Set Theory (MATH 6730)

Background in Logic

Modern (axiomatic) set theory is developed within first-order logic. We will briefly review the basic concepts and facts that we will need. It will be useful to start with sentential logic (also called propositional logic), which can be viewed as 'zeroth-order logic'.

1. Sentential Logic

The symbols used in sentential logic are

- the logical connective symbols $\neg$ and $\to$ (for 'not' and 'if ... then'),
- the sentential variables S_i ($i \in \omega$) where $\omega = \{0, 1, \dots\}$ denotes the set of natural numbers, and
- (,) (parentheses).[1]

Any finite sequence of these symbols will be referred to as an *expression*. Equality of expressions (in particular, equality of symbols) will be denoted by $\equiv$.[2] The 'meaningful' expressions, called *sentential formulas*, are the expressions that can be built up, in finitely many steps, from sentential variables using the formula building operations

$$\psi \mapsto \neg\psi \qquad \text{and} \qquad (\psi, \chi) \mapsto (\psi \to \chi). \tag{1}$$

The precise definition is as follows.

Definition 1.1. A *sentential formula construction* is a finite sequence $\langle \varphi_0, \dots, \varphi_{m-1} \rangle$ of expressions such that $m \geq 1$ and for each $i < m$ one of the following holds:

- $\varphi_i \equiv S_j$ for some $j \in \omega$;
- $\varphi_i \equiv \neg\varphi_k$ for some $k < i$;
- $\varphi_i \equiv (\varphi_k \to \varphi_l)$ for some $k, l < i$.

A *sentential formula* is an expression that occurs in a sentential formula construction.

[1] The use of parentheses can be avoided by using Polish notation.

[2] We use a symbols other than =, because in first order logic = will be a logical symbol.

It is important that sentential formulas are *uniquely readable*. Informally, this means that for every sentential formula φ there is exactly one way to built up φ from sentential variables using the formula building operations in (1). A precise formulation of a formally weaker, but equivalent statement is the following theorem.

Theorem 1.2. *For every sentential formula φ exactly one of the following three conditions holds:*

- $\varphi \equiv S_j$ *for some* $j \in \omega$*;*
- $\varphi \equiv \neg\psi$ *for some sentential formula* ψ*;*
- $\varphi \equiv (\psi \to \chi)$ *for some sentential formulas* ψ *and* χ*.*

Moreover:

- *If* ψ, ψ' *are sentential formulas such that* $\neg\psi \equiv \neg\psi'$*, then* $\psi \equiv \psi'$*.*
- *If* ψ, χ, ψ', χ' *are sentential formulas such that* $(\psi \to \chi) \equiv (\psi' \to \chi')$*, then* $\psi \equiv \psi'$ *and* $\chi \equiv \chi'$*.*

Theorem 1.2 says that for any sentential formula φ which is not a sentential variable, we can uniquely determine

– the last formula-building operation used to obtain φ, and also
– the subformulas to which this formula-building operation was applied, to get φ.

Unique readability follows from this by induction (on the length of φ or on the length of a sentential formula construction for φ). Warning: Sentential formula constructions for sentential formulas are not unique.

The ‘meaning’ of a sentential formula φ is its truth value, given that we know the truth values of the sentence symbols that occur in φ. For a rigorous discussion of evaluating sentential formulas, we introduce the following concepts.

Definition 1.3. We define a *truth assignment* (or *sentential assignment*) to be a function that maps the set $\{S_i : i \in \omega\}$ of sentential variables into the set $\{0, 1\}$ of *truth values* where 0 means ‘false’ and 1 means ‘true’.

Theorem 1.4. *Every truth assignment f can be extended uniquely to a function $\widehat{f}$ that maps the set of all sentential formulas into $\{0, 1\}$ in such a way that*

- $\widehat{f}(S_i) = f(S_i)$ *for all $i \in \omega$ (i.e., $\widehat{f}$ extends f);*
- $\widehat{f}(\neg\psi) = 1 - \widehat{f}(\psi)$ *for all sentential formulas ψ; and*
- $\widehat{f}((\psi \to \chi)) = 0$ *iff $\widehat{f}(\psi) = 1$ and $\widehat{f}(\chi) = 0$ holds for all sentential formulas ψ and χ.*

Unique readability of sentential formulas is crucial in showing that $\widehat{f}$ is well-defined.

It is not hard to show that for any truth assignment f and sentential formula φ, the truth value $\widehat{f}(\varphi)$ depends only on the truth values that f assigns to the sentential variables occurring in φ. Precisely:

Fact 1.5. *Let φ be a sentential formula. If f, g are truth assignments such that $f(S_i) = g(S_i)$ for every sentential variable S_i occurring in φ, then $\widehat{f}(\varphi) = \widehat{g}(\varphi)$.*

We use this fact, for example, when we work with truth tables.

Our sentential formulas use only the logical connective symbols $\neg$ and $\to$. Other logical connectives, like $\vee$ ('or'), $\wedge$ ('and'), and $\leftrightarrow$ ('if and only if') can be expressed using $\neg$ and $\to$. Formally, we introduce the following definition.

Definition 1.6. For arbitrary sentential formulas φ and ψ we define

$$\varphi \vee \psi :\equiv \neg\varphi \to \psi, \quad \varphi \wedge \psi :\equiv \neg(\varphi \to \neg\psi), \quad \text{and} \quad \varphi \leftrightarrow \psi :\equiv (\varphi \to \psi) \wedge (\psi \to \varphi).$$

For any truth assignment f we have $\widehat{f}(\neg\varphi \to \psi) = 0$ iff $\widehat{f}(\varphi) = \widehat{f}(\psi) = 0$. Therefore the first definition is reasonable. Similar arguments apply to the other two definitions.

Definition 1.7. A sentential formula τ is called a *tautology* if $\widehat{f}(\tau) = 1$ holds for all truth assignments f.

Remarks. For the proof of Theorem 1.2, see Proposition 1.2 in [1].[3] Note that in [1], sentential formulas are defined in Polish notation; that is, in the third bullet of Definition 1.1, the right hand side is replaced by the expression $\to\varphi_k\varphi_l$, and $(\varphi_k \to \varphi_l)$ is used only as an alternate notation for this expression. Therefore, adopting our Definition 1.1 requires minor changes in the proof of Proposition 1.2 to obtain a proof of Theorem 1.2.

For the proofs of Theorem 1.4 and Fact 1.5 see Propositions 1.3 and 1.4 in [1].

[3] [1] *Lectures on Set Theory* by J. Donald Monk.

2. First-Order Logic for Set Theory

The language of set theory, which we will denote by $\mathcal{L}$, uses the following symbols:

- the logical symbols $\neg$, $\to$, $\forall$, $=$ (for 'not', 'if ... then', 'for all', and 'equality'),
- the variables v_i $(i \in \omega)$,
- (,) (parentheses),[4] and
- $\in$ (a binary relation symbol).[5]

It will sometimes be useful to add constant symbols to $\mathcal{L}$. For any set C of constant symbols, $\mathcal{L}_C$ will denote the language of set theory, expanded by the constant symbols in C.[6]

Any finite sequence of symbols in $\mathcal{L}_C$ will be referred to as an *expressions* or $\mathcal{L}_C$*-expressions*. As in sentential logic, we will use $\equiv$ to denote equality of symbols and equality of expressions. The 'meaningful expressions', called *formulas* or $\mathcal{L}_C$*-formulas*, are defined as follows.

Definition 2.1. An *atomic formula* is an expression of the form $s = t$ or $s \in t$ where each one of s, t is a variable or a constant symbol.[7]

A *formula construction sequence* is a finite sequence $\langle \varphi_0, \dots, \varphi_{m-1} \rangle$ of expressions such that $m \geq 1$ and for each $i < m$ one of the following holds:

- φ_i is an atomic formula;
- $\varphi_i \equiv \neg \varphi_k$ for some $k < i$;
- $\varphi_i \equiv (\varphi_k \to \varphi_l)$ for some $k, l < i$;
- $\varphi_i \equiv \forall v_l \, \varphi_k$ for some $k < i$ and for some variable v_l.

A *formula* is an expression that occurs in a formula construction sequence.

Informally, an expression is a formula iff it can be built up, in finitely many steps, from atomic formulas by using the formula-building operations

$$\psi \mapsto \neg\psi, \qquad (\psi, \chi) \mapsto (\psi \to \chi), \qquad \text{and} \qquad \psi \mapsto \forall v_i \, \psi \quad (i \in \omega). \tag{2}$$

---

[4]As is sentential logic, the use of parentheses can be avoided by using Polish notation.

[5]Other first-order languages might have relation symbols of arbitrary arities (≥ 1), function symbols of arbitrary arities (≥ 1), and constant symbols.

[6]For most purposes we may assume that C is finite or can be indexed by ω.

[7]In Polish notation an atomic formula would have the form $=st$ or $\in st$.

For the same reason as in sentential logic, $\mathcal{L}_C$-formulas are *uniquely readable*; that is, for every formula φ there is exactly one way to built up φ from the atomic $\mathcal{L}_C$-formulas using the formula building operations in (2). The analogue of Theorem 1.2 expressing this fact is the following theorem.

Theorem 2.2. *For every $\mathcal{L}_C$-formula φ exactly one of the following four conditions holds:*

- *φ is an atomic formula;*
- *$\varphi \equiv \neg\psi$ for some formula ψ;*
- *$\varphi \equiv (\psi \to \chi)$ for some formulas ψ and χ;*
- *$\varphi \equiv \forall v_i\, \psi$ for some formula ψ and some variable v_i.*

Moreover:

- *If $s \lozenge t$ and $s' \blacklozenge t'$ are atomic formulas with $\lozenge, \blacklozenge \in \{=, \in\}$ such that $s \lozenge t \equiv s' \blacklozenge t'$, then $s \equiv s'$, $\lozenge \equiv \blacklozenge$, and $t \equiv t'$.*
- *If ψ, ψ' are formulas such that $\neg\psi \equiv \neg\psi'$, then $\psi \equiv \psi'$.*
- *If ψ, χ, ψ', χ' are formulas such that $(\psi \to \chi) \equiv (\psi' \to \chi')$, then $\psi \equiv \psi'$ and $\chi \equiv \chi'$.*
- *If ψ, ψ' are formulas and v_i, v_j are variable such that $\forall v_i\, \psi \equiv \forall v_j\, \psi'$, then $i = j$ (i.e., $v_i \equiv v_j$) and $\psi \equiv \psi'$.*

A *subformula* of a formula φ is defined to be a formula φ' that appears in every construction sequence for φ. More informally, φ' is a subformula of φ if φ' is constructed in the (unique) process of building up φ from the atomic formulas using the formula building operations in (2). A subformula may have more than one occurrence in φ.

The following fact is not hard to show (by induction).

Fact 2.3. *Every quantifier symbol $\forall$ that occurs in a formula φ occurs as the first symbol of a uniquely determined subformula of the form $\forall v_i \, \psi$ (hence ψ is also a subformula of φ).*

Definition 2.4. Let φ be a formula. For every occurrence of the symbol $\forall$ in φ, if $\forall v_i \, \psi$ is the subformula of φ described in Fact 2.3, then the expression $\forall v_i$ is referred to as a *quantifier* in φ, and ψ is called the *scope* of this occurrence of the quantifier $\forall v_i$.

For every occurrence of a variable v_i in φ, this occurrence of v_i is said to be *bound* in φ if either v_i is preceded by a $\forall$ symbol (i.e., v_i is the variable in a quantifier $\forall v_i$ in φ), or v_i is in the scope of a quantifier $\forall v_i$ in φ. An occurrence of a variable in φ is called *free* if it is not bound. We say that v_i is a *free variable* of φ if v_i has at least one free occurrence in φ.

A formula which has no free variables is called a *sentence*.

To attach 'meaning' to formulas, we need to specify

- the (set of) objects $\forall$ is applied to, and
- the meaning of the relation symbol $\in$ and constant symbols for those objects.

This will be done by specifying a *structure (for the language $\mathcal{L}_C$)*.

Definition 2.5. A *structure* $\overline{A}$ (*for* $\mathcal{L}_C$), also called an $\mathcal{L}_C$*-structure*, is

- a nonempty set A (called the *universe* of $\overline{A}$), together with
- a binary relation $\in^{\overline{A}}$ on A (called *the interpretation of the symbol* $\in$ *in* $\overline{A}$), and
- for every constant symbol c, an element $c^{\overline{A}}$ of A (called *the interpretation of the symbol c in* $\overline{A}$).

Usually, we will write

$$\overline{A} = \big\langle A;\ \in^{\overline{A}}, \langle c^{\overline{A}} \rangle_{c\in C} \big\rangle \qquad \text{or} \qquad \overline{A} = \big\langle A;\ \in^{\overline{A}}, c^{\overline{A}}\ (c \in C) \big\rangle.$$

Simple examples show that a formula which has free variables may be true or false in a given structure, depending on which elements of the structure the variables stand for. Therefore, to determine the truth value of a given formula in a fixed structure $\overline{A}$, we need to consider functions $a\colon \omega \to A$ that specify elements of (the universe of) $\overline{A}$ to be assigned to the variables. If a is such a function, then for any $i \in \omega$ and $b \in A$ we will use the notation a^i_b for the function $\omega \to A$ defined by

$$a^i_b(i) = b \qquad \text{and} \qquad a^i_b(j) = a(j) \ \text{ for all } j \neq i \ (j \in \omega).$$

Definition 2.6. Let $\overline{A}$ be an $\mathcal{L}_C$-structure, and let $a\colon \omega \to A$. Furthermore, let φ be an $\mathcal{L}_C$-formula. We define (by recursion) what it means that φ *holds in* $\overline{A}$ *under the assignment* a (or, $\overline{A}$ *models* φ *under* a, or φ *is satisfied by* a *in* $\overline{A}$); in symbols:

$$\overline{A} \models \varphi\,[a].$$

Case 1: φ is atomic.

$$\overline{A} \models s = t\,[a] \qquad \text{iff} \qquad s^{\overline{A}} = t^{\overline{A}},$$
$$\overline{A} \models s \in t\,[a] \qquad \text{iff} \qquad s^{\overline{A}} \in^{\overline{A}} t^{\overline{A}},$$

where $s^{\overline{A}}$ denotes $a(i)$ if $s \equiv v_i$, and $s^{\overline{A}}$ denotes $c^{\overline{A}}$ if $s \equiv c$; similarly for t.

Case 2: $\varphi \equiv \neg\psi$ for some formula ψ.

$$\overline{A} \models \neg\psi\,[a] \qquad \text{iff} \qquad \overline{A} \not\models \psi\,[a].$$

Case 3: $\varphi \equiv (\psi \to \chi)$ for some formulas ψ and χ.

$$\overline{A} \models (\psi \to \chi)\,[a] \qquad \text{iff} \qquad \overline{A} \not\models \psi\,[a] \ \text{ or } \ \overline{A} \models \chi\,[a].$$

Case 4: $\varphi \equiv \forall v_i\,\psi$ for some formula ψ and some variable v_i.

$$\overline{A} \models \forall v_i\,\psi\,[a] \qquad \text{iff} \qquad \overline{A} \models \psi\,[a^i_b] \ \text{ for all } \ b \in A.$$

Unique readability of formulas is crucial in order to see that this definition is correct, that is, to see that, given $\overline{A}$ and $a\colon \omega \to A$, exactly one of $\overline{A} \models \varphi\,[a]$ and $\overline{A} \not\models \varphi\,[a]$ holds for every formula φ.

It is not hard to deduce from Definition 2.6 that the fact whether or not $\overline{A} \models \varphi\,[a]$, depends only on what elements of A are assigned by a to the free variables of φ. This claim can be stated more precisely as follows.

Fact 2.7. *Let $\overline{A}$ be an $\mathcal{L}_C$-structure, and let φ be an $\mathcal{L}_C$-formula. If $a_1, a_2\colon \omega \to A$ are two assignments such that $a_1(i) = a_2(i)$ holds whenever v_i is a free variable of φ, then*

$$\overline{A} \models \varphi\,[a_1] \qquad \textit{iff} \qquad \overline{A} \models \varphi\,[a_2].$$

Corollary 2.8. *Let $\overline{A}$ be an $\mathcal{L}_C$-structure. For every $\mathcal{L}_C$-sentence φ exactly one of the following holds:*

(†) $\overline{A} \models \varphi\,[a]$ *for all* $a\colon \omega \to A$;
(‡) $\overline{A} \models \varphi\,[a]$ *for no* $a\colon \omega \to A$.

In case (†) we say that *the sentence φ is true in $\overline{A}$*, and write $\overline{A} \models \varphi$. In case (‡) we say that *the sentence φ is false in $\overline{A}$*, and write $\overline{A} \not\models \varphi$.

We can introduce the logical connectives $\vee$, $\wedge$, $\leftrightarrow$ the same way as in sentential logic, and we can also introduce existential quantification $\exists$ ('there exists').

Definition 2.9. For arbitrary formulas φ and ψ and for any variable v_i we define $\varphi \vee \psi$, $\varphi \wedge \psi$ and $\varphi \leftrightarrow \psi$ as in Definition 1.6, and we define

$$\exists v_i\,\varphi :\equiv \neg\forall v_i\,(\neg\varphi).$$

As an easy consequence of Definitions 2.6 and 2.9 we get the fact below, which justifies Definition 2.9.

Fact 2.10. *Let $\overline{A}$ be an $\mathcal{L}_C$-structure, and let $a\colon \omega \to A$. Furthermore, let φ, ψ be $\mathcal{L}_C$-formulas, and let v_i be a variable. Then*

$$\begin{aligned}
\overline{A} \models (\varphi \vee \psi)\,[a] &\quad \textit{iff} \quad \overline{A} \models \varphi\,[a] \ \textit{ or } \ \overline{A} \models \psi\,[a], \\
\overline{A} \models (\varphi \wedge \psi)\,[a] &\quad \textit{iff} \quad \overline{A} \models \varphi\,[a] \ \textit{ and } \ \overline{A} \models \psi\,[a], \\
\overline{A} \models (\varphi \leftrightarrow \psi)\,[a] &\quad \textit{iff} \quad \Big(\overline{A} \models \varphi\,[a] \ \textit{ if and only if } \ \overline{A} \models \psi\,[a]\Big), \quad \textit{and} \\
\overline{A} \models \exists v_i\,\varphi\,[a] &\quad \textit{iff} \quad \overline{A} \models \varphi\,[a^i_b] \ \textit{ for some } \ b \in A.
\end{aligned}$$

Definition 2.11. Let Γ be a set of $\mathcal{L}_C$-formulas, and let φ be an $\mathcal{L}_C$-formula. We say that:

- Γ *logically implies* φ (or φ is a *logical consequence* of Γ), and write $\Gamma \models \varphi$, if for every $\mathcal{L}_C$-structure and every assignment $a\colon \omega \to A$ such that $\overline{A} \models \gamma\,[a]$ holds for each $\gamma \in \Gamma$, we have that $\overline{A} \models \varphi\,[a]$.[8]
- φ is *universally valid* (or φ is a *valid formula*) if $\emptyset \models \varphi$, that is, if $\overline{A} \models \varphi\,[a]$ for every structure $\overline{A}$ and every assignment $a\colon \omega \to A$; notation: $\models \varphi$.
- φ is a *tautology* if there exist a sentential tautology τ and $\mathcal{L}_C$-formulas ψ_i ($i \in \omega$) such that φ is obtained from τ by replacing each sentential variable S_i occurring in τ by ψ_i.

Another easy consequence of Definition 2.6 is the following.

Fact 2.12. *If an $\mathcal{L}_C$-formula is a tautology, then it is universally valid.*

Next we define the concept of logical implication used in [1], which is different from the one introduced in Definition 2.11. Unlike in [1], we will use the notation $\models_\Box$ for this weaker/restricted notion of logical implication.

Definition 2.13. Let Γ be a set of $\mathcal{L}_C$-formulas, let φ be an $\mathcal{L}_C$-formula, and let $\overline{A}$ be an $\mathcal{L}_C$-structure. We say that

- $\overline{A}$ is a *model of* φ if $\overline{A} \models \varphi\,[a]$ for every assignment $a\colon \omega \to A$.
- $\Gamma \models_\Box \varphi$ if for every structure $\overline{A}$ such that $\overline{A}$ is a model of every member of Γ, we have that $\overline{A}$ is a model of φ.

---

[8] This is the definition of logical implication in

[2] *A Mathematical Introduction to Logic* by H. B. Enderton.

To clarify the relationship between $\models$ and $\models_\square$, the following notation will be useful. If φ is an $\mathcal{L}_C$-formula and the free variables of φ are $v_{i_0}, \dots, v_{i_{k-1}}$ with $i_0 < \cdots < i_{k-1}$, then let

$$[\![\varphi]\!] :\equiv \forall v_{i_0} \dots \forall v_{i_{k-1}}\, \varphi \quad \text{(the } \textit{universal closure of } \varphi\text{)}.$$

Clearly, $[\![\varphi]\!]$ is a sentence.

If Γ is a set of formulas, let $[\![\Gamma]\!]$ denote the set of all sentences $[\![\gamma]\!]$ with $\gamma \in \Gamma$.

Facts 2.14. *Let Γ be a set of $\mathcal{L}_C$-formulas and let φ be an $\mathcal{L}_C$-formula.*

(i) *If $\Gamma \models \varphi$, then $\Gamma \models_\square \varphi$.*

(ii) *We have that*

$$\Gamma \models_\square \varphi \qquad \textit{iff} \qquad [\![\Gamma]\!] \models \varphi \qquad \textit{iff} \qquad [\![\Gamma]\!] \models [\![\varphi]\!].$$

(iii) *Consequently, $\Gamma \models \varphi$ and $\Gamma \models_\square \varphi$ are equivalent if Γ is a set of sentences. In particular, $\models \varphi$ if and only if $\models_\square \varphi$.*

Remark. For the proofs of Theorem 2.2, Fact 2.3, the correctness of Definition 2.6, Fact 2.7, Fact 2.10, and Fact 2.12 see Proposition 2.6, Proposition 3.9, Proposition 2.7, Lemma 4.4, Proposition 2.8, and Theorem 2.9 in [1].

3. Proof Systems

In every branch of mathematics, 'theorems' – i.e., logical consequences of the axioms of the given subject – are established by 'proofs', rather than by using Definition 2.11 (or 2.13), unless the 'theorem' is very simple. The advantage of a 'proof' is that it is a step-by-step argument that may use all previously established 'theorems' in additions to the axioms, and in every step of a proof it is easy to see that the reasoning is logically sound.

Now we will discuss a formal proof system for first-order logic (in the languages $\mathcal{L}_C$), which can be viewed as the rigorous mathematical model of 'proofs', and can be studied by mathematical techniques. The word "formal" refers to the fact that a proof system uses formulas only, without reference to 'meaning' (structures, satisfaction, logical implication, etc.).

Ideally, a proof system captures logical implication exactly, that is, a formula φ is 'provable' in the formal proof system from a set Γ of assumptions if and only if φ is a logical consequence of Γ. It is a deep fact that such proof systems exist in first-order logic (see Theorems 3.4–3.5 and Theorems 3.8–3.9 below).

We have introduced two different notions of logical implication, $\models$ and $\models_\Box$, therefore we will briefly discuss proof systems for each one separately.

A proof system for $\models$.[9] The following notation will be useful: if φ is a formula, x is a variable, and t is a variable or constant symbol, then $\mathrm{Subf}_t^x(\varphi)$ will denote the formula obtained from φ by replacing every free occurrence of x by t.

Definition 3.1. The set Λ of *logical axioms* is the set of all $\mathcal{L}_C$-formulas of the form

$$\forall v_{j_0} \ldots \forall v_{j_{l-1}}\, \lambda \quad (\textit{generalization of } \lambda)$$

where λ is one of the following formulas:

(Ax1) a tautology (see Definition 2.11);

(Ax2) $\forall x\, \varphi \to \mathrm{Subf}_t^x(\varphi)$ where φ is a formula, x is a variable, and either t is a constant symbol or t is a variable such that no quantifier $\forall t$ in φ has a free occurrence of x in its scope.

(Ax3) $\forall x\, (\varphi \to \psi) \to (\forall x\, \varphi \to \forall x\, \psi)$ where φ, ψ are arbitrary formulas and x is any variable;

(Ax4) $\varphi \to \forall x\, \varphi$ where φ is any formula and x is any variable that is not free in φ;

(Ax5) $x = x$ where x is a variable;

(Ax6) $x = y \to (\varphi \to \varphi')$ where φ is an atomic formula, x, y are variables, and φ' is obtained from φ by replacing zero or more (but not necessarily all) occurrences of x by y.

[9]This discussion follows Sections 2.4-2.5 of [2], but is restricted to the languages $\mathcal{L}_C$.

Definition 3.2. Let Γ be a set of formulas. A Γ-*proof* (or a *deduction* from Γ) is a finite sequence $\langle \varphi_0, \dots, \varphi_{m-1} \rangle$ of formulas such that for each $i < m$ one of the following holds:

(P1) φ_i is a logical axiom or $\varphi_i \in \Gamma$;

(P2) there exist $j, k < i$ such that $\varphi_j \equiv (\varphi_k \to \varphi_i)$.

If φ_i satisfies condition (P2), we say that φ_i is obtained from the formulas φ_k and $\varphi_j \equiv (\varphi_k \to \varphi_i)$ (which occur earlier in the proof) by *modus ponens* (abbreviated *MP*):

$$\langle \, \dots \; \varphi_k, \; \dots \; \underbrace{(\varphi_k \to \varphi_i)}_{\varphi_j}, \; \dots \; \varphi_i, \; \dots \rangle.$$

Modus ponens allows us to create (to 'deduce') a new entry in a proof using earlier entries; therefore it is referred to as a *rule of inference*.

Definition 3.3. Let Γ be a set of formulas and φ another formula. We say that Γ *proves* φ (or φ *is provable from* Γ, or φ *is deducible from* Γ, or φ *is a theorem of* Γ), and write $\Gamma \vdash \varphi$, if φ appears as a member (equivalently: last member) of a Γ-proof.

By first showing that every formula in Λ is universally valid, one can use induction on the lengths of proofs to get the following result.

Theorem 3.4. (Soundness Theorem) *Let Γ be a set of formulas and φ another formula. If $\Gamma \vdash \varphi$, then $\Gamma \models \varphi$.*

The converse of this theorem is much deeper; it was proved by K. Gödel in 1929.[10]

Theorem 3.5. (Completeness Theorem) *Let Γ be a set of formulas and φ another formula. If $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.*

[10] Gödel proved the version of the completeness theorem when $\Gamma = \emptyset$, for any countable language.

A proof system for $\models_\Box$.[11] We start again with the set of logical axioms.

Definition 3.6. The set $\Lambda_\Box$ of *logical axioms* is the set of all $\mathcal{L}_C$-formulas of one of the following forms:

($\Box$ax1) $\varphi \to (\psi \to \varphi)$,
$\big(\varphi \to (\psi \to \chi)\big) \to \big((\varphi \to \psi) \to (\varphi \to \chi)\big)$,
$(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$ where φ, ψ, χ are arbitrary formulas;
($\Box$ax3) the same formulas as in (Ax3);
($\Box$ax4) the same formulas as in (Ax4) so that x does not occur in φ at all;
($\Box$ax6) $s = t \to (\varphi \to \varphi')$ where s and t is a variable or a constant symbol, φ is an atomic formula involving s, and φ' is obtained from φ by replacing one occurrence of s by t;
($\Box$ax7) $\exists x\, x = s$ where s is a constant symbol or a variable other than x.

The definition of a Γ-proof is similar to the earlier one; the only difference is that we will now have an additional rule of inference.

Definition 3.7. With the same notation as in Definition 3.2, a Γ-*proof* is a finite sequence $\langle \varphi_0, \dots, \varphi_{m-1} \rangle$ of formulas such that for each $i < m$ one of the earlier conditions (P1), (P2), or the new condition (P3) below holds:

(P3) there exists $j < i$ and a variable x such that $\varphi_i \equiv \forall x\, \varphi_j$.

If φ_i satisfies condition (P3), we say that φ_i is obtained from the formula φ_j (which occurs earlier in the proof) by *generalization*.

The concept "Γ *proves* φ" for this proof system can be defined the same way as before (see Definition 3.3), but we will use the notation $\Gamma \vdash_\Box \varphi$ instead of $\Gamma \vdash \varphi$.

Both the Soundness and Completeness Theorems hold for $\models_\Box$ and $\vdash_\Box$:

Theorem 3.8. (Soundness Theorem) *Let Γ be a set of formulas and φ another formula. If $\Gamma \vdash_\Box \varphi$, then $\Gamma \models_\Box \varphi$.*

Theorem 3.9. (Completeness Theorem) *Let Γ be a set of formulas and φ another formula. If $\Gamma \models_\Box \varphi$, then $\Gamma \vdash_\Box \varphi$.*

[11] This discussion follows Sections 3–4 of [1], but is restricted to the languages $\mathcal{L}_C$.

Let's compare the two proof systems.

- It is not hard to show that

(3) $$\vdash \mu \quad \text{for every formula } \mu \in \Lambda_\Box.$$

  In fact, each set ($\Box$ax1), ($\Box$ax3), ($\Box$ax4) of axioms in $\Lambda_\Box$ is a subset of the corresponding set (Ax1), (Ax3), (Ax4) in Λ, the set ($\Box$ax6) is just slightly different from (Ax6), and for the axioms μ in ($\Box$ax7), a short $\emptyset$-proof verifies $\vdash \mu$.
- The proof of the analogous fact with the roles reversed, namely that

(4) $$\vdash_\Box \lambda \quad \text{for every formula } \lambda \in \Lambda,$$

  is much more involved. Since $\vdash_\Box$ allows the use of the rule of inference 'generalization', it suffices to establish (4) for the formulas λ in (Ax1)–(Ax6).
  - For λ in (Ax1), (4) follows by showing that every tautology has an $\emptyset$-proof that uses only the axioms in ($\Box$ax1); in other words, the set of sentential formulas of the form indicated in ($\Box$ax1) is a set of logical axioms for a proof system for sentential logic.[12]
  - For λ in (Ax2), (4) is proved in [1] in several stages[13] which, when combined together, provide an explicit $\emptyset$-proof establishing $\vdash_\Box \lambda$.
  - Even for the simple formulas λ in (Ax5), a nontrivial $\emptyset$-proof is needed to verify (4).[14]
- These facts indicate that it is possible to prove (without relying on the Soundness/Completeness Theorems) that the two sets of logical axioms have the same strength. The main difference in strength between the two proof systems lies in the rules of inferences allowed in Γ-proofs, namely that the proof system $\vdash_\Box$ allows an unrestricted use of 'generalization'.

[12]See pp. 6–11 in [1].
[13]See pp. 34–36 in [1].
[14]See Proposition 3.4 in [1]

Metatheorems. Metatheorems are theorems (stated in a natural language) about Γ-proofs and Γ-theorems (i.e., proofs and theorems in a formal language). We are interested in the metatheorems which express (within first-order logic) that certain methods of proof used in mathematical practice are 'correct'. We will restrict the discussion to the proof system $\vdash$.

Definition 3.10. We say that a set Γ of formulas is *inconsistent* if $\Gamma \vdash \beta$ and $\Gamma \vdash \neg\beta$ both hold for some formula β.

It is easy to see that if Γ is inconsistent, then $\Gamma \vdash \chi$ for every formula χ.

Metatheorems 3.11. *Let Γ, Δ be sets of $\mathcal{L}_C$-formulas and φ, ψ be further $\mathcal{L}_C$-formulas.*

(i) *If $\Gamma' \subseteq \Gamma$ and $\Gamma' \vdash \varphi$, then $\Gamma \vdash \varphi$.*

(ii) *If $\Gamma \cup \Delta \vdash \varphi$ and $\Gamma \vdash \delta$ for every $\delta \in \Delta$, then $\Gamma \vdash \varphi$.*

(iii) (**Deduction Theorem**)
If $\Gamma \cup \{\varphi\} \vdash \psi$, then $\Gamma \vdash \varphi \to \psi$.
(Note: The converse is easily seen to be true by MP.)

(iv) (**Generalization Theorem**)
If $\Gamma \vdash \varphi$ and x is a variable which is not free in any $\gamma \in \Gamma$, then $\Gamma \vdash \forall x\, \varphi$.
(Note: The converse is true by the logical axiom $\forall x\, \varphi \to \varphi$ in (Ax2) and MP.)

(v) (**Proof by Contraposition**)
$\Gamma \cup \{\varphi\} \vdash \neg\psi$ if and only if $\Gamma \cup \{\psi\} \vdash \neg\varphi$.

(vi) (**Reductio ad Absurdum** or **Proof by Contradiction**)
$\Gamma \cup \{\neg\varphi\}$ is inconsistent if and only if $\Gamma \vdash \varphi$.

(vii) (**Generalization on Constants**) *Let d be a constant symbol not in C.*
If $\Gamma \vdash \mathrm{Subf}^x_d(\varphi)$ (in $\mathcal{L}_{C\cup\{d\}}$), then $\Gamma \vdash \forall x\, \varphi$ (in $\mathcal{L}_C$).
(Again, the converse is true by the axiom $\forall x\, \varphi \to \mathrm{Subf}^x_d(\varphi)$ in (Ax2) and MP.)

(viii) (**Existential Instantiation**) *Let d be a constant symbol not in C.*
If $\Gamma \cup \{\mathrm{Subf}^x_d(\varphi)\} \vdash \psi$ (in $\mathcal{L}_{C\cup\{d\}}$), then $\Gamma \cup \{\exists x\, \varphi\} \vdash \psi$ (in $\mathcal{L}_C$).
(The converse is also true, because by Contraposition, this metatheorem reduces to one of the form (vii).)

We conclude by two statements on bound variables, which confirm that — with some reasonable restrictions — the choice of bound variables in formulas is irrelevant. The proofs are not hard if one uses some of the metatheorems above (and induction on formulas).

Corollary 3.12. *For any $\mathcal{L}_C$-formula φ and any variables x, y such that y does not occur in φ, we have that*

$$\mathrm{Subf}\,^y_x\big(\mathrm{Subf}\,^x_y(\varphi)\big) \equiv \varphi,$$

and the two formulas $\forall x\, \varphi$ and $\forall y\, \mathrm{Subf}\,^x_y(\varphi)$ are 'provably equivalent', that is,

$$\forall x\, \varphi \vdash \forall y\, \mathrm{Subf}\,^x_y(\varphi) \quad \textit{and} \quad \forall y\, \mathrm{Subf}\,^x_y(\varphi) \vdash \forall x\, \varphi.$$

Corollary 3.13. *Let χ be an $\mathcal{L}_C$-formula, let $\forall x\, \varphi$ be a subformula of χ, and let y be a variable not occurring in φ. If χ' is obtained from χ by replacing one or more occurrences of the subformula $\forall x\, \varphi$ by $\forall y\, \mathrm{Subf}\,^x_y(\varphi)$, then $\chi \vdash \chi'$ and $\chi' \vdash \chi$.*

Remark. The proof system for $\models$ discussed above, along with the proofs of the Soundness and Completeness Theorems, Theorems 3.4–3.5, can be found in Sections 2.4–2.5 of [2]. The proof system for $\models_\Box$ and the proofs of the Soundness and Completeness Theorems, Theorems 3.8–3.9, can be found in Sections 3–4 of [1]. For the Metatheorems 3.11(iii)–(viii), see Section 2.4 of [2]. The Metaheorems 3.11(i)–(ii) are straightforward to prove.