Natural Deduction

Paul J.E. Dekker

June, 2019

This document provides an outline of the proof-theoretic method of *natural deduction*.

Natural Deduction The method of natural deduction draws from a set of formally specified rules that constrain and possibly guide the *derivation* of conclusions from (possibly empty) series of premises, or assumptions.

Derivations A derivation is a (i) numbered and (ii) annotated (iii) list of (iv) formulas.

(iv) The formulas are for instance propositions from propositional logic.
(iii) The propositions are most conveniently listed vertically.
(ii) The annotations indicate the justification for the annotated propositions.
(i) The numbers allow one to refer back to the propositions so obtained.

Which lists constitute derivations is constrained by *rules* which are specified below.

Rules The rules of deduction tell one what proposition one may conclude on a line in a derivation, mostly on the basis of what one already has established above that line.

- The rules have a label X, which must annotate any proposition so justified.
- If the rules require precedent material, the annotations must unambiguously specify the lines where that has been established.
- Caveat: the precedently required material must still be available, and may not consist of propositions that are, or are derived from, assumptions that have been withdrawn.

Example The following frame presents a partly schematic snapshot of a derivation.

1.	...	
$\vdots$	$\vdots$	
17.	$((p \vee q) \wedge (q \rightarrow r))$	$[I_\wedge, 6, 13]$
$\vdots$	$\vdots$	
n.	...	

The highlighted proposition is the 17-th formula in the framed derivation. It is said to be justified by "$I_\wedge$, 6, 13", which means by a rule named "$I_\wedge$" and with reference to the propositions appearing on the lines 6 en 13 in the list. The rule so-named (cf. below) requires line 6 to host the proposition $(p \vee q)$ and line 13 to host the proposition $(q \rightarrow r)$.

Assumptions One rule is that assumptions need no justification.

- One may *always* make *whatever* assumption.
- On any line you may enter any arbitrary formula.
- But you have to specify that it is an *assumption*.
- This is done by annotating it with "[ass.]".

Withdrawing Assumptions Certain rules may involve the withdrawal of assumptions.

- They summarize the result of hypothetical reasoning. [E.g., that if something is shown to entail the exclusion of everything, then that something is excluded.]
- The assumption that is withdrawn is always the last pending assumption in the preceding list, i.e., the last assumption that has not yet been withdrawn.
- Anything that has been deduced under an assumption that has been withdrawn, cannot be used anymore once the assumption is withdrawn.
- It is therefore good and common practice to bracket out withdrawn material, i.e., the lines including the withdrawn assumption up to, but not including, the line where the assumption is withdrawn.

Example The following frame shows the expansion of a derivation on the left with the withdrawal on line n. of the assumption that r on line i.

$$
\begin{array}{lll}
 & \vdots & \\
\text{i.} & r & [\text{ass.}] \\
 & \vdots & \\
\text{n-1.} & s &
\end{array}
\quad \Rightarrow \quad
\begin{array}{lll}
 & \vdots & \\
\text{i.} & r & [\text{ass.}] \\
 & \vdots & \\
\text{n-1.} & s & \\
\hline
\text{n.} & (r \to s) & [I_{\to}]
\end{array}
$$

Deductions

- If one has a derivation, one may list (horizontally) the pending assumptions, those that have not been withdrawn, and these then count as your *premises*—e.g. $\phi_1, \ldots, \phi_n$.
- Behind this one may write what is on the last line, as the *conclusion*—e.g. ψ.
- One may put '$\vdash$' in between, as in

$$\phi_1, \ldots, \phi_n \vdash \psi.$$

- This last notation actually means:

  There is a derivation of (the conclusion) ψ from (the premises) $\phi_1, \ldots, \phi_n$.

  It means that ϕ can be validly inferred from $\phi_1, \ldots, \phi_n$, i.e., that the conclusion follows from the premises, according to the specified set of rules.
- Interestingly, and trivially, $\phi \vdash \phi$, for any proposition ϕ, since

$$1. \quad \phi \quad [\text{ass.}]$$

  is a derivation.

Propositional Logic

- The remainder of this document presents deduction rules for a language of propositional logic, and a number of sample deductions.
- A language ($\mathcal{L}_{\mathcal{P}}$) of propositional logic is constructed in the usual way from a set ($\mathcal{P}$) of proposition letters ($p, q, r, \ldots, p_1, p_2, \ldots$) using logical constants '$\bot$', '$\neg$', '$\wedge$', '$\vee$' and '$\to$' and auxiliary devices '(' and ')'.
- The following Bachus-Naur style definition specifies what is a formula ϕ of $\mathcal{L}_{\mathcal{P}}$:

$$\boxed{\phi ::= p \mid \bot \mid \neg\phi \mid (\phi \wedge \phi) \mid (\phi \vee \phi) \mid (\phi \to \phi) \qquad [p \in \mathcal{P}]}$$

Rules for Propositional Logic

The derivation rules for propositional logic are displayed in a schematic form, the application of which will be explained where needed. Notice that the set of rules presented here is somewhat redundant, but it is easy to work with.

Assumption	
$\vdots$	
n. ϕ	[ass.]

[Any assumption can be made any time.]

This rule should be read as licensing one under any list ('$\vdots$') to state any assumption ('ϕ') as the next (n-th) item on the list.

Reiteration	
$\vdots$	
m. ϕ	A
$\vdots$	
n. ϕ	[reit., m]

[Any proposition previously established can be reused.]

The rule licenses one to conclude ϕ at line n if ϕ has already been established before. Note that the proposition ϕ on line m may as a rule not be, or depend on, a withdrawn assumption.

Conjunction Elimination (l)	
$\vdots$	
m. $(\phi \wedge \psi)$	A
$\vdots$	
n. ϕ	[$\mathrm{E}_{\wedge_l}$, m]

Conjunction Elimination (r)	
$\vdots$	
m. $(\phi \wedge \psi)$	A
$\vdots$	
n. ψ	[$\mathrm{E}_{\wedge_r}$, m]

If a conjunction has been established before, say at some previous line m, then either conjunct ('left' or 'right') can be inferred from it on the next line n. Of course, the previously established conjunction may not be or depend on assumptions that are withdrawn.

Conjunction Introduction	
$\vdots$	
m. ϕ	A
$\vdots$	
m′. ψ	B
$\vdots$	
n. $(\phi \wedge \psi)$	[$\mathrm{I}_\wedge$, m, m′]

One may derive the conjunction $(\phi \wedge \psi)$ of any two previously established propositions ϕ and ψ. Again, the previously established propositions may not be or depend on assumptions that are withdrawn. (I will henceforth refrain from making this constraint explicit.)

Implication Elimination

$$
\begin{array}{lll}
 & \vdots & \\
\text{m.} & (\phi \rightarrow \psi) & \\
 & \vdots & \\
\text{m}'. & \phi & \\
 & \vdots & \\
\text{n.} & \psi & [\text{E}_{\rightarrow}, \text{m}, \text{m}']
\end{array}
$$

Implication Introduction

$$
\begin{array}{lll}
 & \vdots & \\
\text{m.} & \phi & [\text{ass.}] \\
 & \vdots & \\
\text{n-1.} & \psi & \\
\hline
\text{n.} & (\phi \rightarrow \psi) & [\text{I}_{\rightarrow}]
\end{array}
$$

An implication $(\phi \rightarrow \psi)$ established at some line m can be used if its antecedent ϕ has been established as well, say, at line m$'$, and licenses one to conclude to its consequent ψ, at line n. This rule is generally known as 'Modus Ponens', and often taken to be the heart of the propositional logical inference engine. How can we conclude to an implication? This can be achieved after hypothetically assuming its antecedent ϕ on some line m, and then, in the context of that proposition (i.e., in the context of the lines above m) validly derive some conclusion ψ, and then—withdrawing that assumption, and everything based on it—concluding that $(\phi \rightarrow \psi)$, in that very same context again.

Disjunction Elimination

$$
\begin{array}{lll}
 & \vdots & \\
\text{m.} & (\phi \vee \psi) & \\
 & \vdots & \\
\text{m}'. & (\phi \rightarrow \chi) & \\
 & \vdots & \\
\text{m}''. & (\psi \rightarrow \chi) & \\
 & \vdots & \\
\text{n.} & \chi & [\text{E}_{\vee}, \text{m}, \text{m}', \text{m}'']
\end{array}
$$

If we have established a disjunction $(\phi \vee \psi)$ on some previous line m, and perhaps we don't know which of either one of the two disjuncts (or perhaps even both of them), we can nevertheless draw some conclusion χ at line n, provided that we have in addition established that either one of the two disjuncts implies that conclusion χ, on some earlier lines m$'$ and m$''$.

Disjunction Introduction (l)

$$
\begin{array}{lll}
 & \vdots & \\
\text{m.} & \phi & \\
 & \vdots & \\
\text{n.} & (\phi \vee \psi) & [\text{I}_{\vee_l}, \text{m}]
\end{array}
$$

Disjunction Introduction (r)

$$
\begin{array}{lll}
 & \vdots & \\
\text{m.} & \psi & \\
 & \vdots & \\
\text{n.} & (\phi \vee \psi) & [\text{I}_{\vee_r}, \text{m}]
\end{array}
$$

Any one of two disjuncts entails the disjunction of them, because the disjunction is satisfied by either one of them—no matter which one of them, or even perhaps both.

Negation Elimination

$$\begin{array}{lll} & \vdots & \\ \text{m.} & \neg\phi & \\ & \vdots & \\ \text{m}'. & \phi & \\ & \vdots & \\ \text{n.} & \bot & [\text{E}_\neg, \text{m}, \text{m}'] \end{array}$$

Negation Introduction

$$\begin{array}{lll} & \vdots & \\ \text{m.} & \phi & [\text{ass.}] \\ & \vdots & \\ \text{n-1.} & \bot & \\ \hline \text{n.} & \neg\phi & [\text{I}_\neg] \end{array}$$

The hallmark of a negation $\neg\phi$ is that it excludes that ϕ, so if we establish both, say on some lines m and m′, respectively, then we have reached a dead end, marked by the falsum (‘$\bot$’) in the negation elimination rule. If the falsum (‘$\bot$’) marks the dead end of a line of hypothetical reasoning, we may conclude to the negation of the assumption that it is built on, as the negation introduction rule tells us.

The last rule effectively completes our set of rules and turns it into a classical logical system.

Double Negation

$$\begin{array}{lll} & \vdots & \\ \text{m.} & \neg\neg\phi & \\ & \vdots & \\ \text{n.} & \phi & [\text{E}_{\neg\neg}, \text{m}] \end{array}$$

If it is excluded that ϕ is excluded, as $\neg\neg\phi$ says, we cannot but agree to accept it, even if we fail a direct proof of it. (For the latter reason the rule is not universally agreed upon.)

Reading a Proof

The following is a theorem of propositional logic.

$$(\neg p \to q) \vdash (\neg q \to p)$$

To prove it is to present a derivation of the conclusion $(\neg q \to p)$ from the premise $(\neg p \to q)$. One derivation is presented here, unfolding it comics-strip style.

Assume

1.	$(\neg p \to q)$	[ass.]

Assume

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]

Assume

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]

Use the Implication

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]
4.	q	[$\mathrm{E}_\to$, 1, 3]

Use the Negation

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]
4.	q	[$\mathrm{E}_\to$, 1, 3]
5.	$\bot$	[$\mathrm{E}_\neg$, 2, 4]

Introduce a Negation

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]
4.	q	[$\mathrm{E}_\to$, 1, 3]
5.	$\bot$	[$\mathrm{E}_\neg$, 2, 4]
6.	$\neg\neg p$	[$\mathrm{I}_\neg$]

Eliminate Double Negation

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]
4.	q	[$\mathrm{E}_\to$, 1, 3]
5.	$\bot$	[$\mathrm{E}_\neg$, 2, 4]
6.	$\neg\neg p$	[$\mathrm{I}_\neg$]
7.	p	[$\neg\neg$, 6]

Introduce an Implication

1.	$(\neg p \to q)$	[ass.]
2.	$\neg q$	[ass.]
3.	$\neg p$	[ass.]
4.	q	[$\mathrm{E}_\to$, 1, 3]
5.	$\bot$	[$\mathrm{E}_\neg$, 2, 4]
6.	$\neg\neg p$	[$\mathrm{I}_\neg$]
7.	p	[$\neg\neg$, 6]
8.	$(\neg q \to p)$	[$\mathrm{I}_\to$]

Summarize what has been shown

$(\neg p \to q)$	$\vdash$	$(\neg q \to p)$
(assumption	entails	conclusion)

Constructing a Proof

In the construction of a proof the derivation normally does not evolve step-wise. It is usually guided by the attempt to make two extremes meet: (a) what is given and (b) what is to be shown. If a certain inference is to be proven valid, what is given at the start are the premises of the inference, and what is to be shown is the conclusion. If it has to be shown that, e.g., $\phi, \psi \vdash \chi$, then a proof-agent's initial state is one of wondering how to fill the gap between the premises ϕ and ψ and the conclusion ψ, as is displayed by the following aporetic snaphsot of a derivation under construction.

1.	ϕ	[ass.]
2.	ψ	[ass.]
	?	
n.	χ	[?]

If the given (specified in green on the lines above '?') and the goal (specified in red on the line below '?') match the format of one of the inference rules, then the agent is basically done. She deletes the question line, adjusts the number (sets the number n to 3 above) and supplies the annotation by adding the rule in charge. If the agent is not basically done she may go one of three ways.

1. **Use** and expand the given.
2. **Settle** a required goal.
3. Go **Absurd**.

Adopting option 1. the agent uses what she has, and *expands the list of givens* above '?' with what the *elimination rules* tell her what she can use them for. Adopting option 2. the agent attempts to *settle a new goal* below '?', viz., a proposition required to license the original goal according to the *introduction rule* governing its main logical constant. These two options can be intertwined. To use a given disjunction $(\phi \vee \psi)$ in order to derive a goal χ, one may have to settle two additionally required goals: $(\phi \rightarrow \chi)$ and $(\psi \rightarrow \chi)$. If an implication $(\phi \rightarrow \psi)$ (or a negation $\neg\phi$) is *given*, one may have to adopt a subsidiary and temporary *new goal* of establishing the antecedent (or negated proposition) ϕ, in order to be able to *use* the corresponding elimination rule. Likewise, in case the goal is an implication $(\phi \rightarrow \psi)$ (or negation $\neg\phi$), *the given is expanded* with the antecedent (or negated proposition) ϕ, and *the new goal* is to derive the consequent (ψ or the falsum $\perp$).

If the iterated and intertwined application of options 1. and 2. does not produce a match between the given and the goal, then one may have to resort to the Aristotelian option 3: to show that the negation of the goal produces a dead end. The agent then takes the negation of one's original goal as a given and proceeds to show that it will lead to contradiction (the falsum). All of these options are demonstrated in the following, constructively unfolding, proofs of some theorems.

Consider the intuitively simple theorem:

$$(p \vee (q \wedge r)) \vdash (p \vee r)$$

The initial aporetic state is displayed on the the left below, and invites us to try and employ the given disjunction as displayed on the right.

Initial State.

1.	$(p \vee (q \wedge r))$	[ass.]
	?	
n.	$(p \vee r)$	[?]

Attempt to Use Disjunction 1.

1.	$(p \vee (q \wedge r))$	[ass.]
	?	
i.	$(p \rightarrow (p \vee r))$	[?]
	?	
j.	$((q \wedge r) \rightarrow (p \vee r))$	[?]
n.	$(p \vee r)$	[E$_\vee$, 1, i, j]

(Experience teaches us that there are no prospects in trying to employ the introduction rule for the goal $(p \vee r)$.) From here we proceed as displayed below, without any further comments.

Attempt to Settle Implications i and j.

1.	$(p \vee (q \wedge r))$	[ass.]
2.	p	[ass.]
	?	
i-1.	$(p \vee r)$	[?]
i.	$(p \rightarrow (p \vee r))$	[I$_\rightarrow$]
i+1.	$(q \wedge r)$	[ass.]
	?	
j-1.	$(p \vee r)$	[?]
j.	$((q \wedge r) \rightarrow (p \vee r))$	[I$_\rightarrow$]
n.	$(p \vee r)$	[E$_\vee$, 1, i, j]

Match: Settle Disjunction i-1 = 3.

1.	$(p \vee (q \wedge r))$	[ass.]
2.	p	[ass.]
3.	$(p \vee r)$	[I$_\vee$, 2]
4.	$(p \rightarrow (p \vee r))$	[I$_\rightarrow$]
5.	$(q \wedge r)$	[ass.]
	?	
j-1.	$(p \vee r)$	[?]
j.	$((q \wedge r) \rightarrow (p \vee r))$	[I$_\rightarrow$]
n.	$(p \vee r)$	[E$_\vee$, 1, 4, j]

Use Conjunction 5.

1.	$(p \vee (q \wedge r))$	[ass.]
2.	p	[ass.]
3.	$(p \vee r)$	[I$_\vee$, 2]
4.	$(p \rightarrow (p \vee r))$	[I$_\rightarrow$]
5.	$(q \wedge r)$	[ass.]
6.	r	[E$_\wedge$, 5]
	?	
j-1.	$(p \vee r)$	[?]
j.	$((q \wedge r) \rightarrow (p \vee r))$	[I$_\rightarrow$]
n.	$(p \vee r)$	[E$_\vee$, 1, 4, j]

Match: Settle Disjunction j-1 = 7.

1.	$(p \vee (q \wedge r))$	[ass.]
2.	p	[ass.]
3.	$(p \vee r)$	[I$_\vee$, 2]
4.	$(p \rightarrow (p \vee r))$	[I$_\rightarrow$]
5.	$(q \wedge r)$	[ass.]
6.	r	[E$_\wedge$, 5]
7.	$(p \vee r)$	[I$_\vee$, 6]
8.	$((q \wedge r) \rightarrow (p \vee r))$	[I$_\rightarrow$]
9.	$(p \vee r)$	[E$_\vee$, 1, 4, 5]

Consider the somewhat more involved theorem:

$$(((\neg p \to p) \to \neg p) \vdash \neg p$$

The initial aporetic state is displayed on the the left below, and invites us to try and introduce the negation that is the goal, as is displayed on the right.

Initial State.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
	?	
n.	$\neg p$	[?]

Attempt to Settle Negation n.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
2.	p	[ass.]
	?	
n-1.	$\bot$	
n.	$\neg p$	[$\mathrm{I}_\neg$]

We obtain the falsum if we can obtain something contradicting something given, e.g., $\neg p$. Note that, significantly, the task of deriving $\neg p$ on line n-2 is different from that of deriving $\neg p$ on line n, because something more is given, viz., line 2. From there we proceed again without any further comments.

Attempt to Settle Falsum n-1.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
2.	p	[ass.]
	?	
n-2.	$\neg p$	[?]
n-1.	$\bot$	[$\mathrm{E}_\neg$, n-2, 2]
n.	$\neg p$	[$\mathrm{I}_\neg$]

Attempt to Use Implication 1.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
2.	p	[ass.]
	?	
n-3.	$(\neg p \to p)$	[?]
n-2.	$\neg p$	[$\mathrm{E}_\to$, 1, n-3]
n-1.	$\bot$	[$\mathrm{E}_\neg$, n-2, 2]
n.	$\neg p$	[$\mathrm{I}_\neg$]

Attempt to Settle Implication n-3.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
2.	p	[ass.]
3.	$\neg p$	[ass.]
	?	
n-4.	p	[?]
n-3.	$(\neg p \to p)$	[$\mathrm{I}_\to$]
n-2.	$\neg p$	[$\mathrm{E}_\to$, 1, n-3]
n-1.	$\bot$	[$\mathrm{E}_\neg$, n-2, 2]
n.	$\neg p$	[$\mathrm{I}_\neg$]

Match: Reiteration n-4 = 4 = 2.

1.	$((\neg p \to p) \to \neg p)$	[ass.]
2.	p	[ass.]
3.	$\neg p$	[ass.]
4.	p	[reit., 2]
5.	$(\neg p \to p)$	[$\mathrm{I}_\to$]
6.	$\neg p$	[$\mathrm{E}_\to$, 1, n-3]
8.	$\bot$	[$\mathrm{E}_\neg$, n-2, 2]
8.	$\neg p$	[$\mathrm{I}_\neg$]

Here is classical theorem, and a major challenge for beginning proof theoreists.

$$\vdash (p \vee \neg p)$$

Initial State.

	?	
n.	$(p \vee \neg p)$	[?]

Attempt to Go Absurd.

1.	$\neg(p \vee \neg p)$	[ass.]
	?	
n-1.	$\perp$	
n.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]

Attempt to Use Negation 1.

1.	$\neg(p \vee \neg p)$	[ass.]
	?	
n-2.	$(p \vee \neg p)$	[?]
n-1.	$\perp$	[$\mathrm{E}_\neg$, 1, n-2]
n.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]

Attempt to Settle Disjunction n-2

1.	$\neg(p \vee \neg p)$	[ass.]
	?	
n-3.	p	[?]
n-2.	$(p \vee \neg p)$	[$\mathrm{I}_\vee$, n-3]
n-1.	$\perp$	[$\mathrm{E}_\neg$, 1, n-2]
n.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]

Attempt to Go Absurd

1.	$\neg(p \vee \neg p)$	[ass.]
2.	$\neg p$	[ass.]
	?	
n-4	$\perp$	[?]
n-3.	p	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]
n-2.	$(p \vee \neg p)$	[$\mathrm{I}_\vee$, n-3]
n-1.	$\perp$	[$\mathrm{E}_\neg$, 1, n-2]
n.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]

Attempt to Use Negation 1.

1.	$\neg(p \vee \neg p)$	[ass.]
2.	$\neg p$	[ass.]
	?	
n-5	$(p \vee \neg p)$	[?]
n-4	$\perp$	[$\mathrm{E}_\neg$, 1, n-5]
n-3.	p	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]
n-2.	$(p \vee \neg p)$	[$\mathrm{I}_\vee$, n-3]
n-1.	$\perp$	[$\mathrm{E}_\neg$, 1, n-2]
n.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]

Match: Settle Disjunction n-5 = 3.

1.	$\neg(p \vee \neg p)$	[ass.]
2.	$\neg p$	[ass.]
3	$(p \vee \neg p)$	[$\mathrm{I}_\vee$, 2]
4	$\perp$	[$\mathrm{E}_\neg$, 1, n-5]
5.	p	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]
6.	$(p \vee \neg p)$	[$\mathrm{I}_\vee$, n-3]
7.	$\perp$	[$\mathrm{E}_\neg$, 1, n-2]
8.	$(p \vee \neg p)$	[$\mathrm{I}_\neg$ + $\mathrm{E}_{\neg\neg}$]