1.4 Natural Deduction

In the preceding sections we have adopted the view that propositional logic is based on truth tables, i.e. we have looked at logic from a semantical point of view. This, however, is not the only possible point of view. If one thinks of logic as a codification of (exact) reasoning, then it should stay close to the practice of inference making, instead of basing itself on the notion of truth. We will now explore the non-semantic approach, by setting up a system for deriving conclusions from premises. Although this approach is of a formal nature, i.e. it abstains from interpreting the statements and rules, it is advisable to keep some interpretation in mind. We are going to introduce a number of derivation rules, which are, in a way, the atomic steps in a derivation. These derivations rules are designed (by Gentzen), to render the intuitive meaning of the connectives as faithfully as possible.

There is one minor problem, which at the same time is a major advantage, namely: our rules express the constructive meaning of the connectives. This advantage will not be exploited now, but it is good to keep it in mind when dealing with logic (it is exploited in intuitionistic logic).

One small example: the principle of the excluded third tells us that $\models \varphi \vee \neg\varphi$, i.e., assuming that φ is a definite mathematical statement, either it or its negation must be true. Now consider some unsolved problem, e.g. Riemann's Hypothesis, call it R. Then either R is true, or $\neg R$ is true. However, we do not know which of the two is true, so the constructive content of $R \vee \neg R$ is nil. Constructively, one would require a method to find out which of the alternatives holds.

The propositional connective which has a strikingly different meaning in a constructive and in a non-constructive approach is the disjunction. Therefore we restrict our language for the moment to the connectives $\wedge, \rightarrow$ and $\perp$. This is no real restriction as $\{\rightarrow, \perp\}$ is a functionally complete set.

Our derivations consist of very simple steps, such as "from φ and $\varphi \rightarrow \psi$ conclude ψ", written as:

$$\frac{\varphi \quad \varphi \rightarrow \psi}{\psi}$$

The propositions above the line are *premises* , and the one below the line is the *conclusion* . The above example *eliminated* the connective $\rightarrow$. We can also *introduce* connectives. The derivation rules for $\wedge$ and $\rightarrow$ are separated into

INTRODUCTION RULES ELIMINATION RULES

$$(\wedge I)\quad \frac{\varphi \quad \psi}{\varphi \wedge \psi}\wedge I \qquad\qquad (\wedge E)\quad \frac{\varphi \wedge \psi}{\varphi}\wedge E \quad \frac{\varphi \wedge \psi}{\psi}\wedge E$$

$$(\to I)\quad \begin{array}{c}[\varphi]\\ \vdots \\ \psi \\ \hline \varphi \to \psi\end{array}\to I \qquad\qquad (\to E)\quad \frac{\varphi \quad \varphi \to \psi}{\psi}\to E$$

We have two rules for $\perp$, both of which eliminate $\perp$, but introduce a formula.

$$(\perp)\ \frac{\perp}{\varphi}\perp \qquad\qquad (\text{RAA})\quad \begin{array}{c}[\neg\varphi]\\ \vdots \\ \perp \\ \hline \varphi\end{array}\text{RAA}$$

As usual '$\neg\varphi$' is used here as an abbreviation for '$\varphi \to\perp$'.

The rules for $\wedge$ are evident: if we have φ and ψ we may conclude $\varphi \wedge \psi$, and if we have $\varphi \wedge \psi$ we may conclude φ (or ψ). The introduction rule for implication has a different form. It states that, if we can derive ψ from φ (as a hypothesis), then we may conclude $\varphi \to \psi$ (without the hypothesis φ). This agrees with the intuitive meaning of implication: $\varphi \to \psi$ means "ψ follows from φ". We have written the rule ($\to$ I) in the above form to suggest a derivation. The notation will become clearer after we have defined derivations. For the time being we will write the premises of a rule in the order that suits us best, later we will become more fastidious

The rule $(\to E)$ is also evident on the meaning of implication. If φ is given and we know that ψ follows from φ, then we have also ψ. The *falsum rule*, $(\perp)$, expresses that from an absurdity we can derive everything (ex falso sequitur quodlibet), and the *reductio ad absurdum rule* , (RAA), is a formulation of the *principle of proof by contradiction* : if one derives a contradiction from the hypothesis $\neg\varphi$, then one has a derivation of φ (without the hypothesis $\neg\varphi$, of course). In both $(\to I)$ and (RAA) hypotheses disappear, this is indicated by the striking out of the hypothesis. We say that such a hypothesis is *cancelled*. Let us digress for a moment on the cancellation of hypotheses. We first consider implication introduction. There is a well-known theorem in plane geometry which states that "if a triangle is isosceles, then the angles opposite the equal sides are equal to one another" (Euclid's Elements, Book I, proposition 5). This is shown as follows: we suppose that we have an isosceles triangle and then, in a number of steps, we deduce that the angles at the base are equal. Thence we conclude that *the angles at the base are equal if the triangle is isosceles*.

Query 1: do we still need the hypothesis that the triangle is isosceles? Of course not! We have, so to speak, incorporated this condition in the statement itself. It is precisely the role of conditional statements, such as "if it rains I will use my umbrella", to get rid of the obligation to require (or verify) the condition. In abstracto: if we can deduce ψ using the hypothesis φ, then $\varphi \to \psi$ is the case *without the hypothesis* φ (there may be other hypotheses, of course).

Query 2: is it forbidden to maintain the hypothesis? Answer: no, but it clearly is superfluous. As a matter of fact we usually experience superfluous conditions as confusing or even misleading, but that is rather a matter of the psychology of problem solving than of formal logic. Usually we want the best possible result, and it is intuitively clear that the more hypotheses we state for a theorem, the weaker our result is. Therefore we will as a rule cancel as many hypotheses as possible.

In the case of reductio ad absurdum we also deal with cancellation of hypotheses. Again, let us consider an example.

In analysis we introduce the notion of a *convergent sequence* (a_n) and subsequently the notion "a is a limit of (a_n)". The next step is to prove that for each convergent sequence there is a unique limit; we are interested in the part of the proof that shows that there is at most one limit. Such a proof may run as follows: we suppose that there are two distinct limits a and a', and from this hypothesis, $a \neq a'$, we derive a contradiction. Conclusion: $a = a'$. In this case we of course drop the hypothesis $a \neq a'$, this time it is not a case of being superfluous, but of being in conflict! So, both in the case $(\to I)$ and of (RAA), it is sound practice to cancel all occurrences of the hypothesis concerned.

In order to master the technique of Natural Deduction, and to get familiar with the technique of cancellation, one cannot do better than to look at a few concrete cases. So before we go on to the notion of *derivation* we consider a few examples.

I
$$\dfrac{\dfrac{\dfrac{[\varphi \wedge \psi]^1}{\psi}\wedge E \qquad \dfrac{[\varphi \wedge \psi]^1}{\varphi}\wedge E}{\psi \wedge \varphi}\wedge I}{\varphi \wedge \psi \to \psi \wedge \varphi}\to I_1$$

II
$$\dfrac{\dfrac{\dfrac{[\varphi]^2 \qquad [\varphi \to \bot]^1}{\bot}\to E}{(\varphi \to \bot) \to \bot}\to I_1}{\varphi \to ((\varphi \to \bot) \to \bot)}\to I_2$$

III

$$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi \wedge \psi]^1}{\psi}\wedge E \qquad \dfrac{\dfrac{[\varphi \wedge \psi]^1}{\varphi}\wedge E \qquad [\varphi \to (\psi \to \sigma)]^2}{\psi \to \sigma}\to E}{\sigma}\to E}{\varphi \wedge \psi \to \sigma}\to I_1}{(\varphi \to (\psi \to \sigma)) \to (\varphi \wedge \psi \to \sigma)}\to I_2$$

If we use the customary abbreviation '$\neg\varphi$' for '$\varphi \to\perp$', we can bring some derivations into a more convenient form. (Recall that $\neg\varphi$ and $\varphi \to\perp$, as given in 1.2, are semantically equivalent). We rewrite derivation II using the abbreviation:

II′

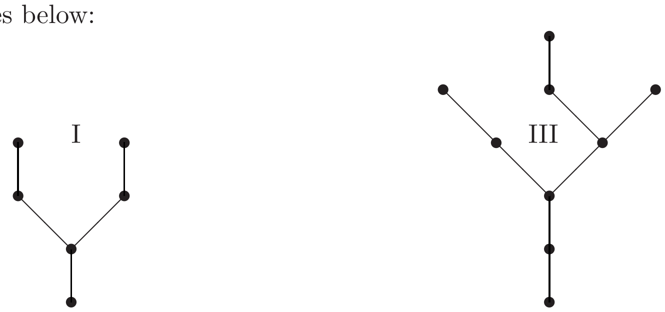
$$\dfrac{\dfrac{\dfrac{[\varphi]^2 \qquad [\neg\varphi]^1}{\perp}\to E}{\neg\neg\varphi}\to I_1}{\varphi \to \neg\neg\varphi}\to I_2$$

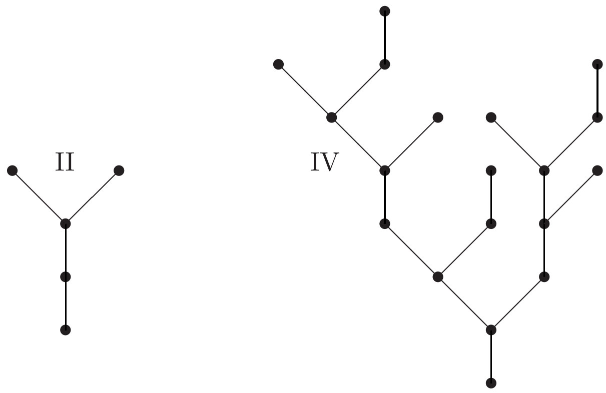
In the following example we use the negation sign and also the bi-implication; $\varphi \leftrightarrow \psi$ for $(\varphi \to \psi) \wedge (\psi \to \varphi)$.

IV

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \qquad \dfrac{[\varphi \leftrightarrow \neg\varphi]^3}{\varphi \to \neg\varphi}\wedge E}{\neg\varphi}\to E \qquad [\varphi]^1}{\perp}\to E}{\neg\varphi}\to I_1 \qquad \dfrac{[\varphi \leftrightarrow \neg\varphi]^3}{\neg\varphi \to \varphi}\wedge E}{\varphi}\to E \qquad \dfrac{\dfrac{\dfrac{[\varphi]^2 \qquad \dfrac{[\varphi \leftrightarrow \neg\varphi]^3}{\varphi \to \neg\varphi}\wedge E}{\neg\varphi}\to E \qquad [\varphi]^2}{\perp}\to E}{\neg\varphi}\to I_2}{\perp}\to E}{\neg(\varphi \leftrightarrow \neg\varphi)}\to I_3$$

The examples show us that derivations have the form of trees. We show the trees below:





One can just as well present derivations as (linear) strings of propositions: we will stick, however, to the tree form, the idea being that what comes naturally in tree form should not be put in a linear straight-jacket.

We now shave to define the notion of *derivation* in general. We will use an inductive definition to produce trees.

Notation

if $\begin{array}{c}\mathcal{D}\\ \varphi\end{array}, \begin{array}{c}\mathcal{D}'\\ \varphi'\end{array}$ are derivations with conclusions φ, φ', then $\dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\end{array}}{\psi}, \quad \dfrac{\begin{array}{cc}\mathcal{D} & \mathcal{D}'\\ \varphi & \varphi'\end{array}}{\psi}$ are derivations obtained by applying a derivation rule to φ (and φ and φ').

The cancellation of a hypothesis is indicated as follows: if $\begin{array}{c}\psi\\ \mathcal{D}\\ \varphi\end{array}$ is a derivation with hypothesis ψ, then $\dfrac{\begin{array}{c}[\psi]\\ \mathcal{D}\\ \varphi\end{array}}{\sigma}$ is a derivation with ψ cancelled.

With respect to the cancellation of hypotheses, we note that one does not necessarily cancel *all* occurrences of such a proposition ψ. This clearly is justified, as one feels that adding hypotheses does not make a proposition underivable (irrelevant information may always be added). It is a matter of prudence, however, to cancel as much as possible. Why carry more hypotheses than necessary?

Furthermore one may apply ($\rightarrow$ I) if there is no hypothesis available for cancellation e.g. $\dfrac{\varphi}{\psi \rightarrow \varphi} \rightarrow I$ is a correct derivation, using just ($\rightarrow I$). To sum it up: given a derivation tree of ψ (or $\perp$), we obtain a derivation tree of $\varphi \rightarrow \psi$ (or φ) at the bottom of the tree and striking out some (or all) occurrences, if any, of φ (or $\neg\varphi$) on top of a tree.

A few words on the practical use of natural deduction: if you want to give a derivation for a proposition it is advisable to devise some kind of strategy, just like in a game. Suppose that you want to show $[\varphi \wedge \psi \to \sigma] \to [\varphi \to (\psi \to \sigma)]$ (Example III), then (since the proposition is an implicational formula) the rule $(\to I)$ suggests itself. So try to derive $\varphi \to (\psi \to \sigma)$ from $\varphi \wedge \psi \to \sigma$.

Now we know where to start and where to go to. To make use of $\varphi \wedge \psi \to \sigma$ we want $\varphi \wedge \psi$ (for $(\to E)$), and to get $\varphi \to (\psi \to \sigma)$ we want to derive $\psi \to \sigma$ from φ. So we may add φ as a hypothesis and look for a derivation of $\psi \to \sigma$. Again, this asks for a derivation of σ from ψ, so add ψ as a hypothesis and look for a derivation of σ. By now we have the following hypotheses available: $\varphi \wedge \psi \to \sigma, \varphi$ and ψ. Keeping in mind that we want to eliminate $\varphi \wedge \psi$ it is evident what we should do. The derivation III shows in detail how to carry out the derivation. After making a number of derivations one gets the practical conviction that one should first take propositions apart from the bottom upwards, and then construct the required propositions by putting together the parts in a suitable way. This practical conviction is confirmed by the *Normalization Theorem*, to which we will return later. There is a particular point which tends to confuse novices:

$$\dfrac{\begin{matrix}[\varphi]\\ \vdots \\ \bot\end{matrix}}{\neg\varphi}\to I \qquad \text{and} \qquad \dfrac{\begin{matrix}[\neg\varphi]\\ \vdots \\ \bot\end{matrix}}{\varphi}\text{RAA}$$

look very much alike. Are they not both cases of Reductio ad absurdum? As a matter of fact the leftmost derivation tells us (informally) that the assumption of φ leads to a contradiction, so φ *cannot be the case*. This is in our terminology the meaning of "not φ". The rightmost derivation tells us that the assumption of $\neg\varphi$ leads to a contradiction, hence (by the same reasoning) $\neg\varphi$ cannot be the case. So, on account of the meaning of negation, we only would get $\neg\neg\varphi$. It is by no means clear that $\neg\neg\varphi$ is equivalent to φ (indeed, this is denied by the intuitionists), so it is an extra property of our logic. (This is confirmed in a technical sense: $\neg\neg\varphi \to \varphi$ is not derivable in the system without RAA.

We now return to our theoretical notions.

Definition 1.4.1 *The set of derivations is the smallest set X such that*

(1) The one element tree φ belongs to X for all $\varphi \in PROP$.

(2$\wedge$) If $\begin{matrix}\mathcal{D}\\ \varphi\end{matrix}, \begin{matrix}\mathcal{D}'\\ \varphi'\end{matrix} \in X$, *then* $\dfrac{\begin{matrix}\mathcal{D}\\ \varphi\end{matrix} \quad \begin{matrix}\mathcal{D}'\\ \varphi'\end{matrix}}{\varphi \wedge \varphi'} \in X$.

$$\textit{If}\ \begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array} \in X,\ \textit{then}\ \ \dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array}}{\varphi},\ \dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array}}{\psi} \in X.$$

$$(2\!\to)\ \textit{If}\ \begin{array}{c}\varphi\\ \mathcal{D}\\ \psi\end{array} \in X,\ \textit{then}\ \ \dfrac{\begin{array}{c}[\varphi]\\ \mathcal{D}\\ \psi\end{array}}{\varphi\to\psi} \in X.$$

$$\textit{If}\ \begin{array}{c}\mathcal{D}\\ \varphi\end{array},\ \begin{array}{c}\mathcal{D}'\\ \varphi\to\psi\end{array} \in X,\ \textit{then}\ \ \dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\end{array}\qquad\begin{array}{c}\mathcal{D}'\\ \varphi\to\psi\end{array}}{\psi} \in X.$$

$$(2\bot)\ \textit{If}\ \begin{array}{c}\mathcal{D}\\ \bot\end{array} \in X,\ \textit{then}\ \ \dfrac{\begin{array}{c}\mathcal{D}\\ \bot\end{array}}{\varphi} \in X.$$

$$\textit{If}\ \begin{array}{c}\neg\varphi\\ \mathcal{D}\\ \bot\end{array} \in X,\ \textit{then}\ \ \dfrac{\begin{array}{c}[\neg\varphi]\\ \mathcal{D}\\ \bot\end{array}}{\varphi} \in X.$$

The bottom formula of a derivation is called its *conclusion* . Since the class of derivations is inductively defined, we can mimic the results of section 1.1.

E.g. we have a *principle of induction on* $\mathcal{D}$: let A be a property. If $A(\mathcal{D})$ holds for one element derivations and A is preserved under the clauses $(2\wedge)$, $(2\to)$ and $(2\bot)$, then $A(\mathcal{D})$ holds for all derivations. Likewise we can define mappings on the set of derivations by recursion (cf. Exercises 6, 7, 9).

Definition 1.4.2 *The relation $\Gamma \vdash \varphi$ between sets of propositions and propositions is defined by: there is a derivation with conclusion φ and with all (uncancelled) hypotheses in Γ. (See also exercise 6).*

We say that φ is *derivable* from Γ. Note that by definition Γ may contain many superfluous "hypotheses". The symbol $\vdash$ is called *turnstile* .

If $\Gamma = \emptyset$, we write $\vdash \varphi$, and we say that φ is a theorem.

We could have avoided the notion of 'derivation' and taken instead the notion of 'derivability' as fundamental, see Exercise 10. The two notions, however, are closely related.

Lemma 1.4.3 *(a)* $\Gamma \vdash \varphi$ *if* $\varphi \in \Gamma$,
(b) $\Gamma \vdash \varphi, \Gamma' \vdash \psi \Rightarrow \Gamma \cup \Gamma' \vdash \varphi \wedge \psi$,
(c) $\Gamma \vdash \varphi \wedge \psi \Rightarrow \Gamma \vdash \varphi$ *and* $\Gamma \vdash \psi$,
(d) $\Gamma \cup \varphi \vdash \psi \Rightarrow \Gamma \vdash \varphi \to \psi$,
(e) $\Gamma \vdash \varphi, \Gamma' \vdash \varphi \to \psi \Rightarrow \Gamma \cup \Gamma' \vdash \psi$,
(f) $\Gamma \vdash \perp \Rightarrow \Gamma \vdash \varphi$,
(g) $\Gamma \cup \{\neg\varphi\} \vdash \perp \Rightarrow \Gamma \vdash \varphi$.

Proof. Immediate from the definition of derivation. □

We now list some theorems. $\neg$ and $\leftrightarrow$ are used as abbreviations.

Theorem 1.4.4 *(1)* $\vdash \varphi \to (\psi \to \varphi)$,
(2) $\vdash \varphi \to (\neg\varphi \to \psi)$,
(3) $\vdash (\varphi \to \psi) \to [(\psi \to \sigma) \to (\varphi \to \sigma)]$,
(4) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$,
(5) $\vdash \neg\neg\varphi \leftrightarrow \varphi$,
(6) $\vdash [\varphi \to (\psi \to \sigma)] \leftrightarrow [\varphi \wedge \psi \to \sigma]$,
(7) $\vdash \perp \leftrightarrow (\varphi \wedge \neg\varphi)$.

Proof.

1. $$\dfrac{\dfrac{[\varphi]^1}{\psi \to \varphi} \to I}{\varphi \to (\psi \to \varphi)} \to I_1$$

2. $$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^2 \quad [\neg\varphi]^1}{\perp} \to E}{\psi} \perp}{\neg\varphi \to \psi} \to I_1}{\varphi \to (\neg\varphi \to \psi)} \to I_2$$

3. $$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad [\varphi \to \psi]^3}{\psi} \to E \qquad [\psi \to \sigma]^2}{\sigma} \to E}{\varphi \to \sigma} \to I_1}{(\psi \to \sigma) \to (\varphi \to \sigma)} \to I_2}{(\varphi \to \psi) \to ((\psi \to \sigma) \to (\varphi \to \sigma))} \to I_3$$

4. For one direction, substitute $\perp$ for σ in 3, then $\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$. Conversely:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\neg\psi]^1 \quad [\neg\psi \to \neg\varphi]^3}{\neg\varphi}\to E \qquad [\varphi]^2}{\bot}\to E}{\psi}\,\mathrm{RAA}_1}{\varphi \to \psi}\to I_2}{(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)}\to I_3$$

So now we have $$\dfrac{\begin{array}{c}\mathcal{D}\\ (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)\end{array} \qquad \begin{array}{c}\mathcal{D}'\\ (\neg\psi \to \neg\varphi) \to (\varphi \to \psi)\end{array}}{(\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)}\wedge I$$

5. We already proved $\varphi \to \neg\neg\varphi$ as an example. Conversely:

$$\dfrac{\dfrac{\dfrac{[\neg\varphi]^1 \quad [\neg\neg\varphi]^2}{\bot}\to E}{\varphi}\,\mathrm{RAA}_1}{\neg\neg\varphi \to \varphi}\to I_2$$

The result now follows. The numbers 6 and 7 are left to the reader. □

The system, outlined in this section, is called the "calculus of natural deduction" for a good reason. That is: its manner of making inferences corresponds to the reasoning we intuitively use. The rules present means to take formulas apart, or to put them together. A derivation then consists of a skilful manipulation of the rules, the use of which is usually suggested by the form of the formula we want to prove.

We will discuss one example in order to illustrate the general strategy of building derivations. Let us consider the converse of our previous example **III**.

To prove $(\varphi \wedge \psi \to \sigma) \to [\varphi \to (\psi \to \sigma)]$ there is just one initial step: assume $\varphi \wedge \psi \to \sigma$ and try to derive $\varphi \to (\psi \to \sigma)$. Now we can either look at the assumption or at the desired result. Let us consider the latter one first: to show $\varphi \to (\psi \to \sigma)$, we should assume φ and derive $\psi \to \sigma$, but for the latter we should assume ψ and derive σ.

So, altogether we may assume $\varphi \wedge \psi \to \sigma$ and φ and ψ. Now the procedure suggests itself: derive $\varphi \wedge \psi$ from φ and ψ, and σ from $\varphi \wedge \psi$ and $\varphi \wedge \psi \to \sigma$.

Put together, we get the following derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^2 \quad [\psi]^1}{\varphi\wedge\psi}\wedge I \qquad [\varphi\wedge\psi\to\sigma]^3}{\sigma}\to E}{\psi\to\sigma}\to I_1}{\varphi\to(\psi\to\sigma)}\to I_2}{(\varphi\wedge\psi\to\sigma)\to(\varphi\to(\psi\to\sigma))}\to I_3$$

Had we considered $\varphi \wedge \psi \to \sigma$ first, then the only way to proceed is to add $\varphi \wedge \psi$ and apply $\to E$. Now $\varphi \wedge \psi$ either remains an assumption, or it is obtained from something else. It immediately occurs to the reader to derive $\varphi \wedge \psi$ from φ and ψ. But now he will build up the derivation we obtained above.

Simple as this example seems, there are complications. In particular the rule of reductio ad absurdum is not nearly as natural as the other ones. Its use must be learned by practice; also a sense for the distinction between *constructive* and *non-constructive* will be helpful when trying to decide on when to use it.

Finally, we recall that $\top$ is an abbreviation for $\neg\bot$ (i.e. $\bot \to \bot$).

Exercises

1. Show that the following propositions are derivable.
   (a) $\varphi \to \varphi$, (d) $(\varphi \to \psi) \leftrightarrow \neg(\varphi \wedge \neg\psi)$,
   (b) $\bot \to \varphi$, (e) $(\varphi \wedge \psi) \leftrightarrow \neg(\varphi \to \neg\psi)$,
   (c) $\neg(\varphi \wedge \neg\varphi)$, (f) $\varphi \to (\psi \to (\varphi \wedge \psi))$.

2. Idem for (a) $(\varphi \to \neg\varphi) \to \neg\varphi$,
   (b) $[\varphi \to (\psi \to \sigma] \leftrightarrow [\psi \to (\varphi \to \sigma]$,
   (c) $(\varphi \to \psi) \wedge (\varphi \to \neg\psi) \to \neg\varphi$,
   (d) $(\varphi \to \psi) \to [(\varphi \to (\psi \to \sigma)) \to (\varphi \to \sigma)]$.

3. Show (a) $\varphi \vdash \neg(\neg\varphi \wedge \psi)$, (d) $\vdash \varphi \Rightarrow \vdash \psi \to \varphi$,
   (b) $\neg(\varphi \wedge \neg\psi), \varphi \vdash \psi$, (e) $\neg\varphi \vdash \varphi \to \psi$.
   (c) $\neg\varphi \vdash (\varphi \to \psi) \leftrightarrow \neg\varphi$,

4. Show $\vdash [(\varphi \to \psi) \to (\varphi \to \sigma)] \to [(\varphi \to (\psi \to \sigma))]$,
   $\vdash ((\varphi \to \psi) \to \varphi) \to \varphi$.

5. Show $\Gamma \vdash \varphi \Rightarrow \Gamma \cup \Delta \vdash \varphi$,
   $\Gamma \vdash \varphi; \Delta, \varphi \vdash \psi \Rightarrow \Gamma \cup \Delta \vdash \psi$.

1.6 The Missing Connectives

The language of section 1.4 contained only the connectives $\wedge, \rightarrow$ and $\perp$. We already know that, from the semantical point of view, this language is sufficiently rich, i.e. the missing connectives can be defined. As a matter of fact we have already used the negation as a defined notion in the preceding sections.

It is a matter of sound mathematical practice to introduce new notions if their use simplifies our labour, and if they codify informal existing practice. This, clearly, is a reason for introducing $\neg, \leftrightarrow$ and $\vee$.

Now there are two ways to proceed: one can introduce the new connectives as abbreviations (of complicated propositions), or one can enrich the language by actually adding the connectives to the alphabet, and providing rules of derivation.

The first procedure was adopted above; it is completely harmless, e.g. each time one reads $\varphi \leftrightarrow \psi$, one has to replace it by $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. So it is nothing but a shorthand, introduced for convenience. The second procedure is of a more theoretical nature. The language is enriched and the set of derivations is enlarged. As a consequence one has to review the theoretical results (such as the Completeness Theorem) obtained for the simpler language.

We will adopt the first procedure and also outline the second approach.

Definition 1.6.1
$$\begin{aligned} \varphi \vee \psi &:= \neg(\neg\varphi \wedge \neg\psi), \\ \neg\varphi &:= \varphi \rightarrow \perp, \\ \varphi \leftrightarrow \psi &:= (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi) \ . \end{aligned}$$

N.B. This means that the above expressions are *not* part of the language, but abbreviations for certain propositions.

The properties of $\vee, \neg$ and $\leftrightarrow$ are given in the following:

Lemma 1.6.2
(i) $\varphi \vdash \varphi \vee \psi, \ \psi \vdash \varphi \vee \psi$,
(ii) $\Gamma, \varphi \vdash \sigma$ *and* $\Gamma, \psi \vdash \sigma \Rightarrow \Gamma, \varphi \vee \psi \vdash \sigma$,
(iii) $\varphi, \neg\varphi \vdash \perp$,
(iv) $\Gamma, \varphi \vdash \perp \Rightarrow \Gamma \vdash \neg\varphi$,
(v) $\varphi \leftrightarrow \psi, \varphi \vdash \psi$ *and* $\varphi \leftrightarrow \psi, \psi \vdash \varphi$,
(vi) $\Gamma, \varphi \vdash \psi$ *and* $\Gamma, \psi \vdash \varphi \Rightarrow \Gamma \vdash \varphi \leftrightarrow \psi$.

Proof. The only non-trivial part is (*ii*). We exhibit a derivation of σ from Γ and $\varphi \vee \psi$ (i.e. $\neg(\neg\varphi \wedge \neg\psi)$), given derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ of $\Gamma, \varphi \vdash \sigma$ and $\Gamma, \psi \vdash \sigma$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}[\varphi]^1\\ \mathcal{D}_1\\ \sigma\end{array}\quad [\neg\sigma]^3}{\bot}\to E}{\neg\varphi}\to I_1 \qquad \dfrac{\dfrac{\begin{array}{c}[\psi]^2\\ \mathcal{D}_2\\ \sigma\end{array}\quad [\neg\sigma]^3}{\bot}\to E}{\neg\psi}\to I_2}{\neg\varphi\wedge\neg\psi}\wedge I \qquad \neg(\neg\varphi\wedge\neg\psi)}{\bot}\to E}{\sigma}\mathrm{RAA}_3$$

The remaining cases are left to the reader. □

Note that (*i*) and (*ii*) read as introduction and elimination rules for $\vee$, (*iii*) and (*iv*) as ditto for $\neg$, (*vi*) and (*v*) as ditto for $\leftrightarrow$.

They legalise the following shortcuts in derivations:

$$\dfrac{\varphi}{\varphi\vee\psi}\vee I \qquad \dfrac{\psi}{\varphi\vee\psi}\vee I \qquad \dfrac{\varphi\vee\psi \quad \begin{array}{c}[\varphi]\\ \vdots\\ \sigma\end{array} \quad \begin{array}{c}[\psi]\\ \vdots\\ \sigma\end{array}}{\sigma}\vee E$$

$$\dfrac{\begin{array}{c}[\varphi]\\ \vdots\\ \bot\end{array}}{\neg\varphi}\neg I \qquad \dfrac{\varphi \quad \neg\varphi}{\bot}\neg E$$

$$\dfrac{\begin{array}{c}[\varphi]\\ \vdots\\ \psi\end{array} \quad \begin{array}{c}[\psi]\\ \vdots\\ \varphi\end{array}}{\varphi\leftrightarrow\psi}\leftrightarrow I \qquad \dfrac{\varphi \quad \varphi\leftrightarrow\psi}{\psi} \qquad \dfrac{\psi \quad \varphi\leftrightarrow\psi}{\varphi}\leftrightarrow E$$

Consider for example an application of $\vee E$

$$\dfrac{\begin{matrix} \\ \mathcal{D}_0 \\ \varphi \vee \psi \end{matrix} \qquad \begin{matrix} [\varphi] \\ \mathcal{D}_1 \\ \sigma \end{matrix} \qquad \begin{matrix} [\psi] \\ \mathcal{D}_2 \\ \sigma \end{matrix}}{\sigma} \vee E$$

This is a mere shorthand for

$$\dfrac{\dfrac{\begin{matrix} \mathcal{D}_0 \\ \neg(\neg\varphi \wedge \neg\psi) \end{matrix} \qquad \dfrac{\dfrac{\dfrac{\begin{matrix} [\varphi]^1 \\ \mathcal{D}_1 \\ \sigma \end{matrix} \quad [\neg\sigma]^3}{\bot}}{\neg\varphi}\,1 \qquad \dfrac{\dfrac{\begin{matrix} [\psi]^2 \\ \mathcal{D}_2 \\ \sigma \end{matrix} \quad [\neg\sigma]^3}{\bot}}{\neg\psi}\,2}{\neg\varphi \wedge \neg\psi}}{\bot}\,1}{\sigma}\,3$$

The reader is urged to use the above shortcuts in actual derivations, whenever convenient. As a rule, only $\vee I$ and $\vee E$ are of importance, the reader has of course recognised the rules for $\neg$ and $\leftrightarrow$ as slightly eccentric applications of familiar rules.

Examples. $\vdash (\varphi \wedge \psi) \vee \sigma \leftrightarrow (\varphi \vee \sigma) \wedge (\psi \vee \sigma)$.

$$\dfrac{\dfrac{(\varphi \wedge \psi) \vee \sigma \qquad \dfrac{\dfrac{[\varphi \wedge \psi]^1}{\varphi}}{\varphi \vee \sigma} \qquad \dfrac{[\sigma]^1}{\varphi \vee \sigma}}{\varphi \vee \sigma}\,1 \qquad \dfrac{(\varphi \wedge \psi) \vee \sigma \qquad \dfrac{\dfrac{[\varphi \wedge \psi]^2}{\psi}}{\psi \vee \sigma} \qquad \dfrac{[\sigma]^2}{\psi \vee \sigma}}{\psi \vee \sigma}\,2}{(\varphi \vee \sigma) \wedge (\psi \vee \sigma)} \tag{1}$$

Conversely

$$\dfrac{\dfrac{(\varphi \vee \sigma) \wedge (\psi \vee \sigma)}{\varphi \vee \sigma} \qquad \dfrac{\dfrac{(\varphi \vee \sigma) \wedge (\psi \vee \sigma)}{\psi \vee \sigma} \quad \dfrac{\dfrac{[\varphi]^2 \quad [\psi]^1}{\varphi \wedge \psi}}{(\varphi \wedge \psi) \vee \sigma} \quad \dfrac{[\sigma]^1}{(\varphi \wedge \psi) \vee \sigma}}{(\varphi \wedge \psi) \vee \sigma}\,{\scriptstyle 1} \qquad \dfrac{[\sigma]^2}{(\varphi \wedge \psi) \vee \sigma}}{(\varphi \wedge \psi) \vee \sigma}\,{\scriptstyle 2} \tag{2}$$

Combining (1) and (2) we get one derivation:

$$\dfrac{\begin{matrix}[(\varphi \wedge \psi) \vee \sigma] \\ \mathcal{D} \\ (\varphi \vee \sigma) \wedge (\psi \vee \sigma)\end{matrix} \qquad \begin{matrix}[(\varphi \vee \sigma) \wedge (\psi \vee \sigma)] \\ \mathcal{D}' \\ (\varphi \wedge \psi) \vee \sigma\end{matrix}}{(\varphi \wedge \psi) \vee \sigma \leftrightarrow (\varphi \vee \sigma) \wedge (\psi \vee \sigma)} \leftrightarrow I$$

$\vdash \varphi \vee \neg\varphi$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1}{\varphi \vee \neg\varphi} \vee I \quad [\neg(\varphi \vee \neg\varphi)]^2}{\bot} \to E}{\neg\varphi} \to I_1}{\varphi \vee \neg\varphi} \vee I \qquad [\neg(\varphi \vee \neg\varphi)]^2}{\bot} \to E}{\varphi \vee \neg\varphi} \mathrm{RAA}_2$$

$\vdash (\varphi \to \psi) \vee (\psi \to \varphi)$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1}{\psi \to \varphi} \to I_1}{(\varphi \to \psi) \vee (\psi \to \varphi)} \vee I \quad [\neg((\varphi \to \psi) \vee (\psi \to \varphi))]^2}{\bot} \to E}{\psi} \bot}{\varphi \to \psi} \to I_1}{(\varphi \to \psi) \vee (\psi \to \varphi)} \vee I \qquad [\neg((\varphi \to \psi) \vee (\psi \to \varphi))]^2}{\bot} \to E}{(\varphi \to \psi) \vee (\psi \to \varphi)} \mathrm{RAA}_2$$

$\vdash \neg(\varphi \wedge \psi) \rightarrow \neg\varphi \vee \neg\psi$

$$
\begin{prooftree}
\AxiomC{$[\neg(\varphi \wedge \psi)]$}
\AxiomC{$[\neg(\neg\varphi \vee \neg\psi)]$}
\AxiomC{$[\neg\varphi]$}
\UnaryInfC{$\neg\varphi \vee \neg\psi$}
\BinaryInfC{$\bot$}
\UnaryInfC{φ}
\AxiomC{$[\neg(\neg\varphi \vee \neg\psi)]$}
\AxiomC{$[\neg\psi]$}
\UnaryInfC{$\neg\varphi \vee \neg\psi$}
\BinaryInfC{$\bot$}
\UnaryInfC{ψ}
\BinaryInfC{$\varphi \wedge \psi$}
\BinaryInfC{$\bot$}
\UnaryInfC{$\neg\varphi \vee \neg\psi$}
\UnaryInfC{$\neg(\varphi \wedge \psi) \rightarrow \neg\varphi \vee \neg\psi$}
\end{prooftree}
$$

□

We now give a sketch of the second approach. We add $\vee, \neg$ and $\leftrightarrow$ to the language, and extend the set of propositions correspondingly. Next we add the rules for $\vee, \neg$ and $\leftrightarrow$ listed above to our stock of derivation rules. To be precise we should now also introduce a new derivability sign, we will however stick to the trusted $\vdash$ in the expectation that the reader will remember that now we are making derivations in a larger system. The following holds:

Theorem 1.6.3 $\vdash \varphi \vee \psi \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)$.
$\vdash \neg\varphi \leftrightarrow (\varphi \rightarrow \bot)$.
$\vdash (\varphi \leftrightarrow \psi) \leftrightarrow (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$.

Proof. Observe that by Lemma 1.6.2 the defined and the primitive (real) connectives obey exactly the same derivability relations (derivation rules, if you wish). This leads immediately to the desired result. Let us give one example.
$\varphi \vdash \neg(\neg\varphi \wedge \neg\psi)$ and $\psi \vdash \neg(\neg\varphi \wedge \neg\psi)$ (1.6.2 (i)), so by $\vee E$ we get
$\varphi \vee \psi \vdash \neg(\neg\varphi \wedge \neg\psi) \ldots (1)$
Conversely $\varphi \vdash \varphi \vee \psi$ (by $\vee I$), hence by 1.6.2 (ii)
$\neg(\neg\varphi \wedge \neg\psi) \vdash \varphi \vee \psi \ldots (2)$
Apply $\leftrightarrow I$, to (1) and (2), then $\vdash \varphi \vee \psi \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)$. The rest is left to the reader. □

For more results the reader is directed to the exercises.

The rules for $\vee, \leftrightarrow$, and $\neg$ capture indeed the intuitive meaning of those connectives. Let us consider disjunction: ($\vee$I) : If we know φ then we certainly know $\varphi \vee \psi$ (we even know exactly which disjunct). The ($\vee$E)-rule captures the idea of "proof by cases": if we know $\varphi \vee \psi$ and in each of both cases we can conclude σ, then we may outright conclude σ. Disjunction intuitively calls for a decision: which of the two disjuncts is given or may be assumed? This constructive streak of $\vee$ is crudely but conveniently blotted out by the identification of $\varphi \vee \psi$ and $\neg(\neg\varphi \wedge \neg\psi)$. The latter only tells us that φ and ψ cannot both be wrong, but not which one is right. For more information on this matter of constructiveness, which plays a role in demarcating the borderline between two-valued classical logic and effective intuitionistic logic, the reader is referred to Chapter 5.

Note that with $\vee$ as a primitive connective some theorems become harder to prove. E.g. $\vdash \neg(\neg\neg\varphi \wedge \neg\varphi)$ is trivial, but $\vdash \varphi \vee \neg\varphi$ is not. The following rule of the thumb may be useful: going from non-effective (or no) premises to an effective conclusion calls for an application of RAA.

Exercises

1. Show $\vdash \varphi \vee \psi \to \psi \vee \varphi$, $\vdash \varphi \vee \varphi \leftrightarrow \varphi$.

2. Consider the full language $\mathcal{L}$ with the connectives $\wedge, \to, \bot, \leftrightarrow\ \vee$ and the restricted language $\mathcal{L}'$ with connectives $\wedge, \to, \bot$. Using the appropriate derivation rules we get the derivability notions $\vdash$ and $\vdash'$. We define an obvious translation from $\mathcal{L}$ into $\mathcal{L}'$:
$$\begin{aligned} \varphi^+ &:= \varphi \text{ for atomic } \varphi \\ (\varphi \Box \psi)^+ &:= \varphi^+ \Box \psi^+ \text{ for } \Box = \wedge, \to, \\ (\varphi \vee \psi)^+ &:= \neg(\neg\varphi^+ \wedge \neg\varphi^+), \text{where} \neg \text{ is an abbreviation,} \\ (\varphi \leftrightarrow \psi)^+ &:= (\varphi^+ \to \psi^+) \wedge (\psi^+ \to \varphi^+), \\ (\neg\varphi)^+ &:= \varphi^+ \to \bot\ . \end{aligned}$$

   Show (i) $\vdash \varphi \leftrightarrow \varphi^+$,
   (ii) $\vdash \varphi \Leftrightarrow \vdash' \varphi^+$,
   (iii) $\varphi^+ = \varphi for \varphi \in \mathcal{L}'$.
   (iv) Show that the full logic, is *conservative* over the restricted logic, i.e. for$\varphi \in \mathcal{L}'$ $\vdash \varphi \Leftrightarrow \vdash' \varphi$.

3. Show that the Completeness Theorem holds for the full logic. Hint: use Exercise 2.

4. Show (a) $\vdash \top \vee \bot$.
   (b) $\vdash (\varphi \leftrightarrow \top) \vee (\varphi \leftrightarrow \bot)$.
   (c) $\vdash \varphi \leftrightarrow (\varphi \leftrightarrow \top)$.

5. Show $\vdash (\varphi \vee \psi) \leftrightarrow ((\varphi \to \psi) \to \psi)$.

6. Show (a) Γ is complete $\Leftrightarrow (\Gamma \vdash \varphi \vee \psi \Leftrightarrow \Gamma \vdash \varphi$ or $\Gamma \vdash \psi$, for all $\varphi, \psi)$,
   (b) Γ is maximally consistent $\Leftrightarrow \Gamma$ is a consistent theory and for all φ, ψ $(\varphi \vee \psi \in \Gamma \Leftrightarrow \varphi \in \Gamma$ or $\psi \in \Gamma)$.

7. Show in the system with $\vee$ as a primitive connective
   $\vdash (\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$,
   $\vdash (\varphi \to \psi) \vee (\psi \to \varphi)$.