Week 5

Notes on natural deduction for minimal logic

Instructor: Matteo de Ceglie

02 April 2020

Contents

Now that we have an idea of how propositional logic works, from both the syntactic and the semantics point of view, we can move on and start learning about one of the main tools in logic/philosophical logic/philosophy of logic: natural deduction. This method is not special for a particular logic, but is common to most logical systems.

The various logical systems are defined by the rules we use to define it. The minimal logic possible is the one described in this notes, and it is the most simple (i.e. the one that has less rules than all the others and thus less proving power). Intuitionistic propositional logic will be define by all these rules, plus another. Adding another rule will give us a classical propositional logic system. Finally, adding the rules for quantifiers will give us a full classical first order natural deduction system (adding the rules for quantifiers without adding the rule for classical logic will give us an intuitionistic first order logic).

In this lecture we will focus only on the minimal logic system. Next class (after Easter) will be instead devoted to the classical propositional logic (that is actually easier than intuitionistic logic, that wll not be covered). At the end of the course we will move to the classical first order logic full system.

1 Introduction

Natural deduction is a *syntactic* method developed *to prove a certain formula* from a set of *assumptions*. It is not the only method available (the other being the tableaux method, Gentzen's sequent calculus, Hilbert style proof system), and there are also some variants. In this course we will focus only on the *tree form natural deduction*, both for propositional and (later) first order logic. All the definitions and rules that I will give you for propositional logic will be just the same for first order logic, so it is quite important that you understand them well. Then, when we will focus on first order logic, I will remind you of these rules and give you some more. For starters, this class and the next class will be devoted to natural deduction for propositional logic, so we can go through everything without going to quickly. In particular, this class will be devoted mostly to introduce the system and gives the rules and definitions, while next week we will mostly see examples on how to use them.

First of all, why it is called "natural deduction"? Remember from the first class on arguments that the main focus of logic is to study *deductive* arguments. So this system is trying to abstract that kind of arguments, i.e. arguments in which the acceptance of the conclusion is related to the acceptance of the premises/assumptions. Moreover, it is called "natural" because tries to mimic how we argue in natural language, using very similar strategies and rules. For example, take the following argument:

(1)	Either apples, bananas and pears are all vegetables or else they are all fruit.
(2)	But it's not the case that apples, bananas and pears are all vegetables.
(3)	They are all fruit.

It is easy to show that such an argument is valid. However, suppose we are presented not with its natural language form, but with the equivalent translation in propositional logic:

$$((((a \wedge b) \wedge c) \vee ((d \wedge e) \wedge f)) \wedge \neg((a \wedge b) \wedge c)) \rightarrow ((d \wedge e) \wedge f)$$

Is this formula a tautology? To check it, we will do a truth table, but this truth table would be quite massive, and a chore to check line by line. Thus, that mechanical method is not *efficient* enough to check these massive cases, and we need something else. What we need is something that is closer to the kind of reasoning we used in the natural language case, where we recognise that the following pattern:

(1)	$((a \wedge b) \wedge c) \vee ((d \wedge e) \wedge f))$
(2)	$\neg((a \wedge b) \wedge c))$
(3)	$((d \wedge e) \wedge f)$

is the form of a valid argument, and this settles the question. This is way the natural deduction system is called as such: it is first of all a *deduction*, just like the ones we studied with the natural language arguments; secondly, it is *natural*, since it mimics as closely as possible the way we reason in natural language. In natural deduction we write proofs of formulas just like we arranged the arguments in the natural language, trying to reach a conclusion that is inevitable from the premises. Performing a natural deduction proof is to arrange a chain of arguments in such a way that accepting all of them compels us to accept the conclusion.

Before proceeding forward, a final word on the name of the system we'll be learning. I have said before that this system is called "tree form" natural deduction. This is because the way we arrange the proof will resemble (in an abstract way) a tree. Also, our definitions will take advantage of this. This is not the only way to perform natural deduction (for example, linear style or Fitch style), however this is the most "user friendly".

2 Definition of the system

To define such a system we need the following building blocks:

- a formal definition of what is an *assumptions*;
- a set of rules, called *derivation* rules or *inference* rules, that we will use to perform the proofs;
- a formal definition of what a derivation (our proof) is.

These definitions and rules will help us in arranging the proof and evaluate it.

2.1 Assumptions

First of all we start with what an assumption is. In layman terms, an assumption is something we do not prove, but we only take as truth “for the sake of the argument”. In our natural language arguments, the premises are assumptions. These are accepted as true without any justification or proof. When we do this, we *discharge* the assumptions, i.e. we “use” that assumptions for that conclusion, and we cannot use them again. However, if the conclusion of an argument is then used as a premise in another argument, then *it is not* an assumption, since it was proved from the previous argument.

Definition 1 (Assumption)**.** An *assumption* is any formula in the topmost position of any branch. We call the set of all these assumptions Γ.

This definition clearly takes advantage of the fact that our derivations will be arranges in a tree form, in such a way that our assumptions will be easily spotted.

Definition 2 (Discharging of an assumption)**.** An assumptions is *discharged* when it is used in the application of a derivation rule, and we write $[\varphi]^n$ to signal it, where the n simply enumerate all the assumptions while they are being discharged.

2.2 Derivations

So what is a derivation? A derivation is just a proof of a formula φ from some assumptions Γ. In other words, we lay down our assumptions, then we apply the inference rules until we end up with our goal formula φ as the conclusion. In the meantime, we pay attention in discharging all the assumptions as needed.

Definition 3 (Derivation)**.** A *derivation* of a formula φ from the set of assumptions Γ, written $\Gamma \vdash \varphi$, is a tree of formulas satisfying the following conditions:

1. the topmost formulas of the tree are either in Γ (i.e. are assumptions) or are discharged by an inference rule in the tree;
2. the bottommost formula of the tree is φ;
3. Every formula in the tree, except the formula φ at the bottom, is a premise of a correct application of an inference rule whose conclusion stands directly below that formula in the tree.

A natural deduction derivation $\Gamma \vdash \varphi$ would then look like this:

$$\dfrac{\dfrac{\dfrac{\rho}{\chi}\langle \mathrm{DR}\rangle \quad \dfrac{\psi}{\kappa}\langle \mathrm{DR}\rangle}{\vartheta}\langle \mathrm{DR}\rangle}{\varphi}\langle \mathrm{DR}\rangle$$

where $\langle \mathrm{DR}\rangle$ describe which inference rule is applied at that step. In this case, the assumptions are ρ and ψ, i.e. $\Gamma = \{\rho, \psi\}$. We can also write $\rho, \psi \vdash \varphi$.

When we discharge *all* of our assumptions in the derivation, then the set of assumption is said to be empty. In that case, we have derived our formula φ with no assumptions pending, and write $\vdash \varphi$. When this is not possible, we write $\nvdash \varphi$. Please not that it could be the case that φ cannot be proved without assumptions, but can still be proved with some of them. In other words, $\nvdash \varphi$ does not imply $\Gamma \nvdash \varphi$.

2.3 Inference rules

We now need to define these derivation rules. If you remember what were the forms of an argument, this is going to be easy, since all the derivation rules highly resemble those forms of *valid* arguments. However, since we are doing all of this in propositional logic, we are dealing here if symbols, and not natural language. We can divide all the derivation rules in two groups:

Introduction Rules Rules that *introduce* a new connective;

Elimination Rules Rulers that *eliminate* a connective.

All the derivation rules will have some assumptions, and some of them will also discharge them (not all of them).

2.3.1 Conjunction introduction

The first derivation rule we will define is the *conjunction introduction.* With this rule we will add a conjunction to our derivation, in a way that resembles the following argument:

(1) Socrates is a philosopher.
(2) Socrates is Greek.

---

(3) Socrates is a philosopher and he is Greek.

Such an argument is valid, and so are all the ones that shares its argument form:

(1) A
(2) B

---

(3) $A \wedge B$

The natural deduction derivation rule $\langle \wedge \mathrm{I} \rangle$ is very similar to that argument form:

$$\frac{\varphi \quad \psi}{\varphi \wedge \psi} \langle \wedge \mathrm{I} \rangle$$

2.3.2 Conjunction Elimination

We then have two other rules regarding conjunction, but these are for the *elimination* of it from the derivation. Let's consider the following arguments:

(1) Socrates is Greek and a philosopher.

---

(2) Socrates is Greek.

(1) Socrates is Greek and a philosopher.

---

(2) Socrates is a philosopher.

Both of them are valid, and they share an argument form that makes them so:

(1) $A \wedge B$

---

(2) A

(1) $A \wedge B$

---

(2) B

And just like in the case of the introduction rule, from these forms of argument we have two derivation rules:

$$\frac{\varphi \wedge \psi}{\varphi} \langle \wedge \mathrm{E} \rangle \qquad \frac{\varphi \wedge \psi}{\psi} \langle \wedge \mathrm{E} \rangle$$

2.3.3 A very minimal example

With only these 3 rules, we can start working on a minimal example of a derivation. Let's say that we want to prove that "Socrates is a dolphin and a philosopher". We can only use the rules already introduced, so we can start like this:

(1) Socrates is Greek and a philosopher.

---

(2) Socrates is a philosopher.

So now we have half of our objective. We can now assume that Socrates is a dolphin and mortal:

(1)	Socrates is a dolphin and a mortal.
(2)	Socrates is a dolphin.

And now we can put the two conclusions together to form the final part of our argument:

(1)	Socrates is a philosopher.
(2)	Socrates is a dolphin.
(3)	Socrates is a philosopher and a dolphin.

We can even put this in a tree form:

$$\dfrac{\dfrac{\text{Socrates is Greek and a philosopher.}}{\text{Socrates is a philosopher.}} \quad \dfrac{\text{Socrates is a dolphin and a mortal.}}{\text{Socrates is a dolphin.}}}{\text{Socrates is a philosopher and a dolphin.}}$$

If we consider the argument forms, this looks very similar to a natural deduction derivation:

$$\dfrac{\dfrac{A \wedge B}{B} \quad \dfrac{C \wedge D}{C}}{B \wedge C}$$

Finally, the following would be the derivation of $\psi \wedge \vartheta$ from assumptions $\varphi \wedge \psi$ and $\vartheta \wedge \chi$:

$$\dfrac{\dfrac{\varphi \wedge \psi}{\psi}\langle\wedge\mathrm{E}\rangle \quad \dfrac{\vartheta \wedge \chi}{\vartheta}\langle\wedge\mathrm{E}\rangle}{\psi \wedge \vartheta}\langle\wedge\mathrm{I}\rangle$$

2.3.4 Disjunction Introduction

The rules to introduce a disjunction in a derivation are very similar to the conjunction elimination. The following arguments are trivially valid:

(1)	Socrates is a philosopher.
(2)	Socrates is a philosopher or a dolphin.

(1)	Socrates is a dolphin.
(2)	Socrates is a philosopher or a dolphin.

since if I am stating that something is true, then stating that the same thing or something else is true still follows. The argument form is the following:

(1)	A
(2)	$A \vee B$

(1)	B
(2)	$A \vee B$

And then the derivations rules looks very similar:

$$\dfrac{\varphi}{\varphi \vee \psi}\langle\vee\mathrm{I}\rangle \qquad \dfrac{\psi}{\varphi \vee \psi}\langle\vee\mathrm{I}\rangle$$

2.3.5 Disjunction Elimination

This rule is the first one in which the assumptions are discharged. This rule aims to eliminate a disjunction from the derivation. Since we cannot directly state one of the two disjuncts (because we cannot directly which one we should choose) like in the conjunction case, we have to take an indirect root. The general idea is that if from two assumptions I can derive a conclusion, and then I have also the disjunction of those two assumptions, then I can definitely state that conclusion. At his point, I can discharge the first two assumptions.

Suppose I have the following two arguments:

(1)	Socrates is a philosopher.	(1)	Socrates is Greek.
(2)	...	(2)	...
(3)	Socrates is a dolphin.	(3)	Socrates is a dolphin.

Note that at this stage it doesn't matter how we go from the first premise to the conclusion. It could be another premise that makes everything valid or another argument. Let's just assume that we have those steps. Then suppose we have the statement "Socrates is a philosopher or is Greek". At this point, we can conclude that, in fact, "Socrates is a dolphin". In tree form, this would be:

$$\dfrac{\text{Socrates is a philosopher or Greek.} \qquad \dfrac{\begin{array}{c}\underline{[\text{Socrates is a philosopher.}]^1}\\ \vdots\end{array}}{\text{Socrates is a dolphin.}} \qquad \dfrac{\begin{array}{c}\underline{[\text{Socrates is Greek.}]^1}\\ \vdots\end{array}}{\text{Socrates is a dolphin}}}{\text{Socrates is a dolphin.}}$$

The idea behind such a proof is that we don't know if Socrates is a philosopher or Greek. We only know that in either case he is a dolphin. Since we also know that he is a philosopher or Greek, then we can dismiss our disjunction, and state that he is totally a dolphin. The form of this argument is the following:

$$\dfrac{A \vee B \qquad \dfrac{\begin{array}{c}\underline{[A]^1}\\ \vdots\end{array}}{C} \qquad \dfrac{\begin{array}{c}\underline{[B]^1}\\ \vdots\end{array}}{C}}{C}$$

and the actual derivation rule:

$$\dfrac{\varphi \vee \psi \qquad \dfrac{\begin{array}{c}\underline{[\varphi]^n}\\ \vdots\end{array}}{\chi} \qquad \dfrac{\begin{array}{c}\underline{[\psi]^n}\\ \vdots\end{array}}{\chi}}{\chi}\ \langle\vee\mathrm{E}\rangle$$

In stating this rule some steps were omitted, namely the derivations of χ from φ and ψ. While this is fine for stating the rule "in a vacuum", when trying to prove something in natural deduction this *cannot be done.* Indeed, *every step must be included, always.*

2.3.6 Implication Introduction

The rule to introduce an implication in our derivation is quite intuitive. If we can derive a formula ψ from another formula φ, then we can deduce that $\varphi \to \psi$. This is also true in the natural deduction arguments. For example if I have the following argument:

(1)	Socrates is a philosopher.
(2)	...
(3)	Socrates is a dolphin.

then I can say that "Socrates is a philosopher" *implies* "Socrates is a dolphin", or "If Socrates is a philosopher, then he is a dolphin":

(1)	Socrates is a philosopher.
(2)	...
(3)	Socrates is a dolphin.
(4)	If Socrates is a philosopher, then he is a dolphin

The form of this kind of arguments is the following, and renders them all valid:

(1)	A
(2)	...
(3)	B
(4)	$A \to B$

Finally, this is the derivation rule:When that is the case, we write $\vdash \varphi$.

$$\dfrac{\dfrac{\dfrac{\varphi}{\vdots}}{\psi}}{\varphi \to \psi}\ \langle \to\text{I} \rangle$$

In this rule like in the disjunction elimination there is an omission. Just like in that case, this cannot be done in a full derivation.

2.3.7 Implication Elimination

This rule is the *modus ponens.* Most of the valid arguments that we have seen in this course use this kind of rule. For example:

(1)	If Socrates is a man, then he is mortal.
(2)	Socrates is a man.
(3)	Socrates is mortal.

And the argument form behind its validity is very simple and well-known:

(1)	$A \to B$
(2)	A
(3)	$B.$

The derivation rule is basically the same:

$$\dfrac{\varphi \to \psi \quad \varphi}{\psi}\ \langle \to\text{E} \rangle$$

2.3.8 Negation Rules

The last two rules needed to define this minimal logic natural deduction system are the negation introduction and elimination rules. The reason these rules are the last to be presented is because are the most important: the various systems (minimal logic, intuitionistic propositional logic, classical propositional logic) are differentiated by these rules. Minimal logic, the system we are currently defining, has only two negation rules (one for introduction and one for elimination). Intuitionistic propositional logic adds to those rules two others: one for contradiction and another elimination rule (double negation elimination). Finally, the classical propositional logic adds to the intuitionistic rules also the full *reductio ad absurdum.* For now, we stick with the minimal logic system.

Before proceeding, we need to add a new symbol to our notation. Indeed, it is very helpful if we can have a symbol for *contradiction.* The symbol we will be using is the following: $\bot$. This is shorthand for $\varphi \land \neg\varphi$, but writing $\bot$ helps with out derivations.

The first rule we are going to introduce is the negation introduction rule. This rules allows us to add a negation to our derivation, deducing it from the fact that the non-negated formula produced a contradiction. For example, let's say that we want to prove that "Socrates it is not a dolphin". To do that, we may start by assuming that Socrates, is, in fact, a dolphin, then derive from it a contradiction (for example, the fact that he does not live in the sea and since we are assuming he is a dolphin he does live in the sea). Then, as the last step, we can negate our first assumption and deduce that Socrates is not, in fact, a dolphin. This would, in natural language, be a *reductio ad absurdum.* However, in formal logic we have two different rules: this one starts with a non-negated formula, while the regular *reductio ad absurdum* starts with a negated formula (incidentally, the latter is way more powerful). The following is the rule for introducing a negation in minimal logic:

$$\cfrac{\cfrac{\cfrac{[\varphi]^n}{\vdots}}{\bot}}{\neg\varphi}\ \langle\neg\mathrm{I}\rangle$$

The last rule we are defining for our minimal logic system is the negation elimination rule. This simple states that if we have a contradiction in our derivation (for example "Socrates is a dolphin" and "Socrates is not a dolphin", then we can delete both formulas and write a contradiction:

$$\frac{\varphi \quad \neg\varphi}{\bot}\ \langle\neg\mathrm{E}\rangle$$

This allows us, for example, to then apply the negation introduction rule to add a negation to a previous formula.

3 A fully fledged example

It is now time to apply all the rules that we have learned to try and produce a complete natural deduction derivation.

Suppose we want to derive $\neg(\varphi \land \neg\psi)$ from the assumption $\varphi \to \psi$. This can be written as $\varphi \to \psi \vdash \neg(\varphi \land \neg\psi)$. How can we proceed?

First of all we note that our last step will have to involve a negation introduction, so we can start by writing this:

$$\frac{}{\neg(\varphi \land \neg\psi)}\ \langle\neg\mathrm{I}\rangle$$

Then, remembering the negation introduction rule, we can add some more steps:

$$\cfrac{\cfrac{\cfrac{[\varphi \land \neg\psi]^1}{\vdots}}{\bot}}{\neg(\varphi \land \neg\psi)}\ \langle\neg\mathrm{I}\rangle$$

Note that we already now that, to correctly apply the rule, we have to discharge our assumption. Now we have to delete those omitting dots and replace them with an actual derivation (remember that we said that we cannot use them in a real derivation, only in the rules definitions). We start by noting that we have only one rule do add the contradiction $\bot$, so we can add a new step:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi \wedge \neg\psi]^1}{\vdots}}{\psi} \qquad \dfrac{\dfrac{[\varphi \wedge \neg\psi]^1}{\vdots}}{\neg\psi}}{\bot}\langle\neg\mathrm{E}\rangle}{\neg(\varphi \wedge \neg\psi)}\langle\neg\mathrm{I}\rangle$$

Here we have put ψ and $\neg\psi$ as the two formulas that allow us to put the contradiction, and not two random formulas. This is because we know not only our final step, but also, “caused” by our final step, our first assumptions. So we aim to stay as closed as possible to those formulas, to avoid an overly complex derivation. For the same reason, we could also use a completely different assumption for one of the branches (now we have two of them). However, this would make everything more complex, so it is better to continue using what we already have.

We still have to replace those omitting dots, so we start looking at the right branch. Here we need a way to derive $\neg\psi$ from $\varphi \wedge \neg\psi$. Since it is a negation, we can apply the same strategy as before, i.e. using the negation introduction rule, and before that the negation elimination rule to get the contradiction etc. etc. But this strategy would not help us (even if it may still be possible), because it doesn’t help us in closing that gap between $\varphi \wedge \neg\psi$ and $\neg\psi$. However, we can use more simply a conjunction elimination rule:

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi \wedge \neg\psi]^1}{\vdots}}{\psi} \qquad \dfrac{[\varphi \wedge \neg\psi]^1}{\neg\psi}\langle\wedge\mathrm{E}\rangle}{\bot}\langle\neg\mathrm{E}\rangle}{\neg(\varphi \wedge \neg\psi)}\langle\neg\mathrm{I}\rangle$$

Remember that the conjunction elimination rule does not need a discharge of the assumptions, so we can apply it even if that assumption has already been discharged.

Now focus on the left branch. Here we have to derive ψ from the assumption $\varphi \wedge \neg\psi$. We cannot use a direct step from the conjunction, so we have to add some more steps. Also, we still have to use our assumption $\varphi \rightarrow \psi$. Notice that with the application of a *modus ponens* we could extract the needed ψ from our assumption. So we add it to our derivation:

$$\dfrac{\dfrac{\langle\rightarrow\mathrm{E}\rangle\dfrac{\varphi \wedge \psi \qquad \dfrac{\dfrac{[\varphi \wedge \neg\psi]^1}{\vdots}}{\varphi}}{\psi} \qquad \dfrac{[\varphi \wedge \neg\psi]^1}{\neg\psi}\langle\wedge\mathrm{E}\rangle}{\bot}\langle\neg\mathrm{E}\rangle}{\neg(\varphi \wedge \neg\psi)}\langle\neg\mathrm{I}\rangle$$

And finally, we add our last step, another simple conjunction elimination to get our φ:

$$\dfrac{\dfrac{\langle\rightarrow\mathrm{E}\rangle\dfrac{\varphi \rightarrow \psi \qquad \langle\wedge\mathrm{E}\rangle\dfrac{[\varphi \wedge \neg\psi]^1}{\varphi}}{\psi} \qquad \dfrac{[\varphi \wedge \neg\psi]^1}{\neg\psi}\langle\wedge\mathrm{E}\rangle}{\bot}\langle\neg\mathrm{E}\rangle}{\neg(\varphi \wedge \neg\psi)}\langle\neg\mathrm{I}\rangle$$

This ends our derivation of $\neg(\varphi \wedge \neg\psi)$ from $\varphi \rightarrow \psi$. The reason why this is not a derivation of $\neg(\varphi \wedge \neg\psi)$ is because we didn’t discharged our last assumption, so our set of assumption is not empty.