# PROOF THEORY & PHILOSOPHY

*Greg Restall*

Philosophy Department
University of Melbourne

restall@unimelb.edu.au
http://consequently.org/writing/ptp

VERSION OF SEPTEMBER 18, 2006

# CONTENTS

1 *Why Proof Theory?* | 9

## Part I Propositional Logic

2 *Propositional Logic: Tools & Techniques* | 15
- 2.1 Natural Deduction for Conditionals · 15
- 2.2 Sequents and Derivations · 53
- 2.3 From Proofs to Derivations and Back · 69
- 2.4 Circuits · 99
- 2.5 Counterexamples · 116
- 2.6 Proof Identity · 128

3 *Propositional Logic: Applications* | 129
- 3.1 Assertion and Denial · 129
- 3.2 Definition and Harmony · 138
- 3.3 Tonk and Non-Conservative Extension · 141
- 3.4 Meaning · 144
- 3.5 Achilles and the Tortoise · 147
- 3.6 Warrant · 147
- 3.7 Gaps and Gluts · 147
- 3.8 Realism · 147

## Part II Quantifiers, Identity and Existence

4 *Quantifiers: Tools & Techniques* | 151
- 4.1 Predicate Logic · 151
- 4.2 Identity and Existence · 151
- 4.3 Models · 152
- 4.4 Arithmetic · 153
- 4.5 Second Order Quantification · 153

5 *Quantifiers: Applications* | 155
- 5.1 Objectivity · 155
- 5.2 Explanation · 155
- 5.3 Relativity · 155
- 5.4 Existence · 155
- 5.5 Consistency · 155
- 5.6 Second Order Logic · 155

# Part III Modality and Truth

6 *Modality and Truth: Tools and Techniques* | 159
- 6.1 Simple Modal Logic · 159
- 6.2 Modal Models · 159
- 6.3 Quantified Modal Logic · 159
- 6.4 Truth as a Predicate · 159

7 *Modality and Truth: Applications* | 161
- 7.1 Possible Worlds · 161
- 7.2 Counterparts · 161
- 7.3 Synthetic Necessities · 161
- 7.4 Two Dimensional Semantics · 161
- 7.5 Truth and Paradox · 161

*References* | 163

# INTRODUCTION

This manuscript is a draft of a guided introduction to logic and its applications in philosophy. The focus will be a detailed examination of the different ways to understand *proof*. Along the way, we will also take a few glances around to the other side of inference, the kinds of *counterexamples* to be found when an inference 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 at least these four aims in mind.

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

A GENTLE INTRODUCTION TO KEY IDEAS IN THE THEORY OF PROOF: The literature on proof theory contains some very good introductions to the topic. Bostock's *Intermediate Logic* [9], Tennant's *Natural Logic* [87], Troelstra and Schwichtenberg's *Basic Proof Theory* [91], and von Plato and Negri's *Structural Proof Theory* [56] 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 somewhat 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 2.1.2 we will see a two-page normalisation proof. In Section 2.2.3 we will see a two-page cut-elimination proof. In each case, the aim is to understand the key concepts behind the central results.

AN INTRODUCTION TO LOGIC FROM A NON-PARTISAN, PLURALIST, PROOF-THEORETIC PERSPECTIVE: We are able to take this liberal approach to introducing proof theory because we take a pluralist attitude to the choice of logical system. This book is designed to be an introduction to *logic* that does not have a distinctive axe to grind in favour of a

particular logical system. Instead of attempting to justify this or that formal system, we will give an overview of the panoply of different accounts of consequence for which a theory of proof has something interesting and important to say. As a result, in Chapter 2 we will examine the behaviour of conditionals from intuitionistic, relevant and linear logic. The system of natural deduction we will start off with is well suited to them. In Chapter 2, we also look at a sequent system for the non-distributive logic of conjunction and disjunction, because this results in a very simple cut elimination proof. From there, we go on to more rich and complicated settings, once we have the groundwork in place.

AN INTRODUCTION TO THE APPLICATIONS OF PROOF THEORY: We will always have our eye on the kinds of concerns others have concerning proof theory. What are the connections between proof theories and theories of meaning? What does an 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 constitutes what things *mean*. For everyone, there is an intimate connection between inference and semantics.

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* [11].

A PRESENTATION OF NEW RESULTS: Recent work in proofnets and other techniques in non-classical logics like linear logic can usefully illuminate the theory of much more traditional logical systems, like classical logic itself. I aim to present these results in an accessible form, and extend them to show how you can give a coherent picture of classical and non-classical propositional logics, quantifiers and modal operators.

An accessible example of this work is Robinson's "Proof nets for Classical Logic" [80].

The book is filled with marginal notes which 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 and each part divides into two chapters. The parts cover different aspects of logical vocabulary. First, *propositional logic*; second, *quantifiers, identity and existence*; third, *modality and truth*. In each part, the first chapter covers logical tools and techniques suited to the topic under examination. The second chapter both discusses the issues that are raised in the tools & techniques chapter, and applies these tools and techniques to different issues in philosophy of language, metaphysics, epistemology, philosophy of mathematics and elsewhere.

Each 'Tools & Techniques' chapter contains many 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, intermediate and advanced ones. 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.

Greg Restall

Melbourne

september 18, 2006

# WHY PROOF THEORY? 1

*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 way 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 as follows:

> ⌜p or q⌝ is true if and only if ⌜p⌝ is true or ⌜q⌝ 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

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

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 3.

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 seem to be very difficult to come to universal conclusions. However, we 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?

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 step, or is there some other account of the logical structure of possibility and necessity?

These are three examples of the kinds of issues that we will consider in the light of proof theory. Before we can broach these topics, we need to learn some proof theory. We will start with proofs for conditional judgements.

# PART I

# PROPOSITIONAL LOGIC

# PROPOSITIONAL LOGIC: TOOLS & TECHNIQUES

# 2

## 2.1 | NATURAL DEDUCTION FOR CONDITIONALS

We start with modest ambitions. In this section we focus on one way of understanding inference and proof—natural deduction, in the style of Gentzen [33]—and we will consider just one kind of judgement: *conditionals*. Conditional judgements are judgements of the form

If ... then ...

To make things precise, we will use a formal language in which we can express conditional judgements. Our language will have an unending supply of *atomic* formulas

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

When we need to refer to the collection of all atomic formulas, we will call it 'ATOM.' Whenever we have two formulas A and B, whether A and B are in ATOM or not, we will say that $(A \to B)$ is also a formula. Succinctly, this *grammar* can be represented as follows:

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

That is, a formula is either an ATOM, or is found by placing an arrow ('→') between two formulas, and surrounding the result with parentheses. So, these are 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:

$$t \quad p \to q \to r \quad p \to p$$

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:

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

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/Mathematicians/Gentzen.html

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

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*.

Our last example of an offending 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 in FORMULA.

If you like, you can think of them as including their outer parentheses very *faintly*, like 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 these formulas in at least two different ways. We can think of them as the sentences in a toy 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 themselves, but as constructions used to display the *form* of expressions in a language. Nothing here will stand on which way you understand formulas. In either case, we use the conditional $p \to q$ to represent the conditional proposition with antecedent $p$ and consequent $q$.

Sometimes, we will want to talk quite generally about all formulas of a particular form. We will want to do this very often, when it comes to logic, because we are interested in the *structures* or *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

> *Given a conditional formula, and its antecedent, its consequent follows.*

This can get very complicated very quickly. It is not at all convenient to say

> *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 like

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

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

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

which seems much more perspicuous and memorable. Now we have the raw formal materials to address the question of deduction using conditional judgements. How may we characterise valid reasoning using conditional constructions? We will look at one way of addressing this topic in this section.

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 variables.

## 2.1.1 | PROOFS FOR CONDITIONALS

Start with a piece of reasoning using conditional judgements. One example might be this:

> *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* or *assumptions*. We also infer *from* these assumptions. 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, $B \to C$ 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*.[1] It's easy to see that

[1] "*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* $A$.

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 assume 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 have a valid *argument* to the conclusion that C if we make the *assumption* that A.

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

So, it seems we can reason like this. 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 infer to C *using* A as an assumption, we can conclude $A \to C$. 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, 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, $A \to C$.

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

Figure 2.1: NATURAL DEDUCTION RULES FOR CONDITIONALS

We have motivated two rules of inference. These rules are displayed in Figure 2.1. 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 discharged in this step. Its *conclusion* is $A \to B$.

definition 2.1.1 [proofs for conditionals] A proof is a *tree*, whose nodes are either formulas, or *bracketed* formulas. The formula at the root of the tree is said to be the *conclusion* of the proof. 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.

» If $\pi_l$ is a proof, with conclusion $A \to B$ and $\pi_r$ is a proof, with conclusion A, then the following tree

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

is a proof with conclusion B, and having the premises consisting of the premises of $\pi_l$ together with the premises of $\pi_r$.

» If $\pi$ is a proof, for which A is one of the premises and B is the conclusion, then the following tree

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

is a proof of $A \to B$.

» Nothing else is a proof.

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

This is a recursive definition, in just the same manner as the recursive definition of the class formula.

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

suffixing (inference)

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

assertion (formula)

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

$$\dfrac{}{(A \to B) \to ((C \to A) \to (C \to B))}\ {\to}I,3$$

prefixing (formula)

Figure 2.2: three implicational proofs

Figure 2.2 gives three proofs of implicational 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

$\rightarrow$ 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.

You can try a number of different strategies when making proofs for yourself. For example, you might like to try your hand at constructing a proof to the conclusion that $B \rightarrow (A \rightarrow C)$ from the assumption $A \rightarrow (B \rightarrow C)$. Here are two ways to piece the proof together.

CONSTRUCTING PROOFS TOP-DOWN: You start with the assumptions and see what you can do with them. In this case, with $A \rightarrow (B \rightarrow C)$ you can, clearly, get $B \rightarrow C$, 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 $B \rightarrow (A \rightarrow C)$ if we discharge our assumptions, A first, and then B.

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

I have been intentionally unspecific when it comes to discharging formulas in proofs. In the examples in Figure 2.2 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 in one go.

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

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

$$\dfrac{A}{B \rightarrow A}\rightarrow I,1$$

Here, we assume A, and then, we infer $B \to A$ 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: $B \to A$.

You might think that this is silly—how, after all, can you discharge a nonexistent assumption? Nonetheless, discharging assumptions that are not there plays a role in what will follow. To give you a foretaste of why, notice that the inference, from A to $B \to A$, is *valid* if we read "$\to$" 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 work on a "pluralist spirit," see my work with JC Beall [4, 5, 77].

DEFINITION 2.1.2 [DISCHARGE POLICY] A DISCHARGE POLICY may either allow or disallow *duplicate* discharges (discharging more than one instance of a formula at once) or *vacuous* discharges (discharging *zero* instances of a formula in a discharge step). We can present names for the different discharge policies possible in a table.

| | VACUOUS OK | VACUOUS NOT OK |
|---|---|---|
| DUPLICATES OK | *Standard* | *Relevant* |
| DUPLICATES NOT OK | *"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 combinations. As I indicated above, you might think that vacuous discharge is a bit silly. It is not merely *silly*: it seems downright *wrong* if you think that a judgement of the form $A \to B$ records the claim that B may be inferred *from* A.[2] 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 [1, 2, 71].

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 [58]. 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 $A \to (A \to B)$ and $A \to B$. We shall see more of the distinctive properties of different discharge policies as the book progresses.

DEFINITION 2.1.3 [DISCHARGE IN PROOFS] An 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*.

[2] You must be careful if you think that more than one discharge policy is OK. Consider Exercise 19 at the end of this section, in which it is shown that if you have two conditionals $\to_1$ and $\to_2$ with different discharge policies, the conditionals *collapse* into one (in effect having the most lax discharge policy of either $\to_1$ or $\to_2$). Consider Exercise 23 to explore how you might have the one logic with more than one conditional connective.

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 $B$:

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.

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

We shall see later that there is *no* linear proof from $A \to (A \to B), A$ to $B$. 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 2.1.4 [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 $f_M(x)$, the number of occurrences of the object $x$ in the multiset $M$. The number $f_M(x)$ is sometimes said to be the *degree* to which $x$ is a member of $M$.

If you like, you could *define* a multiset of formulas to *be* the function $f_M(x)$ function $f$ : FORMULA $\to \omega$ from formulas to counting numbers. Then $f = g$ when $f(A) = g(A)$ for each formula $A$. $f(A)$ is the number of times $A$ is in the multiset $f$. A *finite* multiset is a multiset $f$ such that $f(A) > 0$ for only finitely many objects $A$.

The multiset $M$ is *finite* if $f_M(x) > 0$ for only finitely many objects $x$. The multiset $M$ is identical to the multiset $M'$ if and only if $f_M(x) = f_{M'}(x)$ 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 $[\,]$.

DEFINITION 2.1.5 [COMPARING MULTISETS] When $M$ and $M'$ are multisets and $f_M(x) \leqslant f_{M'}(x)$ 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 52.) 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. So, we will consider arguments in the following way.

definition 2.1.6 [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 2.1.7 [validity facts] *Let $\nu$-validity be any of linear, relevant, affine or standard validity.*

1. $A \therefore A$ *is valid.*
2. $X, A \therefore B$ *is $\nu$-valid if and only if* $X \therefore A \to B$ *is $\nu$-valid.*
3. *If* $X, A \therefore B$ *and* $Y \therefore A$ *are both $\nu$-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*: It is not difficult to verify these claims. The first is given by the proof consisting of $A$ as premise and conclusion. For the second, take a proof $\pi$ from $X, A$ to $B$, and in a single step $\to I$, discharge the (single instance of) $A$ to construct the proof of $A \to B$ from $X$. Conversely, if you have a proof from $X$ to $A \to B$, add a (single) premise $A$ and apply $\to E$ to derive $B$. In both cases here, if the original proofs satisfy a constraint (vacuous or multiple discharge) so do the new proofs.

For the third fact, 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 the fourth fact, if we have a proof $\pi$ from $X$ to $B$, we can extend this as follows

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

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

Finally, if we have a proof $\pi$ 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 $A \to B$ and then reinstate a single instance of A to as a premise to derive B again.

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

Now, we might focus our attention on the distinction between those arguments that are valid and those that are not—to focus on 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 existential fact (for example, if there is a proof from X, A to B, then there is a proof from X to $A \to B$) 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 $A \to B$ in this uniform way).

» «

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. Showing that an argument is not valid seems more difficult. 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. But there are infinitely many proofs! You cannot simply go 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 is not valid. To accomplish this task, subtlety is called for. We will end this section by looking at how we might summon up the required skill.

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 Section 2.5 when we show that all valid arguments *preserve truth in 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 section, on the other hand, we will not go beyond be conceptual bounds of the study of proof. We will find instead a way to show that an argument is invalid, using an analysis of proofs themselves. 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 – might 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 success in showing that there is no proof from X to A.

Happily, this technique works. But to make this work we need to understand what it is for a proof to be "direct" in some salient sense. Direct proofs have a name—they are 'normal'.

I *think* that the terminology 'normal' comes from Prawitz [63], though the idea comes from Gentzen.

## 2.1.2 | NORMAL PROOFS

It is best to introduce normal proofs by contrasting them with non-normal proofs. And non-normal proofs are not difficult to find. Many proofs are quite strange. Take a proof that concludes with an implication introduction: it infers from A to B by way of the sub-proof $\pi_1$. Then we discharge the A to conclude $A \to B$. Imagine that at the very next step, it uses a different proof – call it $\pi_2$ – 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 $\pi_1$ of B, but instead of taking A as an *assumption*, we could use the proof of A we have at hand, namely $\pi_2$. The before-and-after comparison is this:

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

The result is a proof of B from the same premises as our original proof. The premises are the premises of $\pi_1$ (other than the instances of A that were discharged in the other proof) together with the premises of $\pi_2$. This proof does not go through the formula $A \to B$, so it is, in a sense, simpler.

Well ... there are some subtleties with counting, as usual with our proofs. If the discharge of A was vacuous, then we have nowhere to plug in the new proof $\pi_2$, so the premises of $\pi_2$ 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 $\pi_2$, and hence, that many copies of the premises of $\pi_2$. Let's make this discussion more explicit, by considering an example where $\pi_1$ has two instances of A in the premise list. The original proof containing the

introduction and then elimination of $A \to B$ is

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

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.

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

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 formal definition:

DEFINITION 2.1.8 [NORMAL PROOF] A proof is NORMAL if and only if the concluding formula $A \to B$ of an $\to I$ step is not at the same time the major premise of an $\to E$ step.

DEFINITION 2.1.9 [INDIRECT PAIR; DETOUR FORMULA] If a formula $A \to B$ introduced in an $\to I$ step *is* at the same time the major premise of an $\to E$ step, 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 the 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 $\to I/\to E$ step that you

*cannot* prove without it would seem to be quite 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 $A \to B$ 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 2.1.10 [SUBFORMULAS AND PARSE TREES] The PARSE TREE for an atom is that atom itself. The PARSE TREE for a conditional $A \to B$ is the tree containing $A \to B$ 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 $A$. $\mathrm{sf}(p) = \{p\}$, and $\mathrm{sf}(A \to B) = \{A \to B\} \cup \mathrm{sf}(A) \cup \mathrm{sf}(B)$. 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 $(p \to q) \to ((q \to r) \to p)$:

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

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

We may prove the following theorem.

THEOREM 2.1.11 [THE SUBFORMULA THEOREM] *Each normal proof from the premises $X$ to the conclusion $A$ 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 $A$.

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

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 $A$.

*Proof*: To prove the subformula theorem, we need to look carefully at how proofs are constructed. If $\pi$ 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, $\pi$ either is a lone assumption, or $\pi$ 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 $\to I$, $\pi$ is constructed from another normal proof $\pi'$ from $X$ to $B$, with the new step added on (and with the discharge of a number – possibly zero – of assumptions). $\pi$ is a proof from $X'$ to $A \to B$, where $X'$ is $X$ with the deletion of some number of instances of $A$. Since $\pi'$ is normal, we may assume that every formula in $\pi'$ 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 $\pi$ (namely, those formulas in $\pi'$, together with $A \to B$) is in $\mathrm{sf}(X', A \to B)$ as desired.

ELIMINATION In the case of $\to E$, $\pi$ is constructed out of *two* normal proofs: one (call it $\pi_1$) to the conclusion of a conditional $A \to B$ from premises $X$, and the other (call it $\pi_2$) to the conclusion of the antecedent of that conditional $A$ from premises $Y$. Both $\pi_1$ and $\pi_2$ are normal, so we may assume that each formula in $\pi_1$ is in $\mathrm{sf}(X, A \to B)$ and each formula in $\pi_2$ is in $\mathrm{sf}(Y, A)$. We wish to show that every formula in $\pi$ 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: $\pi_1$ cannot end in the *introduction* of the conditional $A \to B$. So, $\pi_1$ is either the assumption $A \to B$ itself (in which case $Y = A \to B$, and clearly in this case each formula in $\pi$ *is* in $\mathrm{sf}(X, A \to B, B)$) or $\pi_1$ ends in a $\to E$ step. But if $\pi_1$ ends in an $\to E$ step, the major premise of that inference is a formula of the form $C \to (A \to B)$. So $\pi_1$ 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, $A \to B \in sf(Y)$. In this case too, every formula in $\pi$ is in $sf(X, Y, B)$, 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 useful to work with. Even though an argument might have very many proofs, it will have many fewer *normal* proofs, and we can exploit this fact.

example 2.1.12 [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 $sf(p, q)$: 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 $q$.

Consider 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 2.1.13 [normalisation theorem] *A proof $\pi$ from $X$ to $A$ reduces in some number of steps to a proof $\pi'$ from $X'$ to $A$.*

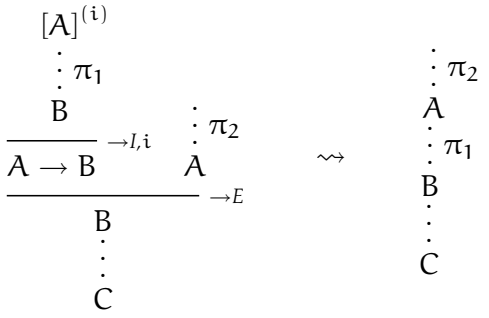
*If $\pi$ is linear, so is $\pi'$, and $X = X'$. If $\pi$ is affine, so is $\pi'$, and $X'$ is a sub-multiset of $X$. If $\pi$ is relevant, then so is $\pi'$, and $X'$ covers the same ground as $X$, and is a super-multiset of $X$. If $\pi$ is standard, then so is $\pi'$, and $X'$ covers no more ground 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 new premises. 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 2.1.14 [REDUCTION] A proof $\pi$ *reduces* to $\pi'$ ($\pi \rightsquigarrow \pi'$) if some indirect pair in $\pi$ is eliminated in the usual way.




If there is no $\pi'$ such that $\pi \rightsquigarrow \pi'$, then $\pi'$ is normal. If $\pi_0 \rightsquigarrow \pi_2 \rightsquigarrow \cdots \rightsquigarrow \pi_n$ we write "$\pi_0 \rightsquigarrow_* \pi_n$" and we say that $\pi_0$ reduces to $\pi_n$ in a number of steps. We aim to show that for any proof $\pi$, there is some normal $\pi^*$ 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 2.1.13*: linear and affine cases*]: If $\pi$ 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 $\pi_2$ for A is inserted into the proof $\pi_1$. (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 $\pi_2$ 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 26 goes from a proof with 11 formulas to another proof with 11 formulas. The result is no smaller, so *size* 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 2.1.15 [COMPLEXITY] A formula's *complexity* is the number of connectives in that formula. In this case, it is the number of instances of '$\rightarrow$' 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 $A \rightarrow B$, and suppose that otherwise, $\pi_1$ (the proof of $B$ from $A$) and $\pi_2$ (the proof of $A$) are normal.

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

Unfortunately, the new proof is not necessarily normal. The new proof is non-normal if $\pi_2$ ends in the introduction of $A$, while $\pi_1$ 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 $A \rightarrow B$, or any other formula that complex. The potential non-normality is with respect to a subformula $A$. This result would still hold if the proofs $\pi_1$ and $\pi_2$ weren't normal themselves, but when they might have $\rightarrow I/\rightarrow E$ pairs for formulas less complex than $A \rightarrow B$. If $A \rightarrow B$ is the most complex detour formula in the original proof, then the new proof has a *smaller* most complex detour formula.

DEFINITION 2.1.16 [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 any infinite descending chain of measures below it. 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 $\langle\rangle$. 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 2.1.17 [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 $\pi$ 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 $\pi'$ 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 2.1.13: *relevant and standard case*]: Start with $\pi$, a proof that isn't normal, and use Lemma 2.1.17 to choose a proof $\pi'$ with a lower measure of non-normality. If $\pi'$ 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.

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

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 damage to the number of times a premise is used.

» «

It is very tempting to view normalisation as a way of eliminating redundancies and making explicit the structure of a proof. However, if that is the case, then it should be the case that the process of normalisation cannot give us two distinct "answers" for the structure of the one proof. Can two different reduction sequences for a single proof result in *different* normal proofs? To investigate this, we need one more notion of reduction.

This passage is the hardest part of Section 2.1. Feel free to skip over the proofs of theorems in this section, until page 38 on first reading.

definition 2.1.18 [parallel reduction] A proof $\pi$ *parallel reduces* to $\pi'$ if some number of indirect pairs in $\pi$ are eliminated in parallel. We write "$\pi \overset{=}{\rightsquigarrow} \pi'$."

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

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

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

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

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

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

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

LEMMA 2.1.19 [DIAMOND PROPERTY FOR $\overline{\rightsquigarrow}$] *If* $\pi \overline{\rightsquigarrow} \pi_1$ *and* $\pi \overline{\rightsquigarrow} \pi_2$ *then there is some* $\pi'$ *where* $\pi_1 \overline{\rightsquigarrow} \pi'$ *and* $\pi_2 \overline{\rightsquigarrow} \pi'$.

**To do**: This proof sketch should be made more precise.

*Proof* [*sketch*]: Take the detour formulas in $\pi$ that are eliminated in the move to $\pi_1$ or to $\pi_2$. For those *not* eliminated in the move to $\pi_1$, mark their corresponding occurrences in $\pi_1$. Similarly, mark the occurrences of formulas in $\pi_2$ that are detour formulas in $\pi$ that are eliminated in the move to $\pi_1$. Now eliminate the marked formulas in $\pi_1$ and those in $\pi_2$ to produce the proof $\pi'$. ∎

THEOREM 2.1.20 [ONLY ONE NORMAL FORM] *Any sequence of reduction steps from a proof* $\pi$ *that terminates, terminates in a* unique *normal proof* $\pi^*$.

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

$$\pi \overline{\rightsquigarrow} \pi'_1 \overline{\rightsquigarrow} \pi'_2 \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} \pi'_n \overline{\rightsquigarrow} \pi'$$

$$\pi \overline{\rightsquigarrow} \pi''_1 \overline{\rightsquigarrow} \pi''_2 \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} \pi''_m \overline{\rightsquigarrow} \pi''$$

By the diamond property, we have a $\pi_{1,1}$ where $\pi'_1 \overline{\rightsquigarrow} \pi_{1,1}$ and $\pi''_1 \overline{\rightsquigarrow} \pi_{1,1}$. Then $\pi''_1 \overline{\rightsquigarrow} \pi_{1,1}$ and $\pi''_1 \overline{\rightsquigarrow} \pi''_2$ so by the diamond property there is some $\pi_{2,1}$ where $\pi''_2 \overline{\rightsquigarrow} \pi_{2,1}$ and $\pi_{1,1} \overline{\rightsquigarrow} \pi_{2,1}$. Continue in this vein, guided by the picture below:

$$\begin{array}{ccccccc}
\pi & \overline{\rightsquigarrow} & \pi'_1 & \overline{\rightsquigarrow} & \pi'_2 & \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} & \pi'_n \\
\wr\| & & \wr\| & & \wr\| & & \wr\| \\
\pi''_1 & \overline{\rightsquigarrow} & \pi_{1,1} & \overline{\rightsquigarrow} & \pi_{1,2} & \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} & \pi_{1,n} \\
\wr\| & & \wr\| & & \wr\| & & \wr\| \\
\pi''_2 & \overline{\rightsquigarrow} & \pi_{2,1} & \overline{\rightsquigarrow} & \pi_{2,2} & \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} & \pi_{2,n} \\
\wr\| & & \wr\| & & \wr\| & & \wr\| \\
\vdots & & \vdots & & \vdots & & \vdots \\
\wr\| & & \wr\| & & \wr\| & & \wr\| \\
\pi''_m & \overline{\rightsquigarrow} & \pi_{m,1} & \overline{\rightsquigarrow} & \pi_{m,2} & \overline{\rightsquigarrow} \cdots \overline{\rightsquigarrow} & \pi^*
\end{array}$$

to find the desired proof $\pi^*$. So, if $\pi'_n$ and $\pi''_n$ are *normal* they must be identical. ∎

So, sequences of reductions from $\pi$ cannot terminate in two different proofs. However, does every reduction process terminate?

DEFINITION 2.1.21 [STRONGLY NORMALISING] A proof $\pi$ is strongly normalising (under a reduction relation $\rightsquigarrow$) if and only if there is no infinite reduction sequence starting from $\pi$.

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 2.1.22 [REDUCTION TREE] The reduction tree (under $\rightsquigarrow$) of a proof $\pi$ is the tree whose branches are the reduction sequences on the relation $\rightsquigarrow$. So, from the root $\pi$ we reach any proof accessible in one $\rightsquigarrow$ step from $\pi$. From each $\pi'$ 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 $\pi$, $\nu(\pi)$ is the size of its reduction tree.

LEMMA 2.1.23 [THE SIZE OF REDUCTION TREES] *The reduction tree of a strongly normalising proof is finite. It follows that not only is every reduction path finite, but there is a* longest *reduction path.*

*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*. It follows that any strongly normalising proof not only has only finite reduction paths, it also has 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?

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**.

DEFINITION 2.1.24 [**red** PROOFS] We define a new predicate '**red**' applying to proofs in the following way.

The term '**red**' should bring to mind 'reducible.' This formulation of strong normalisation is originally due to William Tait [86]. I am following the presentation of Jean-Yves Girard [36, 37].

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

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

$$\dfrac{\begin{array}{cc}\vdots\,\pi & \vdots\,\pi' \\ A \to B & A\end{array}}{B}$$

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

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

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 2.1.25 [PROPERTIES OF **red** PROOFS] *For any proof $\pi$, the following three conditions hold*:

C1 *If $\pi$ is* **red** *then $\pi$ is strongly normalisable.*

C2 *If $\pi$ is* **red** *and $\pi$ reduces to $\pi'$ in one step, then $\pi'$ is* **red** *too.*

C3 *If $\pi$ is a proof not ending in $\rightarrow$I, and whenever we eliminate one indirect pair in $\pi$ we have a* **red** *proof, then $\pi$ is* **red** *too.*

*Proof*: We prove this result by induction on the formula proved by $\pi$. 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 $\pi$ is strongly normalising, then so is any proof to which $\pi$ reduces.

C3 $\pi$ does not end in $\rightarrow I$ as it is a proof of an atomic formula. If whenever $\pi \Rightarrow_1 \pi'$ and $\pi'$ is **red**, since $\pi'$ is a proof of an atomic formula, it is strongly normalising. Since *any* reduction path through $\pi$ must travel through one such proof $\pi'$, each such path through $\pi$ terminates. So, $\pi$ is **red**.

Now we prove the results for a proof $\pi$ 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 $\pi$ is a **red** proof of $A \rightarrow B$, consider the proof

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

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, $\sigma$ is a **red** proof of $B$. Condition C1 tells us that **red** proofs of $B$ are strongly normalising, so any reduction sequence for $\sigma$ must terminate. It follows that any reduction sequence for $\pi$ must terminate too, since if we had a non-terminating reduction sequence for $\pi$, we could apply the same reductions to the proof $\sigma$. But since $\sigma$ is strongly normalising, this cannot happen. It follows that $\pi$ is strongly normalising too.

C2 Suppose that $\pi$ reduces in one step to a proof $\pi'$. Given that $\pi$ is **red**, we wish to show that $\pi'$ is **red** too. Since $\pi'$ is a proof of $A \rightarrow B$, we want to show that for any **red** proof $\pi''$ 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 $\pi$ to $\pi'$), and C2 applies to proofs of $B$.

C3 Suppose that $\pi$ does not end in $\rightarrow I$, and suppose that all of the proofs reached from $\pi$ in one step are **red**. Let $\sigma$ 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 $\pi'$ one step from $\pi$. But $\pi'$ is **red** so $(\pi'\,\sigma)$ is too.

On the other hand, suppose σ is not yet normal, but the result holds for all $\sigma'$ with shorter reduction paths than σ. So, suppose τ reduces to $(\pi\,\sigma')$ with $\sigma'$ one step from σ. $\sigma'$ is **red** by the induction hypothesis C2 for A, and $\sigma'$ has a shorter reduction path, so the induction hypothesis for $\sigma'$ 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 2.1.26 [**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*

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

*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 $\tau'$ 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 $\sigma'$ smaller, and we may appeal to the induction hypothesis. ∎

THEOREM 2.1.27 [ALL PROOFS ARE **red**] *Every proof $\pi$ is* **red**.

LEMMA 2.1.28 [**red** PROOFS BY INDUCTION] *Given proof $\pi$ with assumptions $A_1, \ldots, A_n$, and any* **red** *proofs $\sigma_1, \ldots, \sigma_n$ of the respective formulas $A_1, \ldots, A_n$, it follows that the proof $\pi(\sigma_1, \ldots, \sigma_n)$ in which each assumption $A_i$ is replaced by the proof $\sigma_i$ is* **red**.

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

- If $\pi$ is an assumption $A_1$, the claim is a tautology (if $\sigma_1$ is **red**, then $\sigma_1$ is **red**).
- If $\pi$ 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**. So, the proof $(\pi_1(\sigma_1, \ldots, \sigma_n)\, \pi_2(\sigma_1, \ldots, \sigma_n))$ is **red**, but this is simply $\pi(\sigma_1, \ldots, \sigma_n)$.
- If $\pi$ ends in an application of $\rightarrow I$. This case is dealt with by Lemma 2.1.26: if $\pi$ is a proof of $A \rightarrow B$ ending in $\rightarrow E$, then we may assume that $\pi'$, the proof of $B$ from $A$ inside $\pi$ is **red**, so by Lemma 2.1.26, the result $\pi$ 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.

## 2.1.3 | PROOFS AND λ-TERMS

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 $\rightarrow E$ 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 $\rightarrow I$ 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 $A$s into $B$s 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 $\lambda$-notation, due to Alonzo Church [17]. 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* $\lambda$-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

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

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

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

and abstracting away the choice of $y$, we have

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

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*.

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

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

DEFINITION 2.1.29 [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 2.3.

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

Figure 2.3: RULES FOR λ-TERMS

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 $\lambda y^{A\to(A\to B)}.\lambda x^A((y^{A\to(A\to B)}\,x^A)\,x^A)$.

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 is a tree.

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

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

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

we can label them as follows

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

we could combine them into the proof

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

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 $\lambda$ at that stage).

DEFINITION 2.1.30 [FROM TERMS TO PROOFS AND BACK] For every typed term $M$ (of type $A$), we find PROOF($M$) (of the formula $A$) as follows:

» PROOF($x^A$) is the identity proof $A$.

» If $\textsc{proof}(M^{A\to B})$ is the proof $\pi_1$ of $A \to B$ and $\textsc{proof}(N^A)$ is the proof $\pi_2$ 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 $\pi$ of $B$ and $x^A$ is a variable of type $A$, then extend the proof $\pi$ by discharging each premise in $\pi$ of type $A$ labelled with the variable $x^A$. The result is a the proof $\textsc{proof}((\lambda x.M)^{A\to B})$ of type $A \to B$.

Conversely, for any proof $\pi$, 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 $\pi_l$ 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 $\pi$ is the proof found by extending $\pi_l$ and $\pi_r$ by the $\to E$ step. (Note that if the unbound variables in $M$ have types corresponding to the assumptions in $\pi_l$ and those in $N$ have types corresponding to the assumptions in $\pi_r$, then the unbound variables in $(MN)$ have types corresponding to the variables in $\pi$.)

» Suppose $\pi$ is a proof of $B$, and we extend $\pi$ into the proof $\pi'$ by discharging some set (possibly empty) of instances of the formula $A$, to derive $A \to B$ using $\to I$. Then, in $M$ 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 $\pi'$.

theorem 2.1.31 [relating proofs and 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* $M$.

**Todo**: write the proof out in full.

*Proof*: A simple induction on the construction of $\pi$ in the first case, and $M$ in the second. ∎

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

theorem 2.1.32 [discharge conditions and λ-terms] *A λ-term is a term of a* linear *proof iff each λ expression binds exactly one variable. It is a term of a* relevant *proof iff each λ expression binds at* least *one variable. It is a term of an* affine *proof iff each λ expression bind at* most *one variable.*

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 2.1.33 [β REDUCTION] The term $\lambda x.M\,N$ is said to directly 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

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

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

THEOREM 2.1.34 [NORMALISATION AND β-REDUCTION] PROOF(N) *is normal if and only if the term* N *does not* β*-reduce to any other 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.

## 2.1.4 | 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. | $\quad A$ | Supposition |
| 3. | $\quad A \to (A \to B)$ | 1, Repeat |
| 4. | $\quad A \to B$ | 2, 3, Modus Ponens |
| 5. | $\quad B$ | 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. | $\quad B$ | Supposition |
| 3. | $\quad A$ | 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 ststem gained quite some popularity in undergraduate education in logic in the 1960s and following decades in the United States [31]. Edward Lemmon's text *Beginning Logic* [49] served a similar purpose in British logic education. Lemmon's account of natural deduction is similar to this, except that it does without the need to reiterate by *breaking the box*.

| | | | |
|---|---|---|---|
| 1 | (1) | $A \to (A \to B)$ | Assumption |
| 2 | (2) | $A$ | Assumption |
| 1,2 | (3) | $A \to B$ | 1, 2, Modus Ponens |
| 1,2 | (4) | $B$ | 2,3, Modus Ponens |
| 1 | (5) | $A \to B$ | 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 $A \to B$ 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 $A \to B$ 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 |
| ⋮ | (3) | ⋮ | |

There is no way to attach the assumption number "2" on 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 $A \to (B \to A)$ 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) | $B \to A$ | 2,4, Conditionalisation |
| | (6) | $A \to (B \to A)$ | 1,5, Conditionalisation |

This seems quite unsatisfactory, as it breaks the normalisation property. (The formula $A \to (B \to A)$ 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 [62].

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 bust on the scene through the work of Alan Anderson and Nuel Belnap in the 1960s and 1970s [1, 2], though it had precursors in the work of the Russian logician, I. E. Orlov in the 1920s [23, 57]. 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 [26, 27], Routley and Meyer [81], Read [71] and Mares [51] 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 [35, 37]. A helpful introduction to linear logic is the text of Troelstra [90]. Affine logic is introduced in the tradition of linear logic as a variant on linear implication. Affine implication is quite close, however to the implication in Łukasiewicz's infinitely valued logic—which is slightly stronger, but shares the property of rejecting all *contraction*-related principles [75]. These logics are all *substructural* logics [24, 59, 76]

The definition of normality is due to Prawitz [63], though glimpses of the idea are present in Gentzen's original work [33].

The λ-calculus is due to Alonzo Church [17], and the study of λ-calculi has found many different applications in logic, computer science, type theory and related fields [3, 39, 83]. The correspondence between formulas/proofs and types/terms is known as the Curry–Howard correspondence [43].

**Todo**: find the Curry reference.

### 2.1.5 | EXERCISES

Working through these exercises will help you understand the material. As with all logic exercises, if you want to deepend 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* exercises, and so on. (The *project* exercises are not the kind of thing that can be completed in one sitting.)

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.

#### BASIC EXERCISES

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

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. Find a most general λ-term for each provable formula.

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 $B \to (A \to A)$ (also using vacuous discharge). Combine the two proofs, using $\to E$ to deduce $B \to C$. 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."

| → | t | n | f |
|---|---|---|---|
| t | t | n | f |
| n | t | t | f |
| f | t | t | t |

GD3

| → | t | n | f |
|---|---|---|---|
| t | t | n | n |
| n | t | t | f |
| f | t | t | t |

Ł3

| → | t | n | f |
|---|---|---|---|
| t | t | f | f |
| n | t | n | f |
| f | t | t | t |

RM3

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.

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:

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

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 $p \to (p \to p)$. 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: $A \hat{\vee} B ::= (A \to B) \to B$. 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 Provide introduction and elimination rules for $\hat{\vee}$ that do not involve the conditional connective $\to$.

Q14 Now consider *negation*. Given an ATOM $p$, define the $p$-negation $\neg_p A$ to be $A \to p$. 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 $\dot{\vee}$ 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:

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

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

$$\frac{A \otimes B}{A}\otimes I_1 \qquad \frac{A \otimes B}{B}\otimes I_2$$

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 2.1.35 [TRACK] A sequence $A_0, \ldots, A_n$ of formula instances in the proof $\pi$ is a *track* of length $n+1$ in the proof $\pi$ 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 $\rightarrow E$.

A track whose terminus $A_n$ is the conclusion of the proof $\pi$ 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 $\rightarrow E$ 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.

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

(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 —

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

— 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 $\to E$ whose conclusion is $C$, in the $\heartsuit$ track of order 0. Similarly, the $\diamondsuit$ track is order 2 and the $\clubsuit$ track has order 3.

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

LEMMA 2.1.36 *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 2.1.37 *Let* $t : A_0, \ldots, A_n$ *be a track in a normal proof. Then*

*a*) *The rules applied within the track consist of a sequence (possibly empty) of* [$\to$E] *steps and then a sequence (possibly empty) of* [$\to$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: $\to_l$ for linear conditionals, and $\to_s$ 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 $\to$ is translated as $\to_l$) and whenever an argument is valid for a standard conditional, it is (in some appropriate sense) valid in the system you design (when $\to$ is translated as $\to_s$). What mixed inferences (those using both $\to_l$ and $\to_s$) 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. **Add references**.]

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 $\lambda$-terms. Can this kind if identification be maintained along with $\beta$-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.

## 2.2 | SEQUENTS AND DERIVATIONS

In this section we will look at a different way of thinking about inference: 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, or that the argument from A to B is valid. The symbol used here 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 the 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 the following grammar:

FORMULA ::= ATOM | (FORMULA $\wedge$ FORMULA) | (FORMULA $\vee$ FORMULA)

To characterise consequence relations, we need to figure out 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 we can say $A \vdash B$ 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 it is appropriate to infer *from* a formula featuring that connective, and when it is appropriate to infer *to* a formula featuring that connective. Another way of putting it is that we wish to know how a connective works on the left of the turnstile, and how it works on the right.

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

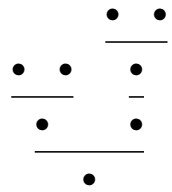
This is a *formal* account of consequence. We look only at the form of propositions and not their content. For atomic propositions (those with no internal form) there is nothing upon which we could pin a claim to consequence. Thus, $p \vdash q$ is never true, unless p and q are the same atom.

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 logical consequence.

### 2.2.1 | DERIVATIONS FOR "AND" AND "OR"

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



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 2.4.

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

DEFINITION 2.2.1 [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 2.4, 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 $\wedge$ and $\vee$.

EXAMPLE 2.2.2 [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.

$$\frac{}{p \vdash p}\ [\mathrm{Id}]$$

$$\frac{L \vdash C \quad C \vdash R}{L \vdash R}\ \mathit{Cut}$$

$$\frac{A \vdash R}{A \wedge B \vdash R}\ \wedge L_1 \qquad \frac{A \vdash R}{B \wedge A \vdash R}\ \wedge L_2 \qquad \frac{L \vdash A \quad L \vdash B}{L \vdash A \wedge B}\ \wedge R$$

$$\frac{A \vdash R \quad B \vdash R}{A \vee B \vdash R}\ \vee L \qquad \frac{L \vdash A}{L \vdash A \vee B}\ \vee R_1 \qquad \frac{L \vdash A}{L \vdash B \vee A}\ \vee R_2$$

Figure 2.4: a simple sequent system

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

$$\dfrac{\dfrac{\dfrac{}{q \vdash q}}{p \wedge q \vdash q}\wedge L_2 \quad \dfrac{\dfrac{}{p \vdash p}}{p \wedge q \vdash p}\wedge L_1}{p \wedge q \vdash q \wedge p}\wedge R \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}}{p \wedge (q \wedge r) \vdash p}\wedge L_1 \quad \dfrac{\dfrac{\dfrac{}{q \vdash q}}{q \wedge r \vdash q}\wedge L_1}{p \wedge (q \wedge r) \vdash q}\wedge L_2}{p \wedge (q \wedge r) \vdash p \wedge q}\wedge R \quad \dfrac{\dfrac{}{r \vdash r}}{q \wedge r \vdash r}\wedge L_2}{p \wedge (q \wedge r) \vdash r}\wedge L_2}{p \wedge (q \wedge r) \vdash (p \wedge q) \wedge r}\wedge R$$

Here are the cases for disjunction. The first derivation is for the commutativity of disjunction, and the second is for associativity. (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 $A \wedge B \vdash B \wedge A$ (for example) since $A \vdash A$ is not an axiom if $A$ is a complex formula. We will see more on this in the next section.)

$$\dfrac{\dfrac{\dfrac{}{p \vdash p}}{p \vdash q \vee p}\vee R_1 \quad \dfrac{\dfrac{}{q \vdash q}}{q \vdash p \vee q}\vee R_2}{p \vee q \vdash q \vee p}\vee L \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}}{p \vdash p \vee (q \vee r)}\vee R_1 \quad \dfrac{\dfrac{\dfrac{}{q \vdash q}}{q \vdash q \vee r}\vee R_1}{q \vdash p \vee (q \vee r)}\vee R_2}{p \vee q \vdash p \vee (q \vee r)}\vee L \quad \dfrac{\dfrac{}{r \vdash r}}{r \vdash q \vee r}\vee R_2}{r \vdash p \vee (q \vee r)}\vee R_2}{(p \vee q) \vee r \vdash p \vee (q \vee r)}\vee L$$

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

Exercise 14 on page 67 asks you to make this duality precise.

more examples of duality between derivations. The first is the dual of the second, and the third is the dual of the fourth.

$$\dfrac{p \vdash p \quad p \vdash p}{p \vee p \vdash p}\vee L \qquad \dfrac{p \vdash p \quad p \vdash p}{p \vdash p \wedge p}\wedge R \qquad \dfrac{p \vdash p \quad \dfrac{p \vdash p}{p \wedge q \vdash p}\wedge L_1}{p \vee (p \wedge q) \vdash p}\vee L \qquad \dfrac{p \vdash p \quad \dfrac{p \vdash p}{p \vdash p \vee q}\vee R}{p \vdash p \wedge (p \vee q)}\wedge R$$

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.

$$\dfrac{\dfrac{p \vdash p \quad \dfrac{p \vdash p}{p \wedge q \vdash p}\wedge L_1}{p \vee (p \wedge q) \vdash p}\vee L \qquad \dfrac{p \vdash p \quad \dfrac{p \vdash p}{p \vdash p \vee q}\vee R}{p \vdash p \wedge (p \vee q)}\wedge R}{p \vee (p \wedge q) \vdash p \wedge (p \vee q)}\mathit{Cut}$$

Notice, too, that in 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:

$$\dfrac{\dfrac{\dfrac{p \vdash p}{p \vdash q \vee p}\vee R_1 \quad \dfrac{q \vdash q}{q \vdash q \vee p}\vee R_2}{p \vee q \vdash q \vee p}\vee L \qquad \dfrac{\dfrac{q \vdash q}{q \vdash p \vee q}\vee R_1 \quad \dfrac{p \vdash p}{p \vdash p \vee q}\vee R_2}{q \vee p \vdash p \vee q}\vee L}{p \vee q \vdash p \vee q}\mathit{Cut}$$

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

This derivation is a complicated way to deduce $p \vee q \vdash p \vee q$, and it includes $q \vee p$, 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.

### 2.2.2 | IDENTITY DERIVATIONS

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

$$\dfrac{\dfrac{p \vdash p}{p \vdash p \vee q}\vee R_1 \quad \dfrac{q \vdash q}{q \vdash p \vee q}\vee R_2}{p \vee q \vdash p \vee q}\vee L \qquad \dfrac{\dfrac{p \vdash p}{p \wedge q \vdash p}\wedge L_1 \quad \dfrac{q \vdash q}{p \wedge q \vdash q}\wedge L_2}{p \wedge q \vdash p \wedge q}\wedge R$$

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

$$\dfrac{\dfrac{p \vdash p}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash p}\wedge L_1 \qquad \dfrac{q \vee (r_1 \wedge r_2) \vdash q \vee (r_1 \wedge r_2)}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash q \vee (r_1 \wedge r_2)}\wedge L_2}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash p \wedge (q \vee (r_1 \wedge r_2))}\wedge R$$

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

$$\dfrac{\dfrac{p \vdash p}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash p} \quad \dfrac{\dfrac{\dfrac{q \vdash q}{q \vdash q \vee (r_1 \wedge r_2)} \vee R_1 \quad \dfrac{\dfrac{\dfrac{r_1 \vdash r_1}{r_1 \wedge r_2 \vdash r_1} \wedge L_1 \quad \dfrac{r_2 \vdash r_2}{r_1 \wedge r_2 \vdash r_2} \wedge L_2}{r_1 \wedge r_2 \vdash r_1 \wedge r_2} \wedge R}{r_1 \wedge r_2 \vdash q \vee (r_1 \wedge r_2)} \vee R_2}{q \vee (r_1 \wedge r_2) \vdash q \vee (r_1 \wedge r_2)} \vee L}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash q \vee (r_1 \wedge r_2)}}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash p \wedge (q \vee (r_1 \wedge r_2))}$$

The derivation of $q \vee (r_1 \wedge r_2) \vdash q \vee (r_1 \wedge r_2)$ itself contains a smaller identity derivation, for $r_1 \wedge r_2 \vdash r_1 \wedge r_2$. 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 2.2.3 [IDENTITY DERIVATIONS] *For each formula* $A$, $A \vdash A$ *has a derivation. A derivation for* $A \vdash A$ *may be systematically constructed from the identity derivations for the subformulas of* $A$.

*Proof*: We define $\mathrm{Id}(A)$, the IDENTITY DERIVATION FOR $A$ by induction on the construction of $A$, as follows. $\mathrm{Id}(p)$ is the axiom $p \vdash p$. For complex formulas, we have

$$\mathrm{Id}(A \vee B): \dfrac{\dfrac{\mathrm{Id}(A)}{A \vdash A \vee B} \vee R_1 \quad \dfrac{\mathrm{Id}(B)}{B \vdash A \vee B} \vee R_2}{A \vee B \vdash A \vee B} \vee L \qquad \mathrm{Id}(A \wedge B): \dfrac{\dfrac{\mathrm{Id}(A)}{A \wedge B \vdash A} \wedge L_1 \quad \dfrac{\mathrm{Id}(B)}{A \wedge B \vdash B} \wedge L_2}{A \wedge B \vdash A \wedge B} \wedge R$$

■

We say that $A \vdash A$ is DERIVABLE in the sequent system. If we think of [Id] as a degenerate *rule* (a rule with no premise), then its generalisation, $[\mathrm{Id}_A]$, is a *derivable rule*.

It might seem *crazy* to have a proof of identity, like $A \vdash A$ where $A$ is a complex formula. Why don't we take $[\mathrm{Id}_A]$ as an axiom? There are a few different reasons we might like to consider for taking $[\mathrm{Id}_A]$ 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 $[\mathrm{Id}_A]$ 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 [38].

THE SYSTEM IS SYSTEMATIC: In the system without $[\mathrm{Id}_A]$ as an axiom, when we consider a sequent like $L \vdash R$ in order to know whether it it is derived (in the absence of Cut, at least), we can ask two separate questions. We can consider $L$. If it is complex perhaps $L \vdash R$ 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 [∧R] or [∨R]. If both are primitive, then $L \vdash R$ 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 $[\mathrm{Id}_A]$, one would have to check if $L = R$ as well as the other conditions.

THE SYSTEM PROVIDES A CONSTRAINT: 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.

## 2.2.3 | CUT IS REDUNDANT

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 2.2.4 [SUBFORMULA PROPERTY] *If $\delta$ is a sequent derivation not containing* Cut*, then the formulas in $\delta$ are all subformulas of the formulas in the endsequent of $\delta$.*

Notice how much simpler this proof is than the proof of Theorem 2.1.11.

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

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 2.2.5 [DERIVATIONS WITH OR WITHOUT CUT] $p \vdash p \vee q$ has only one cut-free derivation, it has infinitely many derivation using Cut.

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

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

$$\dfrac{\dfrac{}{p \vdash p} \quad \dfrac{\dfrac{}{p \vdash p}\,\vee R_1}{p \vdash p \vee q}}{p \vdash p \vee q}\,Cut$$

which contains the cut-free derivation of $p \vdash p \vee q$ inside it. We can nest the cuts with the identity sequent $p \vdash p$ as deeply as we like.

$$\dfrac{\dfrac{}{p \vdash p} \quad \dfrac{\dfrac{\dfrac{}{p \vdash p} \quad \dfrac{\dfrac{}{p \vdash p}\,\vee R_1}{p \vdash p \vee q}}{p \vdash p}\,Cut}{p \vdash p \vee q}\,Cut}{p \vdash p \vee q}\,Cut \qquad \dfrac{\dfrac{}{p \vdash p} \quad \dfrac{\dfrac{\dfrac{}{p \vdash p} \quad \dfrac{\dfrac{\dfrac{}{p \vdash p}\,\vee R_1}{p \vdash p \vee q}}{p \vdash p}\,Cut}{p \vdash p \vee q}\,Cut}{p \vdash p \vee q}\,Cut}{p \vdash p \vee q}\,Cut \quad \cdots$$

However, we can construct quite different derivations of our sequent, and we involve different material in the derivation. For any formula $A$ you wish to choose, we could implicate $A$ (an "innocent bystander") in the derivation as follows:

$$\dfrac{\dfrac{\dfrac{}{p \vdash p}\,\vee R_1}{p \vdash p \vee (q \wedge A)} \quad \dfrac{\dfrac{\dfrac{}{p \vdash p}\,\vee R_1}{p \vdash p \vee q} \quad \dfrac{\dfrac{\dfrac{}{q \vdash q}\,\wedge L_1}{q \wedge A \vdash q}\,\vee R_2}{q \wedge A \vdash p \vee q}}{p \vee (q \wedge A) \vdash p \vee q}\,\vee L}{p \vdash p \vee q}\,Cut$$

In this derivation the cut formula $p \vee (q \wedge A)$ is doing genuine work. It is just repeating either the left formula $p$ or the right formula $q$.

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.

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 it? Take a derivation involving Cut, such as this one:

$$\dfrac{\dfrac{\dfrac{}{p \vdash p}\,\wedge L_1}{p \wedge (q \wedge r) \vdash p} \quad \dfrac{\dfrac{\dfrac{}{q \vdash q}\,\wedge L_1}{q \wedge r \vdash q}\,\wedge L_2}{p \wedge (q \wedge r) \vdash q}\,\wedge R}{p \wedge (q \wedge r) \vdash p \wedge q} \quad \dfrac{\dfrac{\dfrac{}{q \vdash q}\,\wedge L_1}{p \wedge q \vdash q}\,\vee R_1}{p \wedge q \vdash q \vee r}\,Cut}{p \wedge (q \wedge r) \vdash q \vee r}$$

This sequent $p \wedge (q \wedge r) \vdash q \vee r$ 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 $p \wedge q \vdash q$ and left the inference to $q \vee r$ until later:

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

$$\dfrac{\dfrac{\dfrac{\dfrac{p \vdash p}{p \wedge (q \wedge r) \vdash p}\,\wedge L_1 \quad \dfrac{\dfrac{q \vdash q}{q \wedge r \vdash q}\,\wedge L_1}{p \wedge (q \wedge r) \vdash q}\,\wedge L_2}{p \wedge (q \wedge r) \vdash p \wedge q}\,\wedge R \quad \dfrac{q \vdash q}{p \wedge q \vdash q}\,\wedge L_1}{p \wedge (q \wedge r) \vdash q}\,Cut}{p \wedge (q \wedge r) \vdash q \vee r}\,\vee R_1$$

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 $p \wedge (q \wedge r) \vdash r$, 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 $q \vdash q$.)

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

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

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 $L \vdash R$. We can sidestep the move through $A \wedge B$ to cut on the formula $A$, since we have $L \vdash A$ and $A \vdash R$.

$$\dfrac{\dfrac{L \vdash A \quad L \vdash B}{L \vdash A \wedge B}\,\wedge R \quad \dfrac{A \vdash R}{A \wedge B \vdash R}\,\wedge L_1}{L \vdash R}\,Cut$$

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

$$\dfrac{\dfrac{\dfrac{\dfrac{q \vdash q}{q \wedge r \vdash q}\,\wedge L_1}{p \wedge (q \wedge r) \vdash q}\,\wedge L_2 \quad q \vdash q}{p \wedge (q \wedge r) \vdash q}\,Cut}{p \wedge (q \wedge r) \vdash q \vee r}\,\vee R_1 \qquad \dfrac{\dfrac{\dfrac{q \vdash q}{q \wedge r \vdash q}\,\wedge L_1}{p \wedge (q \wedge r) \vdash q}\,\wedge L_2}{p \wedge (q \wedge r) \vdash q \vee r}\,\vee R_1$$

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 the example.

definition 2.2.6 [active and passive formulas] The formulas L and R in each inference in Figure 2.4 are said to be *passive* in the inference (they "do nothing" in the step from top to bottom), while the other formulas are *active*.

definition 2.2.7 [depth of an inference] The depth of an inference in a derivation $\delta$ is the number of nodes in the sub-derivation of $\delta$ 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 cuts from a derivation. First we show that cuts may be moved upward in a derivation. Then we show that this process will terminate in a Cut-free derivation.

lemma 2.2.8 [cut-depth reduction] *Given a derivation $\delta$ of $A \vdash C$, whose final inference is* Cut, *which is otherwise cut-free, and in which that inference has a depth of $n$, we may construct another derivation of $A \vdash C$ which is cut-free, or in which each* Cut *step has a depth less than $n$.*

*Proof*: Our derivation $\delta$ contains two subderivations: $\delta_l$ ending in $A \vdash B$ and $\delta_r$ ending in $B \vdash C$. These subderivations are are cut-free.

$$\dfrac{\begin{array}{cc}\vdots\,\delta_l & \vdots\,\delta_r \\ A \vdash B & B \vdash C\end{array}}{A \vdash C}$$

To find our new derivation, we look at the formula B and its roles in the final inference in $\delta_l$ and $\delta_r$.

case 1: the cut-formula is passive in either inference Suppose that the formula B is *passive* in the last inference in $\delta_l$ or passive in the last inference in $\delta_r$. For example, if $\delta_l$ 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 $A_1$.

$$\text{before:}\quad \dfrac{\dfrac{\vdots\,\delta'_l}{A_1 \vdash B}}{\dfrac{A_1 \wedge A_2 \vdash B}{A_1 \wedge A_2 \vdash C}\,\wedge L_1} \qquad \dfrac{\vdots\,\delta_r}{B \vdash C}\,Cut$$

$$\text{after:}\quad \dfrac{\dfrac{\dfrac{\vdots\,\delta'_l}{A_1 \vdash B} \qquad \dfrac{\vdots\,\delta_r}{B \vdash C}}{A_1 \vdash C}\,Cut}{A_1 \wedge A_2 \vdash C}\,\wedge L_1$$

The resulting derivation has a has a cut-depth lower by one. If, on the other hand, $\delta_l$ 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 $\delta_l$ is split between the two cuts.

$$\textsc{before:}\ \dfrac{\dfrac{\begin{array}{cc}\vdots\,\delta_l^1 & \vdots\,\delta_l^2\\ A_1 \vdash B & A_2 \vdash B\end{array}}{A_1 \vee A_2 \vdash B}\vee L \qquad \begin{array}{c}\vdots\,\delta_r\\ B \vdash C\end{array}}{A_1 \vee A_2 \vdash C}\mathit{Cut} \qquad \textsc{after:}\ \dfrac{\dfrac{\begin{array}{cc}\vdots\,\delta_l^1 & \vdots\,\delta_r\\ A_1 \vdash B & B \vdash C\end{array}}{A_1 \vdash C}\mathit{Cut} \qquad \dfrac{\begin{array}{cc}\vdots\,\delta_l^2 & \vdots\,\delta_r\\ A_2 \vdash B & B \vdash C\end{array}}{A_2 \vdash C}\mathit{Cut}}{A_1 \vee A_2 \vdash C}\vee L$$

The other two ways in which the cut formula could be passive are when $\delta_2$ 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 $\delta_l$ and in $\delta_r$, 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 $\delta_l$ and $\delta_r$.

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.

$$\textsc{before:}\ \dfrac{p \vdash p \quad p \vdash p}{p \vdash p}\mathit{Cut} \qquad \textsc{after:}\ p \vdash p$$

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

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

$$\textsc{before:}\ \dfrac{\dfrac{\begin{array}{cc}\vdots\,\delta_l^1 & \vdots\,\delta_l^2\\ A \vdash B_1 & A \vdash B_2\end{array}}{A \vdash B_1 \wedge B_2}\wedge R \qquad \dfrac{\begin{array}{c}\vdots\,\delta_r'\\ B_1 \vdash C\end{array}}{B_1 \wedge B_2 \vdash C}\wedge L_1}{A \vdash C}\mathit{Cut} \qquad \textsc{after:}\ \dfrac{\begin{array}{cc}\vdots\,\delta_l^1 & \vdots\,\delta_r'\\ A \vdash B_1 & B_1 \vdash C\end{array}}{A \vdash C}\mathit{Cut}$$

CASE 2C: THE CUT-FORMULA IS A DISJUNCTION The case for disjunction is similar. If the cut-formula is a disjunction $B_1 \vee B_2$, then the only inferences in which a conjunction is active in the conclusion are $\vee$R and $\vee$L. Let's suppose that in $\vee$R the disjunction $B_1 \vee B_2$ 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 $B_1 \vee B_2$, with a shallower cut.

$$\text{before:}\quad \dfrac{\dfrac{\begin{array}{c}\vdots\,\delta'_l\\ A \vdash B_1\end{array}}{A \vdash B_1 \vee B_2}\vee R_1 \qquad \dfrac{\begin{array}{cc}\vdots\,\delta^1_r & \vdots\,\delta^2_r\\ B_1 \vdash C & B_2 \vdash C\end{array}}{B_1 \vee B_2 \vdash C}\vee L}{A \vdash C}\,Cut \qquad \text{after:}\quad \dfrac{\begin{array}{cc}\vdots\,\delta'_l & \vdots\,\delta^1_r\\ A \vdash B_1 & B_1 \vdash C\end{array}}{A \vdash C}\,Cut$$

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 $A \vdash C$ with a cut-depth of *zero*: We find a derivation of $A \vdash C$ without a cut. That is,

THEOREM 2.2.9 [CUT ELIMINATION] *If a sequent is derivable with* Cut, *it is derivable without* Cut.

*Proof*: Given a derivation of a sequent $A \vdash C$, take a Cut with no Cuts 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 has a number of fruitful consequences.

COROLLARY 2.2.10 [DECIDABILITY FOR LATTICE SEQUENTS] *There is an* algorithm *for determining whether or not a sequent* $A \vdash B$ *is valid in lattice logic.*

*Proof*: To determine whether or not $A \vdash B$ has a derivation, look for the finitely many different sequents from which this sequent may be derived. Repeat the process until you find atomic sequents. Atomic sequents of the form $p \vdash p$ are derivable, and those of the form $p \vdash q$ are not. ∎

Here is an example:

EXAMPLE 2.2.11 [DISTRIBUTION IS NOT DERIVABLE] The sequent $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$ is not derivable.

*Proof*: Any cut-free derivation of $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$ must end in either a $\wedge L$ step or a $\vee R$ step. If we use $\wedge L$, we infer our sequent from either $p \vdash (p \wedge q) \vee r$, or from $q \vee r \vdash (p \wedge q) \vee r$. None of these are derivable. As you can see, $p \vdash (p \wedge q) \vee r$ is derivable only, using $\vee R$ from either $p \vdash p \wedge q$ or from $p \vdash r$. The latter is not derivable (it is not an axiom, and it cannot be inferred from *anywhere*)

and the former is derivable only when $p \vdash q$ is — and it isn't. Similarly, $q \vee r \vdash (p \wedge q) \vee r$ is derivable only when $q \vdash (p \wedge q) \vee r$ is derivable, and this is only derivable when either $q \vdash p \wedge q$ or when $q \vdash r$ are derivable, and as before, neither of *these* are derivable either.

Similarly, if we use $\vee R$, we infer our sequent from either $p \wedge (q \vee r) \vdash p \wedge q$ or from $p \wedge (q \vee r) \vdash r$. By analogous reasoning, (more precisely, by *dual* reasoning) neither of these sequents are derivable. So, $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$ has no cut-free derivation, and by Theorem 2.2.9 it has no derivation at all. ∎

Searching for derivations in this naïve manner 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 $A \vee B \vdash C \wedge D$ (where A, B, C and D are possibly complex statements). This is derivable in two ways (a) from $A \vdash C \wedge D$ and $B \vdash C \wedge D$ by $\vee L$ or (b) from $A \vee B \vdash C$ and $A \vee B \vdash D$ 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 2.2.12 [INVERTIBILITY] A sequent rule of the form

$$\frac{S_1 \quad \cdots \quad S_n}{S}$$

is *invertible* if and only if whenever the sequent S is derivable, so are the sequents $S_1, \ldots S_n$.

THEOREM 2.2.13 [INVERTIBLE SEQUENT RULES] *The rules* $\vee$L *and* $\wedge$R *are invertible, but the rules* $\vee$R *and* $\wedge$L *are not.*

*Proof*: Consider $\vee L$. If $A \vee B \vdash C$ is derivable, then since we have a derivation of $A \vdash A \vee B$ (by $\vee R$), a use of Cut shows us that $A \vdash C$ is derivable. Similarly, since we have a derivation of $B \vdash A \vee B$, the sequent $B \vdash C$ is derivable too. So, from the conclusion $A \vee B \vdash C$ of a $\vee L$ inference, we may derive the premises. The case for $\wedge R$ is completely analogous.

For $\wedge L$, on the other hand, we have a derivation of $p \wedge q \vdash p$, but no derivation of the premise $q \vdash p$, so this rule is not invertible. Similarly, $p \vdash q \vee p$ is derivable, but $p \vdash q$ is not. ∎

It follows that when searching for a derivation of a sequent, instead of searching for *all* of the ways that a sequent may be derived, if it may be derived from an invertible rule we can look to the premises of that rule *immediately*, and consider those, without pausing to check the other sequents from which our target sequent is constructed.

EXAMPLE 2.2.14 [DERIVATION SEARCH USING INVERTIBILITY] The sequent $(p \wedge q) \vee (q \wedge r) \vdash (p \vee r) \wedge p$ is not derivable. By the invertibility of $\vee L$, it *is* derivable only if (a) $p \wedge q \vdash (p \vee r) \wedge p$ and (b) $q \wedge r \vdash (p \vee r) \wedge p$

are both derivable. Using the invertibility of $\wedge R$, the sequent (b) this is derivable only if ($b_1$) $q \wedge r \vdash p \vee r$ and ($b_2$) $q \wedge r \vdash p$ are both derivable. But ($b_2$) is not derivable because $q \vdash p$ and $r \vdash p$ are underivable.

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*.

COROLLARY 2.2.15 [INTERPOLATION FOR LATTICE SEQUENTS] *If a sequent* $A \vdash B$ *is derivable, then there is a formula* $C$ *containing only atoms present in both* $A$ *and* $B$ *such that* $A \vdash C$ *and* $C \vdash B$ *are derivable.*

This result tells us that if the sequent $A \vdash B$ is derivable then that consequence "factors through" a statement in the vocabulary shared between $A$ and $B$. This means that the consequence $A \vdash B$ 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 $A \vdash B$ provides us with an interpolant.

*Proof*: We prove this by induction on the construction of the derivation of $A \vdash B$. We keep track of the interpolant with these rules:

$$p \vdash_p p \ [\mathrm{Id}]$$

$$\frac{A \vdash_C R}{A \wedge B \vdash_C R} \wedge L_1 \qquad \frac{A \vdash_C R}{B \wedge A \vdash_C R} \wedge L_2 \qquad \frac{L \vdash_{C_1} A \qquad L \vdash_{C_2} B}{L \vdash_{C_1 \wedge C_2} A \wedge B} \wedge R$$

$$\frac{A \vdash_{C_1} R \qquad B \vdash_{C_2} R}{A \vee B \vdash_{C_1 \vee C_2} R} \vee L \qquad \frac{L \vdash_C A}{L \vdash_C A \vee B} \vee R_1 \qquad \frac{L \vdash_C A}{L \vdash_C B \vee A} \vee R_2$$

We show by induction on the length of the derivation that if we have a derivation of $L \vdash_C R$ then $L \vdash C$ and $C \vdash R$ and the atoms in $C$ present in both $L$ and in $R$. These properties are satisfied by the atomic sequent $p \vdash_p p$, and it is straightforward to verify them for each of the rules. ∎

EXAMPLE 2.2.16 [A DERIVATION WITH AN INTERPOLANT] Consider the sequent $p \wedge (q \vee (r_1 \wedge r_2)) \vdash (q \vee r_1) \wedge (p \vee r_2))$. We may annotate a cut-free derivation of it as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{q \vdash_q q}{q \vdash_q q \vee r} \vee R \qquad \dfrac{r_1 \vdash_{r_1} r_1}{r_1 \wedge r_2 \vdash_{r_1} r_1} \wedge L}{q \vee (r_1 \wedge r_2) \vdash_{q \vee r_1} q \vee r_1} \vee L}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash_{q \vee r_1} q \vee r_1} \wedge R \qquad \dfrac{\dfrac{p \vdash p}{p \vdash_p p \vee r_2} \vee R}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash_p \vdash p \vee r_2} \wedge L}{p \wedge (q \vee (r_1 \wedge r_2)) \vdash_{(q \vee r_1) \wedge p} (q \vee r_1) \wedge (p \vee r_2)} \wedge R$$

Notice that the interpolant $(q \vee r_1) \wedge p$ 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

$$p \wedge (q \vee (r_1 \wedge r_2)) \vdash (q \vee r_1) \wedge p, \quad (q \vee r_1) \wedge p \vdash (q \vee r_1) \wedge (p \vee r_2)$$

We can replace the $r_2$ in either derivation derivation with another statement – say $r_3$ – preserving the structure of each derivation. We get the more general fact:

$$p \wedge (q \vee (r_1 \wedge r_2)) \vdash (q \vee r_1) \wedge (p \vee r_3)$$

## 2.2.4 | HISTORY

[To be written.]

## 2.2.5 | EXERCISES

### BASIC EXERCISES

Q1 Show that there is no cut-free derivation of the following sequents

1 : $p \vee (q \wedge r) \vdash p \wedge (q \vee r)$
2 : $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$
3 : $p \wedge (q \vee (p \wedge r)) \vdash (p \wedge q) \vee (p \wedge r)$

Q2 Suppose that there is a derivation of $A \vdash B$. 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 $C(A) \vdash C(B)$. Furthermore, show that a derivation of $C(A) \vdash C(B)$ may be systematically constructed from the derivation of $A \vdash B$ together with the context $C(-)$ (the shape of the formula $C(A)$ with a 'hole' in the place of the subformula $A$).

Q3 Find a derivation of $p \wedge (q \wedge r) \vdash (p \wedge q) \wedge r$. Find a derivation of $(p \wedge q) \wedge r \vdash p \wedge (q \wedge r)$. Put these two derivations together, with a Cut, to show that $p \wedge (q \wedge r) \vdash p \wedge (q \wedge r)$. Then eliminate the cuts from this derivation. What do you get?

Q4 Do the same thing with derivations of $p \vdash (p \wedge q) \vee p$ and $(p \wedge q) \vee p \vdash p$. What is the result when you eliminate this cut?

Q5 Show that (1) $A \vdash B \wedge C$ is derivable if and only if $A \vdash B$ and $A \vdash C$ is derivable, and that (2) $A \vee B \vdash C$ is derivable if and only if $A \vdash C$ and $B \vdash C$ are derivable. Finally, (3) when is $A \vee B \vdash C \wedge D$ derivable, in terms of the derivability relations between $A$, $B$, $C$ and $D$.

Q6 Under what conditions do we have a derivation of $A \vdash B$ when $A$ contains only propositional atoms and *disjunctions* and $B$ contains only propositional atoms and *conjunctions*.

Q7 Expand the system with the following rules for the propositional *constants* $\bot$ and $\top$.

$$A \vdash \top \ [\top R] \qquad \bot \vdash A \ [\bot L]$$

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?)

Q8 Show that lattice sequents including $\top$ and $\bot$ are decidable, following Corollary 2.2.10 and the results of the previous question.

Q9 Show that every formula composed of just $\top$, $\bot$, $\wedge$ and $\vee$ is *equivalent* to either $\top$ or $\bot$. (What does this result remind you of?)

Q10 Prove the interpolation theorem (Corollary 2.2.15) for derivations involving $\wedge, \vee, \top$ and $\bot$.

Q11 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \vdash R}{A \text{ tonk } B \vdash R}\ \text{tonk } L \qquad \frac{L \vdash B}{L \vdash A \text{ tonk } B}\ \text{tonk } R$$

What new things can you derive using tonk? Can you derive $A$ tonk $B \vdash A$ tonk $B$? Is Cut eliminable for formulas involving tonk?

See Arthur Prior's "The Runabout Inference-Ticket" [70] for tonk's first appearance in print.

Q12 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \vdash R}{A \text{ honk } B \vdash R}\ \text{honk } L \qquad \frac{L \vdash A \qquad L \vdash B}{L \vdash A \text{ honk } B}\ \text{honk } R$$

What new things can you derive using honk? Can you derive $A$ honk $B \vdash A$ honk $B$? Is Cut eliminable for formulas involving honk?

Q13 Expand the system with rules for a propositional connective with the following rules:

$$\frac{A \vdash R \qquad B \vdash R}{A \text{ plonk } B \vdash R}\ \text{plonk } L \qquad \frac{L \vdash B}{L \vdash A \text{ plonk } B}\ \text{plonk } R$$

What new things can you derive using plonk? Can you derive $A$ plonk $B \vdash A$ plonk $B$? Is Cut eliminable for formulas involving plonk?

### INTERMEDIATE EXERCISES

Q14 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 $p_1 \wedge (q_1 \vee r_1) \vdash (p_2 \vee q_2) \wedge r_2$ is the sequent $(p_2 \wedge q_2) \vee r_2 \vdash p_1 \vee (q_1 \wedge r_1)$. And then use this definition to prove the following theorem.

THEOREM 2.2.17 [DUALITY FOR DERIVATIONS] *A sequent $A \vdash B$ is derivable if and only if its dual $(A \vdash B)^d$ is derivable. Furthermore, the dual of the derivation of $A \vdash B$ is a derivation of the dual of $A \vdash B$.*

Q15 Even though the distribution sequent $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$ is not derivable (Example 2.2.11), some sequents of the form $A \wedge (B \vee C) \vdash (A \wedge B) \vee C$ *are* derivable. Give an independent characterisation of the triples $\langle A, B, C \rangle$ such that $A \wedge (B \vee C) \vdash (A \wedge B) \vee C$ is derivable.

Q16 Prove the invertibility result of Theorem 2.2.13 without appealing to the Cut rule or to Cut-elimination. (HINT: if a sequent $A \vee B \vdash C$ has a derivation $\delta$, 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 $A \vdash C$? Or to derive $B \vdash C$ instead? Prove this, and a similar result for $\wedge L$.)

ADVANCED EXERCISES

Q17 Define a notion of reduction for lattice derivations. Show that it is strongly normalising and that each derivation reduces to a unique cut-free derivation.

PROJECTS

Q18 Provide sequent formulations for logics intermediate between lattice logic and the logic of *distributive* lattices (those in which $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$). Characterise *which* logics intermediate between lattice logic 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.)

## 2.3 | FROM PROOFS TO DERIVATIONS AND BACK

The goal in this section is to collect together what we have learned so far into a more coherent picture. We will begin to see how natural deduction and sequent systems can be related. It seems clear that there must be connections, as the normalisation theorem and the proof of the redundancy of Cut have a similar "flavour." They both result in a subformula property, they can be proved in similar ways. In this section we will show how close this connection can be.

So, to connect sequent systems and natural deduction, start to think of a derivation of $A \vdash B$ as declaring the existence of a *proof* from $A$ to $B$. A proof from $A$ to $B$ cannot be a derivation: as these proofs contain sequents, not just formulas. A proof from $A$ to $B$ is a natural deduction proof. Thinking of the rules in a sequent system, then, perhaps we can understand them as telling us about a existence (and perhaps the construction) of natural deduction proofs. For example, the step from $L \vdash A$ and $L \vdash B$ to $L \vdash A \wedge B$ might be seen as saying that if we have a proof from $L$ to $A$ and another proof from $L$ to $B$ then these may (somehow) be combined into a proof from $L$ to $A \wedge B$.

The story is not completely straightforward, for we have different vocabularies for our Gentzen system and for natural deduction. By the end of this section we will put them together and look at the logic of conjunction, disjunction and implication. But for now, let us focus on implication alone. Natural deduction proofs in this vocabulary can have many assumptions but always only one conclusion. This means that a natural way of connecting these arguments with sequents is to use sequents of the form $X \vdash A$ where $X$ is a *multiset* and $A$ a single formula.

### 2.3.1 | SEQUENTS FOR LINEAR CONDITIONALS

In this section we will examine *linear* natural deduction, and sequent rules appropriate for it. We need rules for conditionals in a *sequent* context. That is, we want rules that say when it is appropriate to introduce a conditional on the *left* of a sequent, and when it is appropriate to introduce one on the *right*. The rule for conditionals on the *right* seems straightforward:

$$\frac{X, A \vdash B}{X \vdash A \to B} \to R$$

The rule makes *sense* when talking about proofs. If $\pi_1$ is a proof from $X, A$ to $B$, then we can extend it into a proof from $X$ to $A \to B$ by discharging $A$. We use only linear discharge, so we read this rule quite literally. $X, A$ is the multiset containing one more instance of $A$ than $X$ does. We delete that instance of $A$ from $X, A$, and we have the antecedent multiset $X$, from which we can deduce $A \to B$, discharging just that instance of $A$.

The rule for conditionals on the *left*, on the other hand, is not as straightforward as the right rule. Just as with our Gentzen system

for $\wedge$ and $\vee$, we want a rule that introduces our connective in the antecedent of the sequent. This means we are after a rule that indicates when it is appropriate to infer something *from* a conditional formula. The canonical case of inferring something from a conditional formula is by *modus ponens*. This motivates the sequent

$$A \to B, A \vdash B$$

and this should be derivable as a sequent in our system. However, this is surely not the only context in which we may introduce $A \to B$ into the left of a sequent. We may want to infer from $A \to B$ when the minor premise $A$ is not an *assumption* of our proof, but is itself deduced from some other premise set. That is, we at least want to endorse this step:

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

If we have a proof from $A$ on the basis of $X$ then adding to this proof a new assumption of $A \to B$ will lead us to $B$, when we add the extra step of $\to I$. This is straightforward enough. However, we may not only think that the $A$ has been derived from other material — we may also think that the conclusion $B$ has already been used as a premise in another proof. It would be a shame to have to use a Cut to deduce what follows from $B$ (or perhaps, what follows from $B$ together with other premises $Y$.) In other words, we should endorse this inference:

$$\frac{X \vdash A \quad B, Y \vdash C}{A \to B, X, Y \vdash C} \to L$$

which tells us how we can infer from $A \to B$. If we can infer *to* $A$ and *from* $B$, then adding the assumption of $A \to B$ lets us connect the proofs. This is clearly very closely related to the Cut rule, but it satisfies the subformula property, as $A$ and $B$ remain present in the conclusion sequent. The Cut rule is as before, except with the modification for our new sequents. The cut formula $C$ is one of the antecedents in the sequent $C, Y \vdash R$, and it is cut out and replaced by whatever assumptions are required in the proof. This motivates the following four rules for derivations in this sequent calculus in Figure 2.5.

Sequent derivations using these rules can be constructed as follows:

$$\dfrac{\dfrac{\dfrac{p \vdash p \qquad \dfrac{q \vdash q \qquad r \vdash r}{q, q \to r \vdash r} \to L}{p \to q, q \to r, p \vdash r} \to L}{p \to q, q \to r \vdash p \to r} \to R}{p \to q \vdash (q \to r) \to (p \to r)} \to R$$

They have the same structure as the sequent derivations for conjunction and disjunction seen in the previous section: if we do not use the Cut rule, then derivations have the subformula property. Now, however, there are more positions for formulas to appear in a sequent, as

$$p \vdash p \text{ [Id]}$$

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

$$\frac{X \vdash C \quad C, Y \vdash R}{X, Y \vdash R} \mathit{Cut}$$

Figure 2.5: SEQUENTS FOR CONDITIONALS

many formulas may appear on the left hand side of the sequent. In the rules in Figure 2.5, a formula appearing in the spots filled by $p$, $A$, $B$, $A \to B$, or $C$ are *active*, and the formulas in the other positions — filled by $X$, $Y$ and $R$ — are *passive*.

As we've seen, these rules can be understood as "talking about" the natural deduction system. We can think of a derivation of the sequent $X \vdash A$ as a recipe for constructing a proof from $X$ to $A$. We may define a mapping, giving us for each derivation $\delta$ of $X \vdash A$ a proof $nd(\delta)$ from $X$ to $A$.

DEFINITION 2.3.1 [$nd$ : DERIVATIONS → PROOFS] For any sequent derivation $\delta$ of $X \vdash A$, there is a natural deduction proof $nd(\delta)$ from the premises $X$ to the conclusion $A$. It is defined *recursively* by first choosing $nd$ of an *identity* derivation, and then, given $nd$ of simpler derivations, we define $nd$ of a derivation extending those derivations by $\to L$, $\to I$, or Cut:

» If $\delta$ is an identity sequent $p \vdash p$, then $nd(\delta)$ is the proof with the sole assumption $p$. This is a proof from $p$ to $p$.

» If $\delta$ is a derivation

$$\frac{\begin{array}{c}\vdots\,\delta' \\ X, A \vdash B\end{array}}{X \vdash A \to B} \to R$$

then we already have the proof $nd(\delta')$ from $X, A$ to $B$. The proof $nd(\delta)$, from $X$ to $A \to B$ is the following:

$$\frac{\begin{array}{c}X, [A]^{(i)} \\ \vdots\, nd(\delta') \\ B\end{array}}{A \to B} \to I, i$$

» If $\delta$ is a derivation

$$\frac{\vdots\,\delta_1 \quad X \vdash A \qquad \vdots\,\delta_2 \quad B, Y \vdash R}{A \to B, X, Y \vdash R} \to L$$

then we already have the proofs $nd(\delta_1)$ from X to A and $nd(\delta_2)$ from B, Y to R. The proof $nd(\delta)$, from $A \to B, X, Y$ to R is the following:

$$\dfrac{\begin{array}{cc} & \begin{array}{c} X \\ \vdots \\ nd(\delta_1) \\ \vdots \\ A \end{array} \\ A \to B & \end{array}}{B}\,{\to}E \qquad \begin{array}{c} Y \\ \vdots \\ nd(\delta_2) \\ \vdots \\ R \end{array}$$

» If $\delta$ is a derivation

$$\dfrac{\begin{array}{cc} \begin{array}{c} \vdots \\ \delta_3 \\ \vdots \\ X \vdash C \end{array} & \begin{array}{c} \vdots \\ \delta_4 \\ \vdots \\ C, Y \vdash R \end{array} \end{array}}{X, Y \vdash R}\,Cut$$

then we already have the proofs $nd(\delta_3)$ from X to C and $nd(\delta_4)$ from C, Y to R. The proof $nd(\delta)$, from X, Y to R is the following:

$$\begin{array}{c} X \\ \vdots \\ nd(\delta_3) \\ \vdots \\ C \end{array} \qquad \begin{array}{c} Y \\ \vdots \\ nd(\delta_4) \\ \vdots \\ R \end{array}$$

Using these rules, we may read derivations as sets of instructions constructing proofs. If you examine the instructions closely, you will see that we have in fact proved a stronger result, connecting normal proofs and cut-free derivations.

THEOREM 2.3.2 [NORMALITY AND CUT-FREEDOM] *For any cut-free derivation $\delta$, the proof $nd(\delta)$ is normal.*

*Proof*: This can be seen in a close examination of the steps of construction. Prove it by induction on the recursive construction of $\delta$. If $\delta$ is an identity step, $nd(\delta)$ is normal, so the induction hypothesis is satisfied. Notice that whenever an $\to E$ step is added to the proof, the major premise is a new assumption in a proof with a different conclusion. Whenever an $\to I$ step is added to the proof, the conclusion is added at the bottom of the proof, and hence, it cannot be a major premise of an $\to E$ step, which is an assumption in that proof and not a conclusion.

The only way we could introduce an indirect pair in to $nd(\delta)$ would be by the use of the Cut rule, so if $\delta$ is cut-free, then $nd(\delta)$ is normal. ∎

Another way to understand this result is as follows: the connective rules of a sequent system introduce formulas involving that connective either on the *left* or the *right*. Looking at in from the point of view of a *proof*, that means that the new formula is either introduced as an *assumption* or as a *conclusion*. In this way, the new material in the proof is always built on top of the old material, and we never compose

an introduction with an elimination in such a way as to have an indirect pair in a proof. The only way to do this is by way of a Cut step.

This mapping from sequent derivations to proofs brings to light one difference between the systems as we have set up. As we have defined them, there is *no* derivation $\delta$ such that $nd(\delta)$ delivers the simple proof consisting of the sole assumption $p \to q$. It would have to be a derivation of the sequent $p \to q \vdash p \to q$, but the proof corresponding to this derivation is more complicated than the simple proof consisting of the assumption alone:

$$\delta : \dfrac{\dfrac{p \vdash p \quad q \vdash q}{p \to q, p \vdash q}\,{\to}L}{p \to q \vdash p \to q}\,{\to}R \qquad nd(\delta) : \dfrac{\dfrac{p \to q \quad [p]^{(1)}}{q}\,{\to}E}{p \to q}\,{\to}I,1$$

What are we to make of this? If we want there to be a derivation constructing the simple proof for the argument from $p \to p$ to itself, an option is to extend the class of derivations somewhat:

$$\delta^{?} : \; p \to p \vdash p \to p \qquad nd(\delta^{?}) : \; p \to p$$

If $\delta^{?}$ is to be a derivation, we can expand the scope of the identity rule, to allow arbitrary formulas, instead of just atoms.

$$A \vdash A \; [\mathrm{Id}^{+}]$$

This motivates the following distinction:

DEFINITION 2.3.3 [LIBERAL AND STRICT DERIVATIONS] A *strict* derivation of a sequent is one in which [Id] is the only identity rule used. A *liberal* derivation is one in which $[\mathrm{Id}^{+}]$ is permitted to be used.

LEMMA 2.3.4 *If $\delta$ is a liberal derivation, it may be extended into $\delta^{st}$, a strict derivation of the same sequent. Conversely, a strict derivation is already a liberal derivation. A strict derivation featuring a sequent $A \vdash A$ where $A$ is complex, may be truncated into a liberal derivation by replacing the derivation of $A \vdash A$ by an appeal to $[Id^{+}]$.*

*Proof*: The results here are a matter of straightforward surgery on derivations. To transform $\delta$ into $\delta^{st}$, replace each appeal to $[\mathrm{Id}^{+}]$ to justify $A \vdash A$ with the identity derivation $\mathrm{Id}(A)$ to derive $A \vdash A$.

Conversely, in a strict derivation $\delta$, replace each derivation of an identity sequent $A \vdash A$, below which there are no more identity sequents, with an appeal to $[\mathrm{Id}^{+}]$ to find the smallest liberal derivation corresponding to $\delta$. ∎

From now on, we will focus on *liberal* derivations, with the understanding that we may "strictify" our derivations if the need or desire arises.

So, we have $nd$ : DERIVATIONS $\to$ PROOFS. This transformation also sends cut-free derivations to normal proofs. This lends some support to the view that derivations without cut and normal proofs are closely

related, and that cut elimination and normalisation are in some sense the *same* kind of process. Can we make this connection tighter? What about the reverse direction? Is there a map that takes proofs to derivations? There is, but the situation is somewhat more complicated. In the rest of this section we will see how to transform proofs into derivations, and we will examine the way that normal proofs can be transformed into cut-free derivations.

Firstly, note that the map *nd* is many-to-one. There are derivations $\delta \neq \delta'$ such that $nd(\delta) = nd(\delta')$. Here is a simple example:

$$\delta:\ \dfrac{\dfrac{p \vdash p \quad q \vdash q}{p \to q, p \vdash q}\,{\to}L \qquad r \vdash r}{q \to r, p \to q, p \vdash r}\,{\to}L \qquad\qquad \delta':\ \dfrac{p \vdash p \qquad \dfrac{q \vdash q \quad r \vdash r}{q \to r, q \vdash r}\,{\to}L}{q \to r, p \to q, p \vdash r}\,{\to}L$$

Applying *nd* to $\delta$ and $\delta'$, you generate the one proof $\pi$ in two different ways:

$$\pi:\ \dfrac{q \to r \qquad \dfrac{p \to q \quad p}{q}\,{\to}E}{r}\,{\to}E$$

This means that there are at least two different ways to make the reverse trip, from $\pi$ to a derivation. The matter is more complicated than this. There is another derivation $\delta''$, using a cut, such that $nd(\delta'') = \pi$.

$$\dfrac{\dfrac{p \vdash p \quad q \vdash q}{p \to q, \vdash q}\,{\to}L \qquad \dfrac{q \vdash q \quad r \vdash r}{q \to r, q \vdash r}\,{\to}L}{q \to r, p \to q, p \vdash r}\,Cut$$

So, even though *nd* sends cut-free derivations to normal proofs, it also sends some derivations with cut to normal proofs.

To understand what is going on, we will consider two different ways to reverse the trip, to go from a proof $\pi$ to a (possibly liberal) derivation $\delta$.

BOTTOM-UP CONSTRUCTION OF DERIVATIONS: If $\pi$ is a proof from $X$ to $A$ then $sq^{b}(\pi)$ is a derivation of the sequent $X \vdash A$, defined as follows:

- If $\pi$ is an assumption $A$, then $sq^{b}(\pi)$ is the identity derivation $\mathrm{Id}(A)$.
- If $\pi$ is a proof from $X$ to $A \to B$, composed from a proof $\pi'$ from $X, A$ to $B$ by a conditional introduction, then take the derivation $sq^{b}(\pi')$ of the sequent $X, A \vdash B$, and extend it with a $[{\to}R]$ step to conclude $X \vdash A \to B$.

$$\dfrac{\begin{array}{c}\vdots\\ sq^{b}(\pi')\\ \vdots\\ X, A \vdash B\end{array}}{X \vdash A \to B}\,{\to}R$$

» If $\pi$ is composed from a proof $\pi'$ from $X$ to $A \to B$ and another proof $\pi''$ from $Y$ to $A$, then take the derivations $sq^b(\pi')$ of $X \vdash A \to B$ and $sq^b(\pi')$ of $Y \vdash A$,

$$\dfrac{\dfrac{\vdots\, sq^b(\pi')}{X \vdash A \to B} \quad \dfrac{\dfrac{\vdots\, sq^b(\pi'')}{Y \vdash A} \qquad B \vdash B}{A \to B, Y \vdash B}\,{\to}L}{X, Y \vdash B}\,Cut$$

This definition constructs a derivation for each natural deduction proof, from the bottom to the top.

The first thing to notice about $sq^b$ is that it does not always generate a cut-free derivation, even if the proof you start off with is normal. We always use a Cut in the translation of a $\to E$ step, whether or not the proof $\pi$ is normal. Let's look at how this works in an example: we can construct $sq^b$ of the following normal proof:

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

We are going to construct a derivation of the sequent $p \to q \vdash (q \to r) \to (p \to r)$. We start by working back from the last step, using the definition. We have the following part of the derivation:

$$\dfrac{\vdots\, \delta'}{p \to q, q \to r \vdash p \to r}\,{\to}R \qquad \text{where } \delta' \text{ is } sq^b \text{ of } \dfrac{\dfrac{\dfrac{p \to q \quad [p]^{(1)}}{q}\,{\to}E}{r}\,{\to}E}{p \to r}\,{\to}I,1$$

Going back, we have

$$\dfrac{\dfrac{\vdots\, \delta''}{p \to q, q \to r, p \vdash r}\,{\to}R}{p \to q, q \to r \vdash p \to r}\,{\to}R \qquad \text{where } \delta'' \text{ is } sq^b \text{ of } \dfrac{\dfrac{p \to q \quad p}{q}\,{\to}E}{r}\,{\to}E$$

And going back further we get:

$$\dfrac{\dfrac{\dfrac{q \to r \vdash q \to r \qquad \dfrac{\vdots\, \delta'''}{\dfrac{p \to q, p \vdash q \quad r \vdash r}{q \to r, p \to q, p \vdash r}\,{\to}L}}{q \to r, p \to q, p \vdash r}\,Cut}{q \to r, p \to q \vdash p \to r}\,{\to}R}{q \to r \vdash (p \to q) \to (p \to r)}\,{\to}R \qquad \text{where } \delta''' \text{ is } sq^b \text{ of } \dfrac{p \to q \quad p}{q}\,{\to}E$$

Finally, we get

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}\quad \dfrac{}{q \vdash q}}{p \to q \vdash p \to q}\to L \quad \dfrac{}{p \to q, p \vdash q}}{p \to q, p \vdash q}\,Cut \qquad \dfrac{}{r \vdash r}}{p \to r, p \to q, p \vdash r}\to L}{q \to r \vdash q \to r \qquad p \to r, p \to q, p \vdash r}\,Cut}{p \to q, q \to r, p \vdash r}\to R}{p \to q, q \to r \vdash p \to r}\to R}{p \to q \vdash (q \to r) \to (p \to r)}\to R$$

This contains redundant Cut steps (we applied Cuts to identity sequents, and these can be done away with). We can eliminate these, to get a much simpler cutfree derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}\quad \dfrac{}{q \vdash q}}{p \to q, p \vdash r}\to L \qquad \dfrac{}{r \vdash r}}{q \to r, p \to q, p \vdash r}\to L}{p \to q, q \to r \vdash p \to r}\to R}{p \to q \vdash (q \to r) \to (p \to r)}\to R$$

You can check for yourself that when you apply *nd* to this derivation, you construct the original proof.

So, we have transformed the proof $\pi$ into a derivation $\delta$, which contained Cuts, and in this case, we eliminated them. Is there a way to construct a cut-free derivation in the first place? It turns out that there is. We need to construct the proof in a more subtle way than unravelling it from the bottom.

PERIMETER CONSTRUCTION OF DERIVATIONS: If we wish to generate a cut-free derivation from a normal proof, the subtlety is in how we use $\to L$ to encode an $\to E$ step. We want the major premise to appear as an assumption in the natural deduction proof. That means we must defer the decoding of an $\to E$ step until the major premise is an undischarged assumption in the proof. Thankfully, we can always do this if the proof is normal.

LEMMA 2.3.5 [NORMAL PROOF STRUCTURE] *Any normal proof, using the rules* $\to$I *and* $\to$E *alone, is either an assumption, or ends in an* $\to$I *step, or contains an undischarged assumption that is the major premise of an* $\to$E *step.*

*Proof*: We show this by induction on the construction of the proof $\pi$. We want to show that the proof $\pi$ has the property of being either (a) an assumption, (b) ends in an $\to I$ step, or (c) contains an undischarged assumption that is the major premise of an $\to E$ step. Consider how $\pi$ is constructed.

» If $\pi$ is an assumptions, it qualifies under condition (a).

» So, suppose that $\pi$ ends in $\to I$. Then it qualifies under condition (b).

» Suppose that $\pi$ ends in $\to E$. Then $\pi$ conbines two proofs, $\pi_1$ ending in $A \to B$ and $\pi_2$ ending in $A$, and we compose these with an $\to E$ step to deduce $B$. Since $\pi_1$ and $\pi_2$ are normal, we may presume the induction hypothesis, and that either (a), (b) or (c) apply to each proof. Since the whole proof $\pi$ is normal, we know that the proof $\pi_1$ cannot end in an $\to E$ step. So, it must satisfy property (a) or property (c). If it is (c), then one of the undischarged assumptions in $\pi_1$ is the major premise of an $\to E$ step, and it is undischarged in $\pi$, and hence $\pi$ satisfies property (c). If, on the other hand, $\pi_1$ satisfies property (a), then the formula $A \to B$, the major premise of the $\to E$ step concluding $\pi$, is undischarged, and $\pi$ also satisfies property (c). ∎

Now we may define the different map $sq^{\mathrm{p}}$ ("p" for "perimeter") according to which we strip each $\to I$ off the bottom of the proof $\pi$, until we have no more to take, and then, instead of dealing with the $\to E$ at the bottom of the proof, we deal with the the leftmost undischarged major premise of an $\to E$ step, unless there is none.

definition 2.3.6 [$sq^{\mathrm{p}}$] If $\pi$ is a proof from $X$ to $A$ then $sq^{\mathrm{p}}(\pi)$ is a derivation of the sequent $X \vdash A$, defined as follows:

» If $\pi$ is an assumption $A$, then $sq^{\mathrm{p}}(\pi)$ is the identity derivation $\mathrm{Id}(A)$.

» If $\pi$ is a proof from $X$ to $A \to B$, composed from a proof $\pi'$ from $X, A$ to $B$ by a conditional introduction, then take the derivation $sq^{\mathrm{p}}(\pi')$ of the sequent $X, A \vdash B$, and extend it with a $[\to R]$ step to conclude $X \vdash A \to B$.

$$\dfrac{\begin{array}{c}\vdots\, sq^{\mathrm{b}}(\pi') \\ X, A \vdash B\end{array}}{X \vdash A \to B}\ {\to}R$$

» If $\pi$ is a proof ending in a conditional elimination, then if $\pi$ contains an undischarged assumption that is the major premise of an $\to E$ step, choose the leftmost one in the proof. The proof $\pi$ will have the following form:

$$\dfrac{\begin{array}{cc} & Z \\ & \vdots\, \pi_2 \\ C \to D & C \end{array}}{\begin{array}{cc} D & \\ & \vdots\, \pi_3 \\ & A \end{array}}\ {\to}E \quad Y$$

Take the two proofs $\pi_2$ and $\pi_3$, and apply $sq^p$ to them to find derivations $sq^p(\pi_2)$ of $Z \vdash C$ and $sq^p(\pi_3)$ of $Y, D \vdash A$. Compose these with an $\rightarrow L$ step as follows:

$$\dfrac{\begin{array}{cc}\vdots\, sq^p(\pi_2) & \vdots\, sq^p(\pi_3) \\ Z \vdash C & Y, D \vdash A\end{array}}{C \rightarrow D, Z, Y \vdash A}\ {\scriptstyle \rightarrow L}$$

to complete the derivation for $C \rightarrow D, Z, Y \vdash A$.

» If, on the other hand, there is no major premise of an $\rightarrow E$ step that is an undischarged assumption in $\pi$ (in which case, $\pi$ is not normal), use a Cut as in the last part of the definition of $sq^b$ (the BOTTOM-UP translation) to split the proof at the final $\rightarrow E$ step.

This transformation will send a normal proof into a cut-free derivation, since a Cut is only used in the mapping when the source proof is not normal. We have proved the following result.

THEOREM 2.3.7 *For each natural deduction proof $\pi$ from $X$ to $A$, $sq^p(\pi)$ is a derivation of the sequent $X \vdash A$. Furthermore, if $\pi$ is normal, $sq^p(\pi)$ is cut-free.*

This has an important corollary.

COROLLARY 2.3.8 [CUT IS REDUNDANT] *If $\delta$ is a derivation of $X \vdash A$, then there is a cut-free derivation $\delta'$ of $X \vdash A$.*

*Proof*: Given a proof $\pi$, let $norm(\pi)$ be the *normalisation* of $\pi$. Then given $\delta$, consider $sq^p(norm(nd(\delta)))$. This is a cut free derivation of $X \vdash A$. The result is a cut-free derivation. ∎

This is a different way to prove the redundancy of Cut. Of course, we can prove the redundancy of Cut directly, by eliminating it from a derivation.

The crucial steps in the proof of the elimination of Cut are just as they were before. We have a derivation in which the last step is Cut, and we push this Cut upwards towards the leaves, or trade it in for a cut on a simpler formula. The crucial distinction is whether the cut formula is *active* in both sequents in cut step, or *passive* in either one. Consider the case in which the cut formula is active.

CUT FORMULA ACTIVE ON BOTH SIDES: In this case the derivation is as follows:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, \delta_1 \\ X, A \vdash B\end{array}}{X \vdash A \rightarrow B}\ {\scriptstyle \rightarrow R} \qquad \dfrac{\begin{array}{cc}\vdots\, \delta_2 & \vdots\, \delta_2' \\ Y \vdash A & B, Z \vdash C\end{array}}{A \rightarrow B, Y, Z \vdash C}\ {\scriptstyle \rightarrow L}}{X, Y, Z \vdash C}\ {\scriptstyle Cut}$$

Now, the Cut on $A \to B$ may be traded in for two simpler cuts: one on $A$ and the other on $B$.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_2\\ Y \vdash A\end{array} \qquad \begin{array}{c}\vdots\,\delta_1\\ X, A \vdash B\end{array}}{X, Y \vdash B}\,Cut \qquad \begin{array}{c}\vdots\,\delta_2'\\ B, Z \vdash C\end{array}}{X, Y, Z \vdash C}\,Cut$$

CUT FORMULA PASSIVE ON ONE SIDE: There are more cases to consider here, as there are more ways the cut formula can be passive in a derivation. The cut formula can be passive by occuring in $X$, $Y$, or $R$ in either [→L] or [→R]:

$$\dfrac{X \vdash A \quad B, Y \vdash R}{A \to B, X, Y \vdash R}\,{\to}L \qquad \dfrac{X, A \vdash B}{X \vdash A \to B}\,{\to}R$$

So, let's mark all of the different places that a cut formula could occur passively in each of these inferences. The inferences in Figure 2.6 mark the four different locations of a cut formula with $C$.

$$\dfrac{X', C \vdash A \quad B, Y \vdash R}{A \to B, X', C, Y \vdash R}\,{\to}L \qquad \dfrac{X \vdash A \quad B, Y', C \vdash R}{A \to B, X, Y', C \vdash R}\,{\to}L \qquad \dfrac{X \vdash A \quad B, Y \vdash C}{A \to B, X, Y \vdash C}\,{\to}L \qquad \dfrac{X', C, A \vdash B}{X', C \vdash A \to B}\,{\to}R$$

Figure 2.6: FOUR POSITIONS FOR PASSIVE CUT FORMULAS

In each case we want to show that a cut on the presented formula $C$ occuring in the lower sequent could be pushed up to occur on the upper sequent instead. (That is, that we could permute the Cut step and this inference.)

Start with the first example. We want to swap the Cut and the [→L] step in this fragment of the derivation:

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ Z \vdash C\end{array} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ X', C \vdash A\end{array} \qquad \begin{array}{c}\vdots\,\delta_2'\\ B, Y \vdash R\end{array}}{A \to B, X', C, Y \vdash R}\,{\to}L}{A \to B, X', Z, Y \vdash R}\,Cut$$

But the swap is easy to achieve. We do this:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_1\\ Z \vdash C\end{array} \qquad \begin{array}{c}\vdots\,\delta_2\\ X', C \vdash A\end{array}}{X', Z \vdash A}\,Cut \qquad \begin{array}{c}\vdots\,\delta_2'\\ B, Y \vdash R\end{array}}{A \to B, X', Z, Y \vdash R}\,{\to}L$$

The crucial feature of the rule [→L] that allows this swap is that it is closed under the substitution of formulas in passive position. We could

replace the formula C by Z in the inference without disturbing it. The result is still an instance of [→L]. The case of the second position in [→L] is similar. The Cut replaces the C in $A \to B, X, Y', C \vdash R$ by Z, and we could have just as easily deduced this sequent by cutting on the C in the premise sequent $B, Y', C \vdash R$, and then inferring with [→L].

The final case where the passive cut formula is on the left of the sequent is in the [→R] inference. We have

$$\dfrac{\dfrac{\vdots\,\delta_1}{Z \vdash C} \qquad \dfrac{\dfrac{\vdots\,\delta_2}{X', C, A \vdash B}}{X', C \vdash A \to B}\,{\to}R}{X', Z \vdash A \to B}\,Cut$$

and again, we could replace the C in the [→R] step by Z and still have an instance of the same rule. We permute the Cut and the [→R] step to get

$$\dfrac{\dfrac{\vdots\,\delta_1}{Z \vdash C} \qquad \dfrac{\vdots\,\delta_2}{X', C, A \vdash B}}{\dfrac{X', Z, A \vdash B}{X', Z \vdash A \to B}\,{\to}R}\,Cut$$

a proof of the same endsequent, in which the Cut is higher. The only other case is for an [→L] step in which the cut formula C is on the right of the turnstile. This is slightly more complicated. We have

$$\dfrac{\dfrac{\dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\vdots\,\delta_1'}{B, Y \vdash C}}{A \to B, X, Y \vdash C}\,{\to}L \qquad \dfrac{\vdots\,\delta_2}{Z, C \vdash D}}{Z, A \to B, X, Y \vdash D}\,Cut$$

In this case we can permute the cut and the [→L] step:

$$\dfrac{\dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\dfrac{\vdots\,\delta_1'}{B, Y \vdash C} \qquad \dfrac{\vdots\,\delta_2}{Z, C \vdash D}}{Z, B, Y \vdash D}\,Cut}{Z, A \to B, X, Y \vdash D}\,{\to}L$$

Here, we have taken the C in the step

$$\dfrac{X \vdash A \quad B, Y \vdash C}{A \to B, X, Y \vdash C}\,{\to}L$$

and cut on it. In this case, it does not simply mean replacing the C by another formula, or even by a multiset of formulas. Instead, when you cut with the sequent $Z, C \vdash D$, it means that you replace the C by D *and* you add Z to the left side of the sequent. So, we make the following transformation in the [→L] step:

$$\dfrac{X \vdash A \quad B, Y \vdash C}{A \to B, X, Y \vdash C}\,{\to}L \implies \dfrac{X \vdash A \quad B, Y, Z \vdash D}{A \to B, X, Y, Z \vdash D}\,{\to}L$$

The result is also a [→L] step.

We can see the features of [→L] and [→R] rules that allow the permutation with Cut. They are *preserved under cuts on formulas in passive position*. If you cut on a formula in passive position in the endsequent of the rule, then find the corresponding formula in the topsequent of the rule, and cut on it. The resulting inference is also an instance of the same rule. We have proved the following lemma:

"Corresponding formula"? This requires an *analysis* of the formulas in the rules, according to which you can match formulas in the top and bottom of rules to say which formula instance corresponds to which other instance. Nuel Belnap calls this an *analysis* [7] of the rules. It is often left implicit in discussions of sequent systems. The analysis we use here is as follows: Formula occurrences in an instance of a rule are *matched* if and only if they are either represented by the same schematic formula letter (A, B, etc.), or they occur in the corresponding place in a schematic multiset position (X, Y, etc.).

LEMMA 2.3.9 [CUT-DEPTH REDUCTION] *Given a derivation* $\delta$ *of* $X \vdash A$*, whose final inference is* Cut*, which is otherwise cut-free, and in which that inference has a depth of* $n$*, we may construct another derivation of* $X \vdash C$ *in which each cut on* $C$ *has a depth less than* $n$*.*

The only step for which the depth reduction might be in doubt is in the case where the cut formula is active on both sides. Before, we have

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X, A \vdash B\end{array}}{X \vdash A \to B}\,{\to}R \qquad \dfrac{\begin{array}{cc}\vdots\,\delta_2 & \vdots\,\delta_2'\\ Y \vdash A & B, Z \vdash C\end{array}}{A \to B, Y, Z \vdash C}\,{\to}L}{X, Y, Z \vdash C}\,\mathit{Cut}$$

and the depth of the cut is $|\delta_1| + 1 + |\delta_2| + |\delta_2'| + 1$. After pushing the Cut up we have:

We think of the derivation $\delta_1$ as containing its endsequent $X, A \vdash B$, and so on. So, $|\delta_1|$ is the number of sequents in that derivation, including its endsequent.

$$\dfrac{\dfrac{\begin{array}{cc}\vdots\,\delta_2 & \vdots\,\delta_1\\ Y \vdash A & X, A \vdash B\end{array}}{X, Y \vdash B}\,\mathit{Cut} \qquad \begin{array}{c}\vdots\,\delta_2'\\ B, Z \vdash C\end{array}}{X, Y, Z \vdash C}\,\mathit{Cut}$$

The depth of the first cut is $|\delta_2| + |\delta_1|$ (which is significantly shallower than depth of the previous cut), and the depth of the second is $|\delta_2| + |\delta_1| + 1 + |\delta_2'|$ (which is shallower by one). So, we have another proof of the cut elimination theorem, directly eliminating cuts in proofs by pushing them up until they disappear.

And this, gives a distinct proof of *normalisation* too. Go from a proof to a derivation, eliminate cuts, and then pull back with $nd$.

## 2.3.2 | STRUCTURAL RULES

What about non-linear proofs? If we allow vacuous discharge, or duplicate discharge, we must modify the rules of the sequent system in some manner. The most straightforward possibility is to change the rules for $\to R$, as it is the rule $\to I$ that varies in application when we use different policies for discharge. The most direct modification would be this:

$$\dfrac{X \vdash B}{X - A \vdash A \to B}\,{\to}R^-$$

where $X - A$ is a multiset found by deleting instances of $A$ from $X$. Its treatment depends on the discharge policy in place:

» In *linear* discharge, $X - A$ is the multiset $X$ with *one* instance of $A$ deleted. (If $X$ does not contain $A$, there is no multiset $X - A$.)

» In *relevant* discharge, $X - A$ is a multiset $X$ with *one or more* instances of $A$ deleted. (If $X$ contains more than one instance of $A$, then there are different multisets which can count as $X - A$: it is not a *function* of $X$ and $A$.)

» In *affine* discharge, $X - A$ is a multiset $X$ with *at most one* instance of $A$ deleted. (Now, there is always a multiset $X - A$ for any choice of $X$ and $A$. There are two choices, if $X$ actually contains $A$, delete it or not.)

» In *standard* discharge, $X - A$ is a multiset $X$ with *any number* (*including zero*) of instances of $A$ deleted.

The following derivations give examples of the new rule.

$$\dfrac{\dfrac{\dfrac{p \vdash p \quad \dfrac{q \vdash q}{p \to q, p \vdash q}\,{\to}L}{p \to (p \to q), p, p \vdash q}\,{\to}L}{p \to (p \to q) \vdash p \to q}\,{\to}R^{-}} \qquad \dfrac{\dfrac{p \vdash p \quad \dfrac{q \vdash q}{q \vdash r \to q}\,{\to}R^{-}}{p \to q, p \vdash r \to q}\,{\to}L}{p \to q \vdash p \to (r \to q)}\,{\to}R^{-}$$

In the first derivation, we discharge two instances of $p$, so this is a relevant sequent derivation, but not a linear (or affine) derivation. In the second derivation, the last $\to$R step is linear, but the first is not: it discharges a nonexistent instance of $r$.

These rules match our natural deduction system very well. However, they have undesirable properties. The rules for implication vary from system to system. However, the features of the system do not actually involve implication alone: they dictate the *structural* properties of deduction. Here are two examples. Allowing for vacuous discharge, if the argument from $X$ to $B$ is valid, so is the argument from $X, A$ to $B$.

$$\dfrac{\dfrac{\begin{array}{c}X\\ \vdots\\ B\end{array}}{A \to B}\,{\to}I \quad A}{B}$$

In other words, if we have a derivation for $X \vdash B$, then we also should have a derivation for $X, A \vdash B$. We do, if we go through $A \to B$ and a Cut.

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta\\ X \vdash B\end{array}}{X \vdash A \to B}\,{\to}R^{-} \quad \dfrac{A \vdash A \quad B \vdash B}{A \to B, A \vdash B}\,{\to}L}{X, A \vdash B}\,\mathit{Cut}$$

So clearly, if we allow vacuous discharge, the step from $X \vdash B$ to $X, A \vdash B$ is justified. Instead of requiring the dodgy move through $A \to B$, we may allow it the addition of an antecedent as a primitive rule.

'K' for wea**k**ening. We weaken the sequent by trading in a stronger fact (we can get B from X) for a weaker fact (we can get B from X with A).

$$\dfrac{X \vdash B}{X, A \vdash B}\,K$$

In just the same way, we can motivate the structural rule of *contraction*

$$\dfrac{X, A, A \vdash B}{X, A \vdash B}\, W$$

by going through $A \to B$ in just the same way.

Why "W" for contraction and "K" for *w*eakening? It is due to Schönfinkel's original notation for combinators [82].

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta\\ X, A, A \vdash B\end{array}}{X \vdash A \to B}\,{\to}R^{-} \qquad \dfrac{A \vdash A \quad B \vdash B}{A \to B, A \vdash B}\,{\to}L}{X, A \vdash B}\, Cut$$

With $K$ and $W$ we may use the old $\to R$ rule and 'factor out' the different behaviour of discharge:

$$\dfrac{\dfrac{\dfrac{p \vdash p \quad \dfrac{q \vdash q}{p \to q, p \vdash q}\,{\to}L}{p \to (p \to q), p, p \vdash q}\,{\to}L}{p \to (p \to q), p \vdash q}\, W}{p \to (p \to q) \vdash p \to q}\,{\to}R \qquad \dfrac{\dfrac{p \vdash p \quad \dfrac{\dfrac{q \vdash q}{q, r \vdash q}\, K}{q \vdash r \to q}\,{\to}R}{p \to q, p \vdash r \to q}\,{\to}L}{p \to q \vdash p \to (r \to q)}\,{\to}R$$

The structural rules do not *interfere* with the elimination of cut, though contraction does make the elimination of cut more difficult. The first thing to note is that *every* formula occurring in a structural rule is passive. We may commute cuts above structural rules. In the case of the weakening rule, the weakened-in formula appears only in the endsequent. If the cut is made on the weakened-in formula, it disappears, and is replaced by further instances of weakening, like this:

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \vdash A\end{array} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ Y \vdash B\end{array}}{Y, A \vdash B}\, K}{X, Y \vdash B}\, Cut \qquad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ Y \vdash B\end{array}}{X, Y \vdash B}\, K,\ repeated$$

In this case, the effect of the cut step is achieved without any cuts at all. The new derivation is clearly simpler, in that the derivation $\delta_1$ is rendered unnecessary, and the number of cuts decreases. If the cut formula is not the weakened in formula, then the cut permutes trivially with the weakening step:

$$\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \vdash A\end{array} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ Y, A \vdash B\end{array}}{Y, A, C \vdash B}\, K}{X, Y, C \vdash B}\, Cut \qquad \dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_1\\ X \vdash A\end{array} \qquad \begin{array}{c}\vdots\,\delta_2\\ Y, A \vdash B\end{array}}{X, Y \vdash B}\, Cut}{X, Y, C \vdash B}\, K$$

In the case contraction formula matters are not so simple. If the contracted formula is the cut formula, it occurs once in the endsequent

but twice in the topsequent. This means that if this formula is the cut formula, when the cut is pushed upwards it duplicates.

$$\dfrac{\dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\dfrac{\vdots\,\delta_2}{Y, A, A \vdash B}}{Y, A \vdash B}\,W}{X, Y \vdash B}\,Cut$$

$$\dfrac{\dfrac{\dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\vdots\,\delta_2}{Y, A, A \vdash B}}{X, Y, A \vdash B}\,Cut}{X, X, Y \vdash B}\,Cut}{X, Y \vdash B}\,W,\ repeated$$

In this case, the new proof is not less complex than the old one. The depth of the second cut in the new proof ($2|\delta_1| + |\delta_2| + 1$) is *greater* than in the old one ($|\delta_1| + |\delta_2|$). The old proof of cut elimination no longer works in the presence of contraction. There are number of options one might take here. Gentzen's own approach is to prove the elimination of *multiple* applications of cut.

$$\dfrac{\dfrac{\dfrac{\vdots\,\delta_1}{X \vdash A} \qquad \dfrac{\vdots\,\delta_2}{Y, A, A \vdash B}}{X, X, Y \vdash B}\,Multicut}{X, Y \vdash B}\,W,\ repeated$$

Another option is to eliminate contraction as one of the rules of the system (retaining its effect by rewriting the connective rules) [28]. In our treatment of the elimination of Cut we will not take either of these approaches. We will be more subtle in the formulation of the inductive argument, following a proof due to Haskell Curry [19, page 250], in which we trace the occurrences of the Cut formula in the derivation back to the points (if any) at which the formulas become active. We show how Cuts at *these* points derive the endsequent without further Cuts. But the details of the proof we will leave for later. Now we shall look to the behaviour of the other connectives.

### 2.3.3 | CONJUNCTION AND DISJUNCTION

Let's add conjunction and disjunction to our vocabulary. We have a number of options for the rules. One straightforward option would be to use natural deduction, and use the traditional rules. For example, the rules for conjunction in Gentzen's natural deduction are

$$\dfrac{A \quad B}{A \wedge B}\,\wedge I \qquad \dfrac{A \wedge B}{A}\,\wedge E \qquad \dfrac{A \wedge B}{B}\,\wedge E$$

Notice that the structural rule of weakening is implicit in these rules:

$$\dfrac{\dfrac{A \qquad B}{A \wedge B}\,\wedge I}{A}\,\wedge E$$

So, if we wish to add conjunction to a logic in which we don't have weakening, we must modify the rules. If we view these rules as sequents, it is easy to see what has happened:

$$\frac{X \vdash A \quad Y \vdash B}{X, Y \vdash A \wedge B}\,\wedge R? \qquad \frac{X, A \vdash R}{X, A \wedge B \vdash R}\,\wedge L? \qquad \frac{X, B \vdash R}{X, A \wedge B \vdash R}\,\wedge L?$$

The effect of weakening is then found using a Cut.

$$\dfrac{\dfrac{A \vdash A \quad B \vdash B}{A, B \vdash A \wedge B}\,\wedge L? \qquad \dfrac{A \vdash A}{A \wedge B \vdash A}\,\wedge R?}{A, B \vdash A}\,Cut$$

The $\wedge R?$ rule combines two contexts (X and Y) whereas the $\wedge L?$ does not combine two contexts—it merely infers from $A \wedge B$ to $A$ within the one context. The context 'combination' structure (the 'comma' in the sequents, or the structure of premises in a proof) is modelled using conjunction using the $\wedge R?$ rule but the structure is ignored by the $\wedge L?$ rule. It turns out that there are two *kinds* of of conjunction (and disjunction).

The rules in Figure 2.7 are *additive*. These rules do not exploit premise combination in the definition of the connectives. (The "X," in the conjunction left and disjunction left rules is merely a passive 'bystander' indicating that the rules for conjunction may apply regardless of the context.) These rules define conjunction and disjunction, regardless of the presence or absence of structural rules.

$$\frac{X, A \vdash R}{X, A \wedge B \vdash R}\,\wedge L_1 \qquad \frac{X, A \vdash R}{X, B \wedge A \vdash R}\,\wedge L_2 \qquad \frac{X \vdash A \qquad X \vdash B}{X \vdash A \wedge B}\,\wedge R$$

$$\frac{X, A \vdash R \qquad X, B \vdash R}{X, A \vee B \vdash R}\,\vee L \qquad \frac{X \vdash A}{X \vdash A \vee B}\,\vee R_1 \qquad \frac{X \vdash A}{X \vdash B \vee A}\,\vee R_2$$

Figure 2.7: additive conjunction and disjunction rules

These rules are the generalisation of the lattice rules for conjunction seen in the previous section. Every sequent derivation in the old system is a proof here, in which there is only one formula in the antecedent multiset. We may prove many new things, given the interaction of implication and the lattice connectives:

$$\dfrac{\dfrac{\dfrac{p \vdash p \quad q \vdash q}{p \to q, p \vdash q}\,\to L}{(p \to q) \wedge (p \to r), p \vdash q}\,\wedge L \qquad \dfrac{\dfrac{p \vdash p \quad r \vdash r}{p \to r, p \vdash r}\,\to L}{(p \to q) \wedge (p \to r), p \vdash r}\,\wedge L}{(p \to q) \wedge (p \to r), p \vdash q \wedge r}\,\wedge R$$
$$\frac{(p \to q) \wedge (p \to r), p \vdash q \wedge r}{(p \to q) \wedge (p \to r) \vdash p \to (q \wedge r)}\,\to R$$

Just as with the sequents with pairs of formulas, we cannot derive the sequent $p \wedge (q \vee r) \vdash (p \wedge q) \vee (p \wedge r)$—at least, we cannot without any structural rules. It is easy to see that there is no cut-free derivation of the sequent. There is no cut-free derivation using only sequents with single formulas in the antecedent (we saw this in the previous section) and a cut-free derivation of a sequent in $\wedge$ and $\vee$ containing no commas, will itself contain no commas (the additive conjunction and disjunction rules do not introduce commas into a derivation of a conclusion if the conclusion does not already contain them). So, there is no cut-free derivation of distribution. As we will see later, this means that there is no derivation at all.

But the situation changes in the presence of the structural rules. (See Figure 2.8.) This sequent is not derivable without the use of both weakening and contraction.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A, B \vdash A}\,K \quad \dfrac{B \vdash B}{A, B \vdash B}\,K}{A, B \vdash A \wedge B}\,\wedge R}{A, B \vdash (A \wedge B) \vee (A \wedge C)}\,\vee R_1 \qquad \dfrac{\dfrac{\dfrac{A \vdash A}{A, C \vdash A}\,K \quad \dfrac{C \vdash C}{A, C \vdash C}\,K}{A, C \vdash A \wedge C}\,\wedge R}{A, C \vdash (A \wedge B) \vee (A \wedge C)}\,\vee R_2}{A, B \vee C \vdash (A \wedge B) \vee (A \wedge C)}\,\vee L}{A, A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)}\,\wedge L_2}{A \wedge (B \vee C), A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)}\,\wedge L_1}{A \wedge (B \vee C) \vdash (A \wedge B) \vee (A \wedge C)}\,W$$

Figure 2.8: DISTRIBUTION OF $\wedge$ OVER $\vee$, USING K AND W.

Without using weakening, there is no way to derive $A, B \vdash A \wedge B$ using the additive rules for conjunction. If we think of the conjunction of A and B as the thing derivable from both A and B, then this seems to define a different connective. This motivates a different pair of conjunction rules, the so-called *multiplicative rules*. We use a different symbol (the tensor: $\otimes$) for conjunction defined with the multiplicative rules, because in certain proof-theoretic contexts (in the absence of either contraction or weakening), they differ.

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

For example, you can derive the distribution of $\otimes$ over $\vee$ (see Figure 2.9) in the absence of any structural rules. Notice that the derivation is much simpler than the case for additive conjunction.

### 2.3.4 | NEGATION

You can get *some* of the features of negation by defining it in terms of conditionals. If we pick a particular atomic proposition (call it f for

$$\dfrac{\dfrac{\dfrac{A \vdash A \quad B \vdash B}{A, B \vdash A \otimes B}\otimes R}{A, B \vdash (A \otimes B) \vee (A \otimes C)}\vee R_1 \qquad \dfrac{\dfrac{A \vdash A \quad C \vdash C}{A, C \vdash A \otimes C}\otimes R}{A, C \vdash (A \otimes B) \vee (A \otimes C)}\vee R_2}{\dfrac{A, B \vee C \vdash (A \otimes B) \vee (A \otimes C)}{A \otimes (B \vee C) \vdash (A \otimes B) \vee (A \otimes C)}\otimes L}\vee L$$

Figure 2.9: DISTRIBUTION OF $\otimes$ OVER $\vee$.

the moment) then $A \to f$ behaves somewhat like the negation of $A$. For example, we can derive $A \vdash (A \to f) \to f$, $(A \vee B) \to f \vdash (A \to f) \wedge (B \to f)$, and *vice versa*, $(A \to f) \wedge (B \to f) \vdash (A \vee B) \to f$.

$$\dfrac{\dfrac{\dfrac{A \vdash A \quad f \vdash f}{A \to f, A \vdash f}\to L}{(A \to f) \wedge (B \to f), A \vdash f}\wedge L_1 \qquad \dfrac{\dfrac{B \vdash B \quad f \vdash f}{B \to f, B \vdash f}\to L}{(A \to f) \wedge (B \to f), B \vdash f}\wedge L_2}{\dfrac{(A \to f) \wedge (B \to f), A \vee B \vdash f}{(A \to f) \wedge (B \to f) \vdash (A \vee B) \to f}\to R}\vee L$$

Notice that no special properties of $f$ are required for this derivation to work. The propsoition $f$ is completely arbitrary. Now look at the rules for implication in this special case of implying $f$:

$$\dfrac{X \vdash A \quad f, Y \vdash R}{A \to f, X, Y \vdash R}\to L \qquad \dfrac{X, A \vdash f}{X \vdash A \to f}\to R$$

If we want to do this without appealing to the proposition $f$, we could consider what happens if $f$ *goes away*. Write $A \to f$ as $\neg A$, and consider first [$\neg$R]. If we erase $f$, we get

$$\dfrac{X, A \vdash}{X \vdash \neg A}\neg R$$

Now the topsequent has an empty right-hand side. What might this mean? One possible interpretation is that $X, A \vdash$ is a *refutation* of $A$ in the context of $X$. This could be similar to a proof from $X$ and $A$ to a contradiction, except that we have no particular contradiction in mind. To derive $X, A \vdash$ is to refute $A$ (if we are prepared to keep $X$). In other words, from $X$ we can derive $\neg A$, the negation of $A$. If we keep with this line of investigation, consider the rule [$\to$L]. First, notice the right premise sequent $f, Y \vdash R$. A special case of this is $f \vdash$, and we can take this sequent as a given: if a refutation of a statement is a deduction from it to $f$, then $f$ is self-refuting. So, if we take $Y$ and $R$ to be empty, $A \to f$ to be $\neg A$ and $f \vdash$ to be given, then what is left of the [$\to$L] rule is this:

$$\dfrac{X \vdash A}{X, \neg A \vdash}\neg L$$

If we can deduce $A$ from $X$, then $\neg A$ is refuted (given $X$). This seems an eminently reasonable thing to mean by 'not'. And, these are Gentzen's rules for negation. With them, we can prove many of the usual properties of negation, even in the absence of other structural rules. The proof of distribution of negation over conjunction (one of the de Morgan laws) simplifies in the following way:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{\neg A, A \vdash}\,\neg L}{\neg A \wedge \neg B, A \vdash}\,\wedge L_1 \qquad \dfrac{\dfrac{B \vdash B}{\neg B, B \vdash}\,\neg L}{\neg A \wedge \neg B, B \vdash}\,\wedge L_2}{\neg A \wedge \neg B, A \vee B \vdash}\,\vee L}{\neg A \wedge \neg B \vdash \neg(A \vee B)}\,\neg R$$

We may show that $A$ entails its double negation $\neg\neg A$

$$\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}\,\neg L}{A \vdash \neg\neg A}\,\neg R$$

but we cannot prove the converse. There is no proof of $p$ from $\neg\neg p$. Similarly, there is no derivation of $\neg(p \wedge q) \vdash \neg p \vee \neg q$, using all of the structural rules considered so far. What of the other property of negation, that contradictions imply *everything*? We can get this far:

$$\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}\,\neg L}{A \otimes \neg A \vdash}\,\otimes L$$

(using contraction, we can derive $A \wedge \neg A \vdash$ too) but we must stop there in the absence of more rules. To get from here to $A \otimes \neg A \vdash B$, we must somehow add $B$ into the conclusion. But the $B$ is not there! How can we do this? We can come *close* by adding $B$ to the *left* by means of a weakening move:

$$\dfrac{\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}\,\neg L}{A, \neg A, B \vdash}\,K}{A, \neg A \vdash \neg B}\,\neg R$$

This shows us that a contradiction entails any *negation*. But to show that a contradiction entails *anything* we need a little more. We can do this by means of a structural rule operating on the right-hand side of a sequent. Now that we have sequents with empty right-hand sides, we may perhaps add things in that position, just as we can add things on the left by means of a weakening on the *right*. The rule of right weakening is just what is required to derive $A, \neg A \vdash B$.

$$\dfrac{X \vdash}{X \vdash B}\,KR$$

The result is a sequent system for *intuitionistic logic*. Intuitionistic logic arises out of the program of *intuitionism* in mathematics due to L. E. J. Brouwer [12, 50]. The entire family of rules is listed in Figure 2.10.

The sequents take the form $X \vdash R$ where $X$ is a multiset of formulas and $R$ is either a single formula or is empty. We take a derivation of $X \vdash A$ to record a *proof* of $X$ from $A$. Furthermore, a derivation of $X \vdash$ records a *refutation* of $X$. The system of intuitionistic logic is a stable, natural and useful account of logical consequence [20, 42, 79].

We have not presented the entire system of natural deduction in which these proofs may be found — *yet*.

| | |
|---|---|
| *Identity and Cut* | $p \vdash p \ [Id] \qquad \dfrac{X \vdash C \quad C, X' \vdash R}{X, X' \vdash R}\ Cut$ |
| *Conditional Rules* | $\dfrac{X \vdash A \quad B, X' \vdash R}{A \to B, X, X' \vdash R}\ {\to}L \qquad \dfrac{X, A \vdash B}{X \vdash A \to B}\ {\to}R$ |
| *Negation Rules* | $\dfrac{X \vdash A}{X, \neg A \vdash}\ \neg L \qquad \dfrac{X, A \vdash}{X \vdash \neg A}\ \neg R$ |
| *Conjunction Rules* | $\dfrac{X, A \vdash R}{X, A \wedge B \vdash R}\ {\wedge}L_1 \qquad \dfrac{X, A \vdash R}{X, B \wedge A \vdash R}\ {\wedge}L_2 \qquad \dfrac{X \vdash A \quad X \vdash B}{X \vdash A \wedge B}\ {\wedge}R$ |
| *Disjunction Rules* | $\dfrac{X, A \vdash R \quad X, B \vdash R}{X, A \vee B \vdash R}\ {\vee}L \qquad \dfrac{X \vdash A}{X \vdash A \vee B}\ {\vee}R_1 \qquad \dfrac{X \vdash A}{X \vdash B \vee A}\ {\vee}R_2$ |
| *Structural Rules* | $\dfrac{X, A, A \vdash R}{X, A \vdash R}\ WL \qquad \dfrac{X \vdash R}{X, A \vdash R}\ KL \qquad \dfrac{X \vdash}{X \vdash C}\ KR$ |

Figure 2.10: sequent rules for intuitionistic propositional logic

A case could be made for the claim that intuitionistic logic is the *strongest* logic is the *strongest* and most *natural* logic you can motivate using inference rules on sequents of the form $X \vdash R$. It is possible to go further and to add rules to ensure that the connectives behave as one would expect given the rules of *classical* logic: we can add the rule of *double negation elimination*

$$\frac{X \vdash \neg\neg A}{X \vdash A}\ DNE$$

This is equivalent to the natural deduction rule admitting the inference from $\neg\neg A$ to $A$, used in many systems of natural deduction for classical logic.

which strengthens the system far enough to be able to derive all classical tautologies and to derive all classically valid sequents. However, the results are not particularly attractive on proof-theoretical considerations. For example, the rule *DNE* does not satisfy the subformula property: the concluding sequent $X \vdash A$ is derived by way of the

premise sequent involving *negation*, even when negation does not feature in X or in A. This feature of the rule is exploited in the derivation of the sequent $\vdash ((p \to q) \to p) \to p$, of *Peirce's Law*, which is classically derivable but not derivable intuitionistically. This derivation uses negation liberally, despite the fact that the concluding sequent is negation-free.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{p \vdash p}}{p, (p \to q) \to p \vdash p}\,KL}{p \vdash ((p \to q) \to p) \to p}\,{\to}R}{\neg(((p \to q) \to p) \to p), p \vdash}\,{\neg}L}{\neg(((p \to q) \to p) \to p), p \vdash q}\,KR}{\neg(((p \to q) \to p) \to p) \vdash p \to q}\,{\to}R \qquad p \vdash p}{\neg(((p \to q) \to p) \to p), (p \to q) \to p \vdash p}\,{\to}L}{\neg(((p \to q) \to p) \to p) \vdash ((p \to q) \to p) \to p}\,{\to}R}{\neg(((p \to q) \to p) \to p), \neg(((p \to q) \to p) \to p) \vdash}\,{\neg}L}{\neg(((p \to q) \to p) \to p) \vdash}\,WL}{\vdash \neg\neg(((p \to q) \to p) \to p)}\,{\neg}R}{\vdash ((p \to q) \to p) \to p}\,DNE$$

It seems clear that this is not a particularly *simple* proof of Peirce's law. It violates the subformula property, by way of the detour through negation. Looking at the structure of the proof, it seems clear that the contraction step (marked *WL*) is crucial. We needed to duplicate the conditional for Peirce's law so that the inference of $\to L$ would work. Using $\to L$ on an unduplicated Peirce conditional does not result in a derivable sequent. The options for deriving $(p \to q) \to p \vdash p$ are grim:

$$\dfrac{\vdash p \to q \quad p \vdash p}{(p \to q) \to p \vdash p}\,{\to}L$$

No such proof will work. We must go through negation to derive the sequent, unless we can find a way to mimic the behaviour of the *WL* step without passing the formulas over to the left side of the turnstile, using negation.

## 2.3.5 | CLASSICAL LOGIC

Gentzen's great insight in the sequent calculus was that we could get the full power of *classical* logic by way of a small but profound change to the structure of sequents. We allow multisets on *both* sides of the turnstile. For intuitionistic logic we already allow a single formula or *none*. We now allow for more. The rules are trivial modifications of the standard intuitionistic rule, except for this one change. The rules are listed in Figure 2.11. In each sequent in the rules, 'p' is an atomic

proposition, 'A', 'B' and 'C' are formulas, X, X′, Y and Y′ are multisets (possibly empty), of formulas.

| | |
|---|---|
| *Identity and Cut* | $p \vdash p$ [*Id*] $\qquad \dfrac{X \vdash Y, C \quad C, X' \vdash Y'}{X, X' \vdash Y, Y'}\ Cut$ |
| *Conditional Rules* | $\dfrac{X \vdash Y, A \quad B, X' \vdash Y'}{X, X', A \to B \vdash Y, Y'}\ {\to}L \qquad \dfrac{X, A \vdash B, Y}{X \vdash A \to B, Y}\ {\to}R$ |
| *Negation Rules* | $\dfrac{X \vdash A, Y}{X, \neg A \vdash Y}\ \neg L \qquad \dfrac{X, A \vdash Y}{X \vdash \neg A, Y}\ \neg R$ |
| *Conjunction Rules* | $\dfrac{X, A \vdash Y}{X, A \wedge B \vdash Y}\ {\wedge}L_1 \qquad \dfrac{X, A \vdash Y}{X, B \wedge A \vdash Y}\ {\wedge}L_2 \qquad \dfrac{X \vdash A, Y \quad X' \vdash B, Y'}{X, X' \vdash A \wedge B, Y, Y'}\ {\wedge}R$ |
| *Disjunction Rules* | $\dfrac{X, A \vdash Y \quad X, B \vdash Y}{X, A \vee B \vdash Y}\ {\vee}L \qquad \dfrac{X \vdash A, Y}{X \vdash A \vee B, Y}\ {\vee}R_1 \qquad \dfrac{X \vdash A, Y}{X \vdash B \vee A, Y}\ {\vee}R_2$ |
| *Structural Rules* | $\dfrac{X, A, A \vdash Y}{X, A \vdash Y}\ WL \qquad \dfrac{X \vdash A, A, Y}{X \vdash A, Y}\ WR \qquad \dfrac{X \vdash Y}{X, A \vdash Y}\ KL \qquad \dfrac{X \vdash Y}{X \vdash A, Y}\ KR$ |

Figure 2.11: sequent rules for classical propositional logic

Using these rules, we can derive Peirce's Law, keeping the structure of the old derivation intact, other than the deletion of all of the steps involving negation. Instead of having to swing the formula for Peirce's Law onto the *left* to duplicate it in a contraction step, we may keep it on the right of the turnstile to perform the duplication. The negation laws are eliminated, the *WL* step changes into a *WR* step, but the other rules are unchanged.

You might think that this is what we were 'trying' to do in the other derivation, and we had to be sneaky with negation to do what we wished.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p \vdash p}{p, (p \to q) \to p \vdash p}\ KL}{p, (p \to q) \to p \vdash q, p}\ KR}{p \vdash q, ((p \to q) \to p) \to p}\ {\to}R}{\vdash p \to q, ((p \to q) \to p) \to p}\ {\to}R \qquad p \vdash p}{(p \to q) \to p \vdash p, ((p \to q) \to p) \to p}\ {\to}L}{\vdash ((p \to q) \to p) \to p, ((p \to q) \to p) \to p}\ {\to}R}{\vdash ((p \to q) \to p) \to p}\ WR$$

The system of rules is elegant and completely *symmetric* between left and right. The old derivations have mirror images in every respect.

Here are two double negation sequents:

$$\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}\,\neg L}{A \vdash \neg\neg A}\,\neg R \qquad \dfrac{\dfrac{A \vdash A}{\vdash A, \neg A}\,\neg R}{\neg\neg A \vdash A}\,\neg L$$

Here are two derivations of de Morgan laws. Again, they are completely left–right symmetric pairs (with conjunction and disjunction exchanged, but negation fixed).

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{\neg A, A \vdash}\,\neg L}{\neg A \wedge \neg B, A \vdash}\,\wedge L_1 \qquad \dfrac{\dfrac{B \vdash B}{\neg B, B \vdash}\,\neg L}{\neg A \wedge \neg B, B \vdash}\,\wedge L_2}{\neg A \wedge \neg B, A \vee B \vdash}\,\vee L}{\neg A \wedge \neg B \vdash \neg(A \vee B)}\,\neg R$$

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{\vdash \neg A, A}\,\neg R}{\vdash \neg A \vee \neg B, A}\,\vee R_1 \qquad \dfrac{\dfrac{B \vdash B}{\vdash \neg B, B}\,\neg L}{\vdash \neg A \vee \neg B, B}\,\wedge L_2}{\vdash \neg A \vee \neg B, A \wedge B}\,\wedge R}{\neg(A \wedge B) \vdash \neg A \vee \neg B}\,\neg L$$

The sequent rules for classical logic share the 'true–false' duality implicit in the truth-table account of classical validity. But this leads on to an important question. Intuitionistic sequents, of the form $X \vdash A$, record a proof from $X$ to $A$. What do classical sequents mean? Do they mean anything at all about *proofs*? A sequent of the form $A, B \vdash C, D$ does not tell us that $C$ and $D$ *both* follow from $A$ and $B$. (Then it could be replaced by the two sequents $A, B \vdash C$ and $A, B \vdash D$.) No, the sequent $A, B \vdash C, D$ may be valid even when $A, B \vdash C$ and $A, B \vdash D$ are not valid. The combination of the conclusions is *disjunctive* and not *conjunctive* when read 'positively'. We can think of a sequent $X \vdash Y$ as proclaiming that if *each* member of $X$ is true then *some* member of $Y$ is true. Or to put it 'negatively', it tells us that it would be a *mistake* to assert each member of $X$ and to deny each member of $Y$.

This leaves open the important question: is there any notion of *proof* appropriate for structures like these, in which premises and conclusions are collected in exactly the same way? Whatever is suitable will have to be quite different from the tree-structured proofs we have already seen.

## 2.3.6 | CUT ELIMINATION AND COROLLARIES

[To be written: A simple proof of the cut-elimination theorem of the systems we have seen (it *works*, though the presence of contraction on both sides of the turnstile makes things trickier). And a discussion of interpolation and other niceties.]

We will end this section with a demonstration of yet *another* way that we can show that cut is eliminable from a sequent system. We will show that the system of sequents—without the cut rule—suffices to derive every truth-table-valid sequent.

DEFINITION 2.3.10 [TRUTH TABLE VALIDITY] A sequent $X \vdash Y$ is truth-table valid if and only if there is no truth-table evaluation $v$ such that

$v(A)$ is TRUE for each $A$ in $X$ and $v(B)$ is FALSE for each $B$ in $Y$. If $X \vdash Y$ is not truth-table valid, then we say that an evaluation $v$ that makes each member of $X$ TRUE and each member of $Y$ FALSE a *counterexample* to the sequent.

In other words, a sequent $X \vdash Y$ is truth-table valid if and only if there is no way to make each element of $X$ TRUE while making each element of $Y$ FALSE. Or, if you like, if we make each member of $X$ TRUE, we must also make some member of $Y$ TRUE. Or, to keep the story balanced, if we make each member of $Y$ FALSE, we make some member of $X$ FALSE too. To understand the detail of this, we need another definition. It is what you expect.

DEFINITION 2.3.11 [TRUTH-TABLE EVALUATIONS] A function $v$ assigning each ATOM a truth value (either TRUE or FALSE) is said to be a truth-table evaluation. A truth-table evaluation assigns a truth value to each FORMULA as follows:

- $v(\neg A) =$ TRUE if and only if $v(A) =$ FALSE. (Otherwise, if $v(A) =$ TRUE, then $v(\neg A) =$ FALSE.)
- $v(A \wedge B) =$ TRUE if and only if $v(A) =$ TRUE and $v(B) =$ TRUE.
- $v(A \vee B) =$ TRUE if and only if $v(A) =$ TRUE or $v(B) =$ TRUE (or both).
- $v(A \to B) =$ TRUE if and only if either $v(A) =$ FALSE or $v(B) =$ TRUE.

Now, we will show two facts. Firstly, that if the sequent $X \vdash Y$ is derivable (with or without Cut) then it is truth-table valid. Second, we will show that if the sequent $X \vdash Y$ is *not* derivable, (again, with or without Cut) then it is not truth-table valid. Can you see why this proves that anything derivable with Cut may be derived without it?

THEOREM 2.3.12 [TRUTH-TABLE SOUNDNESS OF SEQUENT RULES] *If* $X \vdash Y$ *is derivable (using* Cut *if you wish), then it is truth-table valid.*

*Proof*: Axiom sequents (identities) are clearly truth-table valid. Take an instance of a rule: if the premises of that rule are truth-table valid, then so is the conclusion. We will consider two examples, and leave the rest as an exercise. Consider the rule $\to R$:

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

Suppose that $X \vdash Y, A$ and $B, X' \vdash Y'$ are both truth-table valid, and that we have an evaluation $v$ for which each member of $X, X'$, and $A \to B$ is TRUE. We wish to show that for $v$. Now, since $A \to B$ is TRUE according to $v$, it follows that either $A$ is FALSE or $B$ is TRUE (again, according to $v$). If $A$ is FALSE, then by the truth-table validity of $X \vdash Y, A$, it follows that one member (at least) of $Y$ is TRUE according to $v$. On the other hand, if $B$ is TRUE, then the truth-table validity of

$B, X' \vdash Y'$ tells us that one member (at least) of $Y'$ is TRUE. In either case, at least one member of $Y, Y'$ is TRUE according to $v$, as we desired.

Now consider the rule *KR*:

$$\frac{X \vdash Y}{X \vdash A, Y}\ KR$$

Suppose that $X \vdash Y$ is truth-table valid. Is $X \vdash A, Y$? Suppose we have an evaluation making each element of $X$ TRUE. By the truth-table validity of $X \vdash Y$, some member of $Y$ is TRUE according to $v$. It follows that some member of $Y, A$ is TRUE according to $v$ too. The other rules are no more difficult to verify, so (after you verify the rest of the rules to your own satisfaction) you may declare this theorem proved. ∎

THEOREM 2.3.13 [TRUTH-TABLE COMPLETENESS FOR SEQUENTS] *If* $X \vdash Y$ *is truth-table valid, then it is derivable. In fact, it has a derivation that does not use* Cut.

| | | |
|---|---|---|
| *Identity and Cut* | $X, A \vdash A, Y$ *[Id]* | $\dfrac{X \vdash Y, C \quad C, X \vdash Y}{X \vdash Y}\ Cut$ |
| *Conditional Rules* | $\dfrac{X \vdash Y, A \quad B, X \vdash Y}{A \to B, X \vdash Y}\ {\to}L$ | $\dfrac{X, A \vdash B, Y}{X \vdash A \to B, Y}\ {\to}R$ |
| *Negation Rules* | $\dfrac{X \vdash A, Y}{X, \neg A \vdash Y}\ {\neg}L$ | $\dfrac{X, A \vdash Y}{X \vdash \neg A, Y}\ {\neg}R$ |
| *Conjunction Rules* | $\dfrac{X, A, B \vdash Y}{X, A \wedge B \vdash Y}\ {\wedge}L$ | $\dfrac{X \vdash A, Y \quad X \vdash B, Y}{X \vdash A \wedge B, Y}\ {\wedge}R$ |
| *Disjunction Rules* | $\dfrac{X, A \vdash Y \quad X, B \vdash Y}{X, A \vee B \vdash Y}\ {\vee}L$ | $\dfrac{X \vdash A, B, Y}{X \vdash A \vee B, Y}\ {\vee}R$ |
| *Structural Rules* | $\dfrac{X, A, A \vdash Y}{X, A \vdash Y}\ WL$ | $\dfrac{X \vdash A, A, Y}{X \vdash A, Y}\ WR$ |

Figure 2.12: ALTERNATIVE SEQUENT RULES FOR CLASSICAL LOGIC

*Proof*: We prove the converse: that if the sequent $X \vdash Y$ has no cut-free derivation, then it is not truth-table valid. To do this, we appeal to the result of Exercise 8 to show that if we have a derivation (without Cut) using the sequent system given in Figure 2.12, then we have a derivation (also without Cut) in our original system. This result is not

too difficult to prove: simply show that the new identity axioms of the system in Figure 8 may be derived using our old identity together with instances of weakening; and that if the premises of any of the new rules are derivable, so are the conclusions, using the corresponding rule from the old system, and perhaps using judicious applications of contraction to manipulate the parameters.

The new sequent system has some very interesting properties. Suppose we have a sequent $X \vdash Y$, that has no derivation (not using Cut) in this system. then we may reason in the following way:

» Suppose $X \vdash Y$ contains no complex formulas, and only atoms. Since it is underivable, it is not an instance of the new *Id* rule. That is, it contains no formula common to both X and Y. Therefore, counterexample evaluation $v$: simply take each member of X to be TRUE and Y to be FALSE.

That deals with what we might call *atomic* sequents. We now proceed by induction, with the hypothesis for a sequent $X \vdash Y$ being that if it has no derivation, it is truth-table invalid. And we will show that if the hypothesis holds for *simpler* sequents than $X \vdash Y$ then it holds for $X \vdash Y$ too. What is a simpler sequent than $X \vdash Y$? Let's say that the complexity of a sequent is the number of connectives ($\wedge$, $\vee$, $\rightarrow$, $\neg$) occuring in that sequent. So, we have shown that the hypothesis holds for sequents of complexity zero.

Now to deal with sequents of greater complexity: that is, those containing formulas with connectives.

» Suppose that the sequent contains a negation formula. If this formula occurs on the left, the sequent has the form $X, \neg A \vdash Y$. It follows that if this is underivable in our sequent system, then so is $X \vdash A, Y$. If this *were* derivable, then we could derive our target sequent by $\neg L$. But look! This is a simpler sequent. So we may appeal to the induction hypothesis to give us a counterexample evaluation $v$, making each member of X TRUE and making A FALSE and each member of Y FALSE. Now since this evaluation makes A FALSE, then it makes $\neg A$ TRUE. So, it is a counterexample for $X, \neg A \vdash Y$ too.

If the negation formula occurs on the right instead, then the sequent has the form $X \vdash \neg A, Y$. It follows that $X, A \vdash Y$ is underivable (for otherwise, we could derive our sequent by $\neg R$). This is a simpler sequent, so it has a counterexample $v$, making each member of X, and A TRUE and Y FALSE. This is also a counterexample to $X \vdash \neg A, Y$, since it makes $\neg A$ FALSE.

» Suppose that our sequent contains a conditional formula $A \rightarrow B$. If it occurs on the left, the sequent has the form $A \rightarrow B, X \vdash Y$. If it is not derivable then, using the rules in Figure 2.12, we may conclude that either $X \vdash Y, A$ is underivable, or $B, X \vdash Y$ is underivable. (If they were both derivable, then we could use $\rightarrow L$ to derive our target sequent $A \rightarrow B, X \vdash Y$.) Both of these

sequents are simpler than our original sequent, so we may apply the induction hypothesis. If $X \vdash Y, A$ is underivable, we have an evaluation $v$ making each member of $X$ TRUE and each member of $Y$ FALSE, together with $A$ FALSE. But look! This makes $A \to B$ TRUE, so $v$ is a counterexample for $A \to B, X \vdash Y$. Similarly, if $B, X \vdash Y$ is underivable, we have a counterexample $v$, making each member of $X$ TRUE and each member of $Y$ FALSE, together with making $B$ TRUE. So we are in luck in this case too! The evaluation $v$ makes $A \to B$ true, so it is a counterexample to our target sequent $A \to B, X \vdash Y$. This sequent is truth-table invalid.

Suppose, on the other hand, that an implication formula is on the right hand side of the sequent. If $X \vdash A \to B, Y$ is not derivable, then neither is $X, A \vdash B, Y$, a simpler sequent. The induction hypothesis applies, and we have an evaluation $v$ making the formulas in $X$ TRUE, the formulas in $Y$ FALSE, and $A$ TRUE and $Y$ FALSE. So, it makes $A \to B$ FALSE, and our evaluation is a counterexample to our target sequent $X \vdash A \to B$. This sequent is truth-table invalid.

» The cases for conjunction and disjunction are left as exercises. They pose no more complications than the cases we have seen. ∎

The similarity to rules for *tableaux* is not an accident [85]. See Exercise 10 on page 97.

So, the sequent rules, read *backwards* from bottom-to-top, can be understood as giving instructions for making a counterexample to a sequent. In the case of sequent rules with more than one premise, these instructions provide *alternatives* which can both be explored. If a sequent is underivable, these instructions may be followed to the end, and we finish with a counterexample to the sequent. If following the instructions does not meet with success, this means that all searches have terminated with *derivable* sequents. So we may play this attempt backwards, and we have a derivation of the sequent.

## 2.3.7 | HISTORY

[To be written.]

## 2.3.8 | EXERCISES

### BASIC EXERCISES

Q1 Which of the following sequents can be proved in intuitionistic logic? For those that can, find a derivation. For those that cannot, find a derivation in *classical* sequent calculus:

1 : $p \to (q \to p \wedge q)$
2 : $\neg(p \wedge \neg p)$
3 : $p \vee \neg p$
4 : $(p \to q) \to ((p \wedge r) \to (q \wedge r))$
5 : $\neg\neg\neg p \to \neg p$
6 : $\neg(p \vee q) \to (\neg p \wedge \neg q)$
7 : $(p \wedge (q \to r)) \to (q \to (p \wedge r))$

8 : $p \vee (p \to q)$
9 : $(\neg p \vee q) \to (p \to q)$
10 : $((p \wedge q) \to r) \to ((p \to r) \vee (q \to r))$

Q2 Consider all of the formulas unprovable in Q1 on page 47. Find derivations for these formulas, using classical logic if necessary.

Q3 Define the *dual* of a classical sequent in a way generalising the result of Exercise 14 on page 67, and show that the dual of a derivation of a sequent is a derivation of the dual of a sequent. What is the dual of a formula involving implication?

Q4 Define $A \to^* B$ as $\neg(A \wedge \neg B)$. Show that any classical derivation of $X \vdash Y$ may be transformed into a classical derivation of $X^* \vdash Y^*$, where $X^*$ and $Y^*$ are the multisets $X$ and $Y$ respectively, with all instances of the connective $\to$ replaced by $\to^*$. Take care to explain what the transformation does with the rules for implication. Does this work for intuitionistic derivations?

Q5 Consider the rules for classical propositional logic in Figure 2.11. Delete the rules for negation. What is the resulting logic like? How does it differ from intuitionistic logic, if at all?

Q6 Define the *Double Negation Translation* $d(A)$ of formula $A$ as follows:

$$\begin{aligned} d(p) &= \neg\neg p \\ d(\neg A) &= \neg d(A) \\ d(A \wedge B) &= d(A) \wedge d(B) \\ d(A \vee B) &= \neg(\neg d(A) \wedge \neg d(B)) \\ d(A \to B) &= d(A) \to d(B) \end{aligned}$$

What formulas are $d((p \to q) \vee (q \to p)))$ and $d(\neg\neg p \to p)$? Show that these formulas have intuitionistic proofs by giving a sequent derivation for each.

### INTERMEDIATE EXERCISES

Q7 Using the double negation translation $d$ of the previous question, show how a *classical* derivation of $X \vdash Y$ may be transformed (with a number of intermediate steps) into an intuitionistic derivation of $X^d \vdash Y^d$, where $X^d$ and $Y^d$ are the multisets of the $d$-translations of each element of $X$, and of $Y$ respectively.

Q8 Consider the *alternative* rules for classical logic, given in Figure 2.12. Show that $X \vdash Y$ is derivable using *these* rules iff it is derivable using the *old* rules. Which of these new rules are invertible? What are some distinctive properties of these rules?

Q9 Construct a system of rules for *intuitionistic* logic with as similar as you can to the classical system in Figure 2.12. Is it quite *as* nice? Why, or why not?

Q10 Relate cut-free sequent derivations of $X \vdash Y$ with *tableaux* refutations of $X, \neg Y$ [44, 78, 85]. Show how to transform any cut-free sequent

derivation of $X \vdash Y$ into a corresponding closed tableaux, and vice-versa. What are the differences and similarities between tableaux and derivations?

Q11 Relate natural deduction proofs for intuitionistic logic in Gentzen–Prawitz style with natural deduction proofs in other systems (such as Lemmon [49], or Fitch [31]). Show how to transform proofs in one system into proofs in the other. How do the systems differ, and how are they similar?

### ADVANCED EXERCISES

Q12 Consider what sort of rules make sense in a sequent system with sequents of the form $A \vdash X$, where $A$ is a formula and $X$ a multiset. What connectives make sense? (One way to think of this is to define the *dual* of an intuitionistic sequent, in the sense of Exercise 3 in this section and Exercise 14 on page 67.)

This defines what Igor Urbas has called 'Dual-intuitionistic Logic' [93].

## 2.4 | CIRCUITS

In this section we will look at the kinds of *proofs* motivated by the two-sided classical sequent calculus. Our aim is to "complete the square."

$$\textit{Derivations} \text{ of } X \vdash A \quad \leftrightarrow \quad \textit{Proofs} \text{ from } X \text{ to } A$$

$$\textit{Derivations} \text{ of } X \vdash Y \quad \leftrightarrow \quad ???$$

Just *what* goes in that corner? If the parallel is to work, the structure is not a straightforward *tree* with premises at the top and conclusion at the bottom, as we have in proofs for a single conclusion A. What other structure could it be?

For the first example of proofs with multiple conclusions as well as multiple premises. We will not look at the case of classical logic, for the presence of the structural rules of weakening and contraction complicates the picture somewhat. Instead, we will start with a logic without these structural rules—linear logic.

### 2.4.1 | DERIVATIONS DESCRIBING CIRCUITS

We will start a with a simple sequent system for the multiplicative fragment of linear logic. So, we will do without the structural rules of contraction or weakening. However, sequents have multisets on the left and on the right. In this section we will work with the connectives $\oplus$ and $\otimes$ (multiplicative disjunction and conjunction respectively) and $\neg$ (negation). The sequent rules are as follows. First, negation flips conclusion to premise, and vice versa.

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

Multiplicative conjunction mirrors the behaviour of premise combination. We may trade in the two premises $A, B$ for the single premise $A \otimes B$. On the other hand, if we have a derivation of $A$ (from $X$, and with $Y$ as alternate conclusions) and a derivation of $B$ (from $X'$ and with $Y'$ as alternate conclusions) then we may combine these derivations to form a derivation of $A \otimes B$ from *both* collections of premises, and with *both* collections of alternative conclusions.

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

The case for multiplicative disjunction is dual to the case for conjunction. We swap premise and conclusion, and replace $\otimes$ with $\oplus$.

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

The cut rule is simple:

$$\frac{X \vdash A, Y \quad X', A \vdash Y'}{X, X' \vdash Y, Y'}\,Cut$$

Girard's preferred notation for multiplicative disjunction in linear logic is '⅋', to emphasise the connection between additive conjunction & and multiplicative disjunction ⅋ (and similarly, between his multiplicative disjunction $\oplus$ and additive conjunction &). We prefer to utilise familiar notation for the additive connectives, and use tensor notation for the multiplicatives where their behaviour differs markedly from the expected classical or intuitionistic behaviour.

The cut formula (here it is A) is left out, and all of the other material remains behind. Any use of the cut rule is eliminable, in the usual manner. Notice that this proof system has no conditional connective. Its loss is no great thing, as we could *define* $A \to B$ to be $\neg(A \otimes \neg B)$, or equivalently, as $\neg A \oplus B$. (It is a useful exercise to verify that these definitions are equivalent, and that they both "do the right thing" by inducing appropriate rules [→E] and [→I].) So that is our sequent system for the moment.

Let's try to find a notion of *proof* appropriate for the derivations in this sequent system. It is clear that the traditional many-premise single-conclusion structure does not fit neatly. The cut free derivation of $\neg\neg A \vdash A$ is no simpler and no more complex than the cut free derivation of $A \vdash \neg\neg A$.

$$\dfrac{\dfrac{\dfrac{A \vdash A}{\vdash \neg A, A}\,\neg R}{\neg\neg A \vdash A}\,\neg L} \qquad \dfrac{\dfrac{\dfrac{A \vdash A}{\neg A, A \vdash}\,\neg L}{A \vdash \neg\neg A}\,\neg R}$$

The natural deduction proof from A to $\neg\neg A$ goes through a stage where we have two premises A and $\neg A$ and has no active conclusion (or equivalently, it has the conclusion $\bot$).

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

In this proof, the premise $\neg A$ is then discharged or somehow otherwise converted to the conclusion $\neg\neg A$. The usual natural deduction proofs from $\neg\neg A$ to A are either *simpler* (we have a primitive inference from $\neg\neg A$ to A) or *more complicated*. A proof that stands to the derivation of $\neg\neg A \vdash A$ would require a stage at which there is no premise but two conclusions. We can get a hint of the desired "proof" by turning the proof for double negation introduction on its head:

$$\dfrac{\dfrac{\dfrac{\forall \vdash \vdash}{*}\,\vdash I, 1}{\forall \quad [\forall \vdash]^{(1)}}\,\exists \vdash$$

Let's make it easier to read by turning the formulas and labels the right way around, and swap I labels with E labels:

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

We are after a proof of double negation elimination at least as simple as *this*. However, constructing this will require hard work. Notice that not only does a proof have a different structure to the natural deduction proofs we have seen—there is *downward* branching, not upward—there is also the kind of "reverse discharge" at the bottom of the tree

which seems difficult to interpret. Can we make out a story like this? Can we define proofs appropriate to linear logic?

To see what is involved in answering this question in the affirmative, we will think more broadly to see what might be appropriate in designing our proof system. Our starting point is the behaviour of each rule in the sequent system. Think of a derivation ending in $X \vdash Y$ as having constructed a proof $\pi$ with the formulas in $X$ as premises or *inputs* and the formulas in $Y$ as conclusions, or *outputs*. We could think of a proof as having a shape reminiscent of the traditional proofs from many premises to a single conclusion:

$$\frac{A_1 \quad A_2 \quad \cdots \quad A_n}{B_1 \quad B_2 \quad \cdots \quad B_m}$$

However, chaining proofs together like this is notationally very difficult to depict. Consider the way in which the sequent rule [Cut] corresponds to the composition of proofs. In the single-formula-right sequent system, a Cut step like this:

$$\frac{X \vdash C \quad A, C, B \vdash D}{A, X, B \vdash D}\ \mathit{Cut}$$

corresponds to the composition of the proofs

$$\frac{\dfrac{X}{\pi_1}}{C} \quad \text{and} \quad \frac{A \quad C \quad B}{\dfrac{\pi_2}{D}} \quad \text{to form} \quad \frac{A \quad \dfrac{\dfrac{X}{\pi_1}}{C} \quad B}{\dfrac{\pi_2}{D}}$$

In the case of proofs with multiple premises and multiple conclusions, this notation becomes difficult if not impossible. The cut rule has an instance like this:

$$\frac{X \vdash D, C, E \qquad A, C, B \vdash Y}{A, X, B \vdash D, Y, E}\ \mathit{Cut}$$

This should correspond to the composition of the proofs

$$\frac{\dfrac{X}{\pi_1}}{D \quad C \quad E} \quad \text{and} \quad \frac{A \quad C \quad B}{\dfrac{\pi_2}{Y}}$$

If we are free to rearrange the order of the conclusions and premises, we could manage to represent the cut:

$$\frac{\dfrac{\dfrac{X}{\pi_1}}{D \quad E \quad C} \quad A \quad C}{\dfrac{\pi_2}{Y}}$$

But we cannot always rearrange the cut formula to be on the left of one proof and the right of the other. Say we want to cut with the conclusion E in the next step? What do we do?

It turns out that it is much more flexible to change our notation completely. Instead of representing proofs as consisting of characters on a page, ordered in a tree diagrams, think of proofs as taking inputs and outputs, where we represent the inputs and outputs as *wires*. Wires can be rearranged willy-nilly—we are all familiar with the tangle of cables behind the stereo or under the computer desk—so we can exploit this to represent cut straightforwardly. In our pictures, then, formulas will *label* wires. This change of representation will afford another insight: instead of thinking of the rules as labelling transitions between formulas in a proof, we will think of inference steps (instances of our rules) as *nodes* with wires coming in and wires going out. Proofs are then *circuits* composed of wirings of nodes. Figure 2.13 should give you the idea.

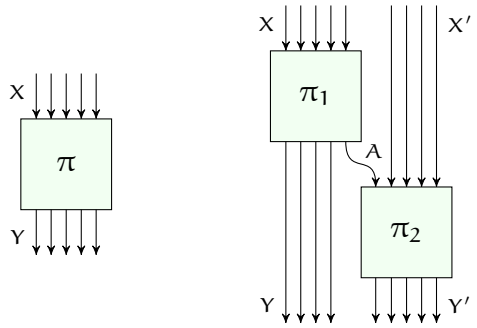
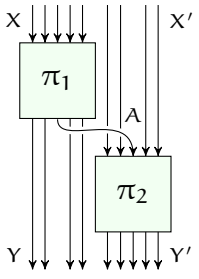



Figure 2.13: a circuit, and chaining together two circuits

Draw for yourself the result of making two cuts, one after another, inferring from the sequents $X_1 \vdash A, Y_1$ and $X_2, A \vdash B, Y_2$ and $X_3, B \vdash, Y_3$ to the sequent $X_1, X_2, X_3 \vdash Y_1, Y_2, Y_3$. You get two different possible derivations with different intermediate steps depending on whether you cut on A first or on B first. Does the order of the cuts matter when these different derivations are represented as circuits?

A proof $\pi$ for the sequent $X \vdash Y$ has premise or *input* wires for each formula in X, and conclusion or *output* wires for each formula in Y. Now think of the contribution of each rule to the development of inferences. The *cut* rule is the simplest. Given two proofs, $\pi_1$ from X to A, Y, and $\pi_2$ from X′, A to Y′, we get a new proof by chaining them together. You can depict this by "plugging in" the A output of $\pi_1$ into the A input of $\pi_2$. The remaining material stays fixed. In fact, this picture still makes sense if the cut wire A occurs in the middle of the output wires of $\pi_1$ and in the middle of the input wires of $\pi_2$.



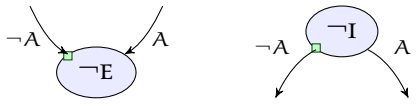

So, we are free to tangle up our wires as much as we like. It is clear from this picture that the conclusion wire A from the proof $\pi_1$ is used

as a premise in the proof $\pi_2$. It is just as clear that *any* output wire in one proof may be used as an input wire in another proof, and we can always represent this fact diagrammatically. The situation is much improved compared with upward-and-downward branching *tree* notation.

As a matter of fact, I will try to make proofs as tangle free as possible, for ease of reading.

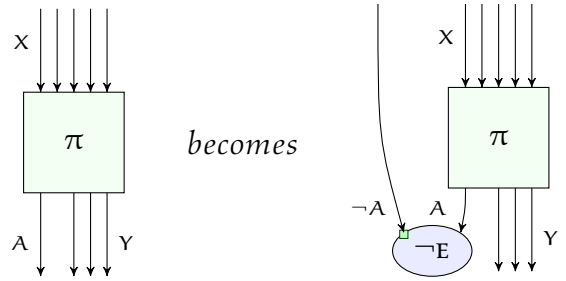
Now consider the behaviour of the connective rules. For negation, the behaviour is simple. An application of a negation rule turns an output $A$ into an input $\neg A$ (this is $\neg$L), or an input $A$ into an output $\neg A$ (this is $\neg$R). So, we can think of these steps as plugging in new *nodes* in the circuit. A [$\neg$E] node takes an input $A$ and input $\neg A$ (and has no outputs), while a [$\neg$I] node has an output $A$ and and output $\neg A$ (and has no inputs). In other words, these nodes may be represented in the following way:

Ignore, for the moment, the little green squares on the surface of the node, and the shade of the nodes. These features have a significance which will be revealed in good time.



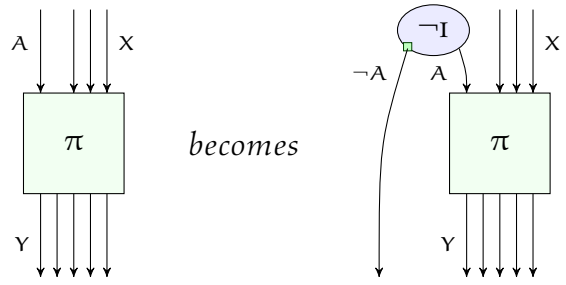

and they can be added to existing proofs to provide the behaviour of the sequent rules $\neg$L and $\neg$R.

$$\frac{X \vdash A, Y}{X, \neg A \vdash Y}\ \neg L$$



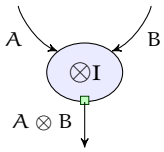

Here, a circuit for $X \vdash A, Y$ becomes, with the addition of a [$\neg$E] node, a circuit for $X, \neg A \vdash Y$. Similarly,

$$\frac{X, A \vdash Y}{X \vdash \neg A, Y}\ \neg R$$

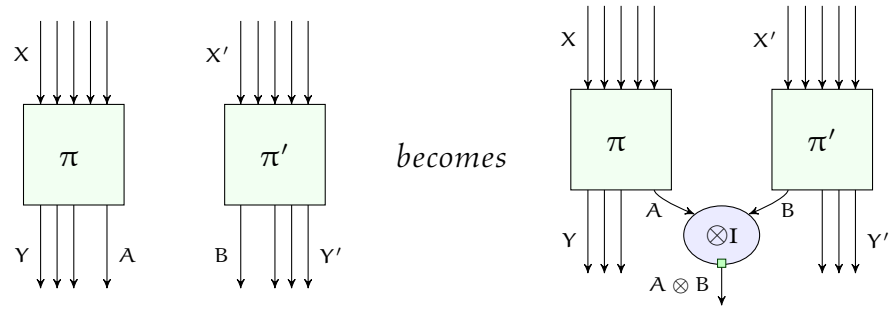



a circuit for the sequent $X, A \vdash Y$ becomes, with the addition of a [$\neg$I] node, a circuit for $X \vdash \neg A, Y$. Notice how these rules (or nodes) are quite simple and local. They do not involve the *discharge* of assumptions (unlike the natural deduction rule $\neg$I we have already seen). Instead, these rules look like straightforward transcriptions of the law of non-contradiction ($A$ and $\neg A$ form a dead-end—don't assert both) and the law of the excluded middle (either $A$ or $\neg A$ is acceptable—don't deny both).

For conjunction, the right rule indicates that if we have a proof $\pi$ with $A$ as one conclusion, and a proof $\pi'$ with $B$ as another conclusion, we can construct a proof by plugging in the $A$ and the $B$ conclusion wires into a new node with a single conclusion wire $A \otimes B$. This motivates a node [$\otimes$I] with two inputs $A$ and $B$ an a single output $A \otimes B$. So, the node looks like this:



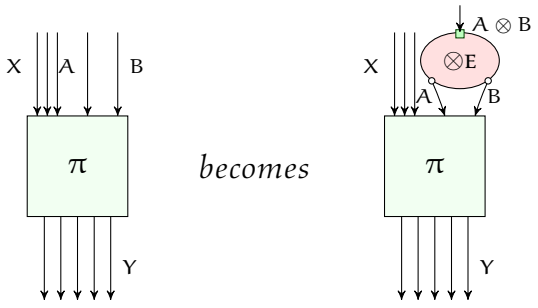

and it can be used to combine circuits in the manner of the [$\otimes$R] sequent rule:




This is no different from the natural deduction rule

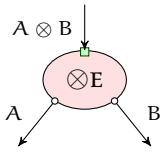
$$\frac{A \quad B}{A \otimes B}$$

except for the notational variance and the possibility that it might be employed in a context in which there are conclusions alongside $A \wedge B$. The rule [$\otimes$E], on the other hand, is novel. This rule takes a single proof $\pi$ with the two premises $A$ and $B$ and modifies it by wiring together the inputs $A$ and $B$ into a node which has a single input $A \otimes B$. It follows that we have a node [$\otimes$E] with a single input $A \otimes B$ and two outputs $A$ and $B$.




Ignore, for the moment, the different colour of this node, and the two small circles on the surface of the node where the $A$ and $B$ wires join. All will be explained in good time.

In this case the relevant node has one input and two *outputs*:

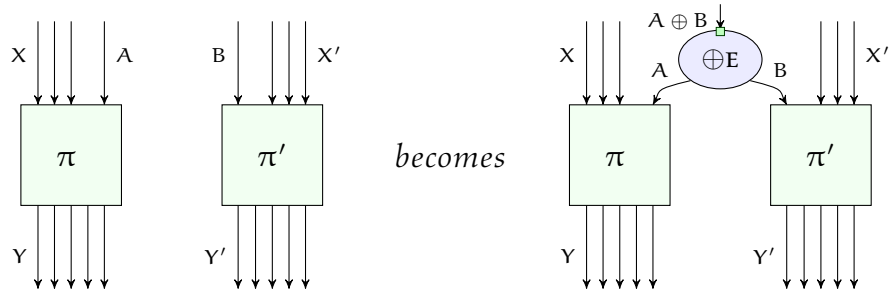


This is not a mere variant of the rules [∧E] in traditional natural deduction. It is novel. It corresponds to the *other* kind of natural deduction rule

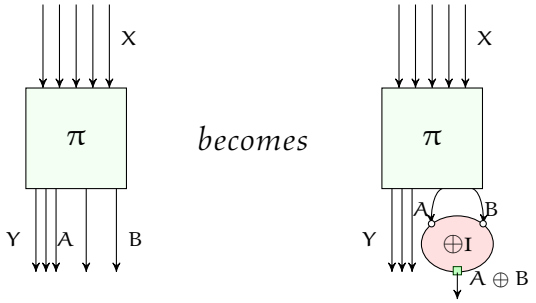
$$\dfrac{\begin{array}{cc} & [A, B] \\ A \otimes B & \vdots \\ & C \end{array}}{C}$$

in which two premises A and B are discharged, and the new premise $A \otimes B$ is used in its place.

The extent of the novelty of this rule becomes apparent when you see that the circuit for [⊕E] also has one input and two outputs, and the two outputs are A and B, if the input is $A \oplus B$. The step for [⊕L] takes two proofs: $\pi_1$ with a premise A and $\pi_2$ with a premise B, and combines them into a proof with the single premise $A \otimes B$. So the node for [⊗E] looks identical. It has a single input wire (in this case, $A \otimes B$), and two output wires, A and B




The same happens with the rule to introduce a disjunction. The sequent step [⊕R] converts the two conclusions A, B into the one conclusion $A \oplus B$. So, if we have a proof π with two conclusion wires A and B, we can plug these into a [⊕I] node, which has two input wires A and B and a single output wire $A \oplus B$.




Notice that this looks just like the node for [⊗I]. Yet ⊗ and ⊕ are very different connectives. The difference between the two nodes is due to the different ways that they are added to a circuit.

## 2.4.2 | CIRCUITS FROM DERIVATIONS

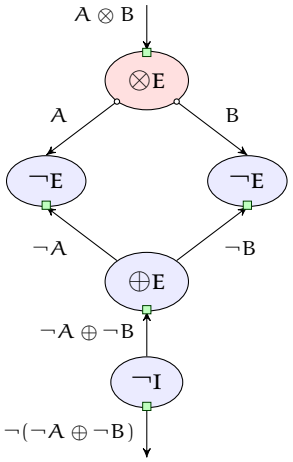
DEFINITION 2.4.1 [INDUCTIVELY GENERATED CIRCUIT] A derivation of $X \vdash Y$ constructs a *circuit* with input wires labelled with the formulas in X and output wires labelled with the formulas in Y in the manner we

have seen in the previous section. We will call these circuits INDUCTIVELY GENERATED.

Here is an example. This derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash} \qquad \dfrac{B \vdash B}{B, \neg B \vdash}}{A, B, \neg A \oplus \neg B \vdash}}{A, B \vdash \neg(\neg A \oplus \neg B)}}{A \otimes B \vdash \neg(\neg A \oplus \neg B)}$$

can be seen as constructing the following circuit:



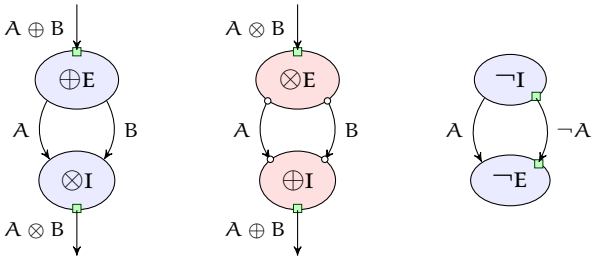

Just as with other natural deduction systems, this is representation of derivations is *efficient*, in that different derivations can represent the one and the same circuit. This derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash} \qquad \dfrac{B \vdash B}{B, \neg B \vdash}}{A, B, \neg A \oplus \neg B \vdash}}{A \otimes B, \neg A \oplus \neg B \vdash}}{A \otimes B \vdash \neg(\neg A \oplus \neg B)}$$

defines exactly the same circuit. The map from derivations to circuits is many-to-one.

Notice that the inductive construction of proof circuits provides for a difference for $\oplus$ and $\otimes$ rules. The nodes [$\otimes$I] and [$\oplus$E] combine *different* proof circuits, and [$\otimes$E] and [$\oplus$I] attach to a single proof circuit. This means that [$\otimes$E] and [$\oplus$I] are *parasitic*. They do not constitute a proof by themselves. (There is no linear derivation that consists merely of the step [$\oplus$R], or solely of [$\otimes$L], since all axioms are of the form $A \vdash A$.) This is unlike [$\oplus$L] and [$\otimes$R] which can make fine proofs on their own.

Not everything that you can make out of the basic nodes is a circuit corresponding to a derivation. Not every "circuit" (in the broad sense) is inductively generated.




You can define 'circuits' for $A \oplus B \vdash A \otimes B$ or $A \otimes B \vdash A \oplus B$, or even worse, $\vdash$, but there are no *derivations* for these sequents. What makes an assemblage of nodes a *proof*?

## 2.4.3 | CORRECT CIRCUITS

When is a circuit inductively generated? There are different *correctness criteria* for circuits. Here are two:

The notion of a SWITCHING is due to Vincent Danos and Laurent Regnier [21], who applied it to give an elegant account of correctness for proofnets.

DEFINITION 2.4.2 [SWITCHED NODES AND SWITCHINGS] The nodes [$\otimes$E] and [$\oplus$I] are said to be *switched nodes*: the two output wires of [$\otimes$E] are its *switched wires* and the two input wires of [$\oplus$I] are its *switched wires*. A *switching of a switched node* is found by breaking one (and one only) of its switched wires. A *switching of a circuit* is found by switching each of its switch nodes.

THEOREM 2.4.3 [SWITCHING CRITERION] *A circuit is inductively generated if and only if each of its switchings is a tree.*

We call the criterion of "every-switching-being-a-tree" the *switching criterion*. There are two ways to fail it. First, by having a switching that contains a loop. Second, by having a switching that contains two disconnected pieces.

*Proof*: The left-to-right direction is a straightforward check of each inductively generated circuit. The basic inductively generated circuits (the single wires) satisfy the switching criterion. Then show that for each of the rules, if the starting circuits satisfy the switching criterion, so do the result of applying the rules. This is a straightforward check.

The right-to-left direction is another thing entirely. We prove it by introducing a new criterion. ∎

The new notion is the concept of *retractability* [21].

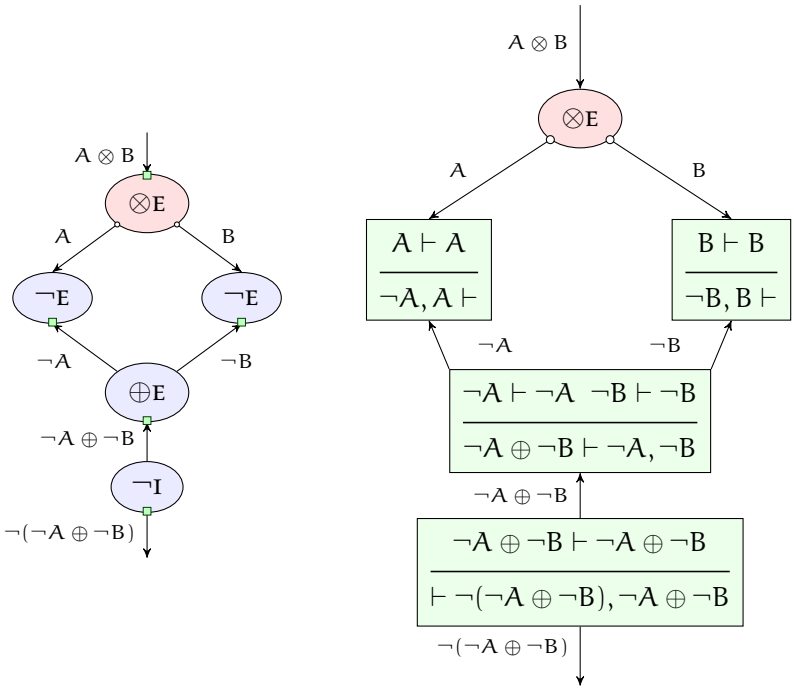
DEFINITION 2.4.4 [RETRACTIONS] A SINGLE-STEP RETRACTION of a circuit [node with a single link to another node, both unswitched, retracts into a single unswitched node], and a [a switched node with its two links into a single unswitched node.]... A circuit $\pi'$ is a RETRACTION of another circuit $\pi$ if there is a sequence $\pi = \pi_0, \pi_1, \ldots, \pi_{n-1}, \pi_n = \pi'$ of circuits such that $\pi_{i+1}$ is a retraction of $\pi_i$.

THEOREM 2.4.5 [RETRACTION THEOREM] *If a circuit satisfies the switching criterion then it retracts to a single node.*

*Proof*: This is a difficult proof. Here is the structure: it will be explained in more detail in class. Look at the number of switched nodes in your circuit. If you have none, it's straightforward to show that the circuit is retractable. If you have more than one, choose a switched node, and look at the subcircuit of the circuit which is *strongly attached* to the switched ports of that node (this is called *empire* of the node). This satisfies the switching criterion, as you can check. So it must either be retractable (in which case we *can* retract away to a single point, and then absorb this switched node and continue) or we cannot. If we cannot, then look at this subcircuit. It must contain a switched node (it would be retractable if it didn't), which must also have an empire, which must be not retractable, and hence, must contain a switched node ... which is impossible. ∎

THEOREM 2.4.6 [CONVERSION THEOREM] *If a circuit is retractable, it can be inductively generated.*

You can use the retraction process to convert a circuit into a derivation. Take a circuit, and replace each unswitched node by a derivation. The circuit on the left becomes the structure on the right:



In this diagram, the inhabitants of a (green) rectangle are *derivations*, whose concluding sequent mirrors exactly the arrows in and out of the box. The in-arrows are on the left, and the out-arrows are on the right. If we have two boxes joined by an arrow, we can merge the two boxes. The effect on the derivation is to cut on the formula in the arrow. The result is in Figure 2.14. After absorbing the two remaining derivations, we get a structure with only one node remaining, the switched ⊗E node. This is in Figure 2.15.

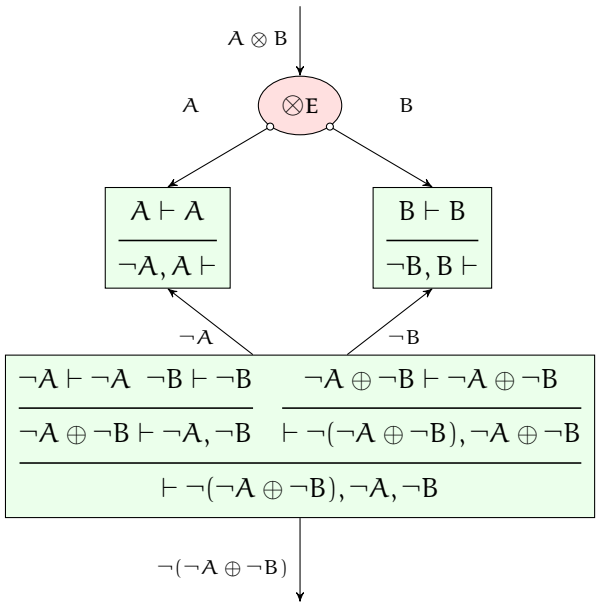



Figure 2.14: a retraction in progress: part 1

Now at last, the switched node ⊗E has both output arrows linked to the one derivation. This means that we have a derivation of a sequent with both A and B on the left. We can complete the derivation with a [⊗L] step. The result is in Figure 2.16.

In general, this process gives us a weird derivation, in which every connective rule, except for ⊗L and ⊕R occurs at the top of the derivation, and the only other steps are cut steps and the inferences ⊗L and ⊕R, which correspond to switched nodes.

[Notice that there is no explicit conception of *discharge* in these circuits. Nonetheless, conditionals may be defined using the vocabulary we have at hand: $A \to B$ is $\neg A \oplus B$. If we consider what it would be to *eliminate* $\neg A \oplus B$, we see that the combination of a ⊕E and a ¬E node allows us to chain together a proof concluding in $\neg A \oplus B$ with a proof concluding in A to construct a proof concluding in B (together with the other conclusions remaining from the original two proofs).

[diagram to go here]

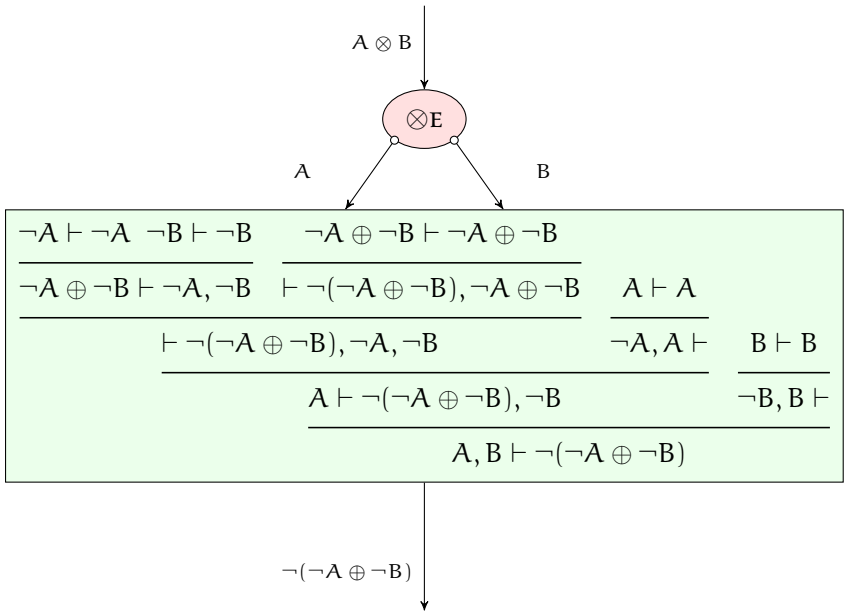



Figure 2.15: A RETRACTION IN PROGRESS: PART 2

For an introduction rule, we can see that a if we have a proof with a premise A and a conclusion B (possibly among other premises and conclusions) we may plug the A input wire into a ¬I node to give us a new ¬A concluding wire, and the two conclusions ¬A and B may be wired up with a ⊕I node to give us the new conclusion ¬A ⊕ B, or if you prefer, A → B.

[diagram to go here]

Expand this, with a discussion of the *locality* of circuits as opposed to the 'action at a distance' of the traditional discharge rules.]

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A \quad \neg B \vdash \neg B}{\neg A \oplus \neg B \vdash \neg A, \neg B}\,{\oplus L} \qquad \dfrac{\neg A \oplus \neg B \vdash \neg A \oplus \neg B}{\vdash \neg(\neg A \oplus \neg B), \neg A \oplus \neg B}\,{\neg L}}{\vdash \neg(\neg A \oplus \neg B), \neg A, \neg B}\,{Cut} \qquad \dfrac{A \vdash A}{\neg A, A \vdash}\,{\neg L}}{A \vdash \neg(\neg A \oplus \neg B), \neg B}\,{Cut} \qquad \dfrac{B \vdash B}{\neg B, B \vdash}\,{\neg L}}{A, B \vdash \neg(\neg A \oplus \neg B)}\,{Cut}}{A \otimes B \vdash \neg(\neg A \oplus \neg B)}\,{\otimes L}$$

Figure 2.16: A DERIVATION FOR $A \otimes B \vdash \neg(\neg A \oplus \neg B)$

## 2.4.4 | NORMAL CIRCUITS

Not every circuit is *normal*. In a natural deduction proof in the system for implication, we said that a proof was normal if there is no step introducing a conditional $A \to B$ which then immediately serves as a major premise in a conditional elimination move. The definition for normality for circuits is completely parallel to this definition. A circuit is *normal* if and only if no wire for $A \otimes B$ (or $A \oplus B$ or $\neg A$) is both the *output* of a [$\otimes$I] node (or a [$\oplus$I] or [$\neg$I] node) and an *input* of a [$\otimes$E] node (or [$\oplus$E] or [$\neg$E]). This is the role of the small dots on the boundary of the nodes: these mark the 'active wire' of a node, and a non-normal circuit has a wire that has this dot at both ends.

It is straightforward to show that if we have a cut-free derivation of a sequent $X \vdash Y$, then the circuit constructed by this derivation is normal. The new nodes at each stage of construction always have their dots facing *outwards*, so a dot is never added to an already existing wire. So, cut-free derivations construct normal circuits.

The process of *normalising* a circuit is simplicity itself: Pairs of introduction and elimination nodes can be swapped out by node-free wires in any circuit in which they occur. The square indicates the "active" port of the node, and if we have a circuit in which two active ports are joined, they can "react" to simplify the circuit. The rules of reaction are presented in Figure 2.17.

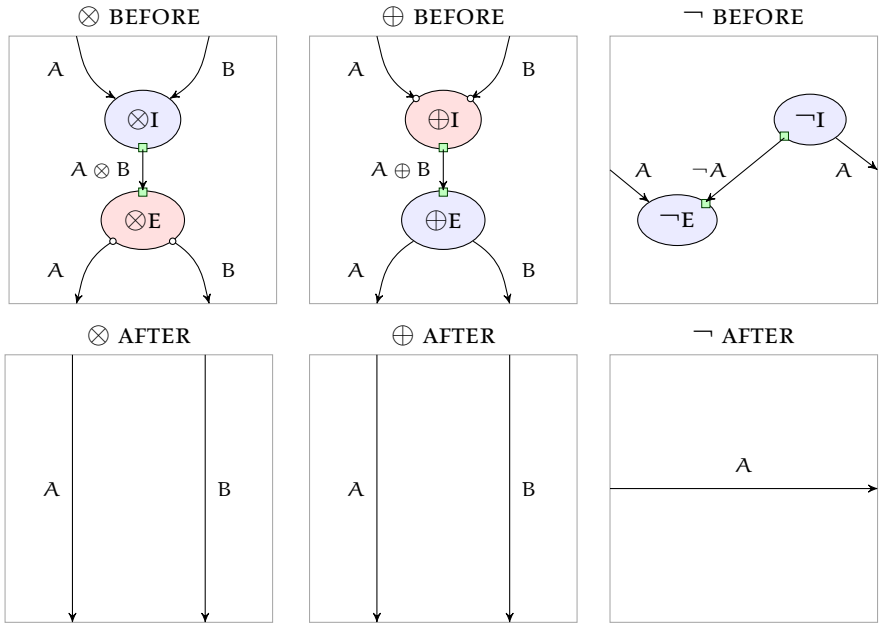



Figure 2.17: NORMALISATION STEPS

The process of normalisation is completely *local*. We replace a region of the circuit by another region with the same *periphery*. At no stage do any global transformations have to take place in a circuit, and so, normalisation can occur in *parallel*. It is clearly *terminating*, as we delete nodes and do not add them. Furthermore, the process of norm-

alisation is *confluent*. No matter what order we decide to process the nodes, we will always end with the same *normal* circuit in the end.

[ADD AN EXAMPLE]

THEOREM 2.4.7 [NORMALISATION FOR LINEAR CIRCUITS]

## 2.4.5 | CLASSICAL CIRCUITS

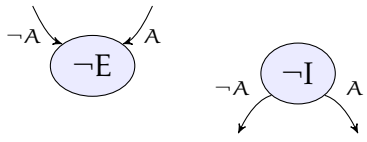
To design circuits for classical logic, you must incorporate the effect of the structural rules in some way. The most straightforward way to do this is to introduce (switched) contraction nodes and (unswitched) weakening nodes. In this way the parallel with the sequent system is completely *explicit*.

The first account of multiple-conclusion proofs is Kneale's "Tables of Development" [47]. Shoesmith and Smiley's *Multiple Conclusion Logic* [84] is an extensive treatment of the topic. The authors explain why Kneale's formulation is not satisfactory due to problems of substitution of one proof into another — the admissibility of *cut*. Shoesmith and Smiley introduce a notation similar to the node and wire diagrams used here. The problem of substitution is further discussed in Ungar's *Normalization, cut-elimination, and the theory of proofs* [92], which proposes a general account of what it is to substitute one proof into another. One account of classical logic that is close to the account given here is Edmund Robinson's, in "Proof Nets for Classical Logic" [80].

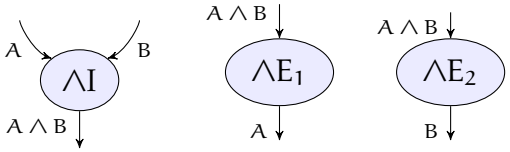
DEFINITION 2.4.8 [CLASSICAL CIRCUITS] These are the *inductively generated circuits*:

- An identity wire: $\xrightarrow{A}$ for any formula A is an inductively generated circuit. The sole input type for this circuit is A and its output type is also (the very same instance) A. As there is only one wire in this circuit, it is near to itself.

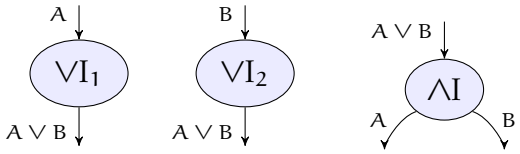
- Each boolean connective node presented below is an inductively generated circuit. The negation nodes:




  The conjunction nodes:



And disjunction nodes:

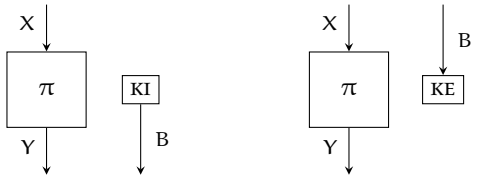



The inputs of a node are those wires pointing *into* the node, and the outputs of a node are those wires pointing *out*.

- Given an inductively generated circuit $\pi$ with an output wire labelled $A$, and an inductively generated circuit $\pi'$ with an input wire labelled $A$, we obtain a *new* inductively generated circuit in which the output wire of $\pi$ is plugged in to the input wire of $\pi'$. The output wires of the new circuit are the output wires of $\pi$ (except for the indicated $A$ wire) and the output wires of $\pi'$, and the input wires of the new circuit are the input wires of $\pi$ together with the input wires of $\pi'$ (except for the indicated $A$ wire).

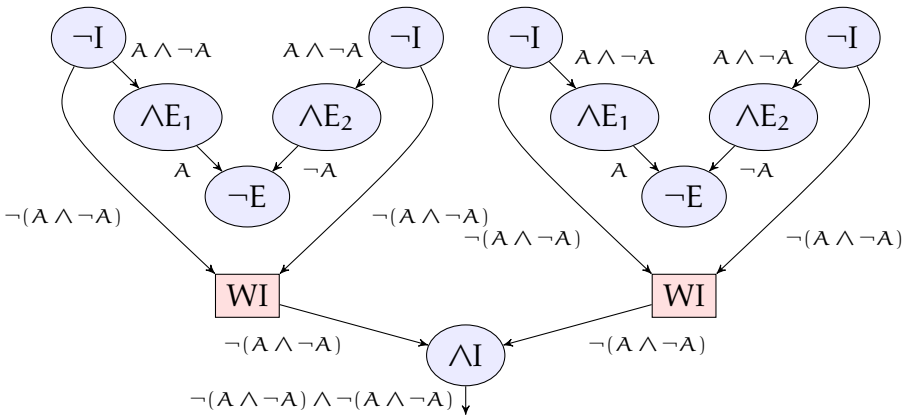
- Given an inductively generated circuit $\pi$ with two input wires $A$, a new inductively generated circuit is formed by plugging both of those input wires into the input contraction node $\boxed{\text{WE}}$. Similarly, two output wires with the same label may be extended with a contraction node $\boxed{\text{WI}}$.

- Given an inductively generated circuit $\pi$, we may form a new circuit with the addition of a new output, or output wire (with an arbitrary label) using a weakening node $\boxed{\text{KI}}$ or $\boxed{\text{KE}}$.[3]




[3]Using an unlinked weakening node like this makes some circuits *disconnected*. It also forces a great number of different sequent derivations to be represented by the same circuit. Any derivation of a sequent of the form $X \vdash Y, B$ in which $B$ is weakened in at the last step will construct the same circuit as a derivation in which $B$ is weakened in at an earlier step. If this identification is not desired, then a more complicated presentation of weakening, using the 'supporting wire' of Blute, Cockett, Seely and Trimble [8] is possible. Here, I opt for a simple presentation of circuits rather than a comprehensive account of "proof identity."

Here is an example circuit:




[Correctness theorem. (Retraction)]

[Translation between circuits and sequents.]

[Normalisation (including strong normalisation). Failure of Church–Rosser with the usual rules. Church-Rosser property for a particular choice of the W/W and K/K rules. Is this desirable?]

## 2.4.6 | HISTORY AND OTHER MATTERS

We can rely on the duality of $\otimes$ and $\oplus$ to do away with half of our rules, if we are prepared to do a little bit of work. Translate the sequent $X \vdash Y$ into $\vdash \neg X, Y \vdash$, and then trade in $\neg(A \otimes B)$ for $\neg A \oplus \neg B$; $\neg(A \oplus B)$ for $\neg A \otimes \neg B$, and $\neg\neg A$ for $A$. The result will be a sequent where the only negations are on atoms. Then we can have rules of the following form:

$$\vdash p, \neg p \quad [\mathrm{Id}]$$

$$\frac{\vdash X, A, B}{\vdash X, A \oplus B}\,\oplus R \qquad \frac{\vdash X, A \quad \vdash X', B}{\vdash X, X', A \otimes B}\,\otimes R$$

The circuits are also much simpler. They only have outputs and no inputs. These are Girard's proofnets [35].

## 2.4.7 | EXERCISES

### BASIC EXERCISES

Q1 Construct circuits for the following sequents:

1 : $\vdash p \oplus \neg p$
2 : $p \otimes \neg p \vdash$
3 : $\neg\neg p \vdash p$
4 : $p \vdash \neg\neg p$
5 : $\neg(p \otimes q) \vdash \neg p \oplus \neg q$
6 : $\neg p \oplus \neg q \vdash \neg(p \otimes q)$
7 : $\neg(p \oplus q) \vdash \neg p \otimes \neg q$
8 : $\neg p \otimes \neg q \vdash \neg(p \oplus q)$

9 : $p \otimes q \vdash q \otimes p$
10 : $p \oplus (q \oplus r) \vdash p \oplus (q \oplus r)$

Q2 Show that every formula $A$ in the language $\oplus, \otimes, \neg$ is equivalent to a formula $n(A)$ in which the only negations are on atomic formulas.

Q3 For every formula $A$, construct a circuit $encode_A$ from $A$ to $n(A)$, and $decode_A$ from $n(A)$ to $A$. Show that $encode_A$ composed with $decode_A$ normalises to the identity arrow $\xrightarrow{A}$, and that $decode_A$ composed with $encode_A$ normalises to $\xrightarrow{n(A)}$. (If this doesn't work for the *encode* and *decode* circuits you chose, then try again.)

Q4 Given a circuit $\pi_1$ for $A_1 \vdash B_1$ and a circuit $\pi_2$ for $A_2 \vdash B_2$, show how to construct a circuit for $A_1 \otimes A_2 \vdash B_1 \otimes B_2$ by adding two more nodes. Call this new circuit $\pi_1 \otimes \pi_2$. Now, suppose that $\tau_1$ is a proof from $B_1$ to $C_1$, and $\tau_2$ is a proof from $B_2$ to $C_2$. What is the relationship between the proof $(\pi_1 \otimes \pi_2) \cdot (\tau_1 \otimes \tau_2)$ (composing the two proofs $\pi_1 \otimes \pi_2$ and $\tau_1 \otimes \tau_2$ with a cut on $B_1 \otimes B_2$) from $A_1 \otimes A_2$ to $C_1 \otimes C_2$ and the proof $(\pi_1 \cdot \tau_1) \otimes (\pi_2 \otimes \tau_2)$, also from $A_1 \otimes A_2$ to $C_1 \otimes C_2$?

Prove the same result for $\oplus$ in place of $\otimes$. Is there a corresponding fact for negation?

Q5 Re-prove the results of all of the previous questions, replacing $\otimes$ by $\wedge$ and $\oplus$ by $\vee$, using the rules for classical circuits. What difference does this make?

Q6 Construct classical circuits for the following sequents

1 : $q \vdash p \vee \neg p$
2 : $p \wedge \neg p \vdash q$
3 : $p \vdash (p \wedge q) \vee (p \wedge \neg q)$
4 : $(p \wedge q) \vee (p \wedge \neg q) \vdash p$
5 : $(p \wedge q) \vee r \vdash p \wedge (q \vee r)$
6 : $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$

### INTERMEDIATE EXERCISES

Q7 The following statement is a tautology:

$$\neg\big((p_{1,1} \vee p_{1,2}) \wedge (p_{2,1} \vee p_{2,2}) \wedge (p_{3,1} \vee p_{3,2}) \wedge \neg(p_{1,1} \wedge p_{2,1}) \wedge \neg(p_{1,1} \wedge p_{3,1}) \wedge \neg(p_{2,1} \wedge p_{3,1}) \wedge \neg(p_{1,2} \wedge p_{3,2}) \wedge \neg(p_{2,2} \wedge p_{3,2})\big)$$

It is the pigeonhole principle for $n = 2$. The general pigeonhole principle is the formula $P_n$.

$$P_n : \quad \neg\Big(\bigwedge_{i=1}^{n+1} \bigwedge_{j=1}^{n} p_{i,j} \wedge \bigwedge_{i=1}^{n+1} \bigwedge_{i'=i+1}^{n+1} \bigwedge_{j=1}^{n} \neg(p_{i,j} \wedge p_{i',j})\Big)$$

$P_n$ says that you cannot fit $n+1$ pigeons in $n$ pigeonholes, no two pigeons in the one hole. Find a proof of the pigeonhole principle for $n = 2$. How *large* is your proof? Describe a proof of $P_n$ for each value of $n$. How does the proof increase in size as $n$ gets larger? Are there non-normal proofs of $P_n$ that are significantly smaller than any non-normal proofs of $P_n$?

Read '$p_{i,j}$' as 'pigeon number $i$ is in pigeonhole number $j$.'

## 2.5 | COUNTEREXAMPLES

Models (truth tables, algebraic models, Kripke frame models for the logics we have seen so far) are introduced as counterexamples to invalid arguments. Search for derivations is construed as a counterexample search, and sequent systems suited to derivation search are presented, and the finite model property and decidability is proved for a range of logical systems.

Natural deduction proofs are structures generated by a recursive definition: we define *atomic proofs* (in this case, assumptions on their own), and given a proof, we provide ways to construct new proofs (conditional introduction and elimination). Nothing is a proof that is not constructed in this way from the atoms. This construction means that we can prove results about them using the standard technique of *induction*. For example, we can show that any standard proof is valid according to the familiar truth-table test. That is, if we assign the values of *true* and *false* to each atom in the language, and then extend the valuation so that $v(A \to B)$ is *false* if and only if $v(A)$ is *true* and $v(B)$ is *false*, and $v(A \to B)$ is *true* otherwise, then we have the following result:

THEOREM 2.5.1 [TRUTH TABLE VALIDITY] *Given a standardly valid argument* $X \therefore A$, *there is no assignment of truth values to the atoms such that each formula in* X *is* true *and the conclusion* A *is* false.

It can be proved systematically from the way that proofs are constructed.

*Proof*: We *first* show that the *simplest* proofs have this property. That is, given a proof that is just an assumption, we show that there is no counterexample in truth tables. But this is obvious. A counterexample for an assumption A would be a valuation such that $v(A)$ was *true* and was at the same time *false*. Truth tables do not allow this. So, mere assumptions have the property of being truth table valid. Now, let's suppose that we have a proof whose last move is an elimination, from $A \to B$ and A to B and let's suppose that its constituent proofs, $\pi_1$ from X to $A \to B$ and $\pi_2$ from Y to A, are truth table valid. It remains to show that our proof from X and Y to B is truth table valid. If it is not, then we have a valuation $v$ that makes each formula in X true, and each formula in Y true, and that makes B false. This cannot be the case, since $v$ must make A either true or false. If it makes A false, then the valuation is a counterexample to the argument from Y to A. (But we have supposed that this argument has no truth table counterexamples.) On the other hand, if it makes A true, then it makes $A \to B$ false (since B is false) and so, it is a counterexample to the argument from X to $A \to B$. (But again, we have supposed that this argument has no truth table counterexamples.) So, we have shown that if our proofs $\pi_1$ and $\pi_2$ are valid, then the result of extending it with an $\to E$ move is also valid.

Let's do the same thing with $\rightarrow I$. Suppose that $\pi'$, from X to B is a valid argument, and let's suppose that we are interested in discharging the assumption A from $\pi'$ to deduce $A \rightarrow B$ from the premise list X′, which is X with a number (possibly zero) of instances of A deleted. Is this new argument truth table valid? Well, let's suppose that it is not. It follows that we have a valuation that makes the list X′ of formulas each *true*, and $A \rightarrow B$ *false*. That is, it makes A *true* and B *false*. So, the valuation makes the formulas in X all *true* (since X is X′ together with some number of instances of A) and B false. So, if our longer proof is invalid truth-table invalid, so is $\pi'$.

It follows, then, that any proof constructs an argument that is valid according to truth tables. There is no counterexample to any one-line proof (simply an assumption), and $\rightarrow I$ and $\rightarrow E$ steps also do not permit counterexamples. So, *no* proof has a counterexample in truth tables. All standardly valid arguments are truth table valid. ∎

However, the converse is not the case. Some arguments that are valid from the perspective of truth tables cannot be supplied with proofs. Truth tables are good for sifting out *some* of the invalid arguments, and for those arguments for which this techniques work, a simple truth table counterexample is significantly more straightforward to work with than a direct demonstration that there is no proof to be found. Regardless, truth tables are a dull instrument. Many arguments with no standard proofs are truth table valid. Here are two examples: $(A \rightarrow B) \rightarrow B \therefore (B \rightarrow A) \rightarrow A$ and $(A \rightarrow B) \rightarrow A \therefore A$. Now we will look at ways to refute these arguments.

## 2.5.1 | COUNTEREXAMPLES FOR CONDITIONALS

In this section we will consider a more discriminating instrument for finding counterexamples to invalid arguments involving conditionals. The core idea is that we will have simple *models* in which we can interpret formulas. Models will consist of *points* at which formulas can be true or false. Points will be able to be *combined*, and this is the heart of the behaviour of implication. If $A \rightarrow B$ is true at x and A is true at y, then B will be true at $x * y$.

DEFINITION 2.5.2 [CONDITIONAL STRUCTURE] A triple $\langle P, *, 0\rangle$ consisting of a nonempty set P of points, an operation $*$ on P and a specific element 0 of P forms a CONDITIONAL STRUCTURE if and only if $*$ is commutative ($a*b = b*a$ for each $a, b \in P$) and associative ($a*(b*c) = (a*b)*c$ for each $a, b, c \in P$) and 0 is an identity for $*$ ($0 * a = a = a * 0$ for each $a \in P$.) A conditional structure is said to be *contracting* if $a * a = a$ for each $a \in P$.

EXAMPLE 2.5.3 [CONDITIONAL STRUCTURES] Here are some simple conditional structures.

[THE TRIVIAL STRUCTURE] The trivial conditional structure has $P = \{0\}$, and $0 * 0 = 0$. It is the only structure in which P has one element.

[TWO-ELEMENT STRUCTURES] There are only two distinct two-element conditional structures. Suppose $P = \{0, 1\}$. We have the addition table settled except for the value of $1 * 1$. So, there are *two* two-element structures. One in which $1 * 1 = 0$ and one in which $1 * 1 = 1$. The second choice provides us with a contracting structure, and the first does not.

[LINEAR STRUCTURES] Consider an conditional structure in which $P = \{0, 1, 2, \ldots\}$ consists of all counting numbers. There are at least two straightforward ways to evaluate $*$ on such a structure. If we wish to have a *contracting* structure, we can set $a * b = \max(a, b)$. If we wish to have a different structure, we can set $a * b = a + b$.

Once we have our structures, we can turn to the task of interpreting formulas in our language.

These models are a slight generalisation of the models introduced by Alasdair Urquhart in 1972 [94], for the relevant logics R and E.

DEFINITION 2.5.4 [CONDITIONAL MODEL] Given a conditional structure $\langle P, *, 0\rangle$, a *model* on that structure is determined by a valuation $\Vdash$ of the atoms of the language at each point in $P$. We will write "$a \Vdash p$" to say that $p$ is true at $a$, and "$a \nVdash p$" to say that $p$ is not true at $a$. Given a model, we may interpret *every* formula at each point as follows:

» $c \Vdash A \to B$ iff for each $a \in P$ where $a \Vdash A$, we have $c * a \Vdash B$.

EXAMPLE 2.5.5 [A SIMPLE VALUATION] Consider the two-element structure in which $P = \{0, 1\}$ and $1 * 1 = 1$. Let's suppose that $0 \nVdash p$, $1 \Vdash p$, $0 \nVdash q$ and $1 \nVdash q$. It follows that $0 \nVdash p \to q$, since $1 \Vdash p$ and $0 * 1 = 1 \nVdash q$. Similarly, $1 \nVdash p \to q$, since $1 \Vdash p$ and $1 * 1 = 1 \nVdash q$. However, $0 \Vdash q \to p$, since there is *no* point at which $q$ is true. Similarly, $0 \Vdash (p \to q) \to p$, since there is no point at which $p \to q$ is true. It follows that $0 \nVdash ((p \to q) \to p) \to p$, since $0 \Vdash (p \to q) \to p$ and $0 * 0 = 0 \nVdash p$.

In this example, we have a refutation of a classically valid formula. $((p \to q) \to p) \to p$ is valid according to truth tables, but it can be refuted at a point in one of our models.

DEFINITION 2.5.6 [ATOM PRESERVATION] A model is said to be *atom preserving* if whenever $a \Vdash p$ for an atom $p$ and a point $a$, then for any point $b$, we have $a * b \Vdash p$ too.

LEMMA 2.5.7 [CONDITIONAL PRESERVATION] *For any points* $a$ *and* $b$ *in an atom preserving model, and for any formula* $A$*, if* $a \Vdash A$ *then* $a * b \Vdash A$ *too. (That is,* $A$ *is preserved across points.)*

*Proof*: By induction on the construction of formulas. The case for atoms holds by our assumption that the model is atom preserving. Now suppose we have a conditional formula $A \to B$, and suppose that $A$ and $B$ are preserved in the model. We will show that $A \to B$ is preserved too. Suppose that $a \Vdash A \to B$. We wish to show that $a * b \Vdash A \to B$ too. So, we take any $c \in P$ where $c \Vdash A$, and we attempt to show that $(a * b) * c \Vdash B$. By the associativity of $*$ we have

$(a * b) * c = a * (b * c)$. We also have that $a \Vdash A \to B$ and that $c \Vdash A$. By the commutativity of $*$, $b * c = c * b$, and by the preservation of $A$, $c * b \Vdash A$. So, $b * c \Vdash A$, and since $a \Vdash A \to B$, $(a * b) * c = a * (b * c) \Vdash B$, as desired. So, $A \to B$ is preserved. ∎

To use these models to evaluate arguments, we need to think of how it is appropriate to evaluate multisets of formulas in a model. Clearly a singleton multiset $A$ is evaluated just as a single formula is. The multiset $A, B$ is to be interpreted as true at a point $x$ when $x = a * b$ and $a \Vdash A$ and $b \Vdash B$. We can generalise this.

Notice that this definition uses the fact that $0$ is an identity for $*$ (this makes $x \Vdash X$ work when $X$ is a singleton) and that $*$ is commutative and associative (this makes it not matter what order you "unwrap" your multiset into elements).

definition 2.5.8 [evaluating multisets] For each point $x \in P$, we say that

- $x \Vdash \square$ iff $x = 0$.
- $x \Vdash Y, A$ iff there are $y, a \in P$ where $x = y * a$, $y \Vdash Y$ and $a \Vdash A$.

So, $A$ is valid iff $\therefore A$ is valid.

definition 2.5.9 [validity in models] A *formula* is valid in a model iff it is true at $0$ in that model. An *argument* $X \therefore A$ is valid in a model iff for each $x \in P$ where $x \Vdash X$, we also have $x \Vdash A$.

theorem 2.5.10 [soundness for conditional models] *If we have a linear proof from $X$ to $A$, then the argument $X \therefore A$ is valid in every model. If the proof utilises duplicate discharge, it is valid in every contracting model. If the proof utilises vacuous discharge, it is valid in every preserving model. So, every standard proof is valid in every contracting preserving model.*

*Proof*: By induction on the construction of the proof from $X$ to $A$. Assumptions are valid in every model. If we have a proof $\pi$ from $X$ to $A \to B$ and a proof $\pi'$ from $Y$ to $A$, we want the argument from $X, Y$ to $B$ to be valid. We may assume that $X \therefore A \to B$ and $Y \therefore A$ are both valid in every model. Is $X, Y \therefore B$? Suppose we have $z \Vdash X, Y$. It follows that $z = x * y$ where $x \Vdash X$ and $y \Vdash Y$. Then by the validity of $X \therefore A \to B$ we have $x \Vdash A \to B$. Similarly, we have $y \Vdash A$. By the valuation clause for $\to$ then, $z = x * y \Vdash B$, as we desired.

Suppose now that $\pi$ is a proof from $X$ to $B$, where $X$ contains some number (possibly zero) of copies of $A$. We may assume that $\pi$ is valid (according to some restriction or other), and hence that if $x \Vdash X$ then $x \Vdash B$ in each model (perhaps these are contracting models, perhaps preserving, depending on whether we are observing or failing to observe the restrictions on duplicating or vacuous discharge). Since $x \Vdash X$, it follows that $x = x_1 * x_2 * \cdots * x_n$ where $x_1 \Vdash P_1$, $x_2 \Vdash P_2$, $\ldots$, $x_n \Vdash P_n$, where $X = P_1, P_2, \ldots, P_n$. Without loss of generality, we may present the list in such a way that the discharged instances of $A$ come last, so $X = Y, A, \ldots, A$ where $i$ instances of $A$ are discharged. We want to show that $Y \therefore A \to B$ is valid. Suppose $y \Vdash Y$. We wish to show that $y \Vdash A \to B$. To do this, we need to show that if $a \Vdash A$, $y * a \Vdash B$. In the simplest case, the proof used a linear discharge at this point, $i = 1$ (the number of instances of $A$ discharged), and hence

We may well have $i = n$, in which case $X = A, \ldots, A$ and then $Y = \square$. On the other hand, we may well have $i = 0$, in which case $X = Y$.

$X = Y, A$. In this case, $y * a \Vdash X = Y, A$, and by our assumption of the validity of $\pi$, $y * a \Vdash B$, as desired.

Suppose the discharge was vacuous, and $X = Y$. In this case, we may assume that our model is a preserving one. Now since $y \Vdash Y$, by preservation $y * a \Vdash Y$ too, but $Y = X$ so, $y * a \Vdash X$ and since $X \therefore B$ is valid (in preserving models) we have $y * a \Vdash B$ as desired.

Suppose the discharge was a duplicate, and $i > n$. In this case, we may assume that our model is a contracting one. Since we have $a \Vdash A$, we also have $a * a \Vdash A$ (since $a * a = a$) and more generally, $a * \cdots * a \Vdash A$ where we choose an $i$-fold repeated $*$-list. So, $y * a = y * a * \cdots * a \Vdash X = Y, A, \ldots, A$ and hence $y * a \Vdash B$ as desired.

So, we have shown that our argument is valid in the appropriate class of models. ∎

Straight away we may use this to provide counterexamples to invalid arguments.

EXAMPLE 2.5.11 The argument $A \to (A \to B) \therefore A \to B$ is linearly invalid (and invalid in "affine" logic too), for example. Take the conditional model on the set of counting numbers, in which we have $n * m = n + m$, and in which we have an evaluation that preserves atoms. Take, for example, $p$ true at 1, 2 and every larger number (but not 0) and $q$ true at 2, 3 and every larger number (but not 0 or 1.) In this model we have $0 \Vdash p \to (p \to q)$, since for each $n$ where $n \Vdash p$ (that is, for each $n > 0$) we have $n \Vdash p \to q$. This is true, since for each $m$ where $m \Vdash p$ (that is, for each $m > 0$) we have $n + m \Vdash q$ (that is, we have $m + n > 1$). However, we do *not* have $0 \Vdash p \to q$, since $1 \Vdash p$ but $0 + 1 = 1 \nVdash q$.

It follows from our soundness theorem and this model that we cannot find a proof satisfying the no-duplicate discharge constraint for the argument $A \to (A \to B) \therefore A \to B$. There is none.

THEOREM 2.5.12 [COMPLETENESS FOR CONDITIONAL MODELS] *If $X \therefore A$ has no linear proof, then it has a counterexample in some model. If it has no proof with duplicate discharge, it has a counterexample in some contracting model. If it has no proof with vacuous discharge, it has a counterexample in a preserving model. If it has no standard proof, it has a counterexample in a standard (contracting and preserving) model.*

*Proof*: Consider the linear discharge policy first. We will construct a model, which will contain a counterexample to *every* linearly invalid argument. The points in this model are the multisets of formulas in the language. $X * Y$ is $X, Y$, the multiset union of $X$ and $Y$. $0$ is $\Box$, the empty multiset. This clearly satisfies the definition of a conditional structure.

Now we will define truth-at-a-point. We will say that $X \Vdash p$ if and only if there is some proof for $X \therefore p$ for any atom $p$. We will show that in general, $X \Vdash A$ if and only if there is a proof for $X \therefore A$. This proceeds by induction on the formation of the formula. We already

have the case for atoms. Now suppose we have $A \to B$ and the result holds for $A$ and $B$. We want to show that there is a proof for $X \therefore A \to B$ if and only if for each $Y$ where $Y \Vdash A$, we have $X, Y \Vdash B$. That is, we wish to show that there is a proof for $X \therefore A \to B$ if and only if for each $Y$ where there's a proof of $Y \therefore A$, there is also a proof of $X, Y \therefore B$. From left-to-right it is straightforward. If there is a proof from $X$ to $A \to B$ and a proof from $Y$ to $A$, then extend it by $\to E$ to form a proof from $X, Y$ to $B$. From right-to-left we may assume that for *any* $Y$, if there's a proof for $Y \therefore A$, then there is a proof $X, Y \therefore B$. Well, there is a proof of $A \therefore A$, so it follows that there's a proof of $X, A \therefore B$. Use that proof and apply $\to I$, to construct a proof of $X \therefore A \to B$.

So, our structure is a model in which $X \Vdash A$ in general if and only if there is a proof for $X \therefore A$. It is an easy induction on the structure of $X$ to show that $X \Vdash X$. It follows, then that if there is no proof for $X \therefore A$, then $X$ itself is a point at which $X \Vdash X$ but $X \nVdash A$. We have a counterexample to any invalid argument.

Consider the other discharge policies. If we allow vacuous discharge, then it is straightforward to show that our model satisfies the preservation condition. If $X \therefore A$ is valid, so is $X, B \therefore A$. If $X \therefore A$ is valid, we may discharge a non-appearing $B$ to find $X \therefore B \to A$. We may then use an assumption of $B$ to deduce $X, B \therefore A$.

$$\dfrac{\dfrac{\begin{array}{c} X \\ \vdots \\ A \end{array}}{B \to A}\,{\scriptstyle \to I,1} \qquad B}{A}\,{\scriptstyle \to E}$$

So, in this model, if vacuous discharge is allowed, the preservation condition is satisfied. So, we have an appropriate model for affine deductions.

If we allow duplicate discharge, we must do a little work. Our model we have constructed so far does not satisfy the contraction condition, since the multiset $A, A$ is not the same multiset as the singleton $A$. Instead, we work simply with *sets* of formulas, and proceed as before. We must do a little more work when it comes to $\to E$. We know that if we have a proof for $X \therefore A \to B$ and one for $Y \therefore A$ then we have a proof from *multiset* union $X, Y$ to $B$. Do we have one for the set union too? We do, because for any proof from a list of premises to a conclusion, if we allow duplicate discharges we can construct a proof in which each premise is used only once.

$$\dfrac{\dfrac{\begin{array}{c} X, [B, B]^{(1)} \\ \vdots \\ A \end{array}}{B \to A}\,{\scriptstyle \to I,1} \qquad B}{A}\,{\scriptstyle \to E}$$

In this example, we trade in two uses of $B$ in a proof from $X, B, B$ to $A$ for one. The rest of the argument goes through just as before. Our

model with *sets* of formulas will do for relevant arguments. It satisfies the preservation condition, if we allow vacuous discharge as required for standard arguments. ∎

## 2.5.2 | COUNTEREXAMPLES THROUGH ALGEBRA

We have seen nothing yet of counterexamples to invalid sequents involving $\wedge$ and $\vee$. A sequent may be underivable, but we have, as yet, no "take-away" representation of that invalidity of the kind we saw in Section 2.5.1 for arguments involving conditionals. Counterexamples for invalid sequents in our logic of conjunction and disjunctions are not going to be straightforward, either. It is not just simple sequents such as $p \vdash q$ that are unprovable. Some sequents that are valid in the *standard* sense (using truth tables for conjunction and disjunction) are also unprovable. For example, we have no derivation of $p \wedge (q \vee r) \vdash (p \wedge q) \vee r$ (see Example 2.2.11 on page 63), a sequent for the *distribution* of conjunction over disjunction. We can show that this has no cut-free derivation, and then appeal to Theorem 2.2.9.

This kind of reasoning is effective enough, but tedious if you need to repeat it every time you want to refute an invalid argument. We want an alternative: a way of *representing* invalidity. Truth-tables, just as in the previous section, are not discriminating enough. They would judge distribution to be valid, not invalid. One alternative, though, is to look for structures that are like truth tables, but more discriminating. Consider truth-tables for conjunction and disjunction:

"Even the sentences of Frege's mature logical system are complex terms; they are terms that denote *truth-values*. Frege distinguished two truth-values, THE TRUE and THE FALSE, which he took to be objects." Edward N. Zalta, "Gottlob Frege," http://plato.stanford.edu/entries/frege/, [98].

| $\wedge$ | t | f |
|---|---|---|
| t | t | f |
| f | f | f |

| $\vee$ | t | f |
|---|---|---|
| t | t | t |
| f | t | f |

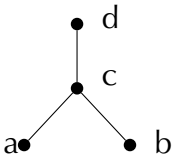

We construe conjunction and disjunction as *operations* on the little set $\{t, f\}$ of truth values. You can encode the information in these tables more compactly, if you are prepared to use a little imagination, and if you are prepared to accept a metaphor. Think of THE TRUE (t) as *higher* than THE FALSE (f). Think of the "under-and-over" relationship as "less than" or "$<$" and you get $f < t$, but not *vice versa*. We also get $t \leqslant t$, $f \leqslant t$ and $f \leqslant f$ but $t \not\leqslant f$. We can then understand the behaviour of conjunction and disjunction on the set of truth values as *minimum* and *maximum* respectively. A sequent $A \vdash B$ is valid (according to truth tables) if and only if the value of the premise A is always less than or equal to the value of the conclusion B, no matter how we evaluate the ATOMS in A and B as truth values. You can think of $\leqslant$ then as a kind of rendering of *entailment* in the small universe of truth values. A conjunction entails the both conjuncts, since its value will be the minimum of the values of either of the conjuncts. A disjunction is entailed by either of the disjuncts, because its value is the maximum of the values of the two disjuncts.

Does the treatment work in the more general context of the logic of conjunction and disjunction generated by our sequent system? The

answer is negative. Minimum and maximum behave quite a lot like conjunction and disjunction, but they do slightly *more* than we can prove with these connectives here. You can show that the distribution law $A \wedge (B \vee C) \vdash (A \wedge B) \vee C$ is valid under the interpretation of conjunction and disjunction as minimum and maximum. Treating disjunction and conjunction as maximum and minimum is too strong for our purposes. Regardless, it points in a helpful direction.

$A \wedge (B \vee C)$ is the smaller of A and (the larger of B and C). $(A \wedge B) \vee C$ is the larger of (the smaller of A and B) and C. The first is always smaller than (or equal to) the second.

Sometimes orderings do not have maxima and minima. Consider the following ordering, in which $a < c$, $b < c$ and $c < d$.




This picture is a Hasse Diagram. This diagram is a way of presenting an ordering relation $<$: a relation that is irreflexive ($x \not< x$ for each $x$) transitive (if $x < y$ and $y < z$ then $x < z$) and antistymmetric (if $x < y$ then $y \not< x$). In the diagram, we represent the objects by dots, and we draw a single line from $x$ *up to* $y$ just when $x < y$ and there is no $z$ where $x < z < y$. Then, in general, $x < y$ if and only if you can follow a path from $x$ up to $y$.

There is a disjunction of a and b in one sense: it is c. $a \leqslant c$ and $b \leqslant c$, so c is an *upper bound* for both a and b. Not only that, but c is a *least* upper bound for a and b. That is, among all of the upper bounds of a and b (that is, among c and d), c is the smallest. A *least upper bound* is a good candidate for disjunction, since if $z$ is a least upper bound for $x$ and $y$ then we have

$$x \leqslant z \text{ and } y \leqslant z$$

(it's an upper bound) and

$$\text{If } x \leqslant z' \text{ and } y \leqslant z' \text{ then } z \leqslant z'$$

(it's the least of the upper bounds). If we write the $z$ here as $x \vee y$, and if we utilise the transitivity of $\leqslant$, we could write $x \leqslant x \vee y$ as "if $v \leqslant x$ then $v \leqslant x \vee z$." Our rules then take the form

$$\frac{v \leqslant x}{v \leqslant x \vee y} \qquad \frac{v \leqslant y}{v \leqslant x \vee y} \qquad \frac{x \leqslant u \quad y \leqslant u}{x \vee y \leqslant u}$$

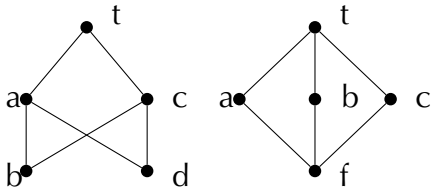
which should look rather familiar. If we think of entailment as an *ordering* among pieces of information (or propositions, or what-have-you), then disjunction forms a least upper bound on that ordering. Clearly the same sort of thing could be said for conjunction. Conjunction is a *greatest lower bound*:

$$\frac{x \leqslant v}{x \wedge y \leqslant v} \qquad \frac{y \leqslant v}{x \wedge y \leqslant v} \qquad \frac{u \leqslant x \quad u \leqslant y}{u \leqslant x \wedge y}$$

Ordered structures in which every pair of elements has a greatest lower bound (or MEET) and least upper bound (or JOIN) are called LATTICES.

DEFINITION 2.5.13 [LATTICE] An ordered set $\langle P, \leqslant, \wedge, \vee \rangle$ with operators $\wedge$ and $\vee$ is said to be a LATTICE iff for each $x, y \in P$, $x \wedge y$ is the greatest lower bound of $x$ and $y$ (with respect to the ordering $\leqslant$) and $x \vee y$ is the least upper bound of $x$ and $y$ (with respect to the ordering $\leqslant$).

Consider the two structures below. The one on the left is not a lattice, but the one on the right is a lattice.




On the left, b and d have no lower bound at all (nothing is below both of them), and while they have an upper bound (both a and c are upper bounds of b and d) they do not have a *least* upper bound. On the other hand, every pair of objects in the structure on the right has a meet and a join. They are listed in the tables below:

| $\wedge$ | f | a | b | c | t |
|---|---|---|---|---|---|
| f | f | f | f | f | f |
| a | f | a | f | f | a |
| b | f | f | b | f | b |
| c | f | f | f | c | c |
| t | f | a | b | c | t |

| $\vee$ | f | a | b | c | t |
|---|---|---|---|---|---|
| f | f | a | b | c | t |
| a | a | a | t | t | t |
| b | b | t | b | t | t |
| c | c | t | t | c | t |
| t | t | t | t | t | t |

Notice that in this lattice, the distribution law fails in the following way:

$$a \wedge (b \vee c) = a \wedge t = a \not\leqslant c = f \vee c = (a \wedge b) \vee c$$

Lattices stand to our logic of conjunction and disjunction in the same sort of way that truth tables stand to traditional classical propositional logic. Given a lattice $\langle P, \leqslant, \wedge, \vee \rangle$ we can define a *valuation* $v$ on FORMULA in the standard way.

DEFINITION 2.5.14 [VALUATION] We assign, for every ATOM $p$, a value $v(p)$ from the lattice $P$. Then we extend $v$ to map each FORMULA into $P$ by setting

$$v(A \wedge B) = v(A) \wedge v(B) \qquad v(A \vee B) = v(A) \vee v(B)$$

Using valuations, we can evaulate sequents.

THEOREM 2.5.15 [SUITABILITY OF LATTICES] *A sequent $A \vdash B$ has a proof if and only if for every lattice and for every valuation $v$, on that lattice, $v(A) \leqslant v(B)$.*

*Proof*: The proof takes two parts, "if" and "only if." For "only if" we need to ensure that if $A \vdash B$ has a proof, then for any valuation on any lattice, $v(A) \leqslant v(B)$. For this, we proceed by induction the construction of the derivation of $A \vdash B$. If the proof is simply the axiom of identity, then $A \vdash B$ is $p \vdash p$, and $v(p) \leqslant v(p)$. Now suppose that the proof is more complicated, and that the hypothesis holds for the prior steps in the proof. We inspect the rules one-by-one. Consider $\wedge L$: from $A \vdash R$ to $A \wedge B \vdash R$. If we have a proof of $A \vdash R$, we know that $v(A) \leqslant v(R)$. We also know that $v(A \wedge B) = v(A) \wedge v(B) \leqslant v(A)$ (since $\wedge$ is a lower bound), so $v(A \wedge B) \leqslant v(R)$ as desired. Similarly for $\vee R$: from $L \vdash A$ to $L \vdash A \vee B$. If we have a proof of $L \vdash A$ we know that $v(L) \vdash v(A)$.

We also know that $v(A) \leqslant v(A \vee B)$ (since $\vee$ is an upper bound), so $v(L) \leqslant v(A \vee B)$ as desired.

For $\wedge R$, we may assume that $L \vdash A$ and that $L \vdash B$ have proofs, so for every $v$, we have $v(L) \leqslant v(A)$ and $v(L) \leqslant v(B)$. Now, let $w$ be any evaluation, since $w(L)$ is a lower bound of $w(A)$ and $w(B)$, it must be less than or equal to the *greatest* lower bound $w(A) \wedge w(B)$. That is, we have $w(L) \leqslant w(A) \wedge w(B) = w(A \wedge B)$, which is what we wanted. For $\vee L$ we reason similarly. Assuming that $A \vdash R$ and $B \vdash R$ we have $w(A) \leqslant w(R)$ and $w(B) \leqslant w(R)$, and hence $w(R)$ is an upper bound for $w(A)$ and $w(B)$. So $w(A \vee B) = w(A) \vee w(B) \leqslant w(R)$, since $w(A) \vee w(B)$ is the least upper bound for $w(A)$ and $w(B)$.

We can even show that if the proof uses Cut it preserves validity in lattices, since if $v(A) \leqslant v(B)$ and $v(B) \leqslant v(C)$ then we have $v(A) \leqslant v(C)$, since $\leqslant$ is a transitive relation.

Now for the "if" part. We need to show that if $A \vdash B$ has *no* proof, then there is some lattice in which $v(A) \not\leqslant v(B)$. Just as in the proof of Theorem 2.5.12, we will construct a model (in this case it is a lattice) from the formulas and proofs themselves. The core idea is that we will assume as little as possible about the relationships between objects in our lattice. We will choose a value $v(A)$ for $A$ such that $v(A) \leqslant v(B)$ if and only if $A \vdash B$. We can do this if let the objects in our lattice be *equivalence classes* of formulas, like this:

The core construction is *really* (as a mathematician will tell you) a construction of the *free lattice* on the generator set ATOM.

$$B \in [A] \text{ if and only if there are proofs } A \vdash B \text{ and } B \vdash A$$

So, $[A]$ contains $A \wedge A$, $A \vee A$, $A \wedge (A \vee B)$, and many other formulas besides. It is the collection of all of the formulas *exactly* as strong as $A$ (and no stronger). These classes of formulas form the objects in our lattice. To define the ordering, we set $[A] \leqslant [B]$ if and only if $A \vdash B$. This definition makes sense, because if we chose *different* representatives for the collections $[A]$ and $[B]$ (say, $A' \in [A]$, so $[A'] = [A]$, and similarly, $[B] = [B']$) we would have $A' \vdash A$ and $B \vdash B'$ so $A' \vdash B'$ too.

The relation $A \sim B$, defined by setting $A \sim B$ iff $A \vdash B$ and $B \vdash A$ is a *congruence*.

It remains to be seen that this ordering has a lattice structure. We want to show that $[A] \wedge [B]$, defined as $[A \wedge B]$, is the meet of $[A]$ and $[B]$, and that $[A] \vee [B]$, defined as $[A \vee B]$, is the join of $[A]$ and $[B]$. This requires first showing that the definitions make sense. (Would we have got a different result for the meet or for the join of two collections had we chosen different representatives?) For this we need to show that if $A' \in [A]$ and $B' \in [B]$ then $A' \wedge B' \in [A \wedge B]$. This is straightforward, using the two proofs below:

$$\dfrac{\dfrac{\dfrac{A' \vdash A}{A' \wedge B' \vdash A}\wedge R_1 \quad \dfrac{B' \vdash B}{A' \wedge B' \vdash B}\wedge R_2}{A' \wedge B' \vdash A \wedge B}\wedge R \qquad \dfrac{\dfrac{\dfrac{A \vdash A'}{A \wedge B \vdash A'}\wedge R_1 \quad \dfrac{B \vdash B'}{A \wedge B \vdash B'}\wedge R_2}{A \wedge B \vdash A' \wedge B'}\wedge R}$$

If $A$ and $A'$ are equivalent, and $B$ and $B'$ are also equivalent, then so are their respective conjunctions. The proof is similar for disjunction, so the operations on equivalence classes are well defined. It remains

to show that they are, respectively, meet and join. For disjunction, we need to show that $[A \vee B]$ is the join of $[A]$ and $[B]$. First, $[A \vee B]$ is an upper bound of $[A]$ and of $[B]$, since $A \vdash A \vee B$ and $B \vdash A \vee B$. It is the least such upper bound, since if $[A] \leqslant [R]$ and $[B] \leqslant [R]$ for any formula $R$, we have a proof of $A \vdash R$ and a proof of $B \vdash R$, which ensures that we have a proof of $A \vee B \vdash R$, and hence $[A \vee B] \leqslant [R]$. The proof that $\wedge$ on the set of equivalence classes is a greatest lower bound has the same form, and we consider it done. We have constructed a lattice out of the set FORMULA.

Now, if $A \nvdash B$, we can construct a refuting valuation $\nu$ for this sequent. Define $\nu$ into our lattice by setting $\nu(p)$ to be $[p]$. For any complex formula $C$, $\nu(C) = [C]$, by tracing through the construction of $C$: $\nu(A \wedge B) = \nu(A) \wedge \nu(B) = [A] \wedge [B] = [A \wedge B]$, and similarly $\nu(A \vee B) = \nu(A) \vee \nu(B) = [A] \vee [B] = [A \vee B]$. So, if $A \nvdash B$, then since $[A] \not\leqslant [B]$, we have $\nu(A) \not\leqslant \nu(B)$. We have a counterexample in our lattice. ∎

This is one example of the way that we can think of our logic in an *algebraic* manner. We will see many more later. Before going on to the next section, let us use a little of what we have seen in order to reflect more deeply on cut and identity.

### 2.5.3 | CONSEQUENCES OF CUT AND IDENTITY

Doing without Cut as well as $[\mathrm{Id}_A]$ puts quite a strong constraint on the rules for a connective. To ensure that identities are provable for conjunctions, we needed to have a proof of $p \wedge q \vdash p \wedge q$. When you think about it, this means that the left and right rules for conjunction must be connected. The rules for a conjunction on the *left* must be strong enough to entail whatever is required to give you the conjunction on the *right*. Using the concepts of algebras, if we think of the left and right rules for conjunction as defining two different connectives, $\wedge_l$ and $\wedge_r$ respectively, the provability of identity ensures that we have $x \wedge_l y \leqslant x \wedge_r y$.

The same holds for the elimination of Cut. Showing that you can eliminate a Cut in which a conjunction is the cut-formula, we need to show now that the rules for conjunction on the *right* are strong enough to ensure that whatever it is that is entailed by a conjunction on the right is "true enough" to be entailed by whatever entails the conjunction on the *left*. Thinking algebraically again, the eliminatibility of Cut ensures that whenever $z \leqslant x \wedge_r y$ and $x \wedge_l y \leqslant z'$ we have $z \leqslant z'$. That is, it ensures that $x \wedge_r y \leqslant x \wedge_l y$.

In other words, you can think of cut-elimination and the provability of identity as ensuring that the left and right rules for a connective are appropriately *coordinated*. You can prove an identity $A \sharp B \vdash A \sharp B$ if the right rules for $\sharp$ make it "more" true than the left rules, but it will not do if $\sharp$-on-the-right is "less" true than $\sharp$-on-the-right. A Cut step from $L \vdash A \sharp B$ and $A \sharp B \vdash R$ to $L \vdash R$, will work when $\sharp$-on-

the-right is "less" true than ♯-on-the-right, but it will not work if the mismatch is in the other direction.

### 2.5.4 | EXERCISES

## 2.6 | PROOF IDENTITY

[When is proof $\pi_1$ *the same proof* as $\pi_2$? There are different things one could mean by this.]

[NORMALISATION ANALYSIS: Proof $\pi_1$ is identical to proof $\pi_2$ when $\pi_1$ and $\pi_2$ have the same normal forms. Examine this analysis of proof identity in the context of intuitionistic logic.]

[GENERALITY ANALYSIS: Proof $\pi_1$ differs from proof $\pi_2$ when they have different *generality*. This can be analysed in terms of *flow graphs* [13, 14, 15].]

[This has an application to category theory, and categories are a natural home for models of proofs, if we think of them as structures with propositions/statements as objects and proofs as arrows. (Provided that we think of proofs as having a single premise and a single conclusion of course. Generalisations of categories are required if we have a richer view of what a proof might be.)]

# PROPOSITIONAL LOGIC: APPLICATIONS

# 3

As promised, we can move from technical details to applications of these results. With rather a lot of proof theory under our collective belts, we can turn our attention to philosophical issues. In this chapter, we will look at questions such as these: How are we to understand the distinctive necessity of logical deduction? What is the distinctively logical '*must*'? In what way are logical rules to be thought of as definitions? What can we say about the epistemology of logical truths? Can there be genuine disagreement between rival logical theories, or are all such discussions a dialogue of in which the participants talk different languages and talk past one another?

In later chapters we will examine other topics, such as generality, predication, objectivity, modality, and truth. For those, we require a little more logical sophistication than we have covered to this point. What we have done so far suffices to equip us to tackle the topics at hand.

The techniques and systems of logic can be used for many different things — we can design electronic circuits using simple boolean logic [97]. We can control washing machines with fuzzy logic [45]. We can use substructural logics to understand syntax [16, 54, 55] — but beyond any of those interesting applications, we can use the techniques of logic to construct arguments, to evaluate them, and to tell us something about how beliefs, conjectures, theories and statements, fit together. It is this role of logic that is our topic.

## 3.1 | ASSERTION AND DENIAL

One distinguishing feature of this application of logical techniques is its bearing on our practices of assertion. We may express beliefs by making assertions. When considering a conjecture, we may tentatively make a hypothetical assertion — to "try it on for size" and to see where it leads. When we use the techniques of logic to evaluate a piece of mathematical reasoning, the components of that reasoning are, at least in part, assertions. If we have an argument leading from a premise A to a conclusion B, then this tells us something significant about an assertion of A (it tells us, in part, where it leads). This also tells us something significant about an assertion of B (it tells us, in part, what leads to it). Can we say more than this? What is the connection between proof (and logical consequence) and assertion? It would seem that there is an intimate connection between proof and assertion, but what is it?

Suppose that A entails B, that there is a proof of B from A. What can we say about assertions of A and of B? If an agent accepts A, then it is tempting to say that the agent also *ought* to accept B, because B follows from A. But this is far too strong a requirement to take seriously. Let's consider why not:

(1) The requirement, as I have expressed it, has many counterexamples. The requirement has the following form:

> *If* A *entails* B, *and I accept* A, *then I ought to accept* B.

Notice that I have a proof from A to A. (It is a very small proof: the identity proof.) It would follow, if I took this requirement seriously, that if I accept A, then I *ought to* accept A. But there are many things – presumably – that I accept that I ought not accept. My beliefs extend beyond my entitled beliefs. The mere fact that I believe A does not in and of itself, give me an entitlement, let alone, an obligation to believe A. So, the requirement that you ought to accept the consequences of your beliefs is altogether too strong.

This error in the requirement is corrected with a straightforward scope distinction. Instead of saying that if A entails B and if you accept A then you ought to accept B, we should perhaps weaken the condition as follows:

> *If* A *entails* B, *then it ought to be that if I accept* A *then I accept* B.

We fix the mistaken scope by holding that what we accept should be closed under entailment. But this, too, is altogether too strong, as the following considerations show.

(2) There are many consequences of which we are unaware. It seems that logical consequence on its own provides no obligation to believe. Here is an example: I accept all of the axioms of Peano arithmetic (PA). However, there are consequences of that theory that I do not accept. I do not accept all of the consequences of those axioms. Goldbach's conjecture (GC) could well be a consequence of those axioms, but I am not aware of this if it is the case, and I do not accept GC. If GC is a consequence of PA, then there is a sense in which I have not lived up to some kind of standard if I fail to accept it. My beliefs are not as comprehensive as they could be. If I believed GC, then in some important sense I would not make any more mistakes than I have already made, because GC is a consequence of my prior beliefs. However, it is by no means clear that comprehensiveness of this kind is desirable.

(3) In fact, comprehensiveness is *undesirable* for limited agents like us. The inference from A to $A \vee B$ is valid, and if our beliefs are always to be closed under logical consequence, then for any belief we must have infinitely many more. But consider a very long disjunction, in which *one* of the disjuncts we already accept. In what sense is it desirable that we accept this long disjunction? The belief may be too complex to even *consider*, let alone, to believe or accept or assert.

Notice that it is not a sufficient repair to demand that we merely accept the *immediate* logical consequences of our beliefs. It may well be true that logical consequence in general may be analysed in terms of chains of immediate inferences we all accept when they are presented to us. The problems we have seen hold for immediate consequence. The inference from the axioms of PA to Goldbach's conjecture might be decomposable into steps of immediate inferences. This would not make Goldbach's conjecture any more rationally obligatory, if we are unaware of that proof. If the inference from A to $A \vee B$ is an immediate inference, then logical closure licenses an infinite collection of (irrelevant) beliefs.

This point is not new. Gilbert Harman, for example, argues for it in *Change in View* [40].

(4) Furthermore, logical consequence is sometimes impossible to check. If I must accept the consequences of my beliefs, then I must accept all tautologies. If logical consequence is as complex as consequence in classical first-order logic, then the demand for closure under logical consequence can easily be *uncomputable*. For very many sets of statements, there is no algorithm to determine whether or not a given statement is a logical consequence of that set. Closure under logical consequence cannot be underwritten by algorithm, so demanding it goes beyond what we could rightly expect for an agent whose capacities are computationally bounded.

So, these arguments show that logical *closure* is too strict a standard to demand, and failure to live up to it is no failure at all. Logical consequence must have some other grip on agents like us. But what could this grip be? Consider again the case of the valid argument from A to B, and suppose, as we did before, that an agent accepts A. What can we say about the agent's attitude to B? One thing that we could clearly say is that the agent is, in some sense, *committed* to B. There is good sense in saying that we are (at least implicitly) committed to the logical consequences of those things we accept. Deduction is one way of drawing out the implicit commitments that we have, and making them explicit. We could hold to the following two norms:

> *If* A *entails* B, *and I accept* A, *then I am committed to* B.

In fact, we could accept the closure of commitment under logical consequence

> *If* A *entails* B, *and I am committed to* A, *then I am committed to* B.

by holding that those claims to which I am committed are all and only the consequences of those things I accept. There are good reasons to think of commitment in this way. However, it remains that for the reasons we have already seen, we need not think of what we accept as conforming to the same kinds of conditions as those things to which we are committed.

I will slightly revise this notion later, but it will do for now.

Do any norms govern what we accept? One plausible constraint is that if some set of propositions is inconsistent, then we should not

accept all of them. If $X \vdash$, then accepting every member of X is some kind of mistake. (It's not just some kind of mistake, it's a *bad* mistake.) This is one plausible constraint on what we accept. Similarly, if we were to *assert* every member of an inconsistent set X, then that act of assertion is, in some sense to be made out, defective. But what about collections of propositions other than inconsistent ones? If we have an argument from A to B, does this argument constrain what we accept in a similar way to the inconsistent collection?

To look at this closely, suppose for a moment that we have an argument from A to B. If $A \vdash B$, then if we have negation in the vocabulary, we can conclude $A, \neg B \vdash$. This tells us that accepting both A and $\neg B$ is a mistake. This points to a general account for valid argument. If an argument is valid then it is a mistake to accept the premise and to reject the conclusion.

If an agent's cognitive state, in part, is measured in terms of those things she accepts and those she rejects, then valid arguments constrain those combinations of acceptance and rejection. As we have seen, a one-premise, one-conclusion argument from A to B constrains acceptance/rejection by ruling out accepting A and rejecting B. This explanation of the grip of valid argument has the advantage of symmetry. A valid argument from A to B does not, except by force of habit, have to be read as *establishing* the conclusion. If the conclusion is unbelievable, then it could just as well be read as *undermining* the premise. Reading the argument as constraining a pattern of accepting and rejecting gives this symmetry its rightful place.

Now consider the connection between negation and rejection. I have introduced rejection by means of negation, but now it is time take away the ladder. Or to change the metaphor, let's turn the picture upside down. We have picked out rejection in using negation, but we do not need to define it in that way. Instead, we will in fact be using rejection to give an account of negation. If there are reasoning and representing agents who do not have the concept of negation, and if it is still appropriate for us to analyse their reasoning using a notion of logical consequence, then we ought to take those agents as possessing the ability to *deny* without having the ability to *negate*. This seems plausible. As an agent accepts and rejects, she filters out information and rules out possibilities. For any possible judgement the agent might consider, there are three possible responses: she might accept it, she might reject it, or she may be undecided. To assert that A is the case is to express your accepting that A is the case. If you prefer, to accept that A is the case is to internalise the assertion that A. To reject A is not merely to fail to accept A, for we may be undecided. To deny A is not merely to fail to assert A, for we may be silent. To accept A is to (in part) close off the possibility of rejecting A. To accept A and then to go on to *reject* A will result in a *revision* of your commitments, and not a mere *addition* to them. Similarly, to reject A is to (in part) close off the possibility of accepting A. To reject A and then to go on to *accept* A will result in a revision of your commitments, and not a mere addition to them. To assert A is to place the denial of A out of bounds—until you

This has nothing to do with a 'three valued logic' as we do not take the attitudes of some agent to a statement to be anything like 'semantic values' let alone 'truth values' of that statement.

revise your commitments by withdrawing the assertion. Similarly, to deny A is to place the assertion of A out of bounds—again, until the denial of A is withdrawn. An agent's ability to do this does not require that she has a concept of *negation*, giving for each statement another statement—its *negation*—which can be asserted or denied, used as a premise in reasoning, as a component of another proposition, and so on.

Before going on to utilise the notions of assertion and denial, I should defend the claim that we ought to treat assertion and denial on a par, and that it is not necessary to treat the denial of a claim as the assertion of the negation of that claim. This matters because in the sections that follow, I will use denial in giving an account of negation. To require of an agent a grasp of negation in order to have the capacity to deny would threaten this account.

I will argue that the speech-act of *denial* is best not analysed in terms of *assertion* and *negation* but rather, that denial is, in some sense, prior to negation. I will provide three different considerations in favour of this position. The first involves the case of an agent with a limited logical vocabulary. The second argument, closely related to the first, involves the case of the proponent of a non-classical logic. The third will rely on general principles about the way logical consequence rationally constrains assertion and denial.

CONSIDERATION ONE: Parents of small children are aware that the ability to *refuse*, *deny* and *reject* arrives very early in life. Considering whether or not something is the case – whether to accept that something is the case or to reject it – at least *appears* to be an ability children acquire quite readily. At face value, it seems that the ability to assert and to deny, to say *yes* or *no* to simple questions, arrives earlier than any ability the child has to form sentences featuring negation as an operator. It is one thing to consider whether or not A is the case, and it is another to take the *negation* ¬A as a further item for consideration and reflection, to be combined with others, or to be supposed, questioned, addressed or refuted in its own right. The case of early development lends credence to the claim that the ability to deny can occur prior to the ability to form negations. If this is the case, the denial of A, in the mouth of a child, is perhaps best not analysed as the assertion of ¬A.

So, we might say that denial may be *acquisitionally prior* to negation. One can acquire the ability to deny before the ability to form negations.

CONSIDERATION TWO: Consider a related case. Sometimes we are confronted with theories which propose non-standard accounts of negation, and sometimes we are confronted with people who endorse such theories. These will give us cases of people who appear to reject A without accepting ¬A, or who appear to accept ¬A without rejecting A. If things are as they appear in these cases, then we have further reason

to reject the analysis of rejection as the acceptance of a negation. I will consider just two cases.

SUPERVALUATIONISM: The supervaluationist account of truth-value gaps enjoins us to allow for claims which are not determinately true, and not determinately false [30, 52, 95]. These claims are those which are true on some valuations and false on others. In the case of the supervaluational account of *vagueness*, borderline cases of vague terms are a good example. If Fred is a borderline case of baldness, then on some valuations "Fred is bald" is true, and on others, "Fred is bald" is false. So, "Fred is bald" is not true under the *super*valuation, and it is to be rejected. However, "Fred is not bald" is similarly true on some valuations and false on others. So, "Fred is not bald" is not true under the supervaluation, and it, too, is to be rejected. Truth value gaps provide examples where denial and the assertion of a negation come apart. The supervaluationist rejects A without accepting ¬A. When questioned, she will deny A, and she will *also* deny ¬A. She will not accept ¬A. The supervaluationist seems to be a counterexample to the analysis of denial as the assertion of a negation.

DIALETHEISM: The dialetheist provides is the dual case [60, 61, 66, 67, 69]. A dialetheist allows for truth-value *gluts* instead of truth-value *gaps*. Dialetheists, on occasion, take it to be appropriate to assert both A and ¬A. A popular example is provided by the semantic paradoxes. Graham Priest's analysis of the liar paradox, for example, enjoins us to accept both the liar sentence and its negation, and to reject neither. In this case, it seems, the dialetheist accepts a negation ¬A *without* rejecting A, the proposition negated. When questioned, he will assert A, and he will *also* assert ¬A. He will not reject ¬A. The dialetheist, too, seems to be a counterexample to the analysis of denial as the assertion of a negation.

In each case, we seem to have reason to take denial to be something other than the assertion of a negation, at least in the mouths of the supervaluationist and the dialetheist. These considerations are not conclusive: the proponent of the analysis of rejection in terms of negation may well say that the supervaluationist and the dialetheist are confused about negation, and that their denials really *do* have the content of a negation (by their own lights), despite their protestations to the contrary. Although this is a possible response, there is no doubt that it does violence to the positions of both the supervaluationist and the dialetheist. We would do better to see if there is an understanding of the connections between assertion, denial, acceptance, rejection and negation which allows us to take these positions at something approaching face value. This example shows that denial may be *conceptually separated* from the assertion of a negation.

CONSIDERATION THREE: The third consideration favouring denial over negation is the fertility of the analysis. Once our agent is able to assert and to deny, the full force of logical consequence will have its grip

on the behaviour of the agent. Asserting A has its consequences for denials (do not deny A, and do not deny any other consequence of A). Denying B has its consequences for assertions (do not assert B, and do not assert any other premise leading to B). This analysis generalises to arguments with multiple premises in the way that you would expect. More interestingly, it also generalises to arguments with multiple conclusions.

There is another consideration in favour of taking denial as prior to negation. See page 141 for the details.

So, let us pass on to the elaboration of the story. We can think of assertions and denials *en masse*, as the normative force of logical consequence will be explained in terms of the proprieties of different combinations of assertions and denials. To fix terminology, we will call collections of assertions and denials POSITIONS. A position [X : Y] is a pair of (finite) sets, X of things *asserted* and Y of things *denied*. Positions are the kinds of things that it is appropriate to evaluate. It is rare that a single assertion or denial deserves a tick of approval or a black mark of opprobrium on the basis of logic alone. It is much more often that collections of assertions and denials — our *positions* — stand before the tribunal of logical evaluation. The most elementary judgement we may make concerning a position is the following:

The 'game' terminology is intentional. A position is a part of a 'scorecard' in Brandom's normative scorekeeping [10]. A scorecard keeps track of commitments and entitlements. Here, a position keeps track merely of explicit commitments made by way of assertion and denial. You find yourself in a position on the basis of *moves* you have made. More on commitment later.

IDENTITY: [A : A] is incoherent.

A position consisting of the solitary assertion of A (whatever claim A might be) together with its denial, is incoherent.

To grasp the import of calling a position incoherent, it is vital to neither understate it, nor to overstate it. First, we should not overstate the claim by taking incoherent positions to be *impossible*. While it might be very difficult for someone to sincerely assert and deny the same statement in the same breath, it is by no means impossible. For example, if we wish to refute a claim, we may proceed by means of a *reductio ad absurdum* by asserting (under an assumption) that claim, deriving others from it, and perhaps leading on to the denial of something we have already asserted. Once we find ourselves in this position (including the assertion and the denial of the one and the same claim) we withdraw the supposition. We may have good reasons to put ourselves in incoherent positions, in order to manage the assertions and denials we wish to make. To call a position incoherent is not to say that the combination of assertions and denials cannot be made.

Conversely, it is important to not *understate* the claim of incoherence. To call the position [X : Y] incoherent is not merely to say that it is *irrational* to assert X and deny Y, or that it is some kind of bad idea. It is much more than that. Consider the case of the position [A : A]. This position is seriously self-defeating in that to take you to assert A is to take you to rule out denials of A (pending a retraction of that assertion), to take you to deny A is to take you to rule out assertions of A (pending a retraction of that denial). The incoherence in the position is due to the connection between assertion and denial, in that to make the one is to preclude the other. The incoherence is not simply due to any external feature of the content of that assertion. As a matter of

fact, it seems that whenever we take someone to have denied A (and to have in the past asserted A) we take this to be a *retraction* of the assertion.

Another claim can be made about the properties of coherence. The incoherence of a position cannot be fixed by merely adding more assertions or denials.

WEAKENING: If [X, A : Y] or [X : A, Y] is coherent, then [X : Y] is too.

This tells us that incoherence is a property preserved when assertions and denials are added to a position, and that coherence is a property preserved when assertions and denials are removed from a position. Contraposing the statement of weakening, we see that if the position [X : Y] is incoherent, then merely adding the assertion of A, or adding the denial of A does not fix things. To take a position to be coherent is to endorse weaker positions (with fewer commitments) as coherent as well. To take a position to be incoherent is to judge stronger positions (with more commitments) as similarly incoherent. It follows that I can judge a position incoherent if it contains the assertion of A and the denial of A, without having to check all of the other assertions and denials in that position.

Our final requirement on the relation of coherence is the converse of the weakening condition. So, we call it

STRENGTHENING: If [X : Y] is coherent, then it either [X, A : Y] or [X : A, Y] is coherent too.

This condition can be viewed as follows: if we have a coherent position [X : Y], then if we cannot add A as an assertion (if [X, A : Y] is incoherent) then the claim A is implicitly denied in that position. An implicit denial is as good as an explicit denial, so since the original position was coherent, the new position in which A is explicitly denied is coherent as well. Taking the other horn of the disjunction, if we cannot add A as a denial (if [X : A, Y] is incoherent) then the claim A is implicitly asserted in that position. An implicit assertion is as good as an explicit assertion, so since the original position was coherent, the new position in which A is explicitly asserted is coherent as well.

These three constraints on coherence, assertion and denial are enough to motivate a definition:

DEFINITION 3.1.1 [COHERENCE] Given a language, we will call a relation on the set of positions in that language a *coherence* relation if and only if it satisfies the conditions of identity, weakening and strengthening. Given a coherence relation, we will indicate the *incoherence* of a position [X : Y] as follows: 'X ⊢ Y'.

The turnstile notation X ⊢ Y makes sense, as a coherence relation is really a consequence relation. The IDENTITY condition is A ⊢ A, the identity sequent. The STRENGTHENING condition tells us that if X ⊢

$A, Y$ and $X, A \vdash Y$ then $X \vdash Y$. This is the *cut* rule. The weakening condition is, well, the weakening structural rule.

The rules are not *exactly* the kinds of rules we have seen before, since the collections on either side of the turnstile here are not multisets of statements, but *sets* of statements. This means that the contraction rule is implicit and does not need to be added as a separate rule: if $X, A, A \vdash Y$ then $X, A \vdash Y$, since the set $X, A, A$ is the same set as $X, A$. The premise and conclusion of a contraction rule are exactly the same sequents.

Viewed one way, the cut rule is a *transitivity* condition: If $A \vdash B$ and $B \vdash C$ then $A \vdash C$. The form that we have here does not look so much like transitivity, but a strange cousin of the law of the excluded middle: it says that it is either ok to assert $A$ or to deny $A$. But it is the same old cut rule that we have already seen. If $A \vdash B$ and $B \vdash C$, then it follows that $A \vdash B, C$ and $A, B \vdash C$ by weakening. The strengthening rule tells us that $A \vdash C$. In other words, since $A \vdash B, C$ (we cannot deny $B$ in the context $[A : C]$) and $A, B \vdash C$ (we cannot assert $B$ in the context $[A : C]$), there is a problem with the context $[A : C]$. It is incoherent: we have $A \vdash C$.

Let us consider, for a moment, the nature of the languages related by a coherence relation, and the kinds of conditions we might find in a coherence relation. In most applications we are concerned with, the coherence relation on our target vocabulary will be rich and interesting, as assertions and denials are made in a vocabulary with robust connections of meaning. It seems to me that all of the kinds of relationships expressible in the vocabulary of coherence make sense. We have $A \vdash B$ if asserting $A$ and denying $B$ is out of bounds. An example might be 'this is a square' and 'this is a parallelogram'. The only way to be a square is by being a parallelogram, and it is a mistake to deny that something is a parallelogram while claiming it to be a square. We have $A, B \vdash$ if asserting both $A$ and $B$ is out of bounds. An example might be 'it is alive' and 'it is dead'. To claim of something that it is both alive and that it is dead is to not grasp what these terms mean. Examples with multiple premises are not difficult to find. We have $A, B, C \vdash D$ when the assertion of both $A$, $B$ and $C$ together with the denial of $D$ is out of bounds: consider 'it is a chess piece', 'it moves forward', 'it captures diagonally', and 'it is a pawn.' Examples with multiple conclusions are more interesting. We have $A \vdash B, C$ if asserting $A$ and denying both $B$ and $C$ is out of bounds. Perhaps the trio 'x is a parent', 'x is a mother' and 'x is a father' is an example, or for a mathematical case, we have 'x is an number', 'x is zero' and 'x is a successor'.

The limiting cases of coherence are the situations in which we have an incoherent position with just *one* assertion or denial. If we have $A \vdash$ then the assertion of $A$ by itself is incoherent. If we have $\vdash B$, then the denial of $B$ is incoherent. It is plausible that there are such claims, but it is less clear that there are *primitive* statements in a vocabulary that deserve to be thought of as inconsistent or undeniable on their own.

Notice, too, that it makes sense to treat any kind of vocabulary

in which assertions and denials are made as one in which assertions and denials may together be coherent or incoherent. We may judge discourse on any subject matter for coherence, even when we are not clear on what the subject is about, or what it would be for the claims in that vocabulary to be justified or warranted. Even though we may be completely unsettled on the matter of what it might be for moral claims to be *true* or *justified*, it can make sense for us to evaluate moral discourse with respect to coherence. The claim that it is wrong to accuse someone of a crime without evidence (for example) may be taken to entail that it is sometimes wrong to accuse someone of a crime. It makes sense to take claims of wrongness and rightness to be coherent or incoherent with other claims, even if we are unsure as to the matter of what makes moral claims true (or even if we are skeptical that they are the kinds of claims that are 'made true' by anything at all). It suffices that they may be asserted and denied.

However, as far as *logic* is concerned what we will haev to say will apply as well to a circumstance in which we have the *smallest* relation satisfying these conditions: in which $X \vdash Y$ only when X and Y share a proposition. In this case, the statements in the vocabulary can be thought of as entirely logically independent. The only way a position in this vocabulary will count as incoherent is if it has managed to assert one thing and to deny it as well — not by meand of asserting or denying something *else*, but by explicitly asserting A and denying A at the very same time. In this case, the statements of the language are entirely logically independent.

## 3.2 | DEFINITION AND HARMONY

Now suppose that we have a vocabulary and the relation of coherence upon it. Suppose that someone wants to understand the behaviour of *negation*. We could incorporate the rules for negation. First, we treat the *syntax* by closing the vocabulary under negation. In other words, we add the condition that if A is a statement, then $\neg A$ is a statement too. This means that not only do we have the original statements, and the new *negations* in our vocabulary, but we also have negations of negations, and so on. If the language is formed by means of other grammatical rules, then these may also compound statements involving negation. (So if we had conjunction, we could now form conjunctions of negations, as well as negations of conjunctions).

To interpret the significance of claims involving negation, we must go beyond the syntax, to semantics. In this case, we will tell you how claims involving negation are to be used. In this case, this will amount to evlauating positions in which negations are asserted, and positions in which negations are denied. We may consider our traditional sequent rules for negation:

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

With our interpretation of sequents in mind, we can see that these rules tell us something important about the coherence of claims involving negation. The ¬L rule tells us that $[X, \neg A : Y]$ is coherent only if $[X : A, Y]$ is coherent. The ¬R rule tells us that $[X : \neg A, Y]$ is coherent only if $[X, A : Y]$ is coherent. In fact, we may show that if these two conditions are the *only* rules governing the behaviour of the operator '¬', then then we have, in some sense, *defined* the behaviour of '¬' clearly and distinctly.

The first fact to notice is that if we extend a coherence relation on the old language with these two rules, the relation on the new language still satisfies the rules of identity, weakening and strengthening. We have identity, for the new formulas: $\neg A \vdash \neg A$, since

$$\dfrac{\dfrac{A \vdash A}{\vdash \neg A, A}\,\neg R}{\neg A \vdash \neg A}\,\neg L$$

This shows us that if we have identity for $A$, we have identity for $\neg A$ too. This gives us identity for each formula consisting of one negation. We can do the same for two, three, or more negations by an inductive argument. For *weakening*, we may argue that if $X \vdash Y$, then since $X, A \vdash Y$ (by weakening for $A$) then $X \vdash \neg A, Y$. Similarly, $X \vdash A, Y$ gives us $X, \neg A, \vdash Y$, and we have weakening for statements with one negation operator. An inductive argument gives us more statements, as usual.

To show that the *other* connectives satisfy identity under the new regime requires a separate argument. Hopefully the other connectives in the old language survive this extension. No general argument can be made for this claim here.

For strengthening, we use the usual argument for the elimination of cut. If the old language does not contain any other connectives, the argument is straightforward: if we have $X \vdash \neg A, Y$ and $X, \neg A \vdash Y$, we reason in the usual manner: we take a derivation $\delta$ of $X \vdash \neg A, Y$ and a derivation $\delta'$ of $X, \neg A \vdash Y$ to consist of the steps to show each sequent from incoherent sequents in the original vocabulary, using ¬L and ¬R. These rules satisfy the conditions for a successful cut elimination argument, and we may derive $X \vdash Y$ in the usual manner in the way that we have seen.

If the old language contains other connectives, then we may need to be a little more subtle, and as a matter of fact, in some cases (where the connectives in the old vocabulary behave in an idiosyncratic manner) the cut rule will not be eliminable. In the case of a primitive vocabulary (or in the case of a vocabulary in which the other connectives are well-behaved, in a sense to be elaborated later), the addition of the operator ¬ is completely straightforward. It the new coherent relation satisfies the constraints of identity, weakening and strengthening.

For a discussion of this phenomenon, see Section 3.3.

In addition, the new connective is *uniquely defined*, in the sense that if we add two negation connectives $\neg_1$ and $\neg_2$ satisfying these rules, then we can reason as follows:

$$\dfrac{\dfrac{A \vdash A}{\vdash \neg_2 A, A}\,\neg_2 R}{\neg_1 A \vdash \neg_2 A}\,\neg_1 L \qquad \dfrac{\dfrac{A \vdash A}{\vdash \neg_1 A, A}\,\neg_1 R}{\neg_2 A \vdash \neg_1 A}\,\neg_2 L$$

and it is *never* coherent to assert $\neg_1 A$ and deny $\neg_2 A$ or *vice versa*. There is no way that they can come apart when it comes to the status of assertion and denial. The result is still a coherence relation satisfying the requirements. This means that we have a *definition*. We have not merely stipulated some rules that the new connective is to satisfy, but we have fixed its behaviour as far as coherence is concerned. If there are two connectives satisfying these rules, they are indistinguishable as far as the norms of coherence of assertion and denial are concerned. It is impossible for them todiffer in that asserting one and denying the other is never permissible.

Of course, one could assert $\neg_1 A$ without at the very same time asserting $\neg_2 A$, but the assertion of $\neg_2 A$ would have been just as good as far as coherence or incoherence is concerned.

Is this definition *available* to us? It seems that again, the answer is 'yes,' or that it is 'yes' if the original vocabulary is well enough behaved. If we can show that the cut rule is admissible in the new coherence relation (without assuming it), then it follows that the only incoherence facts in the old vocabulary are those that are given by the original coherence relation. The new vocabulary brings along with itself new facts concerning coherence, but these merely constrain the new vocabulary. They do not influence the old vocabulary. It follows that if you held that some position $[X : Y]$ (expressed in the original vocabulary) was coherent, then this is still available to you once you add the definition of '$\neg$'. In other words, the presence of negation does not force any new coherence facts on the old vocabulary. The extension of the language to include negation is conservative over the old vocabulary [6], so it may be thought of as not only a *definition* of this concept of negation (as it fixes its behaviour, at least concerning coherence) but it also may be thought of as a *pure* definition, rather than a *creative* one.

A collection of rules for a connective that allow for a unique and pure, conservative definition are said to be in *harmony* [64, 65, 87].

There is a huge literature on harmony. There are fo defences for weaker-than classical logics on grounds of harmony [25, 88, 89], and defences of *classical* logic on similar grounds [53, 72, 96], and discussions of conditions of harmony for the treatment of identity and modal operators [22, 73].

What we have done for negation, we may do for the other connectives of the language of classical logic. The rules for $\wedge$, $\vee$ and $\rightarrow$ also uniquely define the connectives, and are conservative: they are harmonious in just the same way as the rules for negation. We may think of these rules as purely and uniquely *defining* the connectives over a base vocabulary.

Once we have these connectives in our vocabulary, we may use them in characterising coherence. As you can see, the rules for negation convert a denial of $A$ to an assertion of $\neg A$, and an assertion of $A$ to a denial of $\neg A$. So, a sequent $X \vdash Y$ may be converted to an equivalent sequent in which all of the material is on one side of the turnstile. We have $X \vdash Y$ if and only if $X, \neg Y \vdash$ (where $\neg Y$ is the set of all of the statements in $Y$, with a negation prefixed). In other words, if asserting each member of $X$ and denying each member of $Y$ is out of bounds, then so is asserting each member of $X$ and asserting each member of $\neg Y$. But there is no need to privilege assertion. $X \vdash Y$ is also equivalent to $\vdash \neg X, Y$, according to which it is out of bounds to deny each member of $\neg X$ and to deny each member of $Y$.

But with the other connectives, we can go even further than this.

With conjunction, asserting each member of X is equivalent to asserting $\bigwedge X$, the conjunction of each member of X. Denying each member of Y is equivalent to denying $\bigvee Y$. So, we have $X \vdash Y$ if and only if $\bigwedge X \vdash \bigvee Y$, if and only if $\bigwedge X \wedge \neg \bigvee Y \vdash$, if and only if $\vdash \neg \bigwedge X \vee \bigvee Y$. For each position $[X : Y]$ we have a complex statment $\bigwedge X \wedge \neg \bigvee Y$ which is coherent to assert if and only if the position is coherent. We also have a complex statement $\neg \bigwedge X \vee \bigvee Y$ that is coherent to deny if and only if the position is coherent.

Choose you own favourite way of finding a conjunction of each member of a finite set. For an $n$-membered set there are at least $n!$ was of doing this.

Now we can see another reason why it is difficult, but nonetheless important, to distinguish denial and negation. Given a genuine negation connective $\neg$, the denial and the assertion of a negation are interchangeable: we can replace the denial of A with the assertion of $\neg A$ at no change to coherence. The case is completely parallel with conjunction In the presence of the conjunction connective '$\wedge$', we may replace the assertion of both A and B with the assertion of $A \wedge B$. However, there are reasons why we may want the *structural* combination A, B of separate statements alongside the *connective* combination $A \wedge B$. However, we have good reason to prefer the understanding of an argument as having more than one premise, rather than having to make do with a *conjunction* of statements as a premise. We have good reason to prefer to understand the behaviour of conjunction in terms of the validity of the rule $A, B \vdash A \wedge B$, rather than thinking of this rule as a disguised or obfuscated version of the identity sequent $A \wedge B \vdash A \wedge B$. It is genuinely informative to consider the behaviour of conjunction as related to the behaviour of taking assertions together.

This is the *fourth* consideration in favour of taking denial as prior to negation, mentioned on page 135.

If you like the jargon, you can think of conjunction as a way to 'make explicit' what is implicit in making more than one assertion. There is a benefit in allowing an explicit conjunction, in that we can then *reject* a conjunction without committing ourselves to rejecting either conjunct of that conjunction. Without the conjunction connective or something like it, we cannot do express a rejection of a pair of assertions without either rejecting one or both.

The situation with negation is parallel: we have just the same sorts of reasons to prefer to understand the behaviour of negation in terms of the validity of $A, \neg A \vdash$ and $\vdash A, \neg A$ rather than thinking of these rules as obfuscated versions of the rules of identity and cut. Just as conjunction makes explicit the combination of assertions negation makes explicit the relationship between assertions and denials. Negation allows denials to be taken up into assertoric contexts, and assertions to be taken up in the contexts of denials.

… and disjunction makes explicit the combination of denials …

## 3.3 | TONK AND NON-CONSERVATIVE EXTENSION

I have hinted at a way that this nice story can fail: the extension of a coherence relation with a definition of a connective can fail to be conservative. I will start with a discussion of a drastic case, and once we have dealt with that case, we will consider cases that are less dramatic. The drastic case involves Arthur Prior's connective 'tonk' [70]. In the

Prior's original rules were the inferences from A to A tonk B and from A tonk B to B.

context of a sequent system with sets of formulas on the left and the right of the turnstile, we can think of tonk as 'defined' by the following rules:

$$\frac{X, B \vdash Y}{X, A \text{ tonk } B \vdash Y}\ \text{tonk}L \qquad \frac{X \vdash A, Y}{X \vdash A \text{ tonk } B, Y}\ \text{tonk}R$$

This pair of rules has none of the virtues of the rules for the propositional connectives $\wedge$, $\vee$, $\neg$ and $\rightarrow$. If we add them to a coherence relation on a basic vocabulary, they are not strong enough to enable us to derive $A$ tonk $B \vdash A$ tonk $B$. The only way to derive $A$ tonk $B \vdash A$ tonk $B$ using the rules for tonk is to proceed from either $A$ tonk $B \vdash A$ (which is derivable in turn only from $B \vdash A$, in general) or from $B \vdash A$ tonk $B$ (which is derivable in turn only from $B \vdash A$, in general). So, if every sequent $A$ tonk $B \vdash A$ tonk $B$ is derivable, it is only because every sequent $B \vdash A$ is derivable. As a result, the extended relation is not a coherence relation, unless the original coherence relation is trivial in the sense of taking *every* position to be incoherent!

The problem manifests with the strengthening (cut) condition as well. In the new vocabulary we have sequents $A \vdash A$ tonk $B$ and $A$ tonk $B \vdash B$, by weakening we have $A \vdash A$ tonk $B, B$ and $A, A$ tonk $B \vdash B$, so if strengthening is available, we would have $A \vdash B$. So the only way that strengthening is available for our new coherence relation is if absolutely *every* sequent in the old vocabulary is incoherent.

This tells us that the only way that we can *inherit* a coherence relation satisfying identity and strengthening, upon the addition of the tonk rule is if it is trivial at the start. We should consider what happens if we try to impose identity and weakening by fiat. Suppose that we just mandate that the coherence relation satisfies identity and strengthening, in the presence of tonk. What happens?

This happens:

$$\dfrac{\dfrac{A \vdash A}{A \vdash A \text{ tonk } B}\ \text{tonk}L \qquad \dfrac{B \vdash B}{A \text{ tonk } B \vdash B}\ \text{tonk}R}{A \vdash B}\ \mathit{Cut}$$

we have triviality. This time as a *consequence*, rather than a *precondition* of the presence of tonk. The

Given that we use coherence relations to draw *distinctions* between positions, we are not going to wish to use tonk in our vocabulary. Thinking of the rules as constraining assertion and denial, it is clear why. The tonk$L$ rule tells us that it is incoherent to assert $A$ tonk $B$ whenever it is incoherent to assert $B$. The tonk$R$ rule tells us that is is incoherent to deny $A$ tonk $B$ whenever it is incoherent to deny $A$. Now, take a position in which it is incoherent to assert $B$ and in which is is incoherent to deny $A$. The position $[A : B]$ will do nicely. What can we do with $A$ tonk $B$? It is incoherent to assert it and to deny it. The only way this could happen is if the original position is incoherent. In other words, $[A : B]$ is incoherent. That is, $A \vdash B$. We have shown our original result in a slightly different manner: tonk is a defective connective, and the only way to use it is to collapse our coherence relation

into triviality. (This is not to deny that Roy Cook's results about logics in which tonk rules define a connective are not interesting [18]. However, they have little to do with consequence relations as defined here, as they rely on a definition of logical consequence that is essentially not *transitive*—that is, they do not satisfy *strengthening*.)

Now consider a much more interesting case of nonconservative extension. Suppose that our language contains a negationlike connective $\sim$ satisfying the following rules

$$\frac{X \vdash A}{X, \sim A \vdash}\ {\sim}L \qquad \frac{X, A \vdash}{X \vdash \sim A}\ {\sim}R$$

and that otherwise, the vocabulary contains no connectives. This restriction is for simplicity's sake—it eases the presentation, but is not an essential restriction.

This connective satisfies the rules for an *intuitionistic* negation connective. If the primitive vocabulary is governed by a coherence relation satisfying our conditions, then adding this connective will preserve those properties. We can derive $\sim A \vdash \sim A$ as follows:

$$\dfrac{\dfrac{A \vdash A}{\sim A, A \vdash}\ {\sim}L}{\sim A \vdash \sim A}\ {\sim}R$$

The *other* derivation of $\sim A \vdash \sim A$, via $\vdash \sim A, A$, is not available to us becase the position $[\ : \sim A, A]$ is coherent. We cannot show that $\sim\sim A \vdash A$. The rules as we have fixed them are not quite enough to specify the coherence relation, since there is no way to verify that the relation satisfies the condition of weakening on the new vocabulary, given that it does on the old vocabulary. The rules as they stand are not enough to guarantee that $A, \neg A \vdash Y$ for a collection $Y$ of statements from the new vocabulary. So, let us add this condition by fiat. We will say that $X \vdash Y$ holds in this coherence relation if and only if we have $X \vdash A$ for some $A \in Y$, or $X \vdash$. It is not too difficult to verify that this is a genuine coherence relation.

Starting with this relation, the addition of $\neg$ is no longer conservative. If we add the usual rules for $\neg$, we have the following two derivations.

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}\ \neg L}{\neg A \vdash \sim A}\ {\sim}R}{\sim\sim A, \neg A \vdash}\ {\sim}L}{\sim\sim A \vdash \neg\neg A}\ \neg R \qquad \dfrac{\dfrac{A \vdash A}{\vdash A, \neg A}\ \neg R}{\neg\neg A \vdash A}\ \neg L$$

It follows that the new collection of rules either no longer satisfies strengthening, or the new system provides us with a derivation of $\sim\sim A \vdash A$ that we did not have before.

This example shows us that conservative extension, by itself, does not give us a hard-and-fast criterion for choosing between different rules. The classical vocabulary counts as a conservative extension over other vocabulary if all of the connectives are well-behaved, and the cut-elimination theorem holds. This requires not only that the new connectives are well-behaved and harmonious, but also that the *old* connectives behave properly. In this case, it is the rule for $\sim$ that messes up the argument. To eliminate cuts, we must show that if our cut formula is *passive* in some other rule, then we could have performed the cut on the *premises* of the rule, and not the conclusion. In this case, this is not possible. Suppose we have a cut-formula as passive in the rule for $\sim$.

$$\dfrac{\vdash B, C \qquad \dfrac{X, B, A \vdash}{X, B \vdash \sim A}\,{\sim}L}{X \vdash \sim A, C}\,Cut$$

There is no way to get to $X \vdash \sim A, C$ by performing the cut before introducing the $\sim$, since the premise sequent would have to be $X, A \vdash C$, which is no longer appropriate for the application of a the $\sim L$ rule.

$$\dfrac{\dfrac{\vdash B, C \quad X, B, A \vdash}{X, A \vdash C}\,Cut}{X \vdash \sim A, C}\,{\sim}L??$$

From the persepctive of natural deduction, the constraint on $\sim L$ is global and not merely local. If we happen to have a refutation of $X, B, A$, then one step is composing (discharging) the premise $A$ with a $\sim I$ node to derive $\sim A$. Another step is composing the proof with a proof with the two conclusions $B, C$ to add a conclusion $C$ but discharge the premise $B$. From the point of view of circuits, there is no sense in which one of these is done 'before' or 'after' the other.

**todo:** add a picture.

There seems to be a real sense in which the rules for the classical connectives naturally fit the choice of sequent structure. Once we have settled on the 'home' for our connectives — in this case, sequents of the form $X \vdash Y$ — we need to work hard to get *less* than classical logic.

**todo:** expand this cryptic remark.

## 3.4 | MEANING

The idea that the rules of inference confer *meaning* on the logical connectives is a compelling one. What can we say about this idea? We have shown that we can introduce the logical constants into a practice of asserting and denial in such a way that assertions and denials featuring those connectives can be governed for coherence along with the original assertions and denials. We do not have to say anything more about the circumstances in which a negation or a conjunction or a disjunction is *true* or *false*, or to find any thing to which the connectives '*correspond*.' What does this tell us about the meanings of the items that we have introduced?

There are a number of things one could want in a theory of meanings of logical connectives. (1) You could want something that connected meanings to *use*. (2) You could want something that connected meanings to *truth conditions*. (3) You could want something that explains meaning in terms of the proprieties of *translation*. Let us take these considerations in turn.

RULES AND USE: The account of the connectives given here, as operators *introduced* by means of well-behaved rules concerning coherence, gives a clear connection to the way that these connectives are used. To be sure, it is not a *descriptive* account of the use of the connectives. Instead, it is a *normative* account of the use of the connectives, giving us a particular normative category (coherence) with which to judge the practice of assertions and denials in that vocabulary. The rules for the connectives are intimately connected to use in this way. If giving an account of the meaning of a fragment of our vocabulary involves giving a normative account of the proprieties of the use of that vocabulary, the rules for the connectives can at the very least be viewed as a part of the larger story of the meaning of that vocabulary.

RULES AND TRUTH CONDITIONS: Another influential strand (no, let me be honest—it is the overwhelmingly *dominant tradition*) in semantics takes the controlling category in giving a semantic theory for some part of the language to be given in the *truth conditions* for statements expressed in that part of the vocabulary. To be completely concrete, the truth conditional account of the meaning of conjunction goes something like this:

*A conjunction* $A \wedge B$ *is true iff* $A$ *is true and* $B$ *is true.*

More complex accounts say more, by replacing talk of 'true' by 'true relative to a model' or 'true in a world' or 'true given an assignment of values to variables' or any number of embellishments on the general picture.

This looks nothing like the account we have given in terms of rules governing coherence. What is the connection between truth conditions, meaning and the rules for the connectives?

First it must be made clear that whatever we are to make the connection between truth conditions and the inference rules will depend crucially on what we make of the concept of truth. The concept of truth is subtle and difficult. I am not thinking primarily of the philosophical debates over the proper way to analyse the concept and to connect it to concepts of correspondence, coherence, or whatever else. The subtlety of the notion of truth comes from the paradoxes concerning truth. It would be nice to think of truth as the simple expressive concept according to which an assertion of the claim that $A$ is true had no more nor less significance than an assertion of the claim $A$. The liar paradox and its like have put paid to that. Whatever we can say about the concept of truth and the proper analysis of the liar paradox, it is not *simple*, whatever option we take about the paradox.

I am well aware of accounts that attempt to keep the simple expressive correspondence between a claim $A$ and the claim that $A$ is true, at the cost of giving a non-classical account of the logical connectives [68, 74, 75]. This non-classical account has many virtues. *Simplicity* of the concept of truth is one of them. However, the *simplicity* of the overall picture is not so obvious [29].

All of these considerations mean that an extended and judicious discussion of truth and its properties must wait for some more logical sophistication, once we have an account of predicates, names, identity and quantification under our belts. For now, we must make do with a few gestures in the general direction of a more comprehensive account. So, consider simple instances of our rules governing conjunction:

$$A, B \vdash A \wedge B \qquad A \wedge B \vdash A \qquad A \wedge B \vdash B$$

These do not say that a conjunction is true if and only if the conjuncts are both true, but they come very *close* to doing so, given that we are not yet using the truth predicate. These rules tell us that it is never permissible to assert both conjuncts, and to deny the conjunction. This is not expressed in terms of truth conditions, but for many purposes it will have the same consequences. Furthermore, using the completeness proofs of Section ???, they tell us that there is no model in which A and B are satisfied and $A \wedge B$ is not; and that there is no model in which $A \wedge B$ is satisfied, and in which A or B is not. If we think of satisfaction in a model as a model for *truth*, then the truth-conditional account of the meaning of connectives is a consequence of the rules. We may think of a the reification or idealisation of a coherent position (as given to us in the completeness proofs) as a model for what is *true*. We do not need to reject truth-conditional accounts of the meanings of the connectives. They are consequences of the definitions that we have given. Whether or not we take this reading of the truth-conditional account as satisfying or not will depend, of course, on what we need the concept of truth to *do*. We will examine this in more detail later.

**Well, I need to write that bit up.**

RULES AND TRANSLATION: What can we say about the connection between the rules for connectives and how we interpret the assertions and denials of others? Here are some elementary truisms: (a) people do not have to endorse the rules I use for negation for me to take them to mean negation by 'not.' It does not seem that we settle every question of translation by looking at these rules. Nonethless, (b) we can use these rules as a way of making meaning more precise. We can clarify meaning by proposing and adopting rules for connectives. Not every use of 'and' fits the rules for '$\wedge$'. Adopting a precisely delineated coherence relation for an item aids communication, when it is achievable. The rules we have seen are a very good way to be precise about the behaviour of the connectives. (c) Most cruically, what we say about translation depends on the status of claims of coherence. If I take what someone says to be assertoric, I relate what they say and what I am committed to in the one coherence relation. I keep score by keeping track of what (by my lights) you are saying. And you do this for me. I use a coherence relation to keep score of your judgements, and you do the same for me. There can be disputes over the relation: you can take some position to be coherent that I do not. This is made *explicit* by our logical vocabulary. If you and I agree about the rules for classical connectives, then if we disagree over whether or not $X, A \vdash B, Y$

is coherent, then we disagree (in the context $[X : Y]$) over the coherence of asserting $A \to B$. Connectives are a way of articulating this disagreement. Similarly, $\neg$ is a way of making explicit incompatibility judgements of the form $A, B \vdash$ or exhaustiveness judgements of the form $\vdash A, B$.

## 3.5 | ACHILLES AND THE TORTOISE

has gone wrong in the case of the reasoner who is quite prepared to assert $A$ and to assert $A \to B$, but who questions or hesitates over $B$? By no means is this a problem. This is well discussed [Smiley's paper is a very good example, as is Harman.] But it's worth explaining ths story in this vocabulary. Consider the person who endorses $A$, and $B$, and even who endorses $(A \wedge B) \to Z$. Is there a way to force her to endorse $Z$? Of course not. By our lights she is already *committed* to $Z$ in that it is a consequence of what she is already committed to in the position mentioned. However, there is no way to compel her to endorse it.

## 3.6 | WARRANT

Discussion of warrant preservation in proof. In what way is classical inference apt for preservation of warrant, and what is the sense in which intuitionstic logic is appropriate. [Preservation warrant in the case of an argument from $X$ to $A$. These are *verificationally transparent* arguments. Preservation of diswarrant in the case of an argument from $B$ to $Y$. These are the *falsificationally transparent* arguments. Both are nice. But neither is enough.]

## 3.7 | GAPS AND GLUTS

Now, I have asked you to conisder adding to your vocabulary, a connective $\neg$ satisfying the rules for negation. This is actually a very strong requirement. Gaps and gluts, supervaluationism and subvaluationism: a discussion of how the setting here helps explain the significance of these approaches.

## 3.8 | REALISM

Discussion of the status of models and truth this account. Are we realists? Are these merely useful tools, or something more? (Discussion of Blackburn's quasi-realism and its difficulties with logic here. The Frege/Geach problem is discussed at this point, if not before.)

# PART II

## QUANTIFIERS, IDENTITY AND EXISTENCE

# QUANTIFIERS: TOOLS & TECHNIQUES

# 4

## 4.1 | PREDICATE LOGIC

### 4.1.1 | RULES

### 4.1.2 | WHAT DO THE RULES MEAN?

Mark Lance, makes the point in his paper "Quantification, Substitution, and Conceptual Content" [48] that an inference-first account of quantifiers is sensible, but it isn't the kind of "substitutional" account oft mentioned. If we take the meaning of $(\forall x)A$ to be given by what one might infer *from* it, and use to infer *to* it, then one can infer from it to each of its instances (and furthermore, to any other instances we might get as we might further add to the language). What one might use to infer *to* $(\forall x)A$ is not just each of the instances $A[x/n_1]$, $A[x/n_2]$, etc. (Though that might work in some restricted circumstances, clearly it is unwieldy at best and *wrong-headed* at worst.) The idea behind the *rules* is that to infer *to* $(\forall x)A$ you need not just an instance (or all instances). You need something else. You need to have derived an instance in a general way. That is, you need to have a derivation of $A[x/n]$ that applies independently of any information "about $n$." (It would be nice to make some comment about how this sidesteps all of the talk about "general facts" in arguments about what you need to know to know that $(\forall x)A$ apart from knowing that $A[x/n]$ for each particular $n$, but to make that case I would need more space and time than I have at hand at present.)

## 4.2 | IDENTITY AND EXISTENCE

A section on the addition of identity, discussing options for adding an existence predicate if we countenance "non-denoting terms."

$$\frac{\phi(a) \quad a = b}{\phi(b)}\,{=}E \qquad \frac{\begin{array}{c}[Xa]^i \\ \vdots \\ Xb\end{array}}{a = b}\,{=}I,i$$

The crucial side-condition in =I is that the predicate variable $X$ does not appear elsewhere in the premises. In other words, if an *arbitrary* property holds of $a$, it must hold of $b$ too. The only way for this to hold is, of course, for $a$ to be identical to $b$. Here is a proof using these

rules.

$$\dfrac{\dfrac{\dfrac{[Xa]^1 \quad a = b}{Xb}\,{=}E \qquad b = c}{Xc}\,{=}E}{a = c}\,{=}I,1$$

The sequent rules are these:

$$\dfrac{\Gamma \vdash \phi(a), \Delta \quad \Gamma', \phi(b) \vdash \Delta'}{\Gamma, \Gamma', a = b \vdash \Delta, \Delta'}\,{=}L \qquad \dfrac{\Gamma, Xa \vdash Xb, \Delta}{\Gamma \vdash a = b, \Delta}\,{=}R$$

The side condition in $=R$ is that $X$ does not appear in $\Gamma$ or $\Delta$.

Notice that we need to modify the subformula property further, since the predicate variable does not appear in the conclusion of $=R$, and more severely, the predicate $\phi$ does not appear in the conclusion of $=L$ Here is an example derivation:

$$\dfrac{\dfrac{\dfrac{Xa \vdash Xa}{\vdash \neg Xa, Xa}\,\neg R \qquad \dfrac{Xb \vdash Xb}{Xb, \neg Xb \vdash}\,\neg L}{a = b, Xb \vdash Xa}\,{=}L}{a = b \vdash b = a}\,{=}R$$

If we have a cut on a formula $a = b$ which is active in both premises of that rule:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_X\\ \Gamma, Xa \vdash Xb, \Delta\end{array}}{\Gamma \vdash a = b, \Delta}\,{=}R \qquad \dfrac{\begin{array}{c}\vdots\,\delta'\\ \Gamma' \vdash \phi(a), \Delta'\end{array} \qquad \begin{array}{c}\vdots\,\delta''\\ \Gamma'', \phi(b) \vdash \Delta''\end{array}}{\Gamma', \Gamma'', a = b \vdash \Delta', \Delta''}\,{=}L}{\Gamma, \Gamma', \Gamma'' \vdash \Delta, \Delta', \Delta''}\,Cut$$

we can eliminate it in favour of two cuts on the formulas $\phi(a)$ and $\phi(b)$. To do this, we modify the derivation $\delta_X$ to conclude $\Gamma, \phi(a) \vdash \phi(b), \Delta$, which we can by globally replacing $Xx$ by $\phi(x)$. The result is still a derivation. We call it $\delta_\phi$. Then we may reason as follows:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_\phi\\ \Gamma, \phi(a) \vdash \phi(b), \Delta\end{array} \qquad \begin{array}{c}\vdots\,\delta'\\ \Gamma' \vdash \phi(a), \Delta'\end{array}}{\Gamma, \Gamma' \vdash \phi(b), \Delta, \Delta'}\,Cut \qquad \begin{array}{c}\vdots\,\delta''\\ \Gamma'', \phi(b) \vdash \Delta''\end{array}}{\Gamma, \Gamma', \Gamma'' \vdash \Delta, \Delta', \Delta''}\,Cut$$

## 4.3 | MODELS

Traditional Tarski models for classical logic and Kripke intuitionistic logic motivated on the basis of the proof rules we have introduced. A presentation of more "modest" finitist semantics in which the domain is finite at each stage of evaluation, given by the sequent system. A context of evaluation, in this kind of model, is a finite entity, including information about "how to go on."

## 4.4 | ARITHMETIC

Peano and Heyting arithmetic are introduced as a simple example of a rigorous system with enough complexity to be truly interesting. Discussion of the consistency proof for arithmetic. I will point to Gödel's incompleteness results, and show how pa + Con(pa) can be seen as adding to the stock of arithmetic inference principles, in the same way that adding stronger induction principles does. [32]

## 4.5 | SECOND ORDER QUANTIFICATION

Second order quantification introduced, and Girard/Tait normalisation for second order logic.

# QUANTIFIERS: APPLICATIONS 5

## 5.1 | OBJECTIVITY

Substitutional and objectual quantification and objectivity. The account of quantification given here isn't first-and-foremost objectual in the usual sense, but it can be seen as a semantically anti-realist (that is, not truth-first) reading of standard, objectual quantification. A defence of this analysis, and the a discussion of the sense in which this provides properly universal quantification, independently of any consideration of whether the class of "everything" is a set or can constitute a model.

## 5.2 | EXPLANATION

How do we prove a universal claim? By deriving it. Explanation of the reasons why people like "universal facts" and why this is better understood in terms prior to commitment to fact-like entities.

## 5.3 | RELATIVITY

A discussion of ontological relativity, Quine's criterion for commitment to objects. (We discuss the sense in which logic alone does not force the existence of any number of things, and why the choice of ontology depends on the behaviour of names and variables in your theory.)

## 5.4 | EXISTENCE

A discussion of a neo-Carnapian view that to adopt inference principles concerning numbers, say, is *free*. Relating to current discussion of structuralism, plenitudinous platonism and fictionalism in mathematics

## 5.5 | CONSISTENCY

The essential incompleteness and extendibility of our inference principles.

## 5.6 | SECOND ORDER LOGIC

A discussion of the "standard model" of second order logic, and its interpretation as an "ideal" endpoint for language expansion. (Treating

the range of the quantifiers in second order logic as an ideal endpoint of conceptual expansion.) A discussion of why standard Second Order Logic, so construed, is essentially non-axiomatisable.

# PART III

## MODALITY AND TRUTH

# 6 MODALITY AND TRUTH: TOOLS & TECHNIQUES

**6.1** | SIMPLE MODAL LOGIC

**6.2** | MODAL MODELS

**6.3** | QUANTIFIED MODAL LOGIC

**6.4** | TRUTH AS A PREDICATE

# MODALITY AND TRUTH: APPLICATIONS

# 7

**7.1** | POSSIBLE WORLDS

**7.2** | COUNTERPARTS

**7.3** | SYNTHETIC NECESSITIES

**7.4** | TWO DIMENSIONAL SEMANTICS

**7.5** | TRUTH AND PARADOX

# REFERENCES


[1] alan ross anderson and nuel d. belnap. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, Princeton, 1975.

[2] alan ross anderson, nuel d. belnap, and j. michael dunn. *Entailment: The Logic of Relevance and Necessity*, volume 2. Princeton University Press, Princeton, 1992.

[3] 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.

[4] jc beall and greg restall. "Logical Pluralism". *Australasian Journal of Philosophy*, 78:475–493, 2000.

[5] jc beall and greg restall. *Logical Pluralism*. Oxford University Press, Oxford, 2006.

[6] nuel d. belnap. "Tonk, Plonk and Plink". *Analysis*, 22:130–134, 1962.

[7] nuel d. belnap. "Display Logic". *Journal of Philosophical Logic*, 11:375–417, 1982.

[8] richard blute, j. r. b. cockett, r. a. g. seely, and t. h. trimble. "Natural Deduction and Coherence for Weakly Distributive Categories". *Journal of Pure and Applied Algebra*, 13(3):229–296, 1996. Available from ftp://triples.math.mcgill.ca/pub/rags/nets/nets.ps.gz.

[9] david bostock. *Intermediate Logic*. Oxford University Press, 1997.

[10] robert b. brandom. *Making It Explicit*. Harvard University Press, 1994.

[11] robert b. brandom. *Articulating Reasons: an introduction to inferentialism*. Harvard University Press, 2000.

[12] l. e. j. brouwer. "On the Significance of the Principle of Excluded Middle in Mathematics, Especially in Function Theory". In *From Frege to Gödel: a a source book in mathematical logic, 1879–1931*, pages 334–345. Harvard University Press, Cambridge, Mass., 1967. Originally published in 1923.

[13] samuel r. buss. "The Undecidability of k-provability". *Annals of Pure and Applied Logic*, 53:72–102, 1991.

[14] a. carbone. "Interpolants, Cut Elimination and Flow graphs for the Propositional Calculus". *Annals of Pure and Applied Logic*, 83:249–299, 1997.


This bibliography is also available online at http://citeulike.org/user/greg_restall/tag/ptp. citeulike.org is an interesting collaborative site for sharing pointers to the academic literature.

[15] A. CARBONE. "Duplication of directed graphs and exponential blow up of proofs". *Annals of Pure and Applied Logic*, 100:1–67, 1999.

[16] BOB CARPENTER. *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science 32. Cambridge University Press, 1992.

[17] ALONZO CHURCH. *The Calculi of Lambda-Conversion*. Number 6 in Annals of Mathematical Studies. Princeton University Press, 1941.

[18] ROY T. COOK. "What's wrong with tonk(?)". *Journal of Philosophical Logic*, 34:217–226, 2005.

[19] HASKELL B. CURRY. *Foundations of Mathematical Logic*. Dover, 1977. Originally published in 1963.

[20] DIRK VAN DALEN. "Intuitionistic Logic". In DOV M. GABBAY AND FRANZ GÜNTHNER, editors, *Handbook of Philosophical Logic*, volume III. Reidel, Dordrecht, 1986.

[21] VINCENT DANOS AND LAURENT REGNIER. "The Structure of Multiplicatives". *Archive of Mathematical Logic*, 28:181–203, 1989.

[22] NICHOLAS DENYER. "The Principle of Harmony". *Analysis*, 49:21–22, 1989.

[23] KOSTA DOŠEN. "The First Axiomatisation of Relevant Logic". *Journal of Philosophical Logic*, 21:339–356, 1992.

[24] KOSTA DOŠEN. "A Historical Introduction to Substructural Logics". In PETER SCHROEDER-HEISTER AND KOSTA DOŠEN, editors, *Substructural Logics*. Oxford University Press, 1993.

[25] MICHAEL DUMMETT. *The Logical Basis of Metaphysics*. Harvard University Press, 1991.

[26] J. MICHAEL DUNN. "Relevance Logic and Entailment". In DOV M. GABBAY AND FRANZ GÜNTHNER, editors, *Handbook of Philosophical Logic*, volume 3, pages 117–229. Reidel, Dordrecht, 1986.

[27] 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.

[28] ROY DYCKHOFF. "Contraction-Free Sequent Calculi for Intuitionistic Logic". *Journal of Symbolic Logic*, 57:795–807, 1992.

[29] HARTRY FIELD. "Saving the Truth Schema Paradox". *Journal of Philosophical Logic*, 31:1–27, 2002.

[30] KIT FINE. "Vaguness, Truth and Logic". *Synthese*, 30:265–300, 1975. Reprinted in *Vagueness: A Reader* [46].

[31] F. B. FITCH. *Symbolic Logic*. Roland Press, New York, 1952.

[32] TORKEL FRANZÉN. *Inexhaustibility: A Non-Exhaustive Treatment*, volume 16 of *Lecture Notes in Logic*. Association for Symbolic Logic, 2004.

[33] gerhard gentzen. "Untersuchungen über das logische Schliessen". *Math. Zeitschrift*, 39:176–210 and 405–431, 1934. Translated in *The Collected Papers of Gerhard Gentzen* [34].

[34] gerhard gentzen. *The Collected Papers of Gerhard Gentzen*. North Holland, 1969. Edited by M. E. Szabo.

[35] jean-yves girard. "Linear Logic". *Theoretical Computer Science*, 50:1–101, 1987.

[36] jean-yves girard. *Proof Theory and Logical Complexity*. Bibliopolis, Naples, 1987.

[37] jean-yves girard, yves lafont, and paul taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.

[38] ian hacking. "What is Logic?". *The Journal of Philosophy*, 76:285–319, 1979.

[39] chris hankin. *Lambda Calculi: A Guide for Computer Scientists*, volume 3 of *Graduate Texts in Computer Science*. Oxford University Press, 1994.

[40] gilbert harman. *Change In View: Principles of Reasoning*. Bradford Books. MIT Press, 1986.

[41] jean van heijenoort. *From Frege to Gödel: a a source book in mathematical logic, 1879–1931*. Harvard University Press, Cambridge, Mass., 1967.

[42] arend heyting. *Intuitionism: An Introduction*. North Holland, Amsterdam, 1956.

[43] 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.

[44] colin howson. *Logic with Trees: An introduction to symbolic logic*. Routledge, 1996.

[45] rolf isermann. "On Fuzzy Logic Applications for Automatic Control, Supervision, and Fault Diagnosis". *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 28:221–235, 1998.

[46] rosanna keefe and peter smith. *Vagueness: A Reader*. Bradford Books. MIT Press, 1997.

[47] william kneale and martha kneale. *The Development of Logic*. Oxford University Press, 1962.

[48] mark lance. "Quantification, Substitution, and Conceptual Content". *Noûs*, 30(4):481–507, 1996.

[49] e. j. lemmon. *Beginning Logic*. Nelson, 1965.

[50] PAOLA MANCOSU. *From Brouwer to Hilbert*. Oxford University Press, 1998.

[51] EDWIN D. MARES. *Relevant Logic: A Philosophical Interpretation*. Cambridge University Press, 2004.

[52] VANN MCGEE. *Truth, Vagueness and Paradox*. Hackett Publishing Company, Indianapolis, 1991.

[53] PETER MILNE. "Classical Harmony: rules of inference and the meaning of the logical constants". *Synthese*, 100:49–94, 1994.

[54] MICHIEL MOORTGAT. *Categorial Investigations: Logical Aspects of the Lambek Calculus*. Foris, Dordrecht, 1988.

[55] GLYN MORRILL. *Type Logical Grammar: Categorial Logic of Signs*. Kluwer, Dordrecht, 1994.

[56] SARA NEGRI AND JAN VON PLATO. *Structural Proof Theory*. Cambridge University Press, Cambridge, 2001. With an appendix by Aarne Ranta.

[57] I. E. ORLOV. "The Calculus of Compatibility of Propositions (in Russian)". *Matematicheskiĭ Sbornik*, 35:263–286, 1928.

[58] VICTOR PAMBUCCIAN. "Early Examples of Resource-Consciousness". *Studia Logica*, 77:81–86, 2004.

[59] FRANCESCO PAOLI. *Substructural Logics: A Primer*. Trends in Logic. Springer, May 2002.

[60] TERENCE PARSONS. "Assertion, Denial, and the Liar Paradox". *Journal of Philosophical Logic*, 13:137–152, 1984.

[61] TERENCE PARSONS. "True Contradictions". *Canadian Journal of Philosophy*, 20:335–354, 1990.

[62] FRANCIS J. PELLETIER. "A Brief History of Natural Deduction". *History and Philosophy of Logic*, 20:1–31, March 1999.

[63] DAG PRAWITZ. *Natural Deduction: A Proof Theoretical Study*. Almqvist and Wiksell, Stockholm, 1965.

[64] DAG PRAWITZ. "Towards a Foundation of General Proof Theory". In PATRICK SUPPES, LEON HENKIN, ATHANASE JOJA, AND GR. C. MOISIL, editors, *Logic, Methodology and Philosophy of Science IV*. North Holland, Amsterdam, 1973.

[65] DAG PRAWITZ. "Proofs and the Meaning and Completeness of the Logical Constants". In E. SAARINEN J. HINTIKKA, I. NIINILUOTO, editor, *Essays on Mathematical and Philosophical Logic*, pages 25–40. D. Reidel, 1979.

[66] GRAHAM PRIEST. "The Logic of Paradox". *Journal of Philosophical Logic*, 8:219–241, 1979.

[67] GRAHAM PRIEST. "Inconsistencies in Motion". *American Philosophical Quarterly*, 22:339–345, 1985.

[68] graham priest. *In Contradiction: A Study of the Transconsistent*. Martinus Nijhoff, The Hague, 1987.

[69] graham priest, richard sylvan, and jean norman, editors. *Paraconsistent Logic: Essays on the Inconsistent*. Philosophia Verlag, 1989.

[70] arthur n. prior. "The Runabout Inference-Ticket". *Analysis*, 21:38–39, 1960.

[71] stephen read. *Relevant Logic*. Basil Blackwell, Oxford, 1988.

[72] stephen read. "Harmony and Autonomy in Classical Logic". *Journal of Philosophical Logic*, 29:123–154, 2000.

[73] stephen read. "Identity and Harmony". *Analysis*, 64(2):113–115, 2004.

[74] greg restall. "Deviant Logic and the Paradoxes of Self Reference". *Philosophical Studies*, 70:279–303, 1993.

[75] greg restall. *On Logics Without Contraction*. PhD thesis, The University of Queensland, January 1994. Available at http://consequently.org/writing/onlogics.

[76] greg restall. *An Introduction to Substructural Logics*. Routledge, 2000.

[77] greg restall. "Carnap's Tolerance, Meaning and Logical Pluralism". *Journal of Philosophy*, 99:426–443, 2002.

[78] greg restall. *Logic*. Fundamentals of Philosohphy. Routledge, 2006.

[79] fred richman. "Intuitionism as Generalization". *Philosophia Mathematica*, 5:124–128, 1990.

[80] edmund robinson. "Proof Nets for Classical Logic". *Journal of Logic and Computation*, 13(5):777–797, 2003.

[81] richard routley, val plumwood, robert k. meyer, and ross t. brady. *Relevant Logics and their Rivals*. Ridgeview, 1982.

[82] 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* [41].

[83] 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.

[84] d. j. shoesmith and t. j. smiley. *Multiple Conclusion Logic*. Cambridge University Press, Cambridge, 1978.

[85] r. m. smullyan. *First-Order Logic*. Springer-Verlag, Berlin, 1968. Reprinted by Dover Press, 1995.

[86] w. w. tait. "Intensional Interpretation of Functionals of Finite Type I". *Journal of Symbolic Logic*, 32:198–212, 1967.

[87] NEIL TENNANT. *Natural Logic*. Edinburgh University Press, Edinburgh, 1978.

[88] NEIL TENNANT. *Anti-Realism and Logic: Truth as Eternal*. Clarendon Library of Logic and Philosophy. Oxford University Press, 1987.

[89] NEIL TENNANT. *The Taming of the True*. Clarendon Press, Oxford, 1997.

[90] A. S. TROELSTRA. *Lectures on Linear Logic*. CSLI Publications, 1992.

[91] 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.

[92] A. M. UNGAR. *Normalization, cut-elimination, and the theory of proofs*. Number 28 in CSLI Lecture Notes. CSLI Publications, Stanford, 1992.

[93] IGOR URBAS. "Dual-Intuitionistic Logic". *Notre Dame Journal of Formal Logic*, 37:440–451, 1996.

[94] ALASDAIR URQUHART. "Semantics for Relevant Logics". *Journal of Symbolic Logic*, 37:159–169, 1972.

[95] ACHILLÉ VARZI. *An Essay in Universal Semantics*, volume 1 of *Topoi Library*. Kluwer Academic Publishers, Dordrecht, Boston and London, 1999.

[96] ALAN WEIR. "Classical Harmony". *Notre Dame Journal of Formal Logic*, 27(4):459–482, 1986.

[97] J. ELDON WHITESITT. *Boolean algebra and its applications*. Addison–Wesley Pub. Co., Reading, Mass., 1961.

[98] EDWARD N. ZALTA. "Gottlob Frege". In EDWARD N. ZALTA, editor, *Stanford Encyclopedia of Philosophy*. Stanford University, 2005.