Math 4680, Topics in Logic and Computation, Winter 2012

Lecture Notes 2:
Truth, Proofs, Soundness, and Completeness for Sentential Logic

Peter Selinger

1 Truth tables

Fix some two-element set $\mathbb{B} = \{T, F\}$, whose members we call ***truth values***. The truth values T and F are called "truth" and "falsity", respectively.

On the set $\mathbb{B}$, consider the functions $f_\neg : \mathbb{B} \to \mathbb{B}$ and $f_\wedge, f_\neg, f_\rightarrow, f_\leftrightarrow : \mathbb{B} \times \mathbb{B} \to \mathbb{B}$ whose values are defined in the following table:

x	y	$f_\neg(x)$	$f_\wedge(x,y)$	$f_\neg(x,y)$	$f_\rightarrow(x,y)$	$f_\leftrightarrow(x,y)$
T	T	F	T	T	T	T
T	F	F	F	T	F	F
F	T	T	F	T	T	F
F	F	T	F	F	T	T

We will now define the interpretation of the language of sentential logic. Recall that $\mathcal{W}$ is the set of well-formed formulas. Let $\mathscr{S}$ be the set of sentence symbols.

Definition. A ***truth assignment*** is a function $v : \mathscr{S} \to \mathbb{B}$. It assigns a truth value to each sentence symbol.

Given a truth assignment v, we can calculate the truth value of any formula. Formally, we define an ***interpretation function*** $\bar{v} : \mathcal{W} \to \mathbb{B}$ which maps well-formed formulas to truth values, by the following recursive definition:

$$
\begin{array}{lll}
\bar{v}(\alpha) & = & v(\alpha) \qquad \text{if } \alpha \text{ a sentence symbol,} \\
\bar{v}(\top) & = & T, \\
\bar{v}(\bot) & = & F, \\
\bar{v}(\neg\,\alpha) & = & f_\neg(\bar{v}(\alpha)), \\
\bar{v}(\alpha \wedge \beta) & = & f_\wedge(\bar{v}(\alpha), \bar{v}(\beta)), \\
\bar{v}(\alpha \vee \beta) & = & f_\neg(\bar{v}(\alpha), \bar{v}(\beta)), \\
\bar{v}(\alpha \rightarrow \beta) & = & f_\rightarrow(\bar{v}(\alpha), \bar{v}(\beta)), \\
\bar{v}(\alpha \leftrightarrow \beta) & = & f_\leftrightarrow(\bar{v}(\alpha), \bar{v}(\beta)).
\end{array}
$$

The function $\bar{v}$ can be conveniently calculated, for the different possible v, in a ***truth table***. For example, the following truth table shows the values of the formula $\bar{v}(\neg(\neg\,\mathbf{A}_1 \wedge \mathbf{A}_2))$ for various different truth assignments v. Each row in the table corresponds to a different truth assignment.

$v(\mathbf{A}_1)$	$v(\mathbf{A}_2)$	$\bar{v}(\neg\,\mathbf{A}_1)$	$\bar{v}(\neg\,\mathbf{A}_1 \wedge \mathbf{A}_2)$	$\bar{v}(\neg(\neg\,\mathbf{A}_1 \wedge \mathbf{A}_2))$
T	T	F	F	T
T	F	F	F	T
F	T	T	T	F
F	F	T	F	T

The headers in this table are usually written more succinctly as

$\mathbf{A}_1$	$\mathbf{A}_2$	$\neg\,\mathbf{A}_1$	$\neg\,\mathbf{A}_1 \wedge \mathbf{A}_2$	$\neg(\neg\,\mathbf{A}_1 \wedge \mathbf{A}_2)$
T	T	F	F	T
T	F	F	F	T
F	T	T	T	F
F	F	T	F	T

Definition. If v is a truth assignment and α is a formula, we call $\bar{v}\alpha$ the ***interpretation*** of α with respect to v. We say v ***satisfies*** α if $\bar{v}(\alpha) = T$. We say that α is ***valid*** or a ***tautology*** if every truth assignment satisfies α. If Σ is a set of formulas and α is a formula, then we say Σ ***tautologically implies*** α, in symbols $\Sigma \models \alpha$, if for all truth assignments v,

$$(\forall \sigma \in \Sigma . \bar{v}(\sigma) = T) \Rightarrow \bar{v}(\alpha) = T.$$

In words, we say that $\Sigma \models \alpha$ if every truth assignment that satisfies all the formulas in Σ also satisfies α. We often abbreviate $\{\sigma_1, \ldots, \sigma_n\} \models \alpha$ to $\sigma_1, \ldots, \sigma_n \models \alpha$, and $\Sigma \cup \{\sigma\} \models \alpha$ to $\Sigma, \sigma \models \alpha$. Also, instead of $\emptyset \models \alpha$, we simply write $\models \alpha$. Notice that $\models \alpha$ if and only if α is a tautology.

If $\alpha \models \beta$ and $\beta \models \alpha$, then we say α and β are ***tautologically equivalent***, and we write $\alpha \models\!\!\!\dashv \beta$.

It is easy to check whether two formulas are tautologically equivalent with truth tables: α and β are tautologically equivalent if their corresponding truth table columns are identical. Also, α tautologically implies β if in each row where α has a truth table entry of T, β also has a truth table entry of T. For instance, the following truth table shows that $\neg\,\mathbf{A}_1 \cot \mathbf{A}_2 \models\!\!\!\dashv \mathbf{A}_1 \rightarrow \mathbf{A}_2$:

$\mathbf{A}_1$	$\mathbf{A}_2$	$\neg\,\mathbf{A}_1$	$\neg\,\mathbf{A}_1 \vee \mathbf{A}_2$	$\mathbf{A}_1 \rightarrow \mathbf{A}_2$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

Remark. We never use the symbol "=" to denote tautological equivalence. The equality symbol is reserved to denote *syntactic* equality of formulas, i.e., equality of formulas as strings. For tautological equivalence, we always use "$\models\,\dashv$".

Remark. The statements "$\neg\,\alpha$ is a tautology" and "α is not a tautology" are not equivalent. For instance, if α is a sentence symbol, then neither α nor $\neg\,\alpha$ is a tautology.

For a list of tautologies, see Enderton [**?**, p.37].

2 Natural deduction

There are many possible ways of formalizing proofs. For this course, we choose the formalism of *natural deduction*, which was first introduced by Gentzen in 1935 [**?**, **?**], and further developed by Prawitz in the 1960's [**?**].

2.1 An intuitive explanation of natural deduction

The representation of a proof in the natural deduction system is called a ***derivation***. Loosely speaking, a derivation is a tree-like structure whose leaves are labeled by well-formed formulas. The formulas that occur at the leaves are called the ***hypotheses*** or ***assumptions*** of the derivation, and the unique formula that occurs at the root of the derivation is called its ***conclusion***.

Here is a simple example of a derivation:

$$\frac{\alpha \qquad \beta}{\alpha \wedge \beta}\,(\wedge\mathcal{I})$$

The hypotheses are α and β, and the conclusion is $\alpha \wedge \beta$. Thus, this particular derivation formalizes a proof that the formula α and the formula β imply the formula $\alpha \wedge \beta$. The horizontal line corresponds to an instance of an *inference rule*, in this case the rule whose name is $(\wedge\mathcal{I})$ or "and introduction". The formulas that occur immediately above the horizontal line are called the *premises* of the inference rule (as opposed to *hypotheses* of a derivation). Here is another inference rule, which is known as "modus ponens", or "arrow elimination":

$$\frac{\beta \to \alpha \qquad \beta}{\alpha}\,(\to\mathcal{E})$$

From hypotheses $\alpha \to \beta$ and α, we may conclude β. We can put more than one rule together to form a derivation:

$$\frac{\dfrac{\beta \to \alpha \qquad \beta}{\alpha}\,(\to\mathcal{E}) \qquad \gamma}{\alpha \wedge \gamma}\,(\wedge\mathcal{I})$$

This derivation has three hypotheses, α, γ, and $\alpha \to \beta$, and the conclusion $\gamma \wedge \beta$. It is important to remember that each derivation may have multiple hypotheses, but only a single conclusion.

The following two inference rules are collectively known by the name "and elimination". They are very similar, but not quite the same.

$$\frac{\alpha \wedge \beta}{\alpha}\,(\wedge\mathcal{E}_1) \qquad \frac{\alpha \wedge \beta}{\beta}\,(\wedge\mathcal{E}_2)$$

As you see, an inference rule can have two premises or one premise. We will later see other inference rules that have zero or three premises.

The inference rules that we have seen so far employ only *definite reasoning*: from the hypotheses, we derive the conclusion directly without making any additional assumptions. Another mode of logical reasoning that you are familiar with is *hypothetical reasoning*. In hypothetical reasoning, one makes a *temporary* (or *hypothetical*) assumption in order to explore its consequences. Once the consequences have been satisfactorily explored, the temporary assumption may be ***discarded*** or ***canceled***. An example of hypothetical reasoning is shown in the following proof in the English language:

> We want to show that $\alpha \wedge (\alpha \to \beta) \to \beta$. So **assume** $\alpha \wedge (\alpha \to \beta)$ (we make a temporary assumption). Then α holds, as well as $\alpha \to \beta$. By modus ponens, we get β. **Since we have assumed** $\alpha \wedge (\alpha \to \beta)$, it follows that $\alpha \wedge (\alpha \to \beta) \to \beta$ (at this point, $\alpha \wedge (\alpha \to \beta)$ is no longer a current assumption, and we may go on to proving other things).

In natural deduction, we can introduce a temporary hypothesis at any time. Certain rules allow us to cancel such a temporary hypothesis. It is then put into square brackets to indicate that it is no longer active. An example of this is the following rule for “arrow introduction”:

$$\begin{array}{c}[\alpha]_x\\ \vdots\\ \dfrac{\beta}{\alpha\to\beta}\; x\,(\to\mathcal{I})\end{array}$$

Here, the vertical dots denote an arbitrary derivation of β from hypothesis α (and possibly other hypotheses as well). Notice that α is a temporary hypothesis, which gets canceled when the arrow introduction rule is applied. In order to keep track of when a particular hypothesis was canceled, we decorate it with a lower case letter, such as x, y, z, and we put the same letter next to the rule that canceled the hypothesis. A temporary hypothesis that has been canceled is considered taken care of and is no longer considered a hypothesis of the current overall derivation. The above proof from the English language would be translated as a natural deduction derivation as follows.

$$\dfrac{\dfrac{\dfrac{[\alpha\wedge(\alpha\to\beta)]_x}{\alpha\to\beta}(\wedge\mathcal{E}_2) \qquad \dfrac{[\alpha\wedge(\alpha\to\beta)]_x}{\alpha}(\wedge\mathcal{E}_1)}{\beta}\,(\to\mathcal{E})}{\alpha\wedge(\alpha\to\beta)\to\beta}\; x\,(\to\mathcal{I})$$

As you can see, we may use a temporary hypothesis more than once. In fact, we may use a temporary hypothesis any number of times. We even may use it zero times, as the following derivation demonstrates:

$$\dfrac{\alpha}{\beta\to\alpha}\; y\,(\to\mathcal{I})$$

This is a legal derivation of $\beta\to\alpha$ from α, which makes a hypothetical assumption β and uses it zero times!

2.2 The issue of when to introduce a hypothesis

One issue that is often confusing to beginners when they are first introduced to natural deduction, or indeed to mathematical hypothetical reasoning in general, is the question of when one is “allowed” to make a hypothetical assumption. At

first, it seems that if one were allowed to make just any assumption, then one could prove just any statement simply by assuming it.

In fact, the real issue with hypothetical assumptions is not how to make them, but *how to get rid of them*! You may introduce an assumption at any time, but there are very strict limitations on when you may cancel one. Thus, if you think that you have finished a proof, but you still have unwanted assumptions lying around that you couldn’t cancel, then you simply have not proved what you wanted to prove, at least not from the assumptions that you wanted to prove it from. So as a rule, you should only introduce a hypothetical assumption if you intend to cancel it sometime later.

For instance, suppose you want to prove the formula $(\alpha\to\beta\wedge\gamma)\to\alpha\wedge\beta$. You produce the following derivation:

$$\dfrac{\dfrac{\alpha \qquad \dfrac{\dfrac{[\alpha\to\beta\wedge\gamma]_x \quad \alpha}{\beta\wedge\gamma}\,(\to\mathcal{E})}{\beta}\,(\wedge\mathcal{E}_1)}{\alpha\wedge\beta}\,(\wedge\mathcal{I})}{(\alpha\to\beta\wedge\gamma)\to\alpha\wedge\beta}\; x\,(\to\mathcal{I})$$

Notice that you have made an assumption α which you did not cancel. So you have a proof of $(\alpha\to\beta\wedge\gamma)\to\alpha\wedge\beta$ from hypothesis α. Is this what you wanted, or did you intend to prove $(\alpha\to\beta\wedge\gamma)\to\alpha\wedge\beta$ from no assumptions? In the latter case, you have not succeeded.

As this example demonstrates, extra assumptions do not lead to proofs that are invalid. They simply lead to proofs that are not quite what you wanted.

2.3 The natural deduction system

Definition. The rules of natural deduction are shown in Table 1. We omit the formal definition here; for details, see e.g. van Dalen’s book [**?**].

2.4 Remarks

Most rules of natural deduction, with the exception of the axiom and the (contra) rule, are either introduction ($\mathcal{I}$) or elimination ($\mathcal{E}$) rules. The introduction rules

Table 1: The rules of natural deduction

Axiom

$$\alpha$$

And

$$\frac{\alpha \qquad \beta}{\alpha \wedge \beta}\,(\wedge\mathcal{I}) \qquad\qquad \frac{\alpha \wedge \beta}{\alpha}\,(\wedge\mathcal{E}_1) \qquad \frac{\alpha \wedge \beta}{\beta}\,(\wedge\mathcal{E}_2)$$

Arrow

$$\frac{\begin{array}{c}[\alpha]_x\\ \vdots\\ \beta\end{array}}{\alpha \to \beta}\,x\,(\to\mathcal{I}) \qquad\qquad \frac{\beta \to \alpha \qquad \beta}{\alpha}\,(\to\mathcal{E})$$

Not

$$\frac{\begin{array}{c}[\alpha]_x\\ \vdots\\ \bot\end{array}}{\neg\,\alpha}\,x\,(\neg\,\mathcal{I}) \qquad\qquad \frac{\neg\,\alpha \qquad \alpha}{\bot}\,(\neg\,\mathcal{E})$$

Top and **Bot**

$$\frac{}{\top}\,(\top\mathcal{I}) \qquad\qquad \frac{\begin{array}{c}\vdots\\ \bot\end{array}}{\alpha}\,(\bot\mathcal{E})$$

Or

$$\frac{\alpha}{\alpha \vee \beta}\,(\vee\mathcal{I}_1) \qquad \frac{\beta}{\alpha \vee \beta}\,(\vee\mathcal{I}_2) \qquad\qquad \frac{\alpha \vee \beta \qquad \begin{array}{c}[\alpha]_x\\ \vdots\\ \gamma\end{array} \qquad \begin{array}{c}[\beta]_y\\ \vdots\\ \gamma\end{array}}{\gamma}\,x,y\,(\vee\mathcal{E})$$

Contra

$$\frac{\begin{array}{c}[\neg\,\alpha]_x\\ \vdots\\ \bot\end{array}}{\alpha}\,x\,(\text{contra})$$

introduce a connective, and the elimination rules eliminate one. Note, however, that there is no rule for $\bot$ introduction and no rule for $\top$ elimination.

Some of the rules have traditional names. We have already mentioned that the $(\to\mathcal{E})$ rule is known as *modus ponens*. The $(\bot\mathcal{E})$ rule is known as *absurdity* or *ex falsum quodlibet* (from a falsity, what you want). An application of the $(\vee\mathcal{E})$ rule is called a *case distinction*, and an application of the (contra) rule is called a *proof by contradiction*.

If one omits the (contra) rule from the system of natural deduction, one obtains a logic that is called *intuitionistic logic*. It has many applications in computer science, but we do not consider it further in this course. If we allow the (contra) rule, the resulting logic is called *classical logic*. As we will see soon, classical logic corresponds precisely to our boolean truth table semantics, in the sense that a formula is derivable in classical logic if and only if it is valid in the truth table semantics.

Notice the difference between the $(\neg\,\mathcal{I})$ rule and the (contra) rule. Since $\neg\,\neg\,\alpha$ is a different formula from α, these rules are not derivable from each other.

Also notice that in the presence of the (contra) rule, the $(\bot\mathcal{E})$ rule is actually redundant: it corresponds to an application of the (contra) rule in which zero hypotheses are canceled.

The axiom seems strange at first. Why would one need a proof of α from hypothesis α? One example where this is useful is illustrated in the following derivation. Notice that in the first case of the case distinction, we already have the formula α as a hypothesis, and we need α as a conclusion in order to apply the $(\vee\mathcal{E})$ rule.

$$\frac{\alpha \vee (\alpha \wedge \beta) \qquad [\alpha]_x \qquad \dfrac{[\alpha \wedge \beta]_y}{\alpha}\,(\wedge\mathcal{E}_1)}{\alpha}\,x,y\,(\vee\mathcal{E})$$

We have not introduced any rules for "if and only if", or the connective $\leftrightarrow$. One could easily add such rules to the system:

Arrow

$$\frac{\begin{array}{c}[\alpha]_x\\ \vdots\\ \beta\end{array} \qquad \begin{array}{c}[\beta]_y\\ \vdots\\ \alpha\end{array}}{\alpha \leftrightarrow \beta}\,(\leftrightarrow\mathcal{I}) \qquad\qquad \frac{\alpha \leftrightarrow \beta \qquad \alpha}{\beta}\,(\leftrightarrow\mathcal{E}_1) \qquad \frac{\alpha \leftrightarrow \beta \qquad \beta}{\alpha}\,(\leftrightarrow\mathcal{E}_2)$$

However, we do not afford ourselves this luxury, since it is easily seen that these rules are derivable if we regard $\alpha \leftrightarrow \beta$ as an abbreviation for $(\alpha \to \beta) \wedge (\beta \to \alpha)$.

2.5 Examples

For examples of derivations, see the solutions to Problem Set 3, first part.

3 Soundness

Definition. If Σ is a set of formulas, and α a formula, we write $\Sigma \vdash \alpha$ if there is a derivation with conclusion α whose active hypotheses are contained in Σ. In this case, we say that α is ***derivable*** from Σ or that Σ ***entails*** α. We say that α is a ***theorem*** of natural deduction if $\vdash \alpha$, i.e., if there is a derivation of α from no hypotheses.

Notice that entailment $\Sigma \vdash \alpha$ is defined in terms of derivations, whereas tautological implication $\Sigma \models \alpha$, from Section 1, is defined in terms of truth tables. The following soundness theorem shows that the inference rules of natural deduction are correct with respect to truth table semantics, i.e., they do not derive any false conclusions.

Theorem 1 (Soundness)**.** $\Sigma \vdash \alpha \Rightarrow \Sigma \models \alpha$.

Proof. We need to prove the following statement: if $\mathscr{D}$ is a derivation of α from hypotheses Σ, then $\Sigma \models \alpha$. We prove this by complete induction on the size of the derivation $\mathscr{D}$, and by case distinction on the last rule used in $\mathscr{D}$.

Here, the size of a derivation is, per definition, the number of rule applications (i.e., horizontal lines) that occur in it. Notice that any derivation $\mathscr{D}$ of size $\geqslant 1$ has a unique last rule, whose conclusion is the conclusion of $\mathscr{D}$. In the following, we write

$$\begin{array}{c}\Sigma \\ \vdots \\ \alpha\end{array}$$

to denote a derivation of α whose hypotheses are contained in the set Σ. When using this notation, it is understood that not all formulas in Σ may actually be used.

3.1 Case: Axiom

Suppose the derivation $\mathscr{D}$ consists of an axiom only, i.e., of a single formula α. Since the hypotheses of $\mathscr{D}$ are contained in Σ, we have $\alpha \in \Sigma$. It follows that $\Sigma \models \alpha$, as desired.

3.2 Case: $(\wedge\mathcal{I})$

Suppose the last rule used in $\mathscr{D}$ is $(\wedge\mathcal{I})$. Then $\mathscr{D}$ is of the form

$$\dfrac{\begin{array}{c}\Sigma \\ \vdots \\ \alpha\end{array} \qquad \begin{array}{c}\Sigma \\ \vdots \\ \beta\end{array}}{\alpha \wedge \beta}\,(\wedge\mathcal{I})$$

Since the subderivations are of smaller size, by induction hypothesis, $\Sigma \models \alpha$ and $\Sigma \models \beta$. It follows that $\Sigma \models \alpha \wedge \beta$.

3.3 Case: $(\wedge\mathcal{E}_1)$

Suppose the last rule used in $\mathscr{D}$ is $(\wedge\mathcal{E}_1)$. Then $\mathscr{D}$ is of the form

$$\dfrac{\begin{array}{c}\Sigma \\ \vdots \\ \alpha \wedge \beta\end{array}}{\alpha}\,(\wedge\mathcal{E}_1)$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma \models \alpha \wedge \beta$. It follows that $\Sigma \models \alpha$.

3.4 Case: $(\wedge\mathcal{E}_2)$

Similar to the previous case.

3.5 Case: $(\to\mathcal{I})$

Suppose the last rule used in $\mathscr{D}$ is $(\to\mathcal{I})$. Then $\mathscr{D}$ is of the form

$$\begin{array}{c} \Sigma, [\alpha]_x \\ \vdots \\ \dfrac{\beta}{\alpha \to \beta}\; x\,(\to\mathcal{I}) \end{array}$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma, \alpha \models \beta$. It follows from Problem 1.3.3(a) that $\Sigma \models \alpha \to \beta$.

3.6 Case: $(\to\mathcal{E})$

Suppose the last rule used in $\mathscr{D}$ is $(\to\mathcal{E})$. Then $\mathscr{D}$ is of the form

$$\dfrac{\begin{array}{c} \Sigma \\ \vdots \\ \alpha \to \beta \end{array} \qquad \begin{array}{c} \Sigma \\ \vdots \\ \alpha \end{array}}{\beta}\;(\to\mathcal{E})$$

Since the subderivations are of smaller size, by induction hypothesis, $\Sigma \models \alpha \to \beta$ and $\Sigma \models \alpha$. It follows that $\Sigma \models \beta$.

3.7 Case: $(\neg\mathcal{I})$

Suppose the last rule used in $\mathscr{D}$ is $(\neg\mathcal{I})$. Then $\mathscr{D}$ is of the form

$$\begin{array}{c} \Sigma, [\alpha]_x \\ \vdots \\ \dfrac{\bot}{\neg\,\alpha}\; x\,(\neg\mathcal{I}) \end{array}$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma, \alpha \models \bot$. Thus, no truth assignment that satisfies all formulas in Σ satisfies α. It follows that $\Sigma \models \neg\,\alpha$.

3.8 Case: $(\top\mathcal{I})$

Suppose the last rule used in $\mathscr{D}$ is $(\top\mathcal{I})$. Then $\mathscr{D}$ is

$$\dfrac{}{\top}\;(\top\mathcal{I})$$

Any truth assignment satisfies $\top$, and thus $\Sigma \models \top$, as desired.

3.9 Case: $(\bot\mathcal{E})$

Suppose the last rule used in $\mathscr{D}$ is $(\bot\mathcal{E})$. Then $\mathscr{D}$ is of the form

$$\begin{array}{c} \Sigma \\ \vdots \\ \dfrac{\bot}{\alpha}\;(\bot\mathcal{E}) \end{array}$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma \models \bot$. Thus, no truth assignment satisfies all the formulas of Σ, and it follows that $\Sigma \models \alpha$.

3.10 Case: $(\vee\mathcal{I}_1)$

Suppose the last rule used in $\mathscr{D}$ is $(\vee\mathcal{I}_1)$. Then $\mathscr{D}$ is of the form

$$\begin{array}{c} \Sigma \\ \vdots \\ \dfrac{\alpha}{\alpha \vee \beta}\;(\vee\mathcal{I}_1) \end{array}$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma \models \alpha$. It follows that $\Sigma \models \alpha \vee \beta$.

3.11 Case: $(\vee\mathcal{I}_2)$

Similar to the previous case.

3.12 Case: $(\vee\mathcal{E})$

Suppose the last rule used in $\mathscr{D}$ is $(\vee\mathcal{E})$. Then $\mathscr{D}$ is of the form

$$\dfrac{\begin{matrix}\Sigma \\ \vdots \\ \alpha\vee\beta\end{matrix}\qquad\begin{matrix}\Sigma,[\alpha]_x \\ \vdots \\ \gamma\end{matrix}\qquad\begin{matrix}\Sigma,[\beta]_y \\ \vdots \\ \gamma\end{matrix}}{\gamma}\ {}_{x,y}\,(\vee\mathcal{E})$$

Since the subderivations are of smaller size, by induction hypothesis, $\Sigma \models \alpha\vee\beta$, $\Sigma,\alpha \models \gamma$, and $\Sigma,\beta \models \gamma$. Thus, if v is any truth assignment satisfying Σ, then v satisfies either α or β. In either case, v satisfies γ. It follows that $\Sigma \models \gamma$.

3.13 Case: (contra)

Suppose the last rule used in $\mathscr{D}$ is (contra). Then $\mathscr{D}$ is of the form

$$\begin{matrix}\Sigma,[\neg\,\alpha]_x \\ \vdots \\ \dfrac{\bot}{\alpha}\ {}_x\,\text{(contra)}\end{matrix}$$

Since the subderivation is of smaller size, by induction hypothesis, $\Sigma,\neg\,\alpha \models \bot$. Thus, no truth assignment that satisfies all formulas in Σ satisfies $\neg\,\alpha$. It follows that $\Sigma \models \alpha$.

This finishes the proof of soundness. $\square$

Corollary 2. *Every theorem of natural deduction is a tautology.*

Proof. By taking $\Sigma = \emptyset$ in the Soundness Theorem. $\square$

4 Completeness

Completeness is the converse of soundness. While soundness states that the rules of natural deduction are correct, completeness states that there are enough such rules to prove every possible tautological implication. For the purpose of completeness, we assume that our language does not contain the connective $\leftrightarrow$, since we have not included that connective in our natural deduction formalism. In light of the remarks in Section 2.4, this is not a significant restriction.

Theorem 3 (Completeness)**.** $\Sigma \models \alpha \Rightarrow \Sigma \vdash \alpha$.

The proof of the Completeness Theorem will require some preliminaries.

Definition. A set of formulas Σ is ***inconsistent*** if $\Sigma \vdash \bot$. Otherwise, Σ is ***consistent***.

We say that Σ is ***satisfiable*** if there exists a truth assignment v such that for all $\sigma \in \Sigma$, $\bar{v}(\sigma) = T$. If there is no such truth assignment, then Σ is ***unsatisfiable***.

Note that consistency is a concept related to derivations, whereas satisfiability is a concept related to truth assignments.

Lemma 4. *Σ is unsatisfiable if and only if $\Sigma \models \bot$.*

Proof. By definition of tautological implication, we have $\Sigma \models \bot$ iff for all truth assignments v that satisfy all the formulas in Σ, $\bar{v}(\bot) = T$. But no truth assignment satisfies $\bar{v}(\bot) = T$, and thus $\Sigma \models \bot$ iff there is no truth assignment satisfying all the formulas in Σ, which is the case iff Σ is unsatisfiable. $\square$

Lemma 5. *The following are equivalent:*

(a) For all Σ,α, the implication $\Sigma \models \alpha \Rightarrow \Sigma \vdash \alpha$ holds (completeness),

(b) For all Σ, the implication $\Sigma \models \bot \Rightarrow \Sigma \vdash \bot$ holds,

(c) Any consistent set of formulas is satisfiable.

Proof. (b) and (c) are clearly equivalent, since (b) just states that any unsatisfiable set of formulas is consistent, by Lemma 4. Also, clearly (a) implies (b), by taking $\alpha = \bot$. It remains to show that (b) implies (a).

Assume (b), and assume $\Sigma \models \alpha$. Then any truth assignment which satisfies all formulas in Σ satisfies α. Thus, no such truth assignment satisfies $\neg\,\alpha$, and we have $\Sigma,\neg\,\alpha \models \bot$. By (b), this implies $\Sigma,\neg\,\alpha \vdash \bot$. Thus, there is a derivation of

$\bot$ from hypotheses in Σ and $\neg\,\alpha$. Apply the (contra) rule to construct the derivation

$$\begin{array}{c} \Sigma, [\neg\,\alpha]_x \\ \vdots \\ \dfrac{\bot}{\alpha}\; x\ (\text{contra}) \end{array}$$

This proves $\Sigma \vdash \alpha$, as desired. □

We will prove the Completeness Theorem by proving condition (c) in Lemma 5. Thus, when given a consistent set Σ of formulas, we must show that it is satisfiable. This means, we must produce a truth assignment v which satisfies every formula in Σ. In general, it would be difficult to guess such a truth assignment. One would be tempted to define

$$\begin{array}{ll} v(\mathbf{A}_n) = T & \text{if } \Sigma \vdash \mathbf{A}_n, \\ v(\mathbf{A}_n) = F & \text{if } \Sigma \vdash \neg\,\mathbf{A}_n. \end{array}$$

The problem is that, in general, neither $\mathbf{A}_n$ nor $\neg\,\mathbf{A}_n$ needs to be entailed by Σ, so this approach fails. However, if Σ is a *maximally* consistent set of formulas, then Σ entails every formula or its negation. Thus, the crucial idea in the proof of the Completeness Theorem is to prove first that any consistent set of formulas is contained in a maximally consistent set, and then to use the above method for defining v.

Definition. A set Σ of formulas is called ***maximally consistent*** if

(i) Σ is consistent,

(ii) if $\Sigma \subseteq \Sigma'$ and Σ' is consistent, then $\Sigma = \Sigma'$.

Lemma 6. *If Σ is a maximally consistent set, then it is closed under derivability, i.e., $\Sigma \vdash \alpha$ implies $\alpha \in \Sigma$.*

Proof. Suppose Σ is maximally consistent, and $\Sigma \vdash \alpha$. We claim that $\Sigma \cup \{\alpha\}$ is consistent. For otherwise, one would have $\Sigma, \alpha \vdash \bot$, and then then Σ would be

inconsistent by the following derivation:

$$\dfrac{\dfrac{\begin{array}{c}\Sigma, [\alpha]_x \\ \vdots \\ \bot\end{array}}{\neg\,\alpha}\; x\,(\neg\,\mathcal{I}) \qquad \begin{array}{c}\Sigma \\ \vdots \\ \alpha\end{array}}{\bot}\;(\neg\,\mathcal{E}),$$

But this contradicts the assumption that Σ was consistent. It follows that $\Sigma \cup \{\alpha\}$ is consistent. But since Σ is maximally consistent, we have $\Sigma = \Sigma \cup \{\alpha\}$, and thus $\alpha \in \Sigma$. □

Lemma 7. *Suppose Σ is a maximally consistent set. Then the following hold:*

(1) $\neg\,\alpha \in \Sigma$ iff $\alpha \notin \Sigma$.

(2) $\alpha \wedge \beta \in \Sigma$ iff $\alpha \in \Sigma$ and $\beta \in \Sigma$.

(3) $\alpha \vee \beta \in \Sigma$ iff $\alpha \in \Sigma$ or $\beta \in \Sigma$.

(4) $\alpha \to \beta \in \Sigma$ iff $\alpha \notin \Sigma$ or $\beta \in \Sigma$.

(5) $\top \in \Sigma$.

(6) $\bot \notin \Sigma$.

Proof. (1) Clearly, α and $\neg\,\alpha$ cannot both be in Σ, or else Σ would be inconsistent. It remains to show that at least one of them is in Σ. Consider the set $\Sigma' = \Sigma \cup \{\alpha\}$. Either Σ' is consistent, in which case $\alpha \in \Sigma$ by maximality of Σ. Or Σ' is inconsistent, in which case $\Sigma, \alpha \vdash \bot$. By applying the $(\neg\,\mathcal{I})$, we get $\Sigma \vdash \neg\,\alpha$, and hence $\neg\,\alpha \in \Sigma$ by Lemma 6.

(2) "⇒" If $\alpha \wedge \beta \in \Sigma$, then $\Sigma \vdash \alpha$ and $\Sigma \vdash \beta$ by the $(\wedge\mathcal{E})$ rules. Thus, $\alpha \in \Sigma$ and $\beta \in \Sigma$ by Lemma 6. "⇐" If $\alpha \in \Sigma$ and $\beta \in \Sigma$, then $\Sigma \vdash \alpha \wedge \beta$ by the $(\wedge\mathcal{I})$ rule. Thus, $\alpha \wedge \beta \in \Sigma$ by Lemma 6.

(3) "⇒" Suppose $\alpha \vee \beta \in \Sigma$. Suppose, for the sake of contradiction, that neither $\alpha \in \Sigma$ nor $\beta \in \Sigma$. Then $\neg\,\alpha \in \Sigma$ and $\neg\,\beta \in \Sigma$ by (1). Then the following derivation shows that Σ is inconsistent:

$$\dfrac{\alpha \vee \beta \qquad \dfrac{\neg\,\alpha \quad [\alpha]_x}{\bot}(\neg\,\mathcal{E}) \qquad \dfrac{\neg\,\beta \quad [\beta]_y}{\bot}(\neg\,\mathcal{E})}{\bot}\; x,y\,(\vee\mathcal{E})$$

This contradicts the assumption that Σ was consistent. Thus, $\alpha \in \Sigma$ or $\beta \in \Sigma$, as desired. "$\Leftarrow$" If $\alpha \in \Sigma$ or $\beta \in \Sigma$, then $\Sigma \vdash \alpha \vee \beta$ by one of the $(\vee\mathcal{I})$ rules. Thus, $\alpha \vee \beta \in \Sigma$ by Lemma 6.

(4) "$\Rightarrow$" Suppose $\alpha \to \beta \in \Sigma$. If $\alpha \in \Sigma$, then $\Sigma \vdash \beta$ by the $(\to\mathcal{E})$ rule, and thus $\beta \in \Sigma$ by Lemma 6. So either $\alpha \notin \Sigma$ or $\beta \in \Sigma$. "$\Leftarrow$" Case 1: Suppose $\alpha \notin \Sigma$. Then $\neg\alpha \in \Sigma$ by (1). The following derivation shows that $\Sigma \vdash \alpha \to \beta$:

$$\dfrac{\dfrac{\dfrac{\neg\alpha \qquad [\alpha]_x}{\bot}\,(\neg\mathcal{E})}{\beta}\,(\bot\mathcal{E})}{\alpha \to \beta}\,x\,(\to\mathcal{I})$$

Thus, $\alpha \to \beta \in \Sigma$ by Lemma 6. Case 2: Suppose $\beta \in \Sigma$. By $(\to\mathcal{I})$, we have $\Sigma \vdash \alpha \to \beta$, and thus $\alpha \to \beta \in \Sigma$ by Lemma 6.

(5) Suppose, for the sake of contradiction, that $\Sigma \cup \top$ is inconsistent, i.e., $\Sigma, \top \vdash \bot$. Then the following derivation shows that Σ is already inconsistent:

$$\dfrac{\dfrac{\begin{matrix}\Sigma, [\top]_x \\ \vdots \\ \bot\end{matrix}}{\neg\top}\,x\,(\neg\mathcal{I}) \qquad \dfrac{}{\top}\,(\top\mathcal{I})}{\bot}\,(\neg\mathcal{E})$$

This contradicts the assumption that Σ was consistent. Thus, $\Sigma \cup \top$ is consistent. By maximality of Σ, we have $\top \in \Sigma$.

(6) $\bot \notin \Sigma$ because Σ is consistent. $\square$

Lemma 8. *Every maximally consistent set is satisfiable.*

Proof. Let Σ be a maximally consistent set. Define a truth assignment v by

$$\begin{array}{ll} v(\mathbf{A}_n) = T & \text{if } \mathbf{A}_n \in \Sigma, \\ v(\mathbf{A}_n) = F & \text{if } \mathbf{A}_n \notin \Sigma. \end{array}$$

We claim that for all formulas α, one has $\bar{v}(\alpha) = T$ iff $\alpha \in \Sigma$. In particular, v satisfies all the formulas in Σ, and thus Σ is satisfiable. We prove the claim by induction on α.

Base Case: In case α is a sentence symbol, then $\bar{v}(\alpha) = T$ iff $\alpha \in \Sigma$ by definition of v. In case $\alpha = \top$, then $\bar{v}(\alpha) = T$ and $\alpha \in \Sigma$ by Lemma 7(5). In case $\alpha = \bot$, then $\bar{v}(\alpha) = F$ and $\alpha \notin \Sigma$ by Lemma 7(6).

Induction Step: In case $\alpha = \neg\beta$, then $\bar{v}(\alpha) = T$ iff $\bar{v}(\beta) = F$, which is the case, by induction hypothesis, iff $\beta \notin \Sigma$. By Lemma 7(1), this happens iff $\alpha = \neg\beta \in \Sigma$. In case $\alpha = \beta \wedge \gamma$, then $\bar{v}(\alpha) = T$ iff $\bar{v}(\beta) = T$ and $\bar{v}(\gamma) = T$, which is the case, by induction hypothesis, iff $\beta \in \Sigma$ and $\gamma \in \Sigma$. By Lemma 7(2), this happens iff $\alpha = \beta \wedge \gamma \in \Sigma$. In case $\alpha = \beta \vee \gamma$, then $\bar{v}(\alpha) = T$ iff $\bar{v}(\beta) = T$ or $\bar{v}(\gamma) = T$, which is the case, by induction hypothesis, iff $\beta \in \Sigma$ or $\gamma \in \Sigma$. By Lemma 7(3), this happens iff $\alpha = \beta \vee \gamma \in \Sigma$. In case $\alpha = \beta \to \gamma$, then $\bar{v}(\alpha) = T$ iff $\bar{v}(\beta) = F$ or $\bar{v}(\gamma) = T$, which is the case, by induction hypothesis, iff $\beta \notin \Sigma$ or $\gamma \in \Sigma$. By Lemma 7(4), this happens iff $\alpha = \beta \to \gamma \in \Sigma$. $\square$

The following lemma is crucial to the proof of completeness. It makes the connection between consistent sets and maximally consistent sets.

Lemma 9. *Every consistent set Σ of formulas is contained in some maximally consistent set Σ_*.*

Proof. Since our language is made up from finite strings of countably many symbols, there are countably many well-formed formulas. Make a list $\varphi_0, \varphi_1, \varphi_2, \ldots$ of all formulas. We define a sequence of sets $\Sigma_0, \Sigma_1, \Sigma_2, \ldots$ as follows: $\Sigma_0 = \Sigma$, and for any n,

$$\Sigma_{n+1} = \begin{cases} \Sigma_n \cup \{\varphi_n\} & \text{if } \Sigma_n \cup \{\varphi_n\} \text{ consistent,} \\ \Sigma_n & \text{otherwise.} \end{cases}$$

By construction, $\Sigma_n \subseteq \Sigma_{n+1}$ for all n. Also, an easy induction shows that Σ_n is consistent for all n.

Define $\Sigma_* = \bigcup_{n \in \mathbb{N}} \Sigma_n$. We claim that Σ_* is consistent. Suppose it is not. Then $\Sigma_* \vdash \bot$. Thus, there exists some derivation of $\bot$ from hypotheses in Σ_*. This derivation has finitely many hypotheses, say, $\sigma_0, \sigma_1, \ldots, \sigma_k$. Each σ_i is an element of some Σ_{n_i}. Let $n = \max\{n_0, \ldots, n_k\}$, then $\sigma_0, \sigma_1, \ldots, \sigma_k \in \Sigma_n$. It follows that $\Sigma_n \vdash \bot$, contradicting the fact that Σ_n is consistent. Thus, it follows that Σ_* is consistent.

Clearly, $\Sigma \subseteq \Sigma_*$. It remains to show that Σ_* is maximally consistent. Consider any formula φ which is not in Σ_*. Then φ is one of the formulas in our list, say $\varphi = \varphi_i$. Since $\varphi \notin \Sigma_*$, it follows that $\varphi_i \notin \Sigma_{i+1}$, and thus, by definition of Σ_{i+1}, $\Sigma_{i+1} \cup \{\varphi_i\}$ must be inconsistent. It follows that $\Sigma_* \cup \{\varphi_i\}$ is inconsistent.

Thus, no consistent set properly contains Σ_*, which proves that Σ_* is maximally consistent. $\square$

Proof of the Completeness Theorem: It suffices to show, by Lemma 5, that any consistent set is satisfiable. So let Σ be an arbitrary consistent set. By Lemma 9, Σ is contained in a maximally consistent set Σ_*. By Lemma 8, there exists a truth assignment v satisfying all the formulas in Σ_*. Since Σ is a subset of Σ_*, v also satisfies all the formulas in Σ, showing that Σ is satisfiable, as desired. $\square$