these terms by choosing a sequence $t_0, t_1, \ldots$ of elements from the set T. Given this material, we can specify a limit position $[X : Y]$ the language $\mathcal{L}$ in the following way. Here are the conditions for atomic formulas:

- $Fs_i \in X$ iff $t_i \in \mathcal{F}$; otherwise $Fs_i \in Y$.
- $s_i = s_j \in X$ iff $t_i = t_j$; otherwise $s_i = s_j \in Y$.
- $s_i \geqslant_F s_j \in X$ iff $t_i \succsim t_j$; otherwise $s_i \geqslant_F s_j \in Y$.
- $s_i >_F s_j \in X$ iff $t_j \not\succsim t_i$; otherwise $s_i >_F s_j \in Y$.

Then, for conjunctions and disjunctions of formulas, we can use the following recursive clauses, as usual:

- $A \wedge B \in X$ iff $A \in X$ and $B \in X$; otherwise $A \wedge B \in Y$.
- $A \vee B \in X$ iff $A \in X$ or $B \in X$; otherwise $A \vee B \in Y$.

It is not difficult to show that these clauses truly define a limit position $[X : Y]$ for $\mathcal{L}$. To extend this into a limit position for $\mathcal{L}_T$, we need to find ways to interpret terms of the form $\ulcorner A \urcorner$, and the terms λ_i, the truth predicate T. For this, we select for each fixed point term λ_i a formula A_i to which the term will refer. (This formula may *include* the term λ_i or any other lambda term, to allow for self reference and looping reference — if A_0 is, for example, $T\lambda_0$, then A_0 says of itself that it is true, and we will treat the sentence $\lambda_0 = \ulcorner A_0 \urcorner$ as true). If c is a term (whether of the form s_i, from $\mathcal{L}$, or $\ulcorner A \urcorner$, or λ_i, we will take $[\![c]\!]$ to be its referent: so $[\![s_i]\!] = t_i$, $[\![\ulcorner A \urcorner]\!] = A$ and $[\![\lambda_i]\!] = A_i$. To interpret the predicate F and the comparatives, we will presume that no formula A is in our linearly ordered set T, and we will extend $\succsim$ to order not only members of T but also formulas, by setting $t \succ A$ for each formula A (so formulas are treated as equally *less* F than any object in T). This will be enough to evaluate almost all of the new vocabulary, except for the predicate T.

Let's think about how this ought to be understood. If A is in X, and if $[\![c]\!] = A$, then Tc ought to be in X too. Conversely, if Tc is in X and $[\![c]\!] = A$, then A should be in X. We can't, though, simply add a clause like this

- $Tc \in X$ iff for some A, $[\![c]\!] = A$ and $A \in X$; $Tc \in Y$ otherwise.

because the formula A in question might well be a *more complex* formula than Tc itself, and as a result, this evaluation process may not terminate. If, for example, A_0 is $T\lambda_0$, then this clause tells us that $T\lambda_0 \in X$ iff $T\lambda_0 \in X$ — which is not particularly illuminating on the question of whether $T\lambda_0 \in X$ or not. So, to evaluate formulas of the form Tc, we need to be a more careful than that. Instead, we will build up the evaluation for T in *stages*. We will define limit positions $[X_n : Y_n]$ for $n = 0, 1, \ldots$ which will respect the rules T*L*/*R* better and better as the index n increases. For the atomic formulas, we can take these clauses as given, for every stage:

- $Fc \in X_n$ iff $[\![c]\!] \in \mathcal{F}$; otherwise $Fc \in Y_n$.
- $c = c' \in X_n$ iff $[\![c]\!] = [\![c']\!]$; otherwise $c = c' \in Y_n$.
- $c \geqslant_F c' \in X_n$ iff $[\![c]\!] \succsim [\![c']\!]$; otherwise $c \geqslant_F c' \in Y_n$.
- $c >_F c' \in X_n$ iff $[\![c']\!] \not\succsim [\![c]\!]$; otherwise $c >_F c' \in Y_n$.

since the interpretation of these formulas will not change as we approximate T better and better. For T, we start by setting

- $Tc \in X_0$ never; $Tc \in Y_0$ always.

So *nothing* starts off as falling under the truth predicate. Once we have settled on the interpretation of the atomic formulas in our language, we interpret conjunction and disjunction as before, at every stage:

- $A \wedge B \in X_n$ iff $A \in X_n$ and $B \in X_n$; otherwise $A \wedge B \in Y_n$.
- $A \vee B \in X_n$ iff $A \in X_n$ or $B \in X_n$; otherwise $A \vee B \in Y_n$.

Now, to interpret the truth predicate better, we can, at each stage, look back and see what the previous stage has said about the formula in question:

- $Tc \in X_{n+1}$ iff $[\![c]\!] \in X_n$; and otherwise $Tc \in Y_{n+1}$.

So, for example, to evaluate $T\ulcorner Fs_0 \urcorner$, we see that this formula is in Y_0, since *all* formulas of the form Tc are in Y_0. Then, to evaluate it at stage 1, we check Fs_0 (since $[\![\ulcorner Fs_0 \urcorner]\!] = Fs_0$) at stage 0. $Fs_0 \in X_0$ iff $t_0 = [\![s_0]\!] \in \mathcal{F}$. Let's suppose that t_0 actually *is* in $\mathcal{F}$. That would make $Fs_0 \in X_0$, i.e., Fs_0 holds at stage 0. So, $T\ulcorner Fs_0 \urcorner \in X_1$, or $T\ulcorner Fs_0 \urcorner$ holds at stage 1. In fact, since $Fs_0 \in X_n$ for each n, we have $T\ulcorner Fs_0 \urcorner \in X_{n+1}$ for each n. That is, $T\ulcorner Fs_0 \urcorner$ holds from stage 1 onwards.

For another example, let's suppose that A_0 is chosen to be $T\lambda_0$, so this is a "truth-teller" sentence, which says of itself that it is true. Here, as before, $T\lambda_0 \in Y_0$, as is the case for all formulas of the form T_c. Now, to evaluate it at stage 1, we check $T\lambda_0$ at stage 0, since $[\![\lambda_0]\!] = A_0 = T\lambda_0$.

So here, $T\lambda_0 \in Y_1$ too, and the same will hold at each stage: $T\lambda_0 \in Y_n$ for each n.

Here is the limit position for $\mathcal{L}_T$:

$$[\bigcup_{n=0}^{\infty} X_n : \bigcap_{n=0}^{\infty} Y_n]$$

That is, we collect together everything that is counted holding at any stage and select that for the left of our position, and anything that is excluded at every stage is excluded in our position. This is, in fact, a limit position for $\mathcal{L}_T$. To show this, it will be important to show that the sets $X_0, X_1, \ldots$ are *non-decreasing*, that is $X_n \subseteq X_{n+1}$ for each n. This clearly holds for atomic formulas at the *first* stage: every atomic formula in X_0 is also in X_1, since *no* atoms of the form Tc are in X_0, and the other atomic formulas do not change their status at each stage. So, atoms in X_0 are also in X_1. But now we can prove by induction on the structure of complex formulas that *all* formulas are preserved from X_0 to X_1. Suppose $A \wedge B \in X_0$. Then, by construction $A, B \in X_0$, and hence, by the induction hypothesis, $A, B \in X_1$, and hence $A \wedge B \in X_1$. Similarly, suppose $A \vee B \in X_0$. Then, by construction either $A \in X_0$ or $B \in X_0$, and hence, by the induction hypothesis, either $A \in X_1$ or $B \in X_1$, and hence $A \vee B \in X_1$. So, by induction, all formulas are preserved from stage 0 to stage 1.

This is the point at which the argument would not work for negation, or for the conditional.

Now, suppose we have formulas preserved from stage 0 up to stage n, and consider the step to $n+1$. We know that atoms except for those of the form Tc are automatically preserved, as their evaluations are constant. Suppose, now, that $Tc \in X_n$. This means that $[\![c]\!] \in X_{n-1}$ (so we must have $n > 0$), so, by the induction hypothesis, $[\![c]\!]$ is preserved from stage $n-1$ to stage n, and hence $[\![c]\!] \in X_n$, too. This means that $Tc \in X_{n+1}$, as desired. The reasoning for conjunction and disjunction has exactly the same form as in the 0 to 1 case, and so, we see that *all* formulas are preserved from X_n to X_{n+1}.

We can use this fact to show that $[\bigcup_{n=0}^{\infty} X_n : \bigcap_{n=0}^{\infty} Y_n]$ respects the structural rules, the bounds for comparatives, identity, and conjunction and disjunction (since these are respected at each stage) of the construction. For example, to verify the rules for conjunction, note that $A \wedge B \in \bigcup_{n=0}^{\infty} X_n$ if and only if $A \wedge B \in X_n$ for some n. From this it follows that $A \in X_n$ and $B \in X_n$, and hence $A, B \in \bigcup_{n=0}^{\infty} X_n$. If, conversely, $A, B \in \bigcup_{n=0}^{\infty} X_n$, then we have $A \in X_n$ for some n and $B \in X_m$ for some m. It follows that $A, B \in X_{\max(m,n)}$, since these sets are increasing, and hence, that $A \wedge B \in X_{\max(m,n)}$, and thus, that $A \wedge B \in \bigcup_{n=0}^{\infty} X_n$.

This is the point at which this argument would not work if we had *quantifiers* in our language. To model them, we would need to allow for the process to go beyond the ordinal ω.

To verify the rules T*L*/*R*, we can proceed as follows: suppose $A \in \bigcup_{n=0}^{\infty} X_n$. It follows, then that $A \in X_n$ for some n, and hence, that $T\ulcorner A \urcorner \in X_{n+1}$, and so, $T\ulcorner A \urcorner \in \bigcup_{n=0}^{\infty} X_n$, as desired. Conversely, suppose $T\ulcorner A \urcorner \in X_n$. It follows, then, that $A \in X_{n-1}$, and hence, that $A \in \bigcup_{n=0}^{\infty} X_n$, too.

So, we have a limit position for the extended language $\mathcal{L}_T$. We have indeed shown that we can conservatively extend a simple language—

This is a simplified version of a well-known construction, due (independently) to Ross Brady [28], Paul Gilmore [85], Saul Kripke [121], and Robert Martin and Peter Woodruff [138].

without boolean negation, or the material conditional—with a truth predicate satisfying T*L*/*R*, and allowing for fixed points.

End of Excursus

So, we can have a language $\mathcal{L}_\mathsf{T}$ that has a truth predicate and the ability to form self referential statements. It has bounds that are closed under *Identity*, *Contraction*, *Weakening*, and *Cut*, and so, it may be extended conservatively and uniquely with defining rules, as we have seen in the previous chapter. If we add negation using the rule $\neg$*Df*, we have a consistent extension of $\mathcal{L}_\mathsf{T}$, with negation — call this new language $\mathcal{L}_\mathsf{T}^\neg$.

What about the liar sentence? Now that we have negation, don't we have the means to form a sentence λ of the form $\neg\mathsf{T}\lambda$, and hence, derive a contradiction, using the truth rules? It turns out that this does *not* follow at all. The conservative extension proof from Chapter 6 shows that the extension of $\mathcal{L}_\mathsf{T}$ by way of negation is *consistent*, but when we form that extension, we add to the stock of connectives in our new language — but we do not do anything to the stock of *singular terms* of the original language. Our initial language $\mathcal{L}_\mathsf{T}$ has resources to name sentences from $\mathcal{L}_\mathsf{T}$ by way of quotation. We added a term $\ulcorner A \urcorner$ for every sentence in $\mathcal{L}_\mathsf{T}$. The new language $\mathcal{L}_\mathsf{T}^\neg$ has the means to use negation to describe the same things $\mathcal{L}_\mathsf{T}$ can describe — in particular in can describe the language of $\mathcal{L}_\mathsf{T}$. We can now negate sentences of $\mathcal{L}_\mathsf{T}$, but there is no device in the new language $\mathcal{L}_\mathsf{T}^\neg$, as defined, to form names of sentences in $\mathcal{L}_\mathsf{T}^\neg$, and nor is there any requirement that the constants λ_i in $\mathcal{L}_\mathsf{T}^\neg$ name anything other than sentences from $\mathcal{L}_\mathsf{T}$ as they named before. All we have in $\mathcal{L}_\mathsf{T}^\neg$ is new conceptual resources we can use to describe exactly the same things described in $\mathcal{L}_\mathsf{T}$, the sentences of $\mathcal{L}_\mathsf{T}$ itself. The extended language does not have a truth predicate for the *whole* language, and neither does it have the ability to form *self* referential sentences using its new vocabulary.

Excursus: We do not need to believe this merely on the basis of the argument of Chapter 6. We can show it directly, by extending any limit position $[X : Y]$ for $\mathcal{L}_\mathsf{T}$ into a limit position $[X' : Y']$ for the extended language $\mathcal{L}_\mathsf{T}^\neg$, respecting the defining rule $\neg$*Df*. Sentences in $\mathcal{L}_\mathsf{T}^\neg$ are either atoms (and in $\mathcal{L}_\mathsf{T}$) or they are conjunctions, disjunctions or negations of formulas in $\mathcal{L}_\mathsf{T}^\neg$. We can define the limit position as follows:

- If A is an atomic formula, $A \in X'$ iff $A \in X$; $A \in Y'$ iff $A \in Y$.
- $A \land B \in X'$ iff $A \in X'$ and $B \in X'$; otherwise $A \land B \in Y'$.
- $A \lor B \in X'$ iff $A \in X'$ or $B \in X'$; otherwise $A \lor B \in Y'$.
- $\neg A \in X'$ iff $A \in Y'$; and $\neg A \in Y'$ iff $A \in X'$.

The result is a limit position for $\mathcal{L}_\mathsf{T}^\neg$, which respects the bounds of $\mathcal{L}_\mathsf{T}$, and the defining rules for conjunction, disjunction and negation. However, it does not satisfy the truth rules T*L*/*R* *for the new vocabulary*, since our new language does not include singular terms $\ulcorner A \urcorner$ for the vocabulary introduced in $\mathcal{L}_\mathsf{T}^\neg$. *End of Excursus*

So, this is the *first* moral of the story. If you like, you can have a language which has truth predicate with and self reference, and its bounds may be conservatively extended with a negation operator by way of defining rules but when we do so, we no longer have a truth predicate or self reference for the new language. Our conservative extension result holds, but we have moved from a language with a predicate satisfying the rules T*L/R* to a language which does not. If your aim is to have a predicate satisfying T*L/R*, and allowing for self-reference (in the presence of the structural rules *Identity*, *Cut*, *Contraction*, at least) then defining rules do not provide a sufficient criterion for selecting your connectives. They come for free, but they do not satisfy the constraint imposed by these strong rules for truth.

» «

We have explored the first of our three responses to the paradoxes, restricting ¬*Df* in the context of the standard structural rules. Now to consider the *second*. The landscape looks rather different if you are willing to shift to different territory, in particular, to depart from those standard structural rules. Look back at the derivations of the paradoxes in Figures 71 and 72, and you will see one striking thing: *Contraction* plays a role in both derivations. The kinds of fixed points points formed in the paradoxes are much less problematic in deductive systems without contraction — so much so that they are actually *consistent*, even in the presence of defining rules for *all* of our connectives.

Recall the discussion of Read's problematic constant • from page 240 in Chapter 6. The formula •, given by the rules:

$$\frac{X \succ \bullet, Y}{\bullet, X \succ Y}\,\bullet L \qquad \frac{X, \bullet \succ Y}{X \succ \bullet, Y}\,\bullet R$$

is a proof theoretical analogue of the liar sentence, or of the statement $r \in r$ from Russell's paradox. It is a sentence that is indistinguishable from its own negation:

$$\dfrac{\dfrac{\bullet \succ \bullet}{\succ \bullet, \neg\bullet}\,\neg L}{\bullet \succ \neg\bullet}\,\bullet L \qquad \dfrac{\dfrac{\bullet \succ \bullet}{\bullet, \neg\bullet \succ}\,\neg R}{\neg\bullet \succ \bullet}\,\bullet R$$

in just the way that $r \in r$ is indistinguishable from $r \notin r$. We saw in Chapter 6 that while Read's bullet is not given by a *defining* rule, it can at least be conservatively added to our language if we are willing to do without *Contraction*. This phenomenon, it turns out, is completely general, and has nothing, in *particular*, to do with negation, and it need not be restricted to propositional constants given by rules like •*L/R*. It is much more general than this.

Let's call a sentence $F(p)$ in which the atom p occurs a sentential *context*. We can treat p as a 'hole' in which other formulas are placed, substituting A for p, we get $F(A)$. So, for example, $\neg p$ is a context (the *negation* context), but so is $p \to \bot$, or $(p \otimes p) \to (q \vee r)$, or any other formula involving p. Russell's paradox involves finding a set r such that its self-membership statement $r \in r$ is a fixed point for the *negation* context, that is, $r \in r$ is indistinguishable from $\neg(r \in r)$. The set rules $\in L/R$ are much more powerful than this, and they have consequences for *every* context, not just negation. In particular, the set f, defined as

$$\{x : F(x \in x)\}$$

is such that its self-membership statement $f \in f$ is a fixed pont for the $F(p)$ context: that is, $f \in f$ is indistinguishable from $F(f \in f)$. The argument is perfectly general, and very simple:

$$\frac{F(f \in f) \succ F(f \in f)}{f \in f \succ F(f \in f)}\,{\scriptstyle \in L} \qquad \frac{F(f \in f) \succ F(f \in f)}{F(f \in f) \succ f \in f}\,{\scriptstyle \in R}$$

What holds for negation here holds for *every* context. So, for any formula B there is a statement A that is equivalent to $A \rightarrow B$ and it is well known to be rather difficult to avoid the deduction from the presence of any A equivalent to $A \rightarrow B$ to conclude B — and here, B could be anything you like at all. However, in the absence of *Contraction*, such a paradoxical sentence is no problem at all, in fact, *none* of these paradoxical sentences are a problem, in that the addition of these set principles $\in L/R$ is *conservative* if we are happy to do away with *Contraction*. However, the broad outlines can be covered straightforwardly. Consistency is proved by way of a *Cut* elimination proof. The rules $\in L/R$ can be shown to be consistent if we eliminate *Cuts* as follows:

This is the *Curry Paradox* [46, 146, 192].

I will not go through all the details here, because unlike the case of the fixed point construction for models of the truth predicate, getting the specifics correct would take us too far afield into the details of predicate logic without contraction, and we have already seen the outlines of conservative extension by cut elimination. The details are well covered elsewhere [91, 92, 162, 163, 164] if you wish to explore them.

$$\dfrac{\dfrac{\begin{matrix}\delta_1\\ \vdots\\ X \succ \phi(t), Y\end{matrix}}{X \succ t \in \{x : \phi(x)\}, Y}\,{\scriptstyle \in R} \qquad \dfrac{\begin{matrix}\delta_2\\ \vdots\\ X', \phi(t) \succ Y'\end{matrix}}{X', t \in \{x : \phi(x)\} \succ Y'}\,{\scriptstyle \in L}}{X, X' \succ Y, Y'}\,{\scriptstyle Cut}$$

can simplify to

$$\dfrac{\begin{matrix}\delta_1\\ \vdots\\ X \succ \phi(t), Y\end{matrix} \qquad \begin{matrix}\delta_2\\ \vdots\\ X', \phi(t) \succ Y'\end{matrix}}{X, X' \succ Y, Y'}\,{\scriptstyle Cut}$$

Here, the derivation is at the very least *shorter*, once we have eliminated the *Cut*, and if *Contraction* is not among the structural rules we can also see that all *Cut* reduction steps reduce the size of the derivation. It turns out (see the details elsewhere [91, 92, 162, 163, 164]) that the rules $\in L/R$ are conservatively added to the basic logical rules, since *Cut* can be eliminated form any derivation, and any *Cut*-free derivation of a sequent without involving set membership cannot use the rules $\in L/R$, so their addition must be conservative over the earlier vocabulary. Despite their strength, these rules are conservative, and paradoxical derivations do not threaten, even in the presence any connectives given by defining rules. The resulting theory will not be *dialetheist* — we have no derivation of $A \wedge \neg A$ for any formula A — but nonetheless, paradoxical sentences, such as $r \in r$, behave rather strangely. The statement $r \in r$ is indistinguishable from its negation $r \notin r$. Since the theory is a conservative extension of the basic logic without contraction, we can see that there is *no* derivation of $\succ r \in r$, for if there were we could extend it as follows:

$$\dfrac{\begin{array}{c}\delta\\ \vdots\\ \succ r \in r\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\delta\\ \vdots\\ \succ r \in r\end{array}}{r \notin r \succ}\,\neg L}{r \in r \succ}\,\in\! L}{\succ}\,\mathit{Cut}$$

but *Cut* is eliminable and there is no *Cut*-free derivation of the empty sequent, so $r \in r$ cannot be derived. Since $r \notin r$ is indistinguishable from $r \in r$, $\succ r \notin r$ cannot be derived either, and neither can the disjunction $\succ r \in r \vee r \notin r$, for if we had a derivation of for this instance of the law of the excluded middle, we could also derive $r \in r$. So, the logic of sets (in this *Contraction* free setting) is essentially non-classical.

$$\dfrac{\begin{array}{c}\delta\\ \vdots\\ \succ r \in r \vee r \notin r\end{array} \qquad \dfrac{r \in r \succ r \in r \qquad \dfrac{r \in r \succ r \in r}{r \notin r \succ r \in r}\,\in\! L}{r \in r \vee r \notin r \succ r \in r}\,\vee L}{\succ r \in r}\,\mathit{Cut}$$

Simlarly (and dually) there is no derivation of $r \in r \wedge r \notin r \succ$. However, since $A \otimes \neg A \succ$ and $\succ A \oplus \neg A$ are derivable (these are derivable in linear logic), we do have the *multiplicative* forms of the law of non contradiction and the law of the excluded middle, even for $r \in r$: we have $r \in r \otimes r \notin r \succ$ and $\succ r \in r \oplus r \notin r$. Furthermore, we have the following derivations:

$$\dfrac{\dfrac{\dfrac{\dfrac{r \in r \succ r \in r}{r \in r, r \notin r \succ}\,\neg L}{r \in r, r \in r \succ}\,\in\! L}{r \in r \otimes r \in r \succ}\,\otimes L}{\succ \neg(r \in r \otimes r \in r)}\,\neg R \qquad \dfrac{\dfrac{\dfrac{\dfrac{r \in r \succ r \in r}{\succ r \in r, r \notin r}\,\neg R}{\succ r \in r, r \in r}\,\in\! R}{\succ r \in r \oplus r \in r}\,\oplus R}{\neg(r \in r \oplus r \in r) \succ}\,\neg L$$

We know that there is no derivation of the negation of $r \in r$, but there *is* a derivation of the negation of $r \in r$'s self (multiplicative) conjunction, and similarly (and dually) for multiplicative disjunction.

So, all things considered, the second approach is a very different response to the paradoxes than the dialetheist's response, which takes the paradoxical derivations to encourage us to endorse at least some contradictions. The *Contraction*-free approach does no such thing. Having left *Contraction* behind, this analysis of the paradoxes puts no pressure on our account of defining rules—we can accept whatever defining rules we please, with no threat of paradox.

» «

We have considered two responses to the paradoxes that take the proposed rules $\mathsf{T}L/R$ and $\in L/R$ as given, and reflected on what that might mean for defining rules. If we are willing to jettison *Contraction*, then defining rules remain totally unscathed. If we wish to retain *Contraction* while keeping the strong rules for truth or for set membership, the terrain is more difficult. Connectives with defining rules are always a conservative extension of a prior vocabulary, but that extension need not preserve the rules for truth or set membership. In the rest of this section, I will explore a third response to the paradoxes—the response I prefer, all things considered—of *rejecting* those rules for truth and set membership.

It is important to explore this option, because the two alternatives we have seen (dialetheism, or doing without *Contraction*) have their own significant costs. For dialetheists, who claim that the position $[A, \neg A :]$ is not always out of bounds, we are owed an explanation of the connection between assertion, negation and denial. If not all assertions of $\neg A$ count as *denying* A, is there a way to express denial by way of assertion? Does the assertion of a negation sometimes express denial? If the rule $\neg Df$ does not define negation, are there some other axioms that do? For Graham Priest's favoured logic, LP, the logic of paradox, the rules for negation do not characterise the connective [201]. (It is possible for there to be two non-equivalent connectives $\neg_1$ and $\neg_2$ that both satisfy the logical laws for negation.) On what basis could we distinguish one as the real *negation*? On the other hand, if negation does always express denial, this means that assertion and denial do not always clash. If $[A : A]$ is not out of bounds, are *any* positions out of bounds at all? If so, which ones are, and why? If no positions are out of bounds, then what does logical validity constrain when it comes to combinations of assertions and denials? Furthermore, solving the paradoxes by the dialetheist route seems to save us from *some* paradoxes, but it seems that others, like Curry's paradox, require other solutions [146, 196, 200], and the simple revision of the logic of negation is nowhere near enough to maintain non-triviality. Dialetheism, however it is to be made out, comes with significant costs.

On the other hand, solving the paradoxes by rejecting *Contraction* also comes with significant costs. It is well known that representing mathematical reasoning (especially principles of induction and the like) without *Contraction* is extremely difficult, if not impossible [164]. That would be reason enough to motivate us to explore other approaches. Uwe Petersen has also shown [163, Theorem 7.3] that the reasoning that shows that the strong set principles $\in L/R$ ensure that arbitrary propositional contexts have fixed points can be generalised to show that arbitrary *functions* also have fixed points. In other words, in the context of a *Contraction*-free theory with $\in L/R$, we can show that for any function f, there is some term t such that the theory proves $f(t) = t$. If f is the successor function of arithmetic, for example, we have a fixed point, a number that is identical to its own successor. This means that the theory will look nothing like *standard* theories of arithmetic, and life without contraction will lead us to a very strange place indeed. This does not mean, of course, that such lands are not worth exploring. However, we may see fit to explore more comfortable territory, too, to see if there are other ways to develop theories, even if they include notions of *truth*, *set* and *membership*.

Any response to the paradoxes that involves rejecting the rules $\mathsf{T}L/R$ and $\in L/R$ must involve some kind of explanation of their intuitive appeal. For the rules for truth, and some kind of guide to when these rules do not apply. Starting with set membership:

$$\frac{X, \phi(t) \succ Y}{X, t \in \{x : \phi(x)\} \succ Y}\in L \qquad \frac{X \succ \phi(t), Y}{X \succ t \in \{x : \phi(x)\}, Y}\in R$$

there is a certain appeal to the transition between asserting (or denying) $\phi(t)$ and asserting (or denying) $t \in \{x : \phi(x)\}$. If I take it that t is one of the things with property $\lambda x.\phi(x)$, then I seem to thereby take it to be one of the things in the set of all those things with that property. One way to understand the appeal of this conception of set, and the rules $\in L/R$ is to think of the operation of forming sets like this: we take all those things that satisfy the property $\lambda x.\phi(x)$, and we *construct* the set whose members consist of those things. The members are those things we included. If we think of the set as *generated* or *grounded* in its members, we can see the intuitive appeal of these rules — at least in restricted cases. The operation has its appeal (so this story goes) by taking many objects and considering them as one, where the many are *prior* to the *subsequent* collection. On this picture, whenever we are given some objects, there is no problem with selecting some of those objects and forming those into a collection. The collection we thus form is a new thing. In particular, it is not among those things that were collected together, because the things collected are prior (in some sense) to the collection made out of them. We could, for example, given some things, form the collection of the non-self-membered things among them. This new set could not be self-membered, because on this conception of collection, there are *no* self-membered sets. There is no collection of *all* of non-self-membered sets, because for any non-self-membered things we collect together, the set we have thus constructed is subsequent to the things collected, and is not one of its own members.

This notion of sets, for which sets are constituted *by* their members, and hence, come after them in some ordering of dependence, is one way to motivate the appeal of the membership postulates $\in L/R$ while restricting their application to particular cases. This notion of set membership has been axiomatised and studied in great detail, in work going back to Georg Cantor [98], Ernst Zermelo [248] and Abraham Fraenkel [78], and this is the orthodoxy concerning sets in 20th and 21st Century mathematics. This is one way to understand membership, which explains why the rules $\in L/R$ have their appeal, but only in a restricted sense, and in a sense that undercuts the derivation of the paradoxes.

Øystein Linnebo gives a contemporary interpretation of this conception of sets, according to which the hierarchy of sets is always *potential* and never completed [133].

This way to understand sets and membership is just one of many potential ways to articulate the notion. It is the dominant tradition in contemporary mathematics, but other conceptions of 'set' and 'member' are possible. There is nothing essential, for example, on this picture that says that whenever we comprehend some objects, the only way to form a set is to *include* them. According to a parallel construction, we could consider some objects, and introduce a new object that *excludes* those and only those objects. We could well have two kinds of sets: those defined by what they include, and those defined by what they exclude. The *empty* set includes nothing (and so, could be at least as early as all other sets in the order of priority). What it excludes is ever growing as we find out more about the universe. Dually, the *universal* set excludes nothing (and so, could be at least as early as the empty set in the order of priority), and what it *includes* is ever growing as we find out more about the universe. Such a theory allows for sets to be closed under boolean complementation, and allows sets to be self-membered—the universal set excludes nothing, so it certainly includes itself, in exactly the same way that the empty set includes nothing, and so, excludes itself—while nonetheless avoiding the construction of a set that contains all and only the non-self-membered sets. This account of sets is due (independently) to Alonzo Church [40] and Urs Oswald [156] in the 1970s. (See recent papers by Thomas Forster for an accessible introduction to this work [76, 77].)

There are other ways to understand sets, too. The axiom of *foundation* of standard Zermelo–Fraenkel set theory tells us that there are no infinitely descending membership chains:

$$\cdots \in a_{n+1} \in a_n \in \cdots \in a_2 \in a_1 \in a_0$$

It is possible to construct models of sets in which the axiom of foundation is rejected, not because of *large* sets (as in Church–Oswald set theory), but in very small ways, such as solutions to the equation:

$$a = \{a\}$$

a set which is identical to its own singleton, in direct violation of the axiom of foundation. Here, membership cannot go along with a straightforward notion of priority (nothing is prior to itself), and neither does non-membership. Nonetheless, there is remains a notion of priority in the resulting theory of sets, the *membership graphs* may be ordered in priority. For details on non-wellfounded set theory, Peter Aczel's 1988 monograph remains a classic [1]. Yet other consistent notions of sets and membership are possible [9, 75, 116, 178], but we have seen enough to make the point. There is nothing inherently inconsistent with the notions of *set* and *membership* if it is explicated in any of these different ways.

This point we have made here is not restricted to *sets*. Before going on to consider *truth*, let's pause to see another seeming paradox, the paradox of the barber. Consider a small village, where one of the residents is a barber. This barber is free to choose who will be her clients, and who will not. This barber decides that the best course of action for her is that she shave all and only the residents who don't shave themselves. Since she is totally free to choose her clients, there is no restriction on her selection, so given the people who don't shave themselves, she is free to take them on as clients and shave them, and not shave anyone else. What makes this a paradox is what happens in her own case. If she first did *not* shave herself, then according to her principle of selection, she is now one of the people she chooses for shaving. Now she is one of her own clients, and she has taken on someone who (now) shaves herself. If, on the other hand, she *was* shaving herself, she was not one of the people she chose for shaving, and she no longer shaves herself. There is no way she can shave all and only those people who do not (then) shave themselves. But notice: this is no restriction on who she can choose to shave. For any collection of people in the village, she can choose to shave them and them only. What is impossible is that this choice be described in a particular way. If she chooses the non-self-shavers as the ones for her to shave, then either this choice means she transitions from being a self-shaver to a non-self-shaver, or she transitions from being a non-self-shaver to being a self-shaver. Her choice makes a difference in the world, and as a result, it makes a difference in what descriptions (like the predicate "*is not a self shaver*") pick out. At no point is the description stable enough to remain unchanged across the process of selection.

Notice here the similarities and differences between the role of dependence in this account of the Barber Paradox and its role in motivating the cumulative hierarchy of sets of Zermelo–Fraenkel set theory. In the cumulative hierarchy, the dependence relation induces an unbounded formation of more and more collections, each depending on what comes before. The Barber paradox induces no such thing: we do not construct new collections of residents where there were none before. (If the village numbers n people, then there are 2^n different sets of villagers. Nothing will change that.) Rather, if the predicate "*is shaved by the barber*" is determined by membership in a given set, and if we use the predicate "*is not a self shaver*" to determine membership in that set, then these descriptions are not stable, but vary, across the selection.

Perhaps what goes for the barber in the village can go for the truth predicate, too. Consider a language in which we make assertions and denials. Suppose we add to our vocabulary a device for quotation, which means that we can *name* our expressions and *describe* them, as well as assert and deny with them. And let's, then, attempt to add to our vocabulary a special predicate T with the intention that asserting $T\ulcorner A \urcorner$ has the same force as asserting A, and denying $T\ulcorner A \urcorner$ has the same force as denying A:

$$\frac{X, A \succ Y}{X, T\ulcorner A \urcorner \succ Y}\,{\scriptstyle TL} \qquad \frac{X \succ A, Y}{X \succ T\ulcorner A \urcorner, Y}\,{\scriptstyle TR}$$

You might ask: what about objectivity, or correspondence with the world? Isn't this important? Well of course it is, and we have room for it. What it means to say "*the cat sat on the mat*" *is true* depends, on this view, on what it means to say *the cat sat on the mat*. The objectivity of *truth* claims depend on the objectivity of what is said.

When it comes to sentences of our T-free vocabulary, this proposal is natural and straightforward. We can consistently carve our sentences into the T and the non-T sentences whenever we can consistently assert or deny. Our choice of what counts as *in* our *out* when it comes to the predicate T depends on (and is *subsequent to*) whether we assert or deny the sentences being described. However, once we have the facility to describe sentences in different ways, we have the ability to break out of this simple linear dependence relation. Once I can form a sentence like (λ), the liar, this picture becomes more like the barber paradox. If I attempt to govern the predicate T by saying that ruling *in* $T\ulcorner A \urcorner$ has the same force as asserting A, and ruling *out* $T\ulcorner A \urcorner$ has the same force as denying A, then if λ is the very claim $\neg T\lambda$, then enforcing this rule for T involves *revising* my interpretation for T. If I have denied $T\lambda$, this has the same force as asserting $\neg T\lambda$, which is the sentence λ itself. So, governing T by this rule of interpretation, I assign λ to the extension of T, and assert $T\lambda$. This has the same force as *denying* $\neg T\lambda$, but this is just to deny λ. So, again, governing T by the rule of interpretation, I assign λ to the *anti*-extension of T, and deny $T\lambda$, and through the cycle we go again. If we either assert, or deny λ, the cycle continues.

We first encountered the liar sentence on page 264.

What we have in TL/R is an *unstable* rule for the interpretation of the predicate T, which behaves in just the same way as the unstable extension of the "*shaves*" predicate in the village where the barber selects her clientele on the basis of whether they shave themselves. There is nothing contradictory about the behaviour of who is shaved at any particular time, and in the same way, there is no contradiction in what is true, on this interpretation of the predicate T if it is governed like this. Nonetheless, the truth predicate is *unstable*, in the presence of *this* principle of selection for the predicate T. The rules TL/R for the predicate T cannot be seen as flat descriptions which simply consistently hold, in the same way that no barber in the village can shave all and only the non-self-shavers in the village. We can *assign* to the truth predicate (or *ascribe as true*) those things we assert, but doing this could sometimes mean that we change our position on some sentences (like λ). But this is no different to the barber changing her position on *herself* if she selects all and only the non-self-shavers as her own clientele.

What we have seen here is a consistent (but unstable) way of interpreting the predicate T, which coheres with the approach to meaning through the bounds, assertion and denial. The resulting picture of the truth *predicate*, as being interpreted by a rule for interpretation that is unstable in the presence of looping reference, has been studied in the literature in recent decades. Introduced in the early 1980s by Anil Gupta, *revision theory* is a general account of circular definitions, of which the truth predicate is but one [17, 93, 94]. It is *one* way to motivate a consistent, but unstable, interpretation of the truth predicate, and it is an interpretation which coheres nicely with our approach to semantics explored here. There is much scope for further exploration, to be sure, and we will leave some of this to Chapter 8, where we will see more about norms governing singular terms and predication.

As there is with the notion of sets, there is a range of different, consistent ways to explicate the truth predicate [79, 97, 102, 111, 185, 187, 197]. Instead of exploring any of the other prospects at this stage, we will pass on from our discussion of truth back to reflection about the connectives and propositional logic. We have seen, here, that we have room to explore approaches to truth and sets and membership that avoid paradox (at least at first glance), which allow us to maintain the greatest flexibility with the structural rules available to us, and complete freedom in choosing connectives that can be given by defining rules. One reason we have such flexibility is that the notions of sets, membership and truth do not play an explanatorily central role in our semantics. We do not start with models or truth in defining consequence. We start with positions, bounds and defining rules, and use these to define derivations and logical consequence. As a result, we are free to explore how to consistently and coherently explicate the notions of truth and sets, with the tools of our logical vocabulary arising out of a more fundamental practice, that of assertion and denial. The paradoxes of truth and sets do not give us a reason to restrict or revise our logic of negation, let alone the rest of our logic. On the contrary, they confirm the wisdom of starting to account for the semantics of our concepts without having to rely on difficult and unclarified concepts. We can use the sharp tools of the classical propositional connectives to help explore different options to explicate the notions of truth, sets and membership without fear.

7.2 | DEFINING RULES AND MEANING

7.3 | PROOF AND NECESSITY

7.4 | PROOF AND WARRANT

GROUNDS AND PROOF TERMS

POSSESSING A PROOF

COMBINING GROUNDS AND THE PREFACE PARADOX

7.5 | THE EPISTEMOLOGY OF LOGIC

7.6 | LOGIC AND CONCEPTS

REFERENCES

[1] PETER ACZEL. *Non-Well-Founded Sets*. Number 14 in CSLI Lecture Notes. CSLI Publications, Stanford, 1988.

[2] JONATHAN E. ADLER AND LANCE J. RIPS, editors. *Reasoning: Studies of Human Inference and its Foundations*. Cambridge University Press, 2008.

[3] ALAN R. ANDERSON AND NUEL D. BELNAP. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, Princeton, 1975.

[4] ALAN ROSS ANDERSON, NUEL D. BELNAP, AND J. MICHAEL DUNN. *Entailment: The Logic of Relevance and Necessity*, volume 2. Princeton University Press, Princeton, 1992.

[5] PAAL ANTONSEN. "Scorekeeping". *Analysis*, 2018. Advance Access.

[6] ARNON AVRON. "Gentzen-type systems, resolution and tableaux". *Journal of Automated Reasoning*, 10(2):265–281, 1993.

[7] SORIN BANGU, editor. *Naturalizing Logico-Mathematical Knowledge: Approaches From Philosophy, Psychology and Cognitive Science*. Routledge, New York, 2018.

[8] H. P. BARENDREGT. "Lambda Calculi with Types". In SAMSON ABRAMSKY, DOV GABBAY, AND T. S. E. MAIBAUM, editors, *Handbook of Logic in Computer Science*, volume 2, chapter 2, pages 117–309. Oxford University Press, 1992.

[9] JON BARWISE AND LAWRENCE MOSS. *Vicious Circles: on the mathematics of non-wellfounded phenomena*. CSLI Publications, 1996.

[10] G. BATTILOTTI AND G. SAMBIN. "Basic Logic and the Cube of its Extensions". In ANDREA CANTINI, ETTORE CASARI, AND PIERLUIGI MINARI, editors, *Logic and Foundations of Mathematics*, pages 165–185. Kluwer, Dordrecht, 1999.

[11] JC BEALL. *Spandrels of Truth*. Oxford University Press, 2009.

[12] JC BEALL AND GREG RESTALL. "Logical Pluralism". *Australasian Journal of Philosophy*, 78:475–493, 2000. `http://consequently.org/writing/pluralism`.

[13] JC BEALL AND GREG RESTALL. *Logical Pluralism*. Oxford University Press, Oxford, 2006.

[14] NUEL BELNAP. "Declaratives are not enough". *Philosophical Studies*, 59(1):1–30, 1990.

[15] NUEL D. BELNAP. "Tonk, Plonk and Plink". *Analysis*, 22:130–134, 1962.

[16] NUEL D. BELNAP. "Display Logic". *Journal of Philosophical Logic*, 11:375–417, 1982.

[17] NUEL D. BELNAP. "Gupta's Rule of Revision Theory of Truth". *Journal of Philosophical Logic*, 11(1):105–116, 1982.

[18] NUEL D. BELNAP. "Linear Logic Displayed". *Notre Dame Journal of Formal Logic*, 31(1):15–25, 1990.

[19] NUEL D. BELNAP, JR., ANIL GUPTA, AND J. MICHAEL DUNN. "A consecutive calculus for positive relevant implication with necessity". *Journal of Philosophical Logic*, 9(4):343–362, 1980.

[20] MERRIE BERGMANN. *An Introduction to Many-Valued and Fuzzy Logic: Semantics, Algebras, and Derivation Systems*. Cambridge University Press, Cambridge; New York, 2008.

[21] WOLFGANG BIBEL. *Automated Theorem Proving*. Springer Verlag, 1982.

[22] KATALIN BIMBÓ. "$\mathrm{LE}^{t}_{\rightarrow}$, $\mathrm{LR}^{\circ}_{\wedge,\sim}$ and Cutfree Proofs". *Journal of Philosophical Logic*, 36(5):557–570, 2007.

[23] KATALIN BIMBÓ. *Proof Theory: Sequent Calculi and Related Formalisms (Discrete Mathematics and Its Applications)*. Chapman and Hall/CRC, 2014.

[24] GARRETT BIRKHOFF. *Lattice Theory*. American Mathematical Society Colloquium Publications, Providence, Rhode Island, First edition, 1940.

[25] PAUL A. BOGHOSSIAN. "Analyticity". In BOB HALE AND CRISPIN WRIGHT, editors, *A Companion to Philosophy of Language*, pages 331–368. Blackwell, Oxford, 1997.

[26] DANIEL BONEVAC. "Quantity and Quantification". *Noûs*, 19(2):229–247, 1985.

[27] DAVID BOSTOCK. *Intermediate Logic*. Clarendon Press, Oxford, 1997.

[28] ROSS T. BRADY. "The consistency of the axioms of abstraction and extensionality in a three-valued logic". *Notre Dame Journal of Formal Logic*, 12:447–453, 1971.

[29] ROBERT BRANDOM. "Asserting". *Noûs*, 17(4):637–650, 1983.

[30] ROBERT B. BRANDOM. *Making It Explicit*. Harvard University Press, 1994.

[31] ROBERT B. BRANDOM. *Articulating Reasons: an introduction to inferentialism*. Harvard University Press, 2000.

[32] VASCO BRATTKA AND GUIDO GHERARDI. "Weihrauch degrees, omniscience principles and weak computability". *The Journal of Symbolic Logic*, 76(1):143–176, 2011.

[33] JESSICA BROWN AND HERMAN CAPPELEN, editors. *Assertion: New Philosophical Essays*. Oxford University Press, 2011.

[34] RUDOLF CARNAP. *Logical Foundations of Probability*. University of Chicago Press, Chicago, 1950.

[35] ROBYN CARSTON. "Metalinguistic negation and echoic use". *Journal of Pragmatics*, 25(3):309 – 330, 1996.

[36] A. W. CARUS. *Carnap and Twentieth-Century Thought: Explication as Enlightenment*. Cambridge University Press, 2008.

[37] CARLO CELLUCCI. "The Universal Generalization Problem". *Logique et Analyse*, 52:3–20, 2009.

[38] DAVID J. CHALMERS AND BRIAN RABERN. "Two-dimensional semantics and the nesting problem". *Analysis*, 74(2):210–224, 2014.

[39] ALONZO CHURCH. *The Calculi of Lambda-Conversion*. Number 6 in Annals of Mathematical Studies. Princeton University Press, 1941.

[40] ALONZO CHURCH. "Set theory with a Universal Set". In LEON HENKIN, editor, *Proceedings of the Tarski Symposium*, volume XXV of *Proceedings of Symposia in Pure Mathematics*, page 297–308. Providence RI, 1974.

[41] PABLO COBREROS, PAUL EGRÉ, DAVID RIPLEY, AND ROBERT VAN ROOIJ. "Tolerant, Classical, Strict". *Journal of Philosophical Logic*, 41(2):347–385, 2012.

[42] ROY T. COOK. "What's wrong with tonk(?)". *Journal of Philosophical Logic*, 34(2):217–226, 2005.

[43] STEPHEN CRAIN, ANDREA GUALMINI, AND LUISA MERONI. "The acquisition of logical words". *LOGOS and Language*, 1:49–59, 2000.

[44] HASKELL B. CURRY. *A Theory of Formal Deducibility*. Number 6 in Notre Dame Mathematical Lectures. University of Notre Dame, Notre Dame, Indiana, 1950.

[45] HASKELL B. CURRY. *Foundations of Mathematical Logic*. Dover, 1977. Originally published in 1963.

[46] HASKELL B. CURRY AND R. FEYS. *Combinatory Logic*, volume 1. North Holland, 1958.

[47] HASKELL B. CURRY, J. ROGER HINDLEY, AND JONATHAN P. SELDIN. *Combinatory Logic*, volume 2. North Holland, 1972.

[48] DIRK VAN DALEN. "Intuitionistic Logic". In D. GABBAY AND F. GUENTHNER, editors, *Handbook of Philosophical Logic*, volume III. D. Reidel, Dordrecht, 1986.

[49] B. A. DAVEY AND H. A. PRIESTLEY. *Introduction to Lattices and Order*. Cambridge University Press, Cambridge, 1990.

[50] HANS DE NIVELLE, editor. *Automated Reasoning with Analytic Tableaux and Related Methods*. Springer, 2015.

[51] JOSH DEVER. "Low-grade two-dimensionalism". *Philosophical Books*, 48(1):1–16, 2007.

[52] IMOGEN DICKIE. "Negation, anti-realism, and the denial defence". *Philosophical Studies*, to appear.

[53] LEONARD EUGENE DICKSON. "Finiteness of the Odd Perfect and Primitive Abundant Numbers with n Distinct Prime Factors". *American Journal of Mathematics*, 35(4):413–422, 1913.

[54] KOSTA DOSEN. *Logical Constants: An Essay in Proof Theory*. D. Phil Thesis, University of Oxford, 1980.

[55] KOSTA DOŠEN. "Logical Constants as Punctuation Marks". *Notre Dame Journal of Formal Logic*, 30(3):362–381, 1989.

[56] KOSTA DOŠEN. "The first axiomatization of relevant logic". *Journal of Philosophical Logic*, 21(4):339–356, November 1992.

[57] KOSTA DOŠEN. "A Historical Introduction to Substructural Logics". In PETER SCHROEDER-HEISTER AND KOSTA DOŠEN, editors, *Substructural Logics*. Oxford University Press, 1993.

[58] MICHAEL DUMMETT. *Frege: Philosophy of Language*. Duckworth, 1973.

[59] MICHAEL DUMMETT. *The Logical Basis of Metaphysics*. Harvard University Press, 1991.

[60] J. MICHAEL DUNN. "Relevance Logic and Entailment". In D. GABBAY AND F. GUENTHNER, editors, *Handbook of Philosophical Logic*, volume III, pages 117–229. D. Reidel, Dordrecht, 1986.

[61] J. MICHAEL DUNN AND GARY M. HARDEGREE. *Algebraic Methods in Philosophical Logic*. Clarendon Press, Oxford, 2001.

[62] J. MICHAEL DUNN AND GREG RESTALL. "Relevance Logic". In DOV M. GABBAY, editor, *Handbook of Philosophical Logic*, volume 6, pages 1–136. Kluwer Academic Publishers, Second edition, 2002.

[63] J.MICHAEL DUNN AND NUEL D. BELNAP, JR. "The Substitution Interpretation of the Quantifiers". *Noûs*, 2(2):177–185, 1968.

[64] CATARINA DUTILH NOVAES. *Formal Languages in Logic*. Cambridge University Press, 2012.

[65] ROY DYCKHOFF. "Contraction-Free Sequent Calculi for Intuitionistic Logic". *Journal of Symbolic Logic*, 57:795–807, 1992.

[66] CLAUDIA FAGGIAN AND GIOVANNI SAMBIN. "From Basic Logic to Quantum Logics with Cut-Elimination". *International Journal of Theoretical Physics*, 37(1):31–37, 1998.

[67] HARTRY FIELD. *Saving Truth From Paradox*. Oxford University Press, 2008.

[68] HARTRY FIELD. "What is the Normative Role of Logic?". *Proceedings of the Aristotelian Society, Supplementary Volumes*, 83:251–268, 2009.

[69] HARTRY FIELD. "What is Logical Validity?". In COLIN R. CARET AND OLE T. HJORTLAND, editors, *Foundations of Logical Consequence*, pages 33–70. Oxford University Press, 2015.

[70] KIT FINE. "Vagueness, Truth and Logic". *Synthese*, 30(3–4):265–300, 1975. Reprinted in *Vagueness: A Reader* [118].

[71] KAI VON FINTEL. "What Is Presupposition Accommodation, Again?". *Philosophical Perspectives*, 22(1):137–170, 2008.

[72] F. B. FITCH. *Symbolic Logic*. Roland Press, New York, 1952.

[73] MELVIN FITTING. *First-Order Logic and Automated Theorem Proving*. Springer-Verlag, second edition, 1996.

[74] ANDREAS FJELLSTAND. "How a Semantics for Tonk Should Be". *The Review of Symbolic Logic*, 8(3):488–505, 2015.

[75] THOMAS FORSTER. *Set Theory with a Universal Set*, volume 31 of *Oxford Logic Guides*. Oxford University Press, Second edition, 1995.

[76] THOMAS FORSTER. "Church's Set Theory with a Universal Set". In C. ANTHONY ANDERSON AND MICHAEL ZELËNY, editors, *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, pages 109–138. Springer Netherlands, Dordrecht, 2001.

[77] THOMAS FORSTER. "The Iterative Conception of Set". *Review of Symbolic Logic*, 1:97–110, 2008.

[78] ABRAHAM A. FRAENKEL, YEHOSHUA BAR-HILLEL, AND AZRIEL LEVY. *Foundations of Set Theory*. Elsevier, 1958.

[79] HARVEY FRIEDMAN AND MICHAEL SHEARD. "An axiomatic approach to self-referential truth". *Annals of Pure and Applied Logic*, 33:1 – 21, 1987.

[80] GERHARD GENTZEN. "Untersuchungen über das logische Schliessen". *Math. Zeitschrift*, 39, 1934.

[81] GERHARD GENTZEN. "Untersuchungen über das logische Schließen. I". *Mathematische Zeitschrift*, 39(1):176–210, 1935.

[82] GERHARD GENTZEN. "Untersuchungen über das logische Schließen. II". *Mathematische Zeitschrift*, 39(1):405–431, 1935.

[83] GERHARD GENTZEN. *The Collected Papers of Gerhard Gentzen*. North Holland, Amsterdam, 1969.

[84] BART GEURTS. "The Mechanisms of Denial". *Language*, 74(2):274–307, 1998.

[85] PAUL C. GILMORE. "The Consistency of Partial Set Theory without Extensionality". In *Axiomatic Set Theory*, volume 13 of *Proceedings of Symposia in Pure Mathematics*, pages 147–153, Providence, Rhode Island, 1974. American Mathematical Society.

[86] JEAN-YVES GIRARD. "Linear Logic". *Theoretical Computer Science*, 50:1–101, 1987.

[87] JEAN-YVES GIRARD. *Proof Theory and Logical Complexity*. Bibliopolis, Naples, 1987.

[88] JEAN-YVES GIRARD, YVES LAFONT, AND PAUL TAYLOR. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.

[89] GEORGE GRÄTZER. *General Lattice Theory*. Academic Press, 1978.

[90] GEORGE A. GRÄTZER. *General Lattice Theory*. Birkhäuser Verlag, Basel, Second edition, 2003. With appendices by B. A. Davies, R. Freese, B.

Ganter, M. Gerefath, P. Jipsen, H. A. Priestley, H. Rose, E. T. Schmidt, S. E. Schmidt, F. Wehrung, R. Wille.

[91] V. N. GRIŠIN. "A nonstandard logic and its application to set theory". *Studies in Formalized Languages and Nonclassical Logics*, Nakua:135–171, 1974.

[92] V. N. GRIŠIN. "Predicate and Set-Theoretic Calculi Based on Logic Without Contractions". *Mathematics of the USSR-Izvestiya*, 18(1):41, 1982.

[93] ANIL GUPTA. "Truth and Paradox". *Journal of Philosophical Logic*, 11(1):1–60, 1982.

[94] ANIL GUPTA AND NUEL BELNAP. *The Revision Theory of Truth*. MIT Press, 1993.

[95] IAN HACKING. "What is Logic?". *The Journal of Philosophy*, 76:285–319, 1979.

[96] PETR HÁJEK. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publishers, 2001.

[97] VOLKER HALBACH. *Axiomatic Theories of Truth*. Cambridge University Press, 2011.

[98] LARS HALLNÄS. "On the Proof-theoretic Foundation of General Definition Theory". *Synthese*, 148(3):589–602, February 2006.

[99] CHRIS HANKIN. *Lambda Calculi: A Guide for Computer Scientists*, volume 3 of *Graduate Texts in Computer Science*. Oxford University Press, 1994.

[100] GILBERT HARMAN. *Change In View: Principles of Reasoning*. Bradford Books. MIT Press, 1986.

[101] JEAN VAN HEIJENOORT. *From Frege to Gödel: a source book in mathematical logic, 1879–1931*. Harvard University Press, Cambridge, Mass., 1967.

[102] HANS G. HERZBERGER. "Notes on Naive Semantics". *Journal of Philosophical Logic*, 11(1):61–102, 1982.

[103] AREND HEYTING. *Intuitionism: An Introduction*. North Holland, Amsterdam, 1956.

[104] OLE HJORTLAND AND SHAWN STANDEFER. "Inferentialism, Structure, and Conservativeness". In ONDREJ BERAN, VOJTECH KOLMAN, AND LADISLAV KOREN, editors, *From Rules to Meanings: New Essays on Inferentialism*, pages 115–140. Routledge, 2018.

[105] OLE THOMASSEN HJORTLAND. "Anti-exceptionalism about logic". *Philosophical Studies*, 174(3):631–658, Mar 2017.

[106] WILFRED HODGES. *Building Models by Games*. Cambridge University Press, 1985.

[107] WILFRED HODGES. "Truth in a Structure". *Proceedings of the Aristotelian Society*, 86:135–151, 1986.

[108] WILFRID HODGES. "Two Early Arabic Applications of Model-Theoretic Consequence". *Logica Universalis*, 12(1):37–54, May 2018.

[109] LAURENCE R. HORN. *A Natural History of Negation*. CSLI Press, 2001.

[110] LAURENCE R. HORN AND YASUHIKO KATO, editors. *Negation and Polarity: Syntactic and Symantic Perspectives*. Oxford Linguistics. Oxford University Press, 2000.

[111] LEON HORSTEN. *The Tarskian Turn: Deflationism and Axiomatic Truth*. The MIT Press, 2011.

[112] W. A. HOWARD. "The Formulae-as-types Notion of Construction". In J. P. SELDIN AND J. R. HINDLEY, editors, *To H. B. Curry: Essays on*

Combinatory Logic, Lambda Calculus and Formalism, pages 479–490. Academic Press, London, 1980.

[113] LLOYD HUMBERSTONE. "The Revival of Rejective Negation". *Journal of Philosophical Logic*, 29(4):331–381, 2000.

[114] LLOYD HUMBERSTONE. *The Connectives*. The MIT Press, 2011.

[115] LUCA INCURVATI AND JULIAN J. SCHLÖDER. "Weak Rejection". *Australasian Journal of Philosophy*, 95(4):741–760, 2017.

[116] RONALD BJÖRN JENSEN. "On the consistency of a slight (?) Modification of Quine's *New Foundations*". *Synthese*, 19(1):250–263, 1968.

[117] ROSANNA KEEFE. *Theories of Vagueness*. Cambridge University Press, Cambridge, UK; New York, 2000.

[118] ROSANNA KEEFE AND PETER SMITH. *Vagueness: A Reader*. Bradford Books. MIT Press, 1997.

[119] RUTH KEMPSON. "Formal Semantics and Representationalism". In CLAUDIA MAIENBORN, KLAUS VON HEUSINGER, AND PAUL PORTNER, editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 33.1 of *Handbücher zur Sprach- und Kommunikationswissenschaft*, pages 216–241. De Gruyter Mouton, 2011.

[120] G. KREISEL. "Informal Rigour and Completeness Proofs". In IMRE LAKATOS, editor, *Problems in the philosophy of mathematics*. North Holland, Amsterdam, 1967.

[121] SAUL KRIPKE. "Outline of a Theory of Truth". *The Journal of Philosophy*, 72(19):690–716, 1975.

[122] SAUL A. KRIPKE. "The Problem of Entailment". *Journal of Symbolic Logic*, 24:324, 1959. Abstract.

[123] REBECCA KUKLA AND MARK LANCE. *Yo! and Lo!: The Pragmatic Topography of the Space of Reasons*. Harvard University Press, 2009.

[124] JENNIFER LACKEY. "Norms of Assertion". *Noûs*, 41(4):594–626, 2007.

[125] JOACHIM LAMBEK. "The Mathematics of Sentence Structure". *American Mathematical Monthly*, 65(3):154–170, 1958.

[126] JOACHIM LAMBEK. "On the Calculus of Syntactic Types". In R. JACOBSEN, editor, *Structure of Language and its Mathematical Aspects*, Proceedings of Symposia in Applied Mathematics, XII. American Mathematical Society, 1961.

[127] MARK LANCE. "Quantification, Substitution, and Conceptual Content". *Noûs*, 30(4):481–507, 1996.

[128] MARK NORRIS LANCE. "Some Reflections on the Sport of Language". *Philosophical Perspectives*, 12:219–240, 1998.

[129] E. J. LEMMON. *Beginning Logic*. Nelson, 1965.

[130] WILLIAM LEVITAN, STANLEY LOMBARDO, AND BARBARA THORBURN, editors. *Abelard and Heloise: The Letters and Other Writings*. Hackett Classics. Hackett Publishing Company, Inc., 2007.

[131] DAVID LEWIS. "Probabilities of conditionals and conditional probabilities". *Philosophical Review*, 85:297–315, 1976.

[132] DAVID LEWIS. "Scorekeeping in a language game". *Journal of Philosophical Logic*, 8(1):339–359, 1979.

[133] ØYSTEIN LINNEBO. "The Potential Hierarchy of Sets". *The Review of Symbolic Logic*, 6(2):205–228, 2013.

[134] JOHN MACFARLANE. "Pragmatism and Inferentialism". In BERNHARD WEISS AND JEREMY WANDERER, editors, *Reading Brandom: On Making it Explicit*, pages 81–95. Routledge, 2010.

[135] JOHN MACFARLANE. "What Is Assertion?". In JESSICA BROWN AND HERMAN CAPPELEN, editors, *Assertion: New Philosophical Essays*. Oxford University Press, 2011.

[136] ISHANI MAITRA AND BRIAN WEATHERSON. "Assertion, knowledge, and action". *Philosophical Studies*, 149(1):99–118, 2010.

[137] EDWIN D. MARES. *Relevant Logic: A Philosophical Interpretation*. Cambridge University Press, 2004.

[138] R. L. MARTIN AND P. W. WOODRUFF. "On Representing "true-in-L" in L". *Philosophia (Israel)*, 5:213–217, 1975.

[139] VANN MCGEE. "A Counterexample to Modus Ponens". *Journal of Philosophy*, 82:462–471, 1985.

[140] RACHEL MCKINNON. *The Norms of Assertion: Truth, Lies, and Warrant*. Palgrave Innovations in Philosophy. Palgrave Macmillan, 2015.

[141] ECKART MENZLER-TROTT. *Logic's Lost Genius: The Life of Gerhard Gentzen*. American Mathematical Society, Providence, Rhode Island, 2007. Translated by Edward R. Griffor and Craig Smorńyski, with appendices by Craig Smoryński, Gerhard Genzen and Jan von Plato.

[142] ECKART MENZLER-TROTT AND JAN VON PLATO. *Gentzens Problem: Mathematische Logik Im Nationalsozialistischen Deutschland*. Birkhäuser Verlag, Basel; Boston, 2001.

[143] GEORGE METCALFE, NICOLA OLIVETTI, AND DOV M. GABBAY. *Proof Theory for Fuzzy Logics*. Springer, [Dordrecht], 2009.

[144] CONSTANT J. MEWS. *Abelard and Heloise*. Great Medieval Thinkers. Oxford University Press, 2005.

[145] ROBERT K. MEYER. "Improved Decision Procedures for Pure Relevant Logics". Typescript, 1973.

[146] ROBERT K. MEYER, RICHARD ROUTLEY, AND J. MICHAEL DUNN. "Curry's Paradox". *Analysis*, 39:124–128, 1979.

[147] MICHAEL MOORTGAT. *Categorial Investigations: Logical Aspects of the Lambek Calculus*. Foris, Dordrecht, 1988.

[148] GLYN MORRILL. *Type Logical Grammar: Categorial Logic of Signs*. Kluwer, Dordrecht, 1994.

[149] SARAH E. MURRAY. "Varieties of Update". *Semantics & Pragmatics*, 7(2):1–53, 2014.

[150] JULIEN MURZI AND OLE THOMASSEN HJORTLAND. "Inferentialism and the categoricity problem: reply to Raatikainen". *Analysis*, 69(3):480–488, 2009.

[151] REINHARD MUSKENS. "Intensional models for the theory of types". *Journal of Symbolic Logic*, 72(1):98–118, 2007.

[152] ALBERTO NAIBO AND MATTIA PETROLO. "Are Uniqueness and Deducibility of Identicals the Same?". *Theoria*, 81(2):143–181, 2015.

[153] SARA NEGRI AND JAN VON PLATO. *Structural Proof Theory*. Cambridge University Press, Cambridge, 2001.

[154] H. ONO AND Y. KOMORI. "Logic without the Contraction Rule". *Journal of Symbolic Logic*, 50:169–201, 1985.

[155] I. E. ORLOV. "The Calculus of Compatibility of Propositions". *Matematicheskii Sbornik*, 35:263–286, 1928. In Russian.

[156] URS OSWALD. *Fragmente von "New Foundations" und typentheorie*. PhD thesis, ETH, Zürich, 1976.

[157] VICTOR PAMBUCCIAN. "Early Examples of Resource-Consciousness". *Studia Logica*, 77:81–86, 2004.

[158] FRANCESCO PAOLI. *Substructural Logics: A Primer*. Springer, May 2002.

[159] LUDOVIC PATEY. "The reverse mathematics of non-decreasing subsequences". *Archive for Mathematical Logic*, 56(5):491–506, Aug 2017.
[160] CHARLES SANDERS PEIRCE. "Belief and Judgment". In CHARLES HARTSHORNE AND PAUL WEISS, editors, *Collected Papers of Charles Sanders Peirce*, volume V and VI, pages 376–388. Harvard University Press, 1935.
[161] FRANCIS J. PELLETIER. "A Brief History of Natural Deduction". *History and Philosophy of Logic*, 20(1):1–31, March 1999.
[162] UWE PETERSEN. *Die logische Grundlegung der Dialektik*. Wilhelm Fink Verlag, München, 1980.
[163] UWE PETERSEN. "Logic Without Contraction as Based on Inclusion and Unrestricted Abstraction". *Studia Logica*, 64(3):365–403, 2000.
[164] UWE PETERSEN. "$L^iD^Z_\lambda$ as a basis for PRA". *Archive for Mathematical Logic*, 42(7):665–694, 2003.
[165] PLATO. *Philebus*. C. Scribner's Sons, 1871.
[166] JAN VON PLATO. "Rereading Gentzen". *Synthese*, 137(1):195–209, 2003.
[167] JAN VON PLATO. "Gentzen's Proof of Normalization for Natural Deduction". *Bulletin of Symbolic Logic*, 14(2):240–257, 2008.
[168] K. R. POPPER. "On the Theory of Deduction". *Indagationes Mathematicae*, 10:44–54, 111–120, 1948.
[169] DAG PRAWITZ. *Natural Deduction: A Proof Theoretical Study*. Almqvist and Wiksell, Stockholm, 1965.
[170] DAG PRAWITZ. "The epistemic significance of valid inference". *Synthese*, 187(3):887–898, 2012.
[171] HUW PRICE. "Why 'Not'?". *Mind*, 99(394):222–238, 1990.
[172] GRAHAM PRIEST. "The Logic of Paradox". *Journal of Philosophical Logic*, 8(1):219–241, 1979.
[173] GRAHAM PRIEST. "Sense, Entailment and *Modus Ponens*". *Journal of Philosophical Logic*, 9(4):415–435, 1980.
[174] GRAHAM PRIEST. *Doubt Truth to be a Liar*. Oxford University Press, 2006.
[175] GRAHAM PRIEST. *In Contradiction: a study of the transconsistent*. Oxford University Press, Oxford, Second edition, 2006.
[176] ARTHUR N. PRIOR. "The Runabout Inference-Ticket". *Analysis*, 21(2):38–39, 1960.
[177] ARTHUR N. PRIOR AND KIT FINE. *Worlds, Times and Selves*. Duckworth, 1977.
[178] W. V. QUINE. "New Foundations for Mathematical Logic". *American Mathematical Monthly*, 44(2):70–80, 1937.
[179] W. V. O. QUINE. "Two Dogmas of Empiricism". *The Philosophical Review*, 60(1):20–43, 1951.
[180] WILLARD VAN ORMAN QUINE. "Carnap and Logical Truth". In *The Ways of Paradox and Other Essays*, pages 100–134. Random House, New York, 1966. (Published initially in 1954).
[181] WILLARD VAN ORMAN QUINE. "Truth by Convention". In *The Ways of Paradox and Other Essays*, pages 70–99. Random House, New York, 1966. (Published initially in 1936).
[182] PANU RAATIKAINEN. "On rules of inference and the meanings of logical constants". *Analysis*, 68(4):282–287, 2008.
[183] HELENA RASIOWA. *An Algebraic Approach to Non-Classical Logics*. North-Holland Pub. Co., Amsterdam; New York, 1974.

[184] STEPHEN READ. *Relevant logic: a philosophical examination of inference*. Basil Blackwell, Oxford, 1988.

[185] STEPHEN READ. "The Liar Paradox from John Buridan back to Thomas Bradwardine". *Vivarium*, 40(2):189–218, 2002.

[186] STEPHEN READ. "Symmetry and Paradox". *History and Philosophy of Logic*, 27:307–318, 2006.

[187] STEPHEN READ. "Plural signification and the Liar paradox". *Philosophical Studies*, 145(3):363–375, 2009.

[188] STEPHEN READ. "General-Elimination Harmony and the Meaning of the Logical Constants". *Journal of Philosophical Logic*, 39(5):557–576, 2010.

[189] STEPHEN READ. "Paradox, Closure and Indirect Speech Reports". *Logica Universalis*, 9(2):237–251, 2015.

[190] STEPHEN READ. "Anti-Exceptionalism about Logic". *under submission*, to appear.

[191] GREG RESTALL. "Deviant Logic and the Paradoxes of Self Reference". *Philosophical Studies*, 70(3):279–303, 1993.

[192] GREG RESTALL. *On Logics Without Contraction*. PhD thesis, The University of Queensland, January 1994. `http://consequently.org/writing/onlogics`.

[193] GREG RESTALL. *An Introduction to Substructural Logics*. Routledge, 2000.

[194] GREG RESTALL. "Carnap's Tolerance, Meaning and Logical Pluralism". *The Journal of Philosophy*, 99:426–443, 2002. `http://consequently.org/writing/carnap/`.

[195] GREG RESTALL. "Multiple Conclusions". In PETR HÁJEK, LUIS VALDÉS-VILLANUEVA, AND DAG WESTERSTÅHL, editors, *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 189–205. KCL Publications, 2005. `http://consequently.org/writing/multipleconclusions`.

[196] GREG RESTALL. "Curry's Revenge: the costs of non-classical solutions to the paradoxes of self-reference". In JC BEALL, editor, *The Revenge of the Liar*, pages 262–271. Oxford University Press, 2008.

[197] GREG RESTALL. "Modal models for Bradwardine's theory of truth". *Review of Symbolic Logic*, 1(2):225–240, 2008.

[198] GREG RESTALL. "Models for Liars in Bradwardine's Theory of Truth". In SHAHID RAHMAN, TERO TULENHEIMO, AND EMMANUEL GENOT, editors, *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox*, pages 135–147. Springer, 2008.

[199] GREG RESTALL. "Truth Values and Proof Theory". *Studia Logica*, 92(2):241–264, 2009. `http://consequently.org/writing/tvpt/`.

[200] GREG RESTALL. "Assertion, Denial and Non-Classical Theories". In KOJI TANAKA, FRANCESCO BERTO, EDWIN MARES, AND FRANCESCO PAOLI, editors, *Paraconsistency: Logic and Applications*, pages 81–99. Springer, 2013.

[201] GREG RESTALL. "Two Negations are More than One". In CAN BAŞKENT AND THOMAS MACAULAY FERGUSON, editors, *Graham Priest on Dialetheism and Paraconsistency*. Springer, to appear.

[202] DAVID RIPLEY. "Negation, Denial, and Rejection". *Philosophy Compass*, 6(9):622–629, 2011.

[203] DAVID RIPLEY. "Paradoxes and Failures of Cut". *Australasian Journal of Philosophy*, 91(1):139–164, 2013.

[204] DAVID RIPLEY. "Anything Goes". *Topoi*, 34(1):25–36, Apr 2015.

[205] DAVID RIPLEY. *Uncut*. in progress, to appear.

[206] FLORIS ROELOFSEN AND DONKA F. FARKAS. "Polarity particle responses as a window onto the interpretation of questions and assertions". *Language*, 91(2):359–414, 2015.

[207] RICHARD ROUTLEY, VAL PLUMWOOD, ROBERT K. MEYER, AND ROSS T. BRADY. *Relevant Logics and their Rivals*. Ridgeview, 1982.

[208] IAN RUMFITT. ""Yes" and "No"". *Mind*, 109(436):781–823, 2000.

[209] IAN RUMFITT. *The Boundary Stones of Thought: An Essay in the Philosophy of Logic*. Oxford University Press, 2015.

[210] GILLIAN K. RUSSELL. "Metaphysical Analyticity and the Epistemology of Logic". To appear in *Philosophical Studies*, 2013.

[211] NATHAN SALMON. *Frege's Puzzle*. MIT Press, 1983.

[212] GIOVANNI SAMBIN. *The Basic Picture: structures for constructive topology*. Oxford Logic Guides. Oxford University Press, 2012.

[213] GIOVANNI SAMBIN, GIULIA BATTILOTTI, AND CLAUDIA FAGGIAN. "Basic logic: reflection, symmetry, visibility". *Journal of Symbolic Logic*, 65(3):979–1013, 2014.

[214] J. J. SCHLÖDER. *Assertion and rejection*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2018.

[215] MOSES SCHÖNFINKEL. "Über die Bausteine der mathematischen Logik". *Math. Annallen*, 92:305–316, 1924. Translated and reprinted as "On the Building Blocks of Mathematical Logic" in *From Frege to Gödel* [101].

[216] DANA SCOTT. "Completeness and axiomatizability in many-valued logic". In LEON HENKIN, editor, *Proceedings of the Tarski Symposium*, volume 25, pages 411–436. American Mathematical Society, Providence, 1974.

[217] DANA SCOTT. *Rules and Derived Rules*, pages 147–161. Springer, Dordrecht, 1974.

[218] DANA SCOTT. "Lambda Calculus: Some Models, Some Philosophy". In J. BARWISE, H. J. KEISLER, AND K. KUNEN, editors, *The Kleene Symposium*, pages 223–265. North Holland, Amsterdam, 1980.

[219] WILFRID SELLARS. "Empiricism and the Philosophy of Mind". In H. FEIGL AND M. SCRIVEN, editors, *Minnesota Studies in the Philosophy of Science*, volume I, pages 253–329. University of Minnesota Press, Minneapolis, MN, 1956. Reprinted in 1997 [221].

[220] WILFRID SELLARS. "Philosophy and the Scientific Image of Man". In ROBERT COLODNY, editor, *Frontiers of Science and Philosophy*. University of Pittsburgh Press, 1962.

[221] WILFRID SELLARS. *Empiricism and the Philosophy of Mind*. Harvard University Press, Cambridge, MA, 1997. Republication of [219], with an with an Introduction by Richard Rorty and Study Guide by Brandom.

[222] WILFRID SELLARS. "Language as Thought and as Communication". In KEVIN SCHARP AND ROBERT B. BRANDOM, editors, *In the Space of Reasons: Selected Essays of Wilfrid Sellars*, pages 57–80. Harvard University Press, Cambridge, Massachusetts, 2007.

[223] STEWART SHAPIRO. "Logical Consequence: Models and Modality". In MATTHIAS SCHIRN, editor, *The Philosophy of Mathematics*. Oxford University Press, 1998.

[224] STEWART SHAPIRO. "Logical Consequence, Proof Thory, and Model Theory". In STEWART SHAPIRO, editor, *The Oxford Handbook of Philosophy of Mathematics and Logic*, pages 651–670. Oxford University Press, Oxford, 2005.

[225] MANDY SIMONS. "Presupposition and Accommodation: Understanding the Stalnakerian Picture". *Philosophical Studies*, 112(3):251–278, Feb 2003.

[226] JOHN K. SLANEY. "A General Logic". *Australasian Journal of Philosophy*, 68:74–88, 1990.

[227] TIMOTHY SMILEY. "Rejection". *Analysis*, 56:1–9, 1996.

[228] NICHOLAS J. J. SMITH. *Vagueness and Degrees of Truth*. Oxford University Press, Oxford, 2008.

[229] ROBERT STALNAKER. "Assertion". In PETER COLE, editor, *Syntax and Semantics*, volume 9, pages 315–322. Academic Press, New York, 1978.

[230] ROBERT STALNAKER. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford Cognitive Science Series. Oxford University Press, 1999.

[231] ROBERT STALNAKER. "Common Ground". *Linguistics and Philosophy*, 25:701–721, 2002.

[232] FLORIAN STEINBERGER. "What Harmony Could and Could Not Be". *Australasian Journal of Philosophy*, 89(4):617–639, 2011.

[233] KEITH STENNING AND MICHIEL VAN LAMBALGEN. *Human Reasoning and Cognitive Science*. MIT Press, Cambridge, Mass., 2008.

[234] J. T. STEVENSON. "Roundabout the Runabout Inference-Ticket". *Analysis*, 21(6):124–128, 1961.

[235] P. SUPPES. *Introduction to Logic*. van Nostrand, Princeton, 1957.

[236] W. W. TAIT. "Intensional Interpretation of Functionals of Finite Type I". *Journal of Symbolic Logic*, 32:198–212, 1967.

[237] NEIL TENNANT. *Natural Logic*. Edinburgh University Press, Edinburgh, 1978.

[238] PAUL THISTLEWAITE, MICHAEL MCROBBIE, AND ROBERT K. MEYER. *Automated Theorem Proving in Non-Classical Logics*. Wiley, New York, 1988.

[239] A. S. TROELSTRA. *Lectures on Linear Logic*. CSLI Publications, 1992.

[240] A. S. TROELSTRA AND H. SCHWICHTENBERG. *Basic Proof Theory*, volume 43 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, Cambridge, second edition, 2000.

[241] MICHAEL TYE. "Sorites Paradoxes and the Semantics of Vagueness". *Philosophical Perspectives*, 8:189–206, 1994.

[242] FRANK VELTMAN. "Defaults in update semantics". *Journal of Philosophical Logic*, 25(3):221–261, 1996.

[243] HEINRICH WANSING. "Connectives stranger than tonk". *Journal of Philosophical Logic*, 35(6), 2006.

[244] DALLAS WILLARD. "Degradation of logical form". *Axiomathes*, 8(1):31–52, Dec 1997.

[245] TIMOTHY WILLIAMSON. *Vagueness*. Routledge, London; New York, 1994.

[246] TIMOTHY WILLIAMSON AND PETER SIMONS. "Vagueness and Ignorance". *Aristotelian Society Proceedings*, 66:145–162, 1992. suppl. volume 66.

[247] HENK ZEEVAT. "Presupposition and Accommodation in Update Semantics". *Journal of Semantics*, 9(4):379–412, 1992.

[248] ERNST ZERMELO. "Untersuchungen über die Grundlagen der Mengenlehre, I". *Mathematische Annalen*, 65:261–281, 1908. Reprinted in [101, pages 201–215].