CHAPTER 1. SENTENTIAL LOGIC

1. Introduction

In sentential logic we study how given sentences may be combined to form more complicated compound sentences. For example, from the sentences *7 is prime*, *7 is odd*, *2 is prime*, *2 is odd*, we can obtain the following sentences:

- *7 is prime* and *7 is odd*
- *7 is odd* and *2 is odd*
- *7 is odd* or *2 is odd*
- *7 is* not *odd*
- If *2 is odd* then *2 is prime*

Of course, this process can be iterated as often as we want, obtaining also:

- If *7 is* not *odd* then *2 is odd*
- If *7 is odd* and *2 is odd* then *2 is* not *prime*
- (*7 is odd* and *2 is odd*) or *2 is prime*

We have improved on English in the last example by using parentheses to resolve an ambiguity.

And, or, not, if ... then (or implies) are called *(sentential) connectives*. Using them we can define more connectives, for example

2 is prime iff *2 is odd*

can be defined as

(If *2 is prime* then *2 is odd*) and (if *2 is odd* then *2 is prime*).

The truth value of any compound sentence is determined completely by the truth values of its component parts. For example, assuming *2, 7, odd, prime* all have their usual meanings then *7 is odd* and *2 is odd* is false but (*7 is odd* and *2 is odd*) or *2 is prime* is true. We will discuss implication later.

In the formal system of sentential logic we study in this Chapter, we do not use English sentences, but build the compound sentences from an infinite collection of symbols which we think of as referring to sentences — since these are not built up from other sentences we refer to them as *atomic sentences*. The truth values of the compound sentences formed are then functions of the truth values assigned to these atomic sentences. We also use symbols in place of the English names of the connectives.

2. Connectives and Sentences of Sentential Logic

We use symbols for the sentential connectives as in the following table.

English word	symbol	Name
and	$\wedge$	conjunction
or	$\vee$	disjunction
implies	$\rightarrow$	implication
not	$\neg$	negation

To define a formal system we need to specify the *symbols* of the system and the *rules* for constructing meaningful sequences of symbols (sentences, in this case). We refer to the formal system of sentential logic as $\mathcal{S}$.

Definition 2.1. The *symbols* of $\mathcal{S}$ are as follows:

(i) An infinite set of atomic sentences: $S_1, \ldots, S_k, \ldots$ for all $k \in \mathbb{N}$
(ii) the sentential connectives: $\wedge, \vee, \rightarrow, \neg$
(iii) parentheses: (,)

Definition 2.2. The *sentences* of $\mathcal{S}$ are defined as follows:

(i) S_k is a sentence for every $k \in \mathbb{N}$.
(ii) If φ is a sentence then so is $\neg\varphi$.
(iii) If φ and ψ are sentences then so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, and $(\varphi \rightarrow \psi)$.
(iv) Nothing else is a sentence.

So every sentence is a finite sequence of symbols, but not every finite sequence of symbols is a sentence. To show a sequence is a sentence we must show how it is formed from atomic sentences by repeated applications of rules (ii) and (iii) of the definition. For example

$$((S_3 \wedge \neg S_1) \rightarrow S_4) \text{ is a sentence}$$

since it can be constructed as follows:

$$S_1,\ S_3,\ S_4,\ \neg S_1, (S_3 \wedge \neg S_1),\ ((S_3 \wedge \neg S_1) \rightarrow S_4).$$

Such a sequence is called a *history* of the sentence. Histories are not usually unique since the order of sentences in the history is not completely determined.

Two sentences are *equal* iff they are the same sequence, and the *length* of a sentence is its length as a sequence. Note that the atomic sentences are the only sentences of length 1, the sentence $\neg\varphi$ is longer than φ, and similarly for $(\varphi \vee \psi)$, etc.

Note, for example, that $(S_3 \wedge \neg S_1) \rightarrow S_4$ is *not* a sentence, since it is missing the outer pair of parentheses, but in writing examples we will frequently omit the outer parentheses without fear of confusion.

Also, it is seldom important whether a sentence is built from the atomic sentences S_1, S_3, and S_4 or some other choice of three atomic sentences. We will use $A, B, C, \ldots$ (perhaps with subscripts or superscripts) to stand for arbitrary atomic sentences (assumed distinct unless explicitly noted to the contrary). Thus we will frequently refer to $(A \wedge \neg B) \rightarrow C$ rather than $((S_3 \wedge \neg S_1) \rightarrow S_4)$

3. Truth Assignments

In this section we show how to assign a unique *truth value* to every sentence given an assignment of truth values to the atomic sentences used in constructing the sentence. We use T for "true" and F for "false".

Definition 3.1. Let $\mathcal{A}$ be a set of atomic sentences. A *truth assignment* for $\mathcal{A}$ is any function $h : \mathcal{A} \to \{T, F\}$.

For any set $\mathcal{A}$ of atomic sentences we let $\overline{\mathcal{A}}$ be the set of all sentences containing only atomic sentences from $\mathcal{A}$. The following theorem states that every truth assignment for $\mathcal{A}$ extends uniquely to a function assigning truth values to all sentences in $\overline{\mathcal{A}}$ which respects the intuitive meaning of the connectives.

Theorem 3.1. *Let $\mathcal{A}$ be a set of atomic sentences and let h be a truth assignment for $\mathcal{A}$. The there is exactly one function $\bar{h} : \overline{\mathcal{A}} \to \{T, F\}$ satisfying the following:*

(i) $\bar{h}(A) = h(A)$ *for all* $A \in \mathcal{A}$,
(ii) $\bar{h}(\neg\varphi) = T$ *iff* $\bar{h}(\varphi) = F$,
(iii) $\bar{h}(\varphi \wedge \psi) = T$ *iff* $\bar{h}(\varphi) = \bar{h}(\psi) = T$,
(iv) $\bar{h}(\varphi \vee \psi) = T$ *iff* $\bar{h}(\varphi) = T$ *or* $\bar{h}(\psi) = T$ *(or both)*,
(v) $\bar{h}(\varphi \to \psi) = F$ *iff* $\bar{h}(\varphi) = T$ *and* $\bar{h}(\psi) = F$.

Note in particular the truth values assigned to disjunctions and implications.

We illustrate the proof of 3.1 with an example. Suppose that $\mathcal{A} = \{S_1, S_3, S_4\}$ and $h(A) = F$ for all $A \in \mathcal{A}$. Consider the sentence $\theta = (S_3 \wedge \neg S_1) \to S_4$. Then

$$S_1, S_3, S_4, \neg S_1, (S_3 \wedge \neg S_1), (S_3 \wedge \neg S_1) \to S_4$$

is a history of θ, and we see that any $\bar{h}$ satisfying the clauses in 3.1 *must* assign the truth values F, F, F, T, F, T to the sentences in the history.

But is it clear that every history of θ must yield the same result? In this example, yes, but in general you need to prove the uniqueness of histories, up to the order of the terms. To do this rigorously would require us to prove that our system has no ambigous sentences, such as $S_1 \wedge S_2 \vee S_3$, which can be read in different ways. This require an inductive proof (see section 7).

If h is a truth assignment for the set of all atomic sentences, we will simply refer to h as a truth assignment. Thus if h is a truth assignment then $\bar{h}(\theta)$ is defined for all sentences θ. The following fact is clear.

Lemma 3.1. *Let h_1 and h_2 be truth assignments, and assume that $h_1(A) = h_2(A)$ for all $A \in \mathcal{A}$. Then $\overline{h_1}(\theta) = \overline{h_2}(\theta)$ for all $\theta \in \bar{\mathcal{A}}$.*

This Lemma enables us to just speak of truth assignments, rather than truth assignments for some $\mathcal{A}$, in what follows.

Definition 3.2. Let h be a truth assignment and let θ be a sentence. Then h *satisfies* θ, or θ is *true* under h, iff $\bar{h}(\theta) = T$ where $\bar{h}$ is the unique extension of h given in Theorem 3.1. We write $h \models \theta$ (also read h *models* θ) if this occurs.

Note that h does not satisfy θ, $h \not\models \theta$, iff $h \models \neg\theta$, by (ii) of Theorem 3.1.

We extend the satisfaction terminology and notation to sets of sentences in the obvious way.

Definition 3.3. Let h be a truth assignment and let Σ be a set of sentences. Then h *satisfies* or *models* Σ, written $h \models \Sigma$, iff $h \models \theta$ for all $\theta \in \Sigma$.

Note, however, that $h \not\models \Sigma$ iff $h \not\models \theta$ for *some* $\theta \in \Sigma$.

4. Tautologies, Satisfiability, Truth Tables

We can now define what it means for a sentence (of $\mathcal{S}$) to be *logically true* or *valid*. This means that the sentence is always true. We thus obtain the following definition.

Definition 4.1. The sentence θ is *valid*, or a *tautology*, iff $h \models \theta$ for every truth assignment h. We write $\models \theta$ to mean that θ is valid.

Definition 4.2. The sentence θ is *satisfiable* iff $h \models \theta$ for some truth assignment h.

The following important fact enables us to characterize satisfiability in terms of validity and vice versa.

Lemma 4.1. *The sentence θ is satisfiable iff $\neg\theta$ is not valid; θ is valid iff $\neg\theta$ is not satisfiable.*

Proof. θ is satisfiable iff there is some h such that $h \models \theta$ iff there is some h such that $h \not\models \neg\theta$ iff $\neg\theta$ is not valid. □

How could we determine whether or not a sentence is valid? Given θ let $\mathcal{A}$ be the set of atomic sentences occuring in θ. By Lemma 3.1 $\models \theta$ iff $h \models \theta$ for all truth assignments h for $\mathcal{A}$. But $\mathcal{A}$ is finite, hence there are only finitely many such assignments to check.

The information needed to check whether or not θ is a tautology is conveniently represented in a *truth table*. Across the top of the table you write a history of θ, beginning with the atomic sentences in θ, and each line of the table correcponds to a different assignment of truth values to these atomic sentences.

For example, the following truth table shows that $(S_3 \wedge \neg S_1) \rightarrow S_4$ is not a tautology.

S_1	S_3	S_4	$\neg S_1$	$S_3 \wedge \neg S_1$	$(S_3 \wedge \neg S_1) \rightarrow S_4$
T	T	T	F	F	T
T	T	F	F	F	T
T	F	T	F	F	T
T	F	F	F	F	T
F	T	T	T	T	T
F	T	F	T	T	F
F	F	T	T	F	T
F	F	F	T	F	T

Writing down truth tables quickly becomes tedious. Frequently shortcuts can be used to reduce the amount of work needed. For example, to determine whether or not θ is a tautology, suppose that $\bar{h}(\theta) = F$ and work backwards to see what h must be. For example, if $\theta = (S_3 \wedge \neg S_1) \to S_4$ we obtain the following equivalences:

$$\begin{gathered}\bar{h}((S_3 \wedge \neg S_1) \to S_4) = F \\ \text{iff } \bar{h}(S_3 \wedge \neg S_1) = T \text{ and } h(S_4) = T \\ \text{iff } h(S_3) = T,\ h(S_1) = F \text{ and } h(S_4) = F.\end{gathered}$$

Thus this sentence is not a tautology since the last line gives an assignment making it false.

As another example consider $\theta = A \to ((A \to B) \to B)$.

$$\begin{gathered}\bar{h}(\theta) = F \\ \text{iff } h(A) = T \text{ and } \bar{h}((A \to B) \to B) = F \\ \text{iff } h(A) = T,\ \bar{h}(A \to B) = T \text{ and } h(B) = F.\end{gathered}$$

But if $h(A) = T$ and $h(B) = F$ then $\bar{h}(A \to B) = F$, so there can be no such h, so the sentence is a tautology.

Essentially the same argument shows, more generally, that

$$\models \varphi \to ((\varphi \to \psi) \to \psi)$$

for all sentences φ, ψ.

Note that it is essential in such arguments to have equivalences, since otherwise an important case might be lost.

5. Logical Consequence

We can now define what it means for a sentence θ to be a *logical consequence* of a set Σ of sentences. This means that θ is true whenever all of the sentences in Σ are true. We thus obtain the following definition.

Definition 5.1. A sentence θ is a *logical consequence* of a set Σ of sentences iff every truth assignment satisfying Σ also satifies θ. We write $\Sigma \models \theta$ to mean that θ is a logical consequence of Σ.

Note that θ is not a logical consequence of Σ, $\Sigma \not\models \theta$, iff there is some assignment h such that $h \models \Sigma$ but $h \not\models \theta$. Not also that validity is a special case of logical consequence, since $\models \theta$ iff $\emptyset \models \theta$.

The following properties are immediate from the definition.

(i) If $\theta \in \Sigma$ then $\Sigma \models \theta$.
(ii) If $\Sigma \models \varphi$ for all $\varphi \in \Gamma$ and $\Gamma \models \theta$ then $\Sigma \models \theta$.

At least if Σ is finite we can check whether or not $\Sigma \models \theta$ by truth tables or working backwards. For example, to decide whether or not $\{S_3 \wedge \neg S_1\} \models S_4$ we check whether of not every $h \models (S_3 \wedge \neg S_1)$ satisfies S_4. The answer is negative, since $h(S_1) = F, h(S_3) = T$, and $h(S_4) = F$ is a counterexample.

As another example, it is easy to check that $\{A, (A \to B)\} \models B$ — more generally we will have that $\{\varphi, (\varphi \to \psi)\} \models \psi$ for all sentences φ, ψ (explain!).

Corresponding to Lemma 4.1 we have the following.

Lemma 5.1. $\Sigma \models \theta$ *iff* $(\Sigma \cup \{\neg\theta\})$ *is not satisfiable;* $(\Sigma \cup \{\theta\})$ *is satisfiable iff* $\Sigma \not\models \neg\theta$.

The following important result illuminates the connection between implication and consequence.

Theorem 5.1. $(\Sigma \cup \{\theta\}) \models \varphi$ *iff* $\Sigma \models (\theta \to \varphi)$.

Proof. For simplicity, suppose that $\Sigma = \emptyset$. Suppose further that $\{\theta\} \not\models \varphi$. Then there is some h such that $h \models \theta$ but $h \not\models \varphi$. Thus $h \not\models (\theta \to \psi)$ so $\not\models (\theta \to \psi)$. The reverse direction is the same. □

For example, we noted above that $\{A, (A \to B)\} \models B$. By two uses of Theorem 5.1 we can conclude that $\models A \to ((A \to B) \to B)$, which we established in the preceding section.

Using property (ii) above we obtain the following important fact about logical consequence.

Corollary 5.1. *If* $\Sigma \models \varphi$ *and* $\Sigma \models (\varphi \to \psi)$ *then* $\Sigma \models \psi$.

We collect a few useful examples of logical consequence here.

(i) $\varphi \to \psi, \psi \to \theta \models \varphi \to \theta$
(ii) $\varphi \models \psi \to \varphi$
(iii) $\neg\psi \models \psi \to \varphi$
(iv) $\neg\varphi \to \varphi \models \varphi$
(v) $\varphi \to \psi, \neg\varphi \to \psi \models \psi$
(vi) $\varphi \to \psi \models \neg\psi \to \neg\varphi$ and conversely

6. Logical Equivalence

Definition 6.1. Sentences φ and ψ are *(logically) equivalent* iff $\bar{h}(\varphi) = \bar{h}(\psi)$ for every truth assignment h. We write $\varphi \equiv \psi$ to mean that φ and ψ are equivalent.

The following is clear from the definitions.

Lemma 6.1. $\varphi \equiv \psi$ *iff* $\varphi \models \psi$ *and* $\psi \models \varphi$.

Certainly, for example, we always have $\neg\neg\varphi \equiv \varphi$.

The following equivalences enable us to rewrite sentences written using $\wedge, \vee, \neg$.

(i) $(\varphi \wedge \psi) \equiv (\psi \wedge \varphi)$
(ii) $(\varphi \wedge \psi) \wedge \theta \equiv \varphi \wedge (\psi \wedge \theta)$
(iii)(iv) As (i) and (ii) for $\vee$ in place of $\wedge$
(v) $\neg(\varphi \wedge \psi) \equiv (\neg\varphi \vee \neg\psi)$
(vi) $\varphi \wedge (\psi \vee \theta) \equiv (\varphi \wedge \psi) \vee (\varphi \wedge \theta)$
(vii)(viii) As (v) and (vi) with $\wedge$ and $\vee$ interchanged.

For example, $A \wedge \neg(B \wedge C)$ is equivalent to $A \wedge (\neg B \vee \neg C)$ by (v), and this is equivalent to $(A \wedge \neg B) \vee (A \wedge \neg C)$ by (vi).

We also see that $((C \vee D) \wedge A) \vee ((C \vee D) \wedge B)$ is equivalent to $(C \vee D) \wedge (A \vee B)$, using (vi) where $(C \vee D) = \varphi$, $A = \psi$ and $B = \theta$.

We have the following equivalences connecting $\rightarrow$ with $\wedge$, $\vee$ and $\neg$.

(i) $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi$
(ii) $\varphi \vee \psi \equiv \neg\varphi \rightarrow \psi$
(iii) $\varphi \wedge \psi \equiv \neg(\varphi \rightarrow \neg\psi)$

Using (ii) and (iii) we can rewrite any sentence to remove all occurrences of $\wedge$ and $\vee$. We will use the resulting theorem in defining our proof system.

Theorem 6.1. *Given any sentence θ you can find a sentence θ^* with the same atomic sentences and using only the connectives $\neg$ and $\rightarrow$ such that $\theta \equiv \theta^*$.*

For example, $A \wedge \neg(B \wedge C) \equiv A \wedge (B \rightarrow \neg C) \equiv \neg(A \rightarrow \neg(B \rightarrow \neg C))$.

For many practical purposes it is more useful to be able to rewrite a sentence using just $\neg, \wedge, \vee$ — especially in one the following normal forms.

Definition 6.2. (i) A sentence is in *disjunctive normal form (dnf)* if it is a disjunction $(\theta_1 \vee \cdots \vee \theta_n)$ where each disjunct θ_i is a conjunction of atomic sentences and negations of atomic sentences.
(ii) A sentence is in *conjunctive normal form (cnf)* if it is a conjunction $(\theta_1 \wedge \cdots \wedge \theta_n)$ where each conjunct θ_i is a disjunction of atomic sentences and negations of atomic sentences.

For example, $(A \wedge \neg B) \vee (B \wedge \neg C)$ is in dnf. The advantage of having a sentence in dnf is that you can read off the assignments which satisfy it. For example the preceding sentence is satisfied by h iff either $h(A) = T$ and $h(B) = F$, or $h(B) = T$ and $h(C) = F$. Thus it is satisfiable but not a tautology.

Using the equivalences given earlier we can rewrite any sentence in dnf and also in cnf. For example
$(A \rightarrow B) \rightarrow \neg(B \rightarrow C) \equiv \neg(\neg A \vee B) \vee \neg(\neg B \vee C) \equiv (A \wedge \neg B) \vee (B \wedge \neg C)$, which is in dnf. Further
$(A \wedge \neg B) \vee (B \wedge \neg C) \equiv (A \vee B) \wedge (A \vee \neg C) \wedge (\neg B \vee B) \wedge (\neg B \vee \neg C) \equiv$
$(A \vee B) \wedge (A \vee \neg C) \wedge (\neg B \vee \neg C)$, since $\neg B \vee B$ is a tautology.

There is a useful shortcut which enables you to write down the dnf and cnf of a sentence directly from its truth table. For example suppose that θ has the following truth table.

A	B	C	θ
T	T	T	F
T	T	F	T
T	F	T	T
T	F	F	T
F	T	T	F
F	T	F	T
F	F	T	F
F	F	F	F

Thus $h(\theta) = T$ iff either $h(A \wedge B \wedge \neg C) = T$ or $h(A \wedge \neg B \wedge C) = T$ or $h(A \wedge \neg B \wedge \neg C) = T$ or $h(\neg A \wedge B \wedge \neg C) = T$. Thus θ is equivalent to the disjunction of these four conjunctions, which is in dnf. This is not the simplest dnf of θ — using our equivalences you can show it is equivalent to $(A \wedge \neg B) \vee (B \wedge \neg C)$.

We can also define equivalence of sets of sentences.

Definition 6.3. Sets Σ and Γ are *equivalent*, $\Sigma \equiv \Gamma$, iff they are satisfied by precisely the same truth assignments.

Lemma 6.2. *$\Sigma \equiv \Gamma$ iff $\Sigma \models \psi$ for all $\psi \in \Gamma$ and $\Gamma \models \varphi$ for all $\varphi \in \Sigma$.*

7. Proof by Induction

Recall that every sentence has a unique positive integer as its length, and that compound sentences are built up from shorter sentences. This enables us to prove facts about all sentences using ordinary mathematical induction on $\mathbb{N}$. The following theorem rephrases this process to avoid explicit use of mathematical induction.

Theorem 7.1. *Let $\mathcal{X}$ be a set of sentences. Assume the following:*

(i) *every atomic sentence belongs to $\mathcal{X}$*
(ii) *if $\varphi \in \mathcal{X}$ then $\neg\varphi \in \mathcal{X}$*
(iii) *if $\varphi, \psi \in \mathcal{X}$ then $(\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \to \psi) \in \mathcal{X}$*

Then every sentence belongs to $\mathcal{X}$.

Proof. (outline) We prove, by induction on $n \in \mathbb{N}$, that every sentence of length $\leq n$ belongs to $\mathcal{X}$. The base of the induction is $n = 1$, which is given by property (i) since the atomic sentences are the only sentences of length ≤ 1. Given the result for n we must show it for $n + 1$. Let θ be a sentence of length $\leq n+1$ — we may assume that θ has length exacly $n+1$, since otherwise the inductive hypothesis implies that $\theta \in \mathcal{X}$. There are then several cases according to the form of θ. If, for example, $\theta = \neg\psi$ for some ψ, then $\psi \in \mathcal{X}$ since the length of ψ is $\leq n$, hence $\theta \in \mathcal{X}$ by property (ii). The other cases are similar. □

We give one example of the use of induction.

Theorem 7.2. *Fix $n \in \mathbb{N}$ and a sentence φ_n. For any sentence θ we define θ^* to be the result of replacing all occurrences in θ of S_n by φ_n.*

(a) *Let h be an assignment and define h^* by $h^*(S_n) = \bar{h}(\varphi_n)$, $h^*(S_i) = h(S_i)$ for all $i \neq n$. Then*

$$(*)\ \bar{h}(\theta^*) = \overline{h^*}(\theta).$$

(b) *If $\models \theta$ then $\models \theta^*$.*

Proof. (outline) (a) We let $\mathcal{X}$ be the set of all sentences θ satisfying (*). We show all sentences belong to $\mathcal{X}$ using Theorem 7.1. All atomic sentences belong to $\mathcal{X}$, since ${S_n}^* = \varphi_n$ and so $\bar{h}({S_n}^*) = \overline{h^*}(S_n)$. There are several cases in the inductive step. For example, if $\psi \in \mathcal{X}$ and $\theta = \neg\psi$ then $\theta^* = \neg\psi^*$ and so (*) for ψ implies (*) for θ. The other cases are similar.

(b) is proved using (a). Details are left to the reader. □

8. A Formal Proof System

In this section we define a proof system so that the deducible sentences are precisely the tautologies — this is the Completeness Theorem, proved in section 11 below. We need to specify a set of sentences caled the *logical axioms* and one or more *rules of inference*; a sentence is *deducible* in the proof system iff it is obtained from the axioms by a finite number of applications of the rule(s).

We simplify our task by restricting to sentences using only the connectives $\neg$ and $\rightarrow$. There is no loss in doing this since we know by Theorem 6.1 that every sentence is logically equivalent to some such sentence.

Definition 8.1. The set Λ_0 of *(logical) axioms* of $\mathcal{S}$ consists of all sentences of the following forms:

(1) $\varphi \rightarrow (\psi \rightarrow \varphi)$
(2) $(\varphi \rightarrow (\psi \rightarrow \theta)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \theta))$
(3) $(\neg\varphi \rightarrow \psi) \rightarrow ((\neg\varphi \rightarrow \neg\psi) \rightarrow \varphi))$.

Lemma 8.1. *If $\theta \in \Lambda_0$ then $\models \theta$.*

We need only one rule.

Definition 8.2. *Modus ponens* is the rule which allows us to conclude ψ from φ and $\varphi \rightarrow \psi$.

Corollary 5.1 implies that the rule modus ponens preserves validity.

Finally we define deductions, and deducibility, as follows.

Definition 8.3. A *(logical) deduction* in $\mathcal{S}$ is a finite sequence $\varphi_1, \ldots, \varphi_n$ of sentences such that for each $i \leq n$ one of the following holds:

(i) $\varphi_i \in \Lambda_0$,
(ii) there are $j, k < i$ such that φ_i follows from φ_j and φ_k by modus ponens, that is, $\varphi_k = (\varphi_j \rightarrow \varphi_i)$.

Definition 8.4. φ is *(logically) deducible*, written $\vdash \varphi$, iff there is a deduction whose last sentence is φ.

A deduction of length one consists just of some axiom. Any deduction using modus ponens must have length at least three. For example, the following is a deduction:

1. $A \to (A \to A)$ – an instance of Axiom (1)
2. $(A \to (A \to A)) \to ((A \to A) \to (A \to A))$ – an instance of Axiom (2)
3. $(A \to A) \to (A \to A)$ – by modus ponens on lines 1, 2.

This shows that $\vdash (A \to A) \to (A \to A)$.

By being a little bit more clever we can show that $\vdash A \to A$. Here is the deduction.

1. $A \to ((A \to A) \to A)$ – Axiom(1)
2. $(A \to ((A \to A) \to A)) \to ((A \to (A \to A)) \to (A \to A))$ – Axiom(2)
3. $(A \to (A \to A)) \to (A \to A)$ – modus ponens on lines 1, 2
4. $A \to (A \to A)$ – Axiom(1)
5. $A \to A$ – modus ponens on lines 3, 4.

Of course, the atomic sentence A could be replaced by any sentence, and so we have shown that $\vdash \varphi \to \varphi$ for every sentence φ.

We can now show that every deducible sentence is valid.

Theorem 8.1. *(Soundness) If* $\vdash \varphi$ *then* $\models \varphi$.

Proof. Let $\varphi_1, \ldots, \varphi_n$ be a deduction of φ. We show, by induction on $i \leq n$, that $\models \varphi_i$ for every $i \leq n$ — since $\varphi = \varphi_n$ this establishes the theorem. The base of the induction is $i = 1$; necessarily $\varphi_1 \in \Lambda_0$, so $\models \varphi_1$ by Lemma 8.1. For the inductive step, let $1 < i \leq n$ and assume that $\models \varphi_j$ for all $j < i$. We must show $\models \varphi_i$. If $\varphi_i \in \Lambda_0$ this is exactly as $i = 1$. Otherwise there are $j, k < i$ such that φ_i follows from φ_j and φ_k by modus ponens. Since $\models \varphi_j$ and $\models \varphi_k$ by inductive hypothesis, we conclude that $\models \varphi_i$ by Corollary 5.1. □

9. Deductions from Hypotheses and the Deduction Theorem

It is not easy to construct deductions. It is very helpful to allow deductions from sets of additional hypotheses, as we shall see.

Definition 9.1. Let Σ be a set of sentences. A *deduction from* Σ is a sequence $\varphi_1, \ldots, \varphi_n$ of sentences such that for each $i \leq n$ one of the following holds:

(i) $\varphi_i \in (\Lambda_0 \cup \Sigma)$,
(ii) there are $j, k < i$ such that φ_i follows from φ_j and φ_k by modus ponens.

We will say that φ is *deducible from* Σ, written $\Sigma \vdash \varphi$, if there is a deduction of φ from Σ. For example, A has a one line deduction from $\Sigma = \{A\}$, and thus we know that $A \vdash A$. Clearly, if $\Sigma \subseteq \Sigma'$ and $\Sigma \vdash \varphi$ then $\Sigma' \vdash \varphi$. Note that we have the following, corresponding to Corollary 5.1.

Lemma 9.1. *If* $\Sigma \vdash \varphi$ *and* $\Sigma \vdash (\varphi \to \psi)$ *then* $\Sigma \vdash \psi$.

The following is an important fact about deductions from hypotheses.

Lemma 9.2. *(Finiteness) Assume that* $\Sigma \vdash \varphi$. *Then there is some finite* $\Sigma_0 \subseteq \Sigma$ *such that* $\Sigma_0 \vdash \varphi$.

Proof. Let $\varphi_1, \ldots, \varphi_n$ be a deduction of φ from Σ. Let $\Sigma_0 = \Sigma \cap \{\varphi_1, \ldots, \varphi_n\}$ be the set of all formulas in Σ occurring in the deduction. Then $\varphi_1, \ldots, \varphi_n$ is a deduction from Σ_0, and Σ_0 is finite. □

We also have Soundness for deductions from hypotheses.

Theorem 9.1. *(Soundness) If* $\Sigma \vdash \varphi$ *then* $\Sigma \models \varphi$.

The importance of deductions from hypotheses to deducibility without hypotheses is shown by the following theorem.

Theorem 9.2. *(The Deduction Theorem) Assume that* $(\Sigma \cup \varphi) \vdash \psi$. *Then* $\Sigma \vdash (\varphi \to \psi)$.

Proof. For simplicity we assume that $\Sigma = \emptyset$. So we are given that $\varphi \vdash \psi$ and we want to show that $\vdash (\varphi \to \psi)$. Let $\psi_1, \ldots, \psi_n$ be a deduction of ψ from φ. We show, by induction on $i \leq n$, that $\vdash (\varphi \to \psi_i)$ for all $i \leq n$.

The base case is $i = 1$. Either $\psi_1 \in \Lambda_0$ or $\psi_1 \in \{\varphi\}$, i.e. $\psi_1 = \varphi$. In the first case we have $\vdash \psi_1$, but also $\vdash \psi_1 \to (\varphi \to \psi_1)$ by Axiom (1) hence $\vdash (\varphi \to \psi_1)$ by modus ponens, as desired. In the second case we have $\vdash (\varphi \to \psi_1)$, since $\psi_1 = \varphi$ and we know that $\vdash \varphi \to \varphi$.

In the inductive step we have $1 < i \leq n$ and the inductive hypothesis is that $\vdash (\varphi \to \psi_j)$ for all $j < i$. The only case we need consider is when ψ_i is obtained from ψ_j and ψ_k by modus ponens. The inductive hypothesis for j and k and an appropriate instance of Axiom (2) yields the conclusion. □

Note that the converse to Theorem 9.2 is immediate from Lemma 9.1.

We illustrate the use of the Deduction Theorem by showing that

$$\vdash (A \to B) \to ((B \to C) \to (A \to C)).$$

By the Deduction Theorem it suffices to show that

$$(A \to B) \vdash (B \to C) \to (A \to C).$$

By the Deduction Theorem again, it suffices to show that

$$(A \to B), (B \to C) \vdash (A \to C).$$

By the Deduction Theorem a third time, it suffices to show that

$$(A \to B), (B \to C), A \vdash C.$$

But this is immediate by two uses of Lemma 9.1.

We will require the following for the proof of the Completeness Theorem.

Lemma 9.3. *(a)* $\vdash (\neg\varphi \to \varphi) \to \varphi$.
(b) $\vdash \varphi \to (\neg\varphi \to \psi)$.

Proof. (a) By the Deduction Theorem it suffices to show

$$(\neg\varphi \to \varphi) \vdash \varphi.$$

We know that $\vdash \neg\varphi \to \neg\varphi$, hence also

$$(\neg\varphi \to \varphi) \vdash \neg\varphi \to \neg\varphi.$$

But $(\neg\varphi \to \varphi) \vdash (\neg\varphi \to \neg\varphi) \to \varphi$ by Axiom (3) and Lemma 9.1, and so $\neg\varphi \to \varphi \vdash \varphi$ by Lemma 9.1 again.

(b) is left to the reader. □

10. Consistent Sets of Sentences

Definition 10.1. A set Σ of sentences is *consistent* provided there is no sentence φ such that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$. Σ is *inconsistent* if it is not consistent.

It follows from Theorem 9.1 that Σ is consistent provided Σ is satisfiable. Our goal in this section is to prove the converse, and thus that consistency is equivalent to satisfiability.

We first note two important facts about consistency.

Lemma 10.1. *(Finiteness) Σ is consistent if (and only if) every finite $\Sigma_0 \subseteq \Sigma$ is consistent.*

Proof. We show the contrapositive. Assume that Σ is inconsistent. Then there is some φ such that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$. By Lemma 9.2 there are finite subsets Σ_1 and Σ_2 of Σ such that $\Sigma_1 \vdash \varphi$ and $\Sigma_2 \vdash \neg\varphi$. Then both φ and $\neg\varphi$ are deducible from $\Sigma_0 = \Sigma_1 \cup \Sigma_2$, which is therefore inconsistent. □

Lemma 10.2. *Σ is inconsistent iff $\Sigma \vdash \psi$ for all sentences ψ.*

Proof. From right to left is clear. For the other direction, suppose that $\Sigma \vdash \varphi$ and $\Sigma \vdash \neg\varphi$. By Lemma 9.3(b) we know $\vdash \varphi \to (\neg\varphi \to \psi)$ for every ψ, hence $\Sigma \vdash \psi$ by two uses of Lemma 9.1. □

We will want to extend consistent sets to maximal consistent sets. This will use the following Lemma.

Lemma 10.3. *Assume that Σ is consistent and let φ be any sentence. Then either $(\Sigma \cup \{\varphi\})$ or $(\Sigma \cup \{\neg\varphi\})$ is consistent (or both).*

Proof. Assume that $(\Sigma \cup \{\neg\varphi\})$ is inconsistent. Then, by Lemma 10.2

$$(\Sigma \cup \{\neg\varphi\}) \vdash \varphi.$$

Thus

$$\Sigma \vdash \neg\varphi \to \varphi,$$

by the converse to the Deduction Theorem. But

$$\Sigma \vdash (\neg\varphi \to \varphi) \to \varphi$$

by Lemma 9.3(a), and so $\Sigma \vdash \varphi$ by two uses of Lemma 9.1. Therefore $(\Sigma \cup \{\varphi\})$ is consistent since Σ is consistent. □

Definition 10.2. The set Γ of sentences is *maximal consistent* iff it is consistent and for every sentence φ either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.

We outline the proofs of two theorems about maximal consistent sets which we need to conclude that consistent sets are satisfiable.

Theorem 10.1. *Assume that Σ is consistent. Then there is some maximal consistent Γ such that $\Sigma \subseteq \Gamma$.*

Proof. (outline) Enumerate the set of all sentences as $\varphi_1, \ldots, \varphi_n, \ldots$ for $n \in \mathbb{N}$. Let $\Gamma_1 = \Sigma$ and for each n define $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_n\}$, if this is consistent, and $\Gamma_{n+1} = \Gamma_n \cup \{\neg\varphi_n\}$ otherwise. By Lemma 10.3, Γ_{n+1} is consistent is either case. Let $\Gamma = \bigcup_{n\in\mathbb{N}} \Gamma_n$. Then Γ is consistent, by Lemma 10.1, and maximal, by construction. □

Theorem 10.2. *Assume that Γ is maximal consistent. Then Γ is satisfiable.*

Proof. (outline) Define h on the atomic sentences by $h(S_i) = T$ iff $S_i \in \Gamma$. We want to show the following for every sentence φ

$$(*)\ h \models \varphi \text{ iff } \varphi \in \Gamma.$$

Now suppose that (*) fails. Let φ be the shortest sentence which falsifies (*). By definition of h, φ can't be an atomic sentence. Therefore it either has the form $\neg\psi$ or $\psi \to \theta$, where the equivalence in (*) holds for ψ and θ. In the first case $\varphi = \neg\psi$ where $h \models \psi$ iff $\psi \in \Gamma$. Now $\varphi \in \Gamma$ iff $\psi \notin \Gamma$, since Γ is consistent and maximal. Therefore we also have $h \models \varphi$ iff $\varphi \in \Gamma$, contradicting the hypothesis that (*) fails for φ. The other case is similar. □

The desired conclusion follows immediately.

Theorem 10.3. *Σ is consistent iff Σ is satisfiable.*

11. Completeness and Compactness Theorems

In this section we use the equivalence of consistency and satisfibility to obtain the equivalence of deducibility and logical consequence. Recall from Lemma 5.1 that $\Sigma \models \varphi$ iff $\Sigma \cup \{\neg\varphi\}$ is unsatisfiable. We will now establish the same relation between deducibility and consistency.

Lemma 11.1. *$\Sigma \vdash \varphi$ iff $\Sigma \cup \{\neg\varphi\}$ is inconsistent.*

Proof. Assume that $\Sigma \vdash \varphi$. Then $\Sigma \cup \{\neg\varphi\} \vdash \varphi, \neg\varphi$ so $\Sigma \cup \{\neg\varphi\}$ is inconsistent.

For the other direction, assume that $\Sigma \cup \{\neg\varphi\}$ is inconsistent. Then $\Sigma \cup \{\neg\varphi\} \vdash \varphi$, by Lemma 10.2, hence $\Sigma \vdash (\neg\varphi \to \varphi)$ by the Deduction Theorem. But by Lemma 9.3(a) we know $\Sigma \vdash (\neg\varphi \to \varphi) \to \varphi$, hence $\Sigma \vdash \varphi$, as desired. □

The Completeness Theorem is immediate from Lemma 11.1 and Theorem 10.3.

Theorem 11.1. *(Completeness) $\Sigma \models \varphi$ iff $\Sigma \vdash \varphi$.*

As a consequence of Theorem 11.1 and Lemma 10.1 we obtain the following, which does not mention deductions at all.

Theorem 11.2. *(Compactness) (a)* Σ *is satisfiable if (and only if) every finite* $\Sigma_0 \subseteq \Sigma$ *is satisfiable.*

(b) If $\Sigma \models \varphi$ *then there is some finite* $\Sigma_0 \subseteq \Sigma$ *such that* $\Sigma_0 \models \varphi$ *(and conversely).*

Compactness has many applications, such as the following.

Corollary 11.1. *Assume that every assignment satisfies at least one sentence in* Σ*. Then there is a finite set* $\{\varphi_1, \ldots, \varphi_n\} \subseteq \Sigma$ *such that*

$$\models (\varphi_1 \vee \cdots \vee \varphi_n).$$