Mathematical Logic

A summary of "A Concise Introduction to Mathematical Logic", W. Rautenberg

Ari Feiglin

Contents

1 Propositional Logic

1.1 Semantics of Propositional Logic

Propositional logic is the study of logic removed from interpretation of individual variables and context. I will assume that the reader already has experience with propositional logic, as this is something an undergraduate will cover in one of their first courses. While this subsection will focus mainly on the semantics of propositional logic, we will begin by defining its *syntax*,

1.1.1 Definition

Let PV be an arbitrary set of **propositional variables** (which are regarded as arbitrary symbols). **Propositional formulas** are formulas defined recursively by the following rules,

(1) Propositional variables in PV are formulas, called **prime** or **atomic** formulas.

(2) If α and β are formulas, then so are $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$, and $\neg\alpha$. $(\alpha \wedge \beta)$ is called the **conjunction** of α and β, $(\alpha \vee \beta)$ their **disjunction**, and $\neg\alpha$ the **negation** of α.

The set of all the formulas constructed in this matter is denoted $\mathcal{F}$.

We can generalize this definition; instead of utilizing only the symbols $\wedge$ and $\vee$, we can take a general *logical signature* σ consisting of logical connectives of differing arities. We then recursively define σ-formulas as following: if c is an n-ary logical connective in σ, and $\alpha_1, \dots, \alpha_n$ are formulas, then so is

$$(\mathsf{c}\alpha_1, \dots, \alpha_n)$$

Aleternatively, if we only consider binary and unary connectives, then if c is a unary connective, we define $\mathsf{c}\alpha$ to be a formula, and if $\circ$ is a binary connective, then $(\alpha \circ \beta)$ is a formula. But we don't have much need for such generalizations, as $\{\wedge, \vee, \neg\}$ is complete, in the sense that all connectives can be defined using them. This is a fact we will discuss soon.

We can define other connectives, for example $\rightarrow$ and $\leftrightarrow$ are used as shorthands:

$$(\alpha \rightarrow \beta) := \neg(\alpha \wedge \neg\beta), \qquad (\alpha \leftrightarrow \beta) := \big((\alpha \rightarrow \beta) \wedge (\beta \rightarrow \alpha)\big)$$

We similarly define symbols for false and true:

$$\bot := (p_1 \wedge \neg p_1), \qquad \top = \neg\bot$$

For readability, we will use the following conventions when writing formulas (this is not a change to the definition of a formula, rather conventions for writing them in order to enhance readability)

(1) We will omit the outermost parentheses when writing formulas, if there are any.

(2) The order of operations for logical connectives is as follows, from first to last: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$.

(3) We associate $\rightarrow$ from the right, meaning $\alpha \rightarrow \beta \rightarrow \gamma$ is to be read as $\alpha \rightarrow (\beta \rightarrow \gamma)$. All other connectives associate from the left, for example $\alpha \wedge \beta \wedge \gamma$ is to be read as $(\alpha \wedge \beta) \wedge \gamma$.

(4) Instead of writing $\alpha_0 \wedge \alpha_1 \wedge \cdots \alpha_n$, we write $\bigwedge_{i=0}^{n} \alpha_i$, similar for $\vee$.

Since formulas are constructed in a recursive manner, most of our proofs about them are inductive.

1.1.2 Principle (Principle of Formula Induction)

Let $\mathcal{E}$ be a property of strings which satisfies the following conditions:

(1) $\mathcal{E}\pi$ for all prime formulas π,

(2) If $\mathcal{E}\alpha$ and $\mathcal{E}\beta$, then $\mathcal{E}(\alpha \wedge \beta)$, $\mathcal{E}(\alpha \vee \beta)$, and $\mathcal{E}\neg\alpha$ for all formulas $\alpha, \beta \in \mathcal{F}$.

Then $\mathcal{E}\varphi$ is true for all formulas φ.

An example of this is that every formula $\varphi \in \mathcal{F}$ is either prime, or of one of the following forms

$$\varphi = \neg\alpha, \quad \varphi = (\alpha \wedge \beta), \quad \varphi = (\alpha \vee \beta)$$

The proof of this is straightforward: let $\mathcal{E}$ be this property. Then trivially, $\mathcal{E}\pi$ for all prime formulas π. And if $\mathcal{E}\alpha$ and $\mathcal{E}\beta$, then of course we have

$$\mathcal{E}\neg\alpha, \quad \mathcal{E}(\alpha \wedge \beta), \quad \mathcal{E}(\alpha \vee \beta)$$

This is the first step in showing the *unique formula reconstruction property*. Let us prove a lemma before proving the property itself,

1.1.3 Lemma

Proper initial segments of formulas are not formulas. Equivalently (by contrapositive), if α and β are formulas and $\alpha\xi = \beta\eta$ for arbitrary strings ξ and η, then $\alpha = \beta$.

Let us prove this by induction on α. If α is a prime formula, suppose that β is not a prime formula, then its first character is either (or $\neg$, but then $\alpha =$ (or $\alpha = \neg$, in contradiction. Thus β is a prime formula and so $\alpha = \beta$ as they are both a single character. Now if $\alpha = (\alpha_1 \circ \alpha_2)$, then the first character of β must too be (, so β is of the form $(\beta_1 * \beta_2)$. Thus

$$\alpha_1 \circ \alpha_2)\xi = \beta_1 * \beta_2)\eta$$

and so by our inductive assumption, $\alpha_1 = \beta_1$, and so $\circ = *$, and thus $\alpha_2 = \beta_2$ by our inductive assumption again. And so $\alpha = \beta$ as required. The proof for the case that $\alpha = \neg\alpha'$ is similar. ■

1.1.4 Proposition (Unique Formula Reconstruction Property)

Every compound formula $\varphi \in \mathcal{F}$ is of one of the following forms:

$$\varphi = \neg\alpha, \quad \varphi = (\alpha \wedge \beta), \quad \varphi = (\alpha \vee \beta)$$

For some formulas $\alpha, \beta \in \mathcal{F}$.

We have already shown existence. We will now show that this is unique, meaning that φ can be written uniquely as one of these strings. Using the lemma proven above, the proof for uniqueness of the reconstruction property is immediate. For example, if $\varphi = (\alpha_1 \wedge \beta_1)$ then obviously φ cannot be written as $\neg\alpha_2$ since $(\neq \neg$, and if $\varphi = (\alpha_2 \vee \beta_2)$ then by the lemma $\alpha_1 = \alpha_2$, and so we get that $\wedge = \vee$ in contradiction. And finally if $\varphi = (\alpha_2 \wedge \beta_2)$, then again by the lemma, $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$ as required. The proof for $\neg$ and $\vee$ are similar. ■

Utilizing formula recursion, we can define functions on formulas. For example,

1.1.5 Definition

For a formula φ, we define $\mathit{Sf}\varphi$ to be the set of all subformulas of φ. This is done recursively:

$$\mathit{Sf}\pi = \{\pi\} \text{ for prime formulas } \pi,$$
$$\mathit{Sf}\neg\alpha = \mathit{Sf}\alpha \cup \{\alpha\}, \quad \mathit{Sf}(\alpha \circ \beta) = \mathit{Sf}\alpha \cup \mathit{Sf}\beta \cup \{(\alpha \circ \beta)\} \text{ for a binary logical connective } \circ$$

Similarly, we can define the **rank** of a formula φ,

$$\mathit{rank}\pi = 0 \text{ for prime formulas } \pi,$$
$$\mathit{rank}\neg\alpha = \mathit{rank}\alpha + 1, \quad \mathit{rank}(\alpha \circ \beta) = \max\{\mathit{rank}\alpha, \mathit{rank}\beta\} + 1 \text{ for a binary logical connective } \circ$$

And we can also define the set of variables in φ,

$$\mathit{var}\pi = \{\pi\} \text{ for prime formulas } \pi,$$
$$\mathit{var}\neg\alpha = \mathit{var}\alpha, \quad \mathit{var}(\alpha \circ \beta) = \mathit{var}\alpha \cup \mathit{var}\beta \text{ for a binary logical connective } \circ$$

In all definitions $\circ$ is either $\wedge$ or $\vee$.

So now that we have discussed the syntax of propositional logic, it is time to discuss its semantics; how we assign to formulas truth values. Recall the truth tables for $\wedge$, $\vee$, and $\neg$:

α	β	$\alpha\wedge\beta$
1	1	1
1	0	0
0	1	0
0	0	0

α	β	$\alpha\vee\beta$
1	1	1
1	0	1
0	1	1
0	0	0

α	$\neg\alpha$
1	0
0	1

These define how the logical connectives function as functions on $\{0,1\}$.

1.1.6 Definition

A **propositional valuation**, or a **propositional model**, is a function

$$w\colon PV \longrightarrow \{0,1\}$$

We can extend it to a function $w\colon PV \longrightarrow \mathcal{F}$ as follows:

$$w(\alpha\wedge\beta)=w\alpha\wedge w\beta,\quad w(\alpha\vee\beta)=w\alpha\vee w\beta,\quad w\neg\alpha=\neg w\alpha$$

Notice that we would need to define, for example, $w(\alpha\to\beta)=w\alpha\to w\beta$ had $\to$ been an element of our logical signature. But since $\to$ is defined using $\wedge$ and $\neg$, we must prove this identity:

$$w(\alpha\to\beta)=w\neg(\alpha\wedge\neg\beta)=\neg w(\alpha\wedge\neg\beta)=\neg(w\alpha\wedge\neg w\beta)=w\alpha\to w\beta$$

This is of course not a coincidence, but a result of the fact that $\alpha\to\beta=\neg(\alpha\wedge\neg\beta)$ (where $\alpha,\beta\in\{0,1\}$). Notice that furthermore,

$$w\top=1,\quad w\bot=0$$

1.1.7 Proposition

The valuation of a formula is dependent only on its variables. Meaning if φ is a formula and w and w' are two valuations where $w\pi=w'\pi$ for all $\pi\in var\varphi$, then $w\varphi=w'\varphi$.

We will prove this by induction on φ. For prime formulas, this is obvious as $var\varphi=\{\varphi\}$ and then $w\varphi=w'\varphi$ by the proposition's assumption. For $\varphi=\alpha\wedge\beta$, we have that

$$w\varphi=w\alpha\wedge w\beta=w'\alpha\wedge w'\beta=w'\varphi$$

where the second equality is our inductive assumption. The proof for $\varphi=\alpha\vee\beta$ and $\varphi=\neg\alpha$ is similar.

■

Let us suppose that $PV=\{p_1,p_2,\dots,p_n,\dots\}$, then we define $\mathcal{F}_n$ to be the set of formulas φ such that $var\varphi\subseteq\{p_1,\dots,p_n\}$.

1.1.8 Definition

A **boolean function** is a function

$$f\colon\{0,1\}^n\longrightarrow\{0,1\}$$

for some $n\geq 0$. The set of boolean functions of arity n is denoted $\mathbf{B}_n$. A formula $\varphi\in\mathcal{F}_n$ **represents** a boolean function $f\in\mathbf{B}_n$ (similarly, f is represented by φ), if for all valuations w,

$$w\varphi=f(w\vec{p})\qquad(w\vec{p}=(wp_1,\dots,wp_n))$$

So for example, $\alpha=p_1\wedge p_2$ represents the function $f(p,q)=p\wedge q$. This is since

$$f(wp_1,wp_2)=wp_1\wedge wp_2=w(p_1\wedge p_2)=w\alpha$$

Since valuations of $\varphi \in \mathcal{F}_n$ are defined by their values on $p_1, \dots, p_n$, φ represents at most a single function f. In fact, it represents the function

$$\varphi^{(n)}(x_1, \dots, x_n) = w\varphi$$

where w is any valuation such that $wp_i = x_i$ (all of these valuations valuate φ the same). Now, notice that $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ and $\mathbf{B}_n \subset \mathbf{B}_{n+1}$ and so $\varphi \in \mathcal{F}_n$ represents a function in $\mathcal{B}_{n+1}$ as well. But this function is not essentially in $\mathcal{B}_n$ in the sense that its last argument does not impact its value. Formally we say that given a function $f\colon M^n \longrightarrow M$, we call its ith argument *fictional* if for all $x_1, \dots, x_i, \dots, x_n \in M$ and $x_i' \in M$:

$$f(x_1, \dots, x_i, \dots, x_n) = f(x_1, \dots, x_i', \dots, x_n)$$

An *essentially n-ary* function is a function with no fictional arguments.

1.1.9 Definition

Two formulas α and β are **equivalent** if for every valuation w, $w\alpha = w\beta$. This is denoted $\alpha \equiv \beta$.

It is immediate that α and β are equivalent if and only if they represent the same function. A simple example of equivalence is $\alpha \equiv \neg\neg\alpha$. The following equivalences are easy to verify and the reader should already be familiar with them (α, β, and γ are formulas):

$$\begin{array}{ccl}
\alpha \wedge (\beta \wedge \gamma) \equiv \alpha \wedge \beta \wedge \gamma & \alpha \vee (\beta \vee \gamma) \equiv \alpha \vee \beta\gamma & \text{(associativity)} \\
\alpha \wedge \beta \equiv \beta \wedge \alpha & \alpha \vee \beta \equiv \beta \vee \alpha & \text{(commutativity)} \\
\alpha \wedge \alpha \equiv \alpha & \alpha \vee \alpha \equiv \alpha & \text{(idempotency)} \\
\alpha \wedge (\alpha \vee \beta) \equiv \alpha & \alpha \vee \alpha \wedge \beta \equiv \alpha & \text{(absorption)} \\
\multicolumn{2}{c}{\alpha \wedge (\beta \vee \gamma) \equiv \alpha \wedge \beta \vee \alpha \wedge \gamma} & (\wedge\text{-distributivity}) \\
\multicolumn{2}{c}{\alpha \vee \beta \wedge \gamma \equiv (\alpha \vee \beta) \wedge (\alpha \vee \gamma)} & (\vee\text{-distributivity}) \\
\neg(\alpha \wedge \beta) \equiv \neg\alpha \vee \neg\beta & \neg(\alpha \vee \beta) \equiv \neg\alpha \wedge \neg\beta & \text{(de Morgan rules)}
\end{array}$$

Furthermore,

$$\alpha \vee \neg\alpha \equiv \top, \quad \alpha \wedge \neg\alpha \equiv \bot, \quad \alpha \wedge \top \equiv \alpha \vee \bot \equiv \alpha$$

Since $\alpha \to \beta = \neg(\alpha \wedge \neg\beta)$, by de Morgan rules, this is equivalent to

$$\equiv \neg\alpha \vee \neg\neg\beta \equiv \neg\alpha \vee \beta$$

Notice that

$$\alpha \to \beta \to \gamma \equiv \neg\alpha \vee (\beta \to \gamma) \equiv \neg\alpha \vee \neg\beta \vee \gamma \equiv \neg(\alpha \wedge \beta) \vee \gamma \equiv \alpha \wedge \beta \to \gamma$$

Inductively, we see that

$$\alpha_1 \to \cdots \to \alpha_n \to \gamma \equiv \alpha_1 \wedge \cdots \wedge \alpha_n \to \gamma$$

We could go on, but I assume you get the point.

$\equiv$ is obviously reflexive, symmetric, and transitive: therefore it is an equivalence relation on $\mathcal{F}$. But moreso it is a *congruence relation*, meaning it respects connectives. Explicitly, for all formulas $\alpha, \beta, \alpha', \beta' \in \mathcal{F}$:

$$\alpha \equiv \alpha',\, \beta \equiv \beta' \implies \alpha \wedge \beta \equiv \alpha' \wedge \beta',\, \alpha \vee \beta \equiv \alpha' \vee \beta',\, \neg\alpha \equiv \neg\alpha'$$

Congruence relations will be discussed in more generality in later sections. Inductively, we can prove the following result:

1.1.10 Theorem (The Replacement Theorem)

Suppose α and α' are equivalent formulas. Let φ be some other formula, and define φ' to be the result of replacing all occurences of α within φ by α'. Then $\varphi \equiv \varphi'$.

This will be proven more generally later.

1.1.11 Definition

Prime formulas and their negations are called **literals**. A formula of the form $\alpha_1 \vee \cdots \vee \alpha_n$ where each

α_i is a conjunction of literals is called a **disjunctive normal form**. And similarly a formula of the form $\alpha_1 \wedge \cdots \wedge \alpha_n$ where each α_i is a disjunction of literals is called a **conjunctive normal form**. We will use the abreviations DNF and CNF for disjunctive and conjunctive normal forms, respectively.

So a DNF is a formula of the form

$$\bigvee_{i=1}^{n} \bigwedge_{j=1}^{m_i} \ell_{i,j}$$

where for every i, j, $\ell_{i,j}$ is a literal: a formula of the form $p_{i,j}$ or $\neg p_{i,j}$ for some prime formula $p_{i,j}$. Similarly a CNF is a formula of the form

$$\bigwedge_{i=1}^{n} \bigvee_{j=1}^{m_i} \ell_{i,j}$$

Let us temporarily introduce the following notation: for a prime formula p, let

$$p^1 := p, \qquad p^0 := \neg p$$

This allows us to more concisely state and prove the following theorem:

1.1.12 Theorem

Every boolean function $f \in \mathbf{B}_n$ for $n > 0$ is representable by the DNF

$$\alpha_f := \bigvee_{f(\vec{x})=1} p_1^{x_1} \wedge \cdots \wedge p_n^{x_n}$$

and a CNF

$$\beta_f := \bigwedge_{f(\vec{x})=0} p_1^{\neg x_1} \wedge \cdots \wedge p_n^{\neg x_n}$$

Let w be a valuation and $\vec{p} = (p_1, \ldots, p_n)$ then

$$w\alpha_f = \bigvee_{f(\vec{x})=1} wp_1^{x_1} \wedge \cdots \wedge wp_n^{x_n}$$

Notice that wp^x is equal to 1 if and only if $wp = x$: suppose $x = 0$ then $wp^x = \neg wp$, which is equal to 1 if and only if $wp = 0 = x$, and similar for $x = 1$. Thus $w\alpha_f = 1$ if and only if there exists a $\vec{x}$ such that $f(\vec{x}) = 1$ and $wp_1^{x_1} \wedge \cdots \wedge wp_n^{x_n} = 1$, meaning that for each i, $wp_i = x_i$. This means that $\vec{x} = w\vec{p}$, and so $f(w\vec{p}) = f(\vec{x}) = 1$. Similarly if $f(w\vec{p}) = 1$ then let $\vec{x} = w\vec{p}$, and then $wp_1^{x_1} \wedge \cdots \wedge wp_n^{x_n} = 1$ and $f(\vec{x}) = 1$, so $w\alpha_f = 1$. So $w\alpha_f = f(w\vec{p})$ for all valuations w, which means that f is represented by α_f, as required. The proof for β_f is similar. ■

Notice that since every formula represents a boolean function, which by above can be represented by a DNF and a CNF, we get that every formula is equivalent to a DNF and a CNF.

1.1.13 Corollary

Every formula is equivalent to a DNF and a CNF.

1.1.14 Definition

A logical signature σ is **functional complete** if every boolean function is representable by a formula in this signature.

By corollary 1.1.13, $\{\neg, \wedge, \vee\}$ is functional complete. Since

$$\alpha \vee \beta \equiv \neg(\neg\alpha \wedge \neg\beta), \quad \alpha \wedge \beta \equiv \neg(\neg\alpha \vee \neg\beta)$$

$\{\neg, \wedge\}$ and $\{\neg, \vee\}$ are both functional complete. Thus in order to show that a logical signature σ is functional complete, it is sufficient to show that $\neg$ and $\wedge$ or $\neg$ and $\vee$ can be represented by σ.

Note

If f is a function, instead of writing $f(x)$ or fx, many times we will instead write x^f. This is more concise and may reduce confusion in the case that x itself is a string wrapped in parentheses.

Let us define the function $\delta\colon \mathcal{F} \longrightarrow \mathcal{F}$ on formulas recursively by $p^\delta = p$ for prime formulas p and

$$(\neg\alpha)^\delta = \neg\alpha^\delta, \quad (\alpha \wedge \beta)^\delta = \alpha^\delta \vee \beta^\delta, \quad (\alpha \vee \beta)^\delta = \alpha^\delta \wedge \beta^\delta$$

Alternatively, α^δ is simply the result of swapping all occurences of $\wedge$ with $\vee$, and all occurences of $\vee$ with $\wedge$. α^δ is called the *dual formula* of α. Notice that the dual formula of a DNF is a CNF, and vice versa.

Now, suppose $f \in \mathbf{B}_n$, then let us define the *dual* of f,

$$f^\delta(\vec{x}) := \neg f(\neg\vec{x})$$

where $\neg\vec{x} = (\neg x_1, \dots, \neg x_n)$. Notice that δ is idempotent:

$$f^{\delta^2}(\vec{x}) = \neg f^\delta(\neg\vec{x}) = \neg\neg f(\neg\neg\vec{x}) = f(\vec{x})$$

1.1.15 Theorem (The Duality Principle for Two-Valued Logic)

If α represents the function f, then α^δ represents f^δ.

We will prove this by induction on α. If $\alpha = p$ is prime, then this is trivial. Now suppose that α and β represent f_1 and f_2 respectively. Then $\alpha \wedge \beta$ represents $f(\vec{x}) = f_1(\vec{x}) \wedge f_2(\vec{x})$, and $(\alpha \wedge \beta)^\delta = \alpha^\delta \vee \beta^\delta$ represents $g(\vec{x}) = f_1^\delta(\vec{x}) \vee f_2^\delta(\vec{x})$ by the induction hypothesis. Now,

$$f^\delta(\vec{x}) = \neg f(\neg\vec{x}) = \neg(f_1(\neg\vec{x}) \wedge f_2(\neg\vec{x})) = \neg f_1(\neg\vec{x}) \vee \neg f_2(\neg\vec{x}) = f_1^\delta(\vec{x}) \vee f_2^\delta(\vec{x}) = g(\vec{x})$$

So $f^\delta = g$, meaning that $(\alpha \wedge \beta)^\delta$ does indeed represent f^δ. The proof for $\alpha \vee \beta$ is similar. Now suppose α represents f, then $\neg\alpha$ represents $\neg f$, and α^δ represents f^δ by the induction hypothesis. And so $(\neg\alpha)^\delta = \neg\alpha^\delta$ represents $\neg f^\delta$, which is equal to $(\neg f)^\delta$ since

$$(\neg f)^\delta(\vec{x}) = (\neg\neg f)(\neg\vec{x}) = \neg(\neg f(\neg\vec{x})) = \neg f^\delta(\vec{x})$$

And so $(\neg\alpha)^\delta$ represents $(\neg f)^\delta$, as required. ∎

1.1.16 Definition

Suppose α is a formula and w is a valuation. Instead of writing $w\alpha = 1$, we now write $w \vDash \alpha$, and this is read as "w *satisfies* α". If X is a set of formulas, we write $w \vDash X$ if $w \vDash \alpha$ for all $\alpha \in X$, and w is called a **propositional model** for X. A formula α (respectively a set of formulas X) is **satisfiable** if there is some valuation w such that $w \vDash \alpha$ (respectively $w \vDash X$). $\vDash$ is called the **satisfiability relation**.

$\vDash$ has the following immediate properties:

$$\begin{array}{rlrl} w \vDash p \iff & wp = 1\ (p \in PV) & \qquad w \vDash \alpha \iff & w \nvDash \alpha \\ w \vDash \alpha \wedge \beta \iff & w \vDash \alpha \text{ and } w \vDash \beta & \qquad w \vDash \alpha \vee \beta \iff & w \vDash \alpha \text{ or } w \vDash \beta \end{array}$$

These properties uniquely define $\vDash$, meaning we could have defined $\vDash$ recursively by these properties.

Notice that

$$w \vDash \alpha \to \beta \iff \text{if } w \vDash \alpha \text{ then } w \vDash \beta$$

This is due to the definition of $\to$ coinciding with our common usage of implication. Had we not defined $\to$, but instead added it to our logical signature, this above equivalence would have to be taken in the definition of the satisfiability relation (when axiomized by the above properties).

1.1.17 Definition

α is **logically valid**, or a **tautology**, if $w \vDash \alpha$ for all valuations w. This is abbreviated by $\vDash \alpha$. A formula which cannot be satisfied; ie. for all valuations w, $w \nvDash \alpha$; is called a **contradiction**.

For example, $\alpha \vee \neg\alpha$ is a tautology, while $\alpha \wedge \neg\alpha$ and $\alpha \leftrightarrow \neg\alpha$ are contradictions for all formulas α. Notice that the negation of a tautology is a contradiction and vice versa. $\top$ is a tautology and $\bot$ is a contradiction.

The following are important tautologies of implication (keep in mind how $\to$ associates from the right):

$$\begin{array}{ll}
p \to p & \text{(self-implication)} \\
(p \to q) \to (q \to r) \to (p \to r) & \text{(chain rule)} \\
(p \to q \to r) \to (q \to p \to r) & \text{(exchange of premises)} \\
p \to q \to p & \text{(premise change)} \\
(p \to q \to r) \to (p \to q) \to (p \to r) & \text{(Frege's formula)} \\
((p \to q) \to p) \to p & \text{(Peirce's formula)}
\end{array}$$

1.1.18 Definition

Suppose X is a set of formulas and α a formula, we say that α is a **logical consequence** if $w \vDash \alpha$ for every model w of X. In other words,

$$w \vDash X \implies w \vDash \alpha$$

This is denoted $X \vDash \alpha$.

Notice that $\vDash$ here is used for logical consequence (the consequence relation), and we used it before as the symbol for the satisfiability relation. Context will make it clear as to its meaning. We use the notation $\alpha_1, \dots, \alpha_n \vDash \beta$ to mean $\{\alpha_1, \dots, \alpha_n\} \vDash \beta$. This justifies the notation for tautologies: α is a tautology if and only if $\varnothing \vDash \alpha$ (since every valuation models $\varnothing$), which is shortened by the above notation to $\vDash \alpha$.

And we also use $X \vDash \alpha, \beta$ to mean $X \vDash \alpha$ and $X \vDash \beta$. And $X, \alpha \vDash \beta$ to mean $X \cup \{\alpha\} \vDash \beta$.

The following are examples of logical consequences

$$\alpha, \beta \vDash \alpha \wedge \beta, \quad \alpha \wedge \beta \vDash \alpha, \beta$$
$$\alpha, \alpha \to \beta \vDash \beta$$
$$X \vDash \bot \implies X \vDash \alpha \quad \text{for all formulas } \alpha$$
$$X, \alpha \vDash \beta,\ X, \neg\alpha \vDash \beta \implies X \vDash \beta$$

The final example is true because if $w \vDash X$ then either $w \vDash \alpha$ or $w \vDash \neg\alpha$, and in either case $w \vDash \beta$.

Let us now state some obvious properties of the consequence relation:

$$\begin{array}{ll}
\alpha \in X \implies X \vDash \alpha & (\textit{reflexivity}) \\
X \vDash \alpha \text{ and } X \subseteq X' \implies X' \vDash \alpha & (\textit{monotonicity}) \\
X \vDash Y \text{ and } Y \vDash \alpha \implies X \vDash \alpha & (\textit{transitivity})
\end{array}$$

1.1.19 Definition

A **propositional substitution** is a mapping from prime formulas to formulas, $\sigma\colon PV \longrightarrow \mathcal{F}$, which is extended to a mapping between formulas $\sigma\colon \mathcal{F} \longrightarrow \mathcal{F}$ recursively:

$$(\alpha \wedge \beta)^\sigma = \alpha^\sigma \wedge \beta^\sigma, \quad (\alpha \vee \beta)^\sigma = \alpha^\sigma \vee \beta^\sigma, \quad (\neg\alpha)^\sigma = \neg\alpha^\sigma$$

If X is a set of formulas, we define

$$X^\sigma = \{\varphi^\sigma \mid \varphi \in X\}$$

Besides being intuitively important, the following proposition gives more insight into the usefulness of substitutions:

1.1.20 Proposition

Let X be a set of formulas, and α a formula. Then

$$X \vDash \alpha \implies X^\sigma \vDash \alpha^\sigma$$

Thus in a sense consequence is invariant under substitution.

Let w be a valuation, then we define w^σ as follows:

$$w^\sigma p = wp^\sigma$$

for prime formulas p. Now we claim that

$$w \vDash \alpha^\sigma \iff w^\sigma \vDash \alpha$$

We will prove this by induction on α. In the case that $\alpha = p$ is prime, then $w \vDash p^\sigma$ if and only if $wp^\sigma = w^\sigma p = 1$, and so this is if and only if $w^\sigma \vDash p$. Now by induction,

$$w \vDash (\alpha \wedge \beta)^\sigma \iff w \vDash \alpha^\sigma \text{ and } w \vDash \beta^\sigma \iff w^\sigma \vDash \alpha \text{ and } w^\sigma \vDash \beta \iff w^\sigma \vDash \alpha \wedge \beta$$

where the second equivalence is due to the induction hypothesis. The proof for formulas of the form $\alpha \vee \beta$ and $\neg\alpha$ proceed in a similar fashion.

Now, suppose $w \vDash X^\sigma$. This is if and only if $w \vDash \varphi^\sigma$ for all $\varphi \in X$, which is if and only if $w^\sigma \vDash \varphi$ by above. So $w \vDash X^\sigma$ if and only if $w^\sigma \vDash X$. And so if $X \vDash \alpha$ then let $w \vDash X^\sigma$, then $w^\sigma \vDash X$ meaning $w^\sigma \vDash \alpha$ and so $w \vDash \alpha^\sigma$ by above. So $X^\sigma \vDash \alpha^\sigma$ as required. ∎

These four properties,

$$\begin{array}{ll}
\alpha \in X \implies X \vDash \alpha & (\textit{reflexivity}) \\
X \vDash \alpha \text{ and } X \subseteq X' \implies X' \vDash \alpha & (\textit{monotonicity}) \\
X \vDash Y \text{ and } Y \vDash \alpha \implies X \vDash \alpha & (\textit{transitivity}) \\
X \vDash \alpha \implies X^\sigma \vDash \alpha^\sigma & (\textit{substitution invariance})
\end{array}$$

are what define general consequence relations, and form the basis for a general theory of logical systems. Another property is

$$X \vDash \alpha \implies X_0 \vDash \alpha \text{ for some finite } X_0 \subseteq X \quad (\textit{finitary})$$

We will show in the next subsection that this is a property of our consequence relation.

Another property is the property

$$X, \alpha \vDash \beta \iff X \vDash \alpha \to \beta$$

termed the *semantic deduction theorem*. Let us prove the first direction: suppose w is a model of X, then if $w \vDash \alpha$ it is a model of $X \cup \{\alpha\}$ and so $w \vDash \beta$. So we have shown that if $w \vDash \alpha$, then $w \vDash \beta$, meaning $w \vDash \alpha \to \beta$ and so $X \vDash \alpha \to \beta$. end for the converse, if $w \vDash X, \alpha$ then it is a model of X and so $w \vDash \alpha, \alpha \to \beta$ and thus $w \vDash \beta$.

We can show by induction a generalization of this:

$$X, \alpha_1, \dots, \alpha_n \vDash \beta \iff X \vDash \alpha_1 \to \cdots \to \alpha_n \to \beta \iff X \vDash \alpha_1 \wedge \cdots \wedge \alpha_n \to \beta$$

The induction step is simple: take $X' = X \cup \{\alpha_1\}$ we get by our induction hypothesis,

$$\begin{aligned}
X, \alpha_1, \dots, \alpha_n \vDash \beta \iff X', \alpha_2, \dots, \alpha_n \vDash \beta \iff X, \alpha_1 \vDash \alpha_2 \to \cdots \to \alpha_n \to \beta \\
\iff X \vDash \alpha_1 \to \alpha_2 \to \cdots \to \alpha_n \to \beta
\end{aligned}$$

as required. The deduction theorem makes proving many tautologies relating to implication much easier.

1.2 Gentzen Calculi

To begin this subsection, we will define a derivability relation $\vdash$ which axiomatizes the important properties of the consequence relation $\vDash$. Our goal is to show that by using these axioms, $\vdash$ is equivalent to $\vDash$, and this will allow us to prove important facts about $\vDash$, namely its finitaryness.

1.2.1 Definition

We define **Gentzen style sequent calculus** of $\vdash$ as follows: $X \vdash \alpha$ is to be read as "α is derivable from X" where α is a formula and X is a set of formulas. A pair $(X, \alpha) \in \mathcal{P}(\mathcal{F}) \times \mathcal{F}$, or more suggestively written $X \vdash \alpha$, is called a **sequent**. Gentzen-style rules have the form

$$\frac{X_1 \vdash \alpha_1 \mid \cdots \mid X_n \vdash \alpha_n}{X \vdash \alpha}$$

Which is to be understood as meaning that if for every i, $X_i \vdash \alpha_i$, then $X \vdash \alpha$.

Gentzen calculus has the following basic rules:

$$\begin{array}{llll}
\text{(IS)} & \dfrac{}{\alpha \vdash \alpha} & \text{(MR)} & \dfrac{X \vdash \alpha}{X' \vDash \alpha} \quad (X \subseteq X') \\
(\wedge 1) & \dfrac{X \vdash \alpha, \beta}{X \vdash \alpha \wedge \beta} & (\wedge 2) & \dfrac{X \vdash \alpha \wedge \beta}{X \vdash \alpha, \beta} \\
(\neg 1) & \dfrac{X \vdash \alpha, \neg\alpha}{X \vdash \beta} & (\neg 2) & \dfrac{X, \alpha \vdash \beta \mid X, \neg\alpha \vdash \beta}{X \vdash \beta}
\end{array}$$

(IS means "initial sequent", MR means monotonacity rule.)

Now we say that α is derivable from X, in short $X \vdash \alpha$, if $S_n = X \vdash \alpha$ and there exists a sequence of sequents $(S_0; \ldots; S_n)$ where for every S_i, S_i is either an initial sequent (IS) or derivable using the basic rules from previous sequents in the sequence.

For example, we can derive $\alpha \wedge \beta$ from $\{\alpha, \beta\}$, meaning $\alpha, \beta \vdash \alpha \wedge \beta$. This can be done by the sequence:

$$\begin{pmatrix} \alpha \vdash \alpha & ; & \alpha, \beta \vdash \alpha & ; & \beta \vdash \beta & ; & \alpha, \beta \vdash \beta & ; & \alpha, \beta \vdash \alpha \wedge \beta \\ \text{IS} & ; & \text{MR} & ; & \text{IS} & ; & \text{MR} & ; & \wedge 1 \end{pmatrix}$$

Let us prove some more useful rules

$$\frac{X, \neg\alpha \vdash \alpha}{X \vdash \alpha}$$
(¬-elimination)

1	$X, \alpha \vdash \alpha$	(IS), (MR)
2	$X, \neg\alpha \vdash \alpha$	supposition
3	$X \vdash \alpha$	$(\neg 2)$

$$\frac{X, \neg\alpha \vdash \beta, \neg\beta}{X \vdash \alpha}$$
(reductio ad absurdum)

1	$X, \neg\alpha \vdash \beta, \neg\beta$	supposition
2	$X, \neg\alpha \vdash \alpha$	$(\neg 1)$
3	$X \vdash \alpha$	¬-elimination

$$\frac{X \vdash \alpha \to \beta}{X, \alpha \vdash \beta}$$
(→-elimination)

1	$X, \alpha, \neg\beta \vdash \alpha, \neg\beta$	(IS), (MR)
2	$X, \alpha, \neg\beta \vdash \alpha \wedge \neg\beta$	$(\wedge 1)$
3	$X \vdash \neg(\alpha \wedge \neg\beta)$	supposition
4	$X, \alpha, \neg\beta \vdash \neg(\alpha \wedge \neg\beta)$	(MR)
5	$X, \alpha, \neg\beta \vdash \beta$	$(\neg 1)$ on 2 and 4
6	$X, \alpha \vdash \beta$	¬-elimination

$$\frac{X \vdash \alpha \mid X, \alpha \vdash \beta}{X \vdash \beta}$$
(cut rule)

1	$X, \neg\alpha \vdash \alpha$	supposition, (MR)
2	$X, \neg\alpha \vdash \neg\alpha$	(IS), (MR)
3	$X, \neg\alpha \vdash \beta$	$(\neg 1)$
4	$X, \alpha \vdash \beta$	supposition
5	$X \vdash \beta$	$(\neg 2)$ on 3 and 4

$$\frac{X, \alpha \vdash \beta}{X \vdash \alpha \to \beta}$$
($\to$-introduction)

1	$X, \alpha \wedge \neg\beta, \alpha \vdash \beta$	supposition, (MR)
2	$X, \alpha \wedge \neg\beta \vdash \alpha$	(IS), (MR), ($\wedge$2)
3	$X, \alpha \wedge \neg\beta \vdash \beta$	cut rule
4	$X, \alpha \wedge \neg\beta \vdash \neg\beta$	(IS), (MR), ($\wedge$2)
5	$X, \alpha \wedge \neg\beta \vdash \alpha \to \beta$	($\neg$1)
6	$X, \neg(\alpha \wedge \neg\beta) \vdash \alpha \to \beta$	(IS), (MR)
7	$X \vdash \alpha \to \beta$	($\neg$2) on 5 and 6

$$\frac{X \vdash \alpha, \alpha \to \beta}{X \vdash \beta}$$
(modus ponens)

1	$X \vdash \alpha \to \beta$	supposition
2	$X, \alpha \to \beta$	$\to$-elimination
3	$X \vdash \alpha$	supposition
4	$X \vdash \beta$	cut rule

$\to$-elimination and $\to$-introduction give us the *syntactic deduction theorem*:

$$X, \alpha \vdash \beta \iff X \vdash \alpha \to \beta$$

Let R be a rule of the form

$$R\colon \frac{X_1 \vdash \alpha_1 \mid \cdots \mid X_n \vdash \alpha_n}{X \vdash \alpha}$$

Then we say that a property of sequents $\mathcal{E}$ is *closed under* R if $\mathcal{E}(X_1, \alpha_1), \ldots, \mathcal{E}(X_n, \alpha_n)$ implies $\mathcal{E}(X, \alpha)$.

1.2.2 Proposition (Principle of Rule Induction)

Let $\mathcal{E}$ be a property of sequents which is closed under all the basic rules of $\vdash$. Then $X \vdash \alpha$ implies $\mathcal{E}(X, \alpha)$.

We will prove this by induction on the length of the derivation of $S = X \vdash \alpha$, n. If $n = 1$ then $X \vdash \alpha$ must be an initial sequent and so by assumption $\mathcal{E}(X, \alpha)$. For the induction step, suppose the derivation is $(S_0; \ldots; S_n)$, so $S = S_n$. Then by our inductive hypothesis $\mathcal{E}S_i$ for all $i < n$. If S is an initial sequent then $\mathcal{E}S$ holds by assumption. Otherwise S is obtained by applying a basic rule on some of the sequents S_i for $i < n$. And since $\mathcal{E}S_i$ and $\mathcal{E}$ is closed under basic rules, we have that $\mathcal{E}S$ as required. ■

1.2.3 Lemma (Soundness of $\vdash$)

If $X \vdash \alpha$ then $X \vDash \alpha$. More suggestively,

$$\vdash \subseteq \vDash$$

Using the principle of rule induction, let $\mathcal{E}(X, \alpha)$ mean $X \vDash \alpha$ (formally this means $\mathcal{E} = \{(X, \alpha) \mid X \vDash \alpha\}$). Then we must show that $\mathcal{E}$ is closed under all the basic rules of $\vdash$. This means that we must show that

$$\alpha \vDash \alpha, \quad X \vDash \alpha \implies X' \vDash \alpha \text{ for } X \subseteq X', \quad X \vDash \alpha, \beta \iff X \vDash \alpha \wedge \beta,$$
$$X \vDash \alpha, \neg\alpha \implies X \vDash \beta, \quad X, \alpha \vDash \beta \text{ and } X, \neg\alpha \vDash \beta \implies X \vDash \beta$$

These are all readily verifiable (and some we have already shown). So $\mathcal{E}$ is indeed closed under all the basic rules of $\vdash$, and so $\mathcal{E}(X, \alpha)$ (meaning $X \vDash \alpha$) implies $X \vdash \alpha$. ■

The property above is called *soundness*, meaning $\vdash$ does not derive anything "incorrect".

1.2.4 Theorem

If $X \vdash \alpha$ then there exists a finite subset $X_0 \subseteq X$ such that $X_0 \vdash \alpha$.

Let $\mathcal{E}(X, \alpha)$ be the property that there exists a finite subset $X_0 \subseteq X$ such that $X_0 \vdash \alpha$. We will show that $\mathcal{E}$ is closed under the basic rules of $\vdash$. Trivially, $\mathcal{E}(X, \alpha)$ holds for $X = \{\alpha\}$, meaning $\mathcal{E}$ holds for (IS). And similarly if $\mathcal{E}(X,\alpha)$ and $X \subseteq X'$, since there exists a finite $X_0 \subseteq X$ such that $X_0 \vdash \alpha$, this same X_0 is a subset of X' and so $\mathcal{E}(X',\alpha)$ so $\mathcal{E}$ is closed under (MR).

Now if $\mathcal{E}(X,\alpha)$ and $\mathcal{E}(X,\beta)$ then suppose $X_1 \vdash \alpha$ and $X_2 \vdash \beta$ where $X_1, X_2 \subseteq X$ are finite. Then $X_0 = X_1 \cup X_2$ is finite, $X_0 \vdash \alpha, \beta$ and so $X_0 \vdash \alpha \wedge \beta$, and since $X_0 \subseteq X$ is finite, $\mathcal{E}(X, \alpha\wedge\beta)$ so $\mathcal{E}$ is closed under ($\wedge$1). Closure under the rest of the basic rules can be shown similarly. ■

1.2.5 Definition

A set of formulas X is **inconsistent** if $X \vdash \alpha$ for every formula α. If X is not inconsistent, it is termed **consistent**. X is **maximally consistent** if X is consistent but for every proper superset $X \subset Y$, Y is inconsistent.

Notice that X is inconsistent if and only if $X \vdash \bot$. Obviously if X is inconsistent, $X \vdash \bot$. Conversely, if $X \vdash \bot$ then $X \vdash p_1 \wedge \neg p_1$ and so by ($\wedge$2), $X \vdash p_1, \neg p_2$ and thus by ($\neg$1) for all formulas α, $X \vdash \alpha$.

Furthermore, if X is consistent it is maximally consistent if and only if for every formula α, either $\alpha \in X$ or $\neg\alpha \in X$ exclusively. If neither α nor $\neg\alpha$ are in X, then since X is maximally consistent, $X, \alpha \vdash \bot$ and $X, \neg\alpha \vdash \bot$ and therefore by ($\neg$2), $X \vdash \bot$ contradicting X's consistency. And if X contains α or $\neg\alpha$ for every formula α, then it is maximal: adding another formula α would mean that $\alpha, \neg\alpha \in X$ and so by (IS), (MR), and ($\neg$2), X would be inconsistent.

This means that maximally consistent sets X are *deductively closed*:

$$X \vdash \alpha \iff \alpha \in X$$

Obviously if $\alpha \in X$ then by (IS) and (MR), $X \vdash \alpha$. Now suppose that $X \vdash \alpha$, then since $\alpha \in X$ or $\neg\alpha \in X$, we cannot have $\neg\alpha \in X$ since X is consistent. Therefore $\alpha \in X$.

1.2.6 Lemma

The derivability relation has the following properties:

$$\texttt{C}^+\text{:}\quad X \vdash \alpha \iff X, \neg\alpha \vdash \bot, \qquad \texttt{C}^-\text{:}\quad X \vdash \neg\alpha \iff X, \alpha \vdash \bot$$

Meaning α is derivable from X if and only if $X \cup \{\neg\alpha\}$ is inconsistent. And similarly $\neg\alpha$ is derivable from X if and only if $X \cup \{\alpha\}$ is inconsistent.

We will prove $\texttt{C}^+$. Suppose $X \vdash \alpha$, then $X, \neg\alpha \vdash \alpha$ by (MR) and $X, \neg\alpha \vdash \neg\alpha$ by (IS) and (MR). Thus by ($\neg$1), $X, \neg\alpha \vdash \beta$ for all formulas β by ($\neg$1) and in particular, $X, \neg\alpha \vdash \bot$. Now suppose $X, \neg\alpha \vdash \bot$ then by ($\wedge$2) and ($\neg$1), we have $X, \neg\alpha \vdash \alpha$ then by $\neg$-elimination, $X \vdash \alpha$. $\texttt{C}^-$ is proven similarly. ■

1.2.7 Lemma (Lindenbaum's Theorem)

Every consistent set of formulas $X \subseteq \mathcal{F}$ can be extended to a maximally consistent set of formulas $X \subseteq X' \subseteq \mathcal{F}$.

Let us define the set

$$H = \{Y \subseteq \mathcal{F} \mid Y \text{ is consistent and } X \subseteq Y\}$$

This is partially ordered with respect to $\subseteq$, and since $X \in H$, H is not empty. Let $C \subseteq H$ be a chain, meaning that for every $Z, Y \in C$, either $Z \subseteq Y$ or $Y \subseteq Z$. Now we claim that $U = \bigcup C$ is an upper bound for C. So we must show that $U \in H$. Suppose not, then U is not consistent meaning $U \vdash \bot$. But then there must exist a finite $U_0 = \{\alpha_1, \dots, \alpha_n\} \subseteq U$ such that $U_0 \vdash \bot$. Now suppose $\alpha_i \in Y_i \in C$ then since C is linearly ordered, we can assume that every Y_i is contained within Y_n. But then by (MR), $Y_n \vdash \bot$ which conradicts $Y_n \in H$ being consistent.

So U is consistent and so $U \in H$, and obviously for every $Y \in C$, $Y \subseteq U$. So U is an upper bound for C, meaning that every chain in H has an upper bound in H, and so by Zorn's Lemma, H has a maximal element. This maximal element, call it X', is precisely a maximally consistent set containing X: it is consistent and contains X since it is in H, and it is maximal in H so for every $X \subseteq Y$, $Y \notin H$ so Y is inconsistent. ■

1.2.8 Lemma

A maximally consistent set of formulas X has the following property:

$$X \vdash \neg\alpha \iff X \nvdash \alpha$$

for all formulas α.

If $X \vdash \neg\alpha$ then $X \nvdash \alpha$ due to X's consistency. If $X \nvdash \alpha$ then $X \cup \{\neg\alpha\}$ is consistent in lieu of $\mathtt{C}^+$. But since X is maximal, $X \cup \{\neg\alpha\} = X$ meaning $\neg\alpha \in X$ and so by (IS) and (MR), $X \vdash \neg\alpha$. ■

1.2.9 Lemma

Maximally consistent sets are satisfiable.

Suppose X is maximally consistent, then let us define the valuation w by $w \vDash p \iff X \vdash p$. Then we claim that

$$X \vdash \alpha \iff w \vDash \alpha$$

This is trival for prime formulas. Now if $X \vdash \alpha \wedge \beta$:

$$X \vdash \alpha \wedge \beta \iff X \vdash \alpha, \beta \iff w \vDash \alpha, \beta \iff w \vDash \alpha \wedge \beta$$

where the second equivalence is due to the induction hypothesis. And if $X \vdash \neg\alpha$:

$$X \vdash \neg\alpha \iff X \nvdash \alpha \iff w \nvDash \alpha \iff w \vDash \neg\alpha$$

The first equivalence is due to the previous lemma, and the second is due to the induction hypothesis. And therefore $w \vDash X$, meaning X is satisfiable. ■

1.2.10 Theorem (The Completeness Theorem)

Let X and α be an arbitrary set of formulas and formula respectively. Then $X \vdash \alpha$ if and only if $X \vDash \alpha$. More suggestively,

$$\vdash = \vDash$$

We have already shown that $\vdash \subseteq \vDash$ and so all that remains is to show the converse. Suppose that $X \nvdash \alpha$, then $X, \neg\alpha$ is consistent by $\mathtt{C}^+$. Thus it can be extended to a maximally consistent set $X, \neg\alpha \subseteq X'$ which is satisfiable. Therefore so is $X, \neg\alpha$, which means that $X \nvDash \alpha$. ■

We get the following theorem as an immediate result from The Completeness Theorem and theorem 1.2.4:

1.2.11 Theorem

$X \vDash \alpha$ if and only if $X_0 \vDash \alpha$ for a finite $X_0 \subseteq X$.

1.2.12 Theorem (The Compactness Theorem)

A set $X \subseteq \mathcal{F}$ is satisfiable if and only if every finite $X_0 \subseteq X$ is satisfiable.

Obviously if X is satisfiable, so is $X_0 \subseteq X$. Now if X is not satisfiable, then $X \vdash \bot$ and so there exists a finite $X_0 \subseteq X$ such that $X_0 \vdash \bot$ (and so $X_0 \vDash \bot$) by the previous theorem. And so if X is not satisfiable, there exists a finite $X_0 \subseteq X$ which is not satisfiable. ■

Let us now give some examples of applications of the compactness theorem.

1.2.13 Proposition

Every set M can be linearly (also known as totally) ordered.

If M is finite, this is trivial: if $M = \{m_1, \ldots, m_n\}$ simply define $m_1 < \cdots < m_n$. Now let M be any set, let us define the propositional variable (aka prime formula) p_{ab} for every $(a, b) \in M \times M$. This will represent $a < b$. So we define X to be the set of the following formulas, which represents M being linearly ordered,

$$\begin{aligned} \neg p_{aa} &\quad (a \in M), \\ p_{ab} \wedge p_{bc} \to p_{ac} &\quad (a, b, c \in M), \\ p_{ab} \vee p_{ba} &\quad (a \neq b \in M) \end{aligned}$$

If X is satisfiable, suppose $w \vDash X$, then we define the linear order $a < b$ if and only if $w \vDash p_{ab}$. Thus X is precisely the set of conditions necessary for $<$ to be a linear order: the first condition is irreflexivity, the second is transitivity, and the third totality (antisymmetry is gained through the combination of irreflexitivity and transitivity).

So if X is satisfiable, then M can be linearly ordered. By the compactness theorem, we need only to show that every finite $X_0 \subseteq X$ is satisfiable. If $X_0 \subseteq X$ is finite, then let us define M_0 to be the set of all symbols in M which occur in formulas in X_0. Since X_0 is finite, so is M_0 and therefore M_0 can be linearly ordered. Let us define $w_0 \vDash p_{ab} \iff a < b$ in M_0, then $w_0 \vDash X_0$. So by the compactness theorem X is satisfiable, as required. ■

Recall that showing that every set can be well-ordered (the well-ordering theorem) is equivalent to the axiom of choice. Since the compactness theorem is actually weaker than the axiom of choice, the linear ordering theorem (what we just showed) is weaker than the well-ordering theorem. Which is not surprising.

1.2.14 Proposition

A graph is k-colorable if and only if every finite subgraph is k-colorable.

A *graph* is a pair $G = (V, E)$ where V is a set of *vertices* and E is a set of *edges*. E is a subset of $\{\{v, u\} \mid v \neq u \in V\}$. The graph G is k-colorable if V can be partitioned into k *color classes*: $V = C_1 \cup \cdots \cup C_k$ such that if $a, b \in C_i$ then $\{a, b\} \notin E$, meaning two neighboring vertices do not have the same color.

Obviously if a graph is k-colorable, so is every subgraph. To show the converse, let $G = (V, E)$ be a graph, then let us define the set of formulas X, where prime formulas are of the form $p_{a,i}$ where $a \in V$ and $1 \leq i \leq k$:

$$\begin{aligned} &p_{a,1} \vee \cdots \vee p_{a,k} (a \in V) \\ &\quad \neg(p_{a,i} \wedge p_{a,j})(a \in V,\, 1 \leq i < j \leq k) \\ &\quad \neg(p_{a,i} \wedge p_{b,i})(\{a, b\} \in E,\, 1 \leq i \leq k) \end{aligned}$$

If X is satisfiable, $w \vDash X$, then we define $C_i = \{a \in V \mid w \vDash p_{a,i}\}$, ie. we color $a \in V$ with the color i if and only if $p_{a,i}$ is satisfied. Then $V = C_1 \cup \cdots \cup C_k$ since for every $a \in V$, $w \vDash p_{a,1} \vee \cdots \vee p_{a,k}$, so for every $a \in V$ there exists an $1 \leq i \leq k$ such that $w \vDash p_{a,i}$ so $a \in C_i$. And $C_i \cap C_j = \varnothing$ in lieu of $\neg(p_{a,i} \wedge p_{a,j})$. And if $\{a, b\} \in E$ then a and b cannot be in the same color class by $\neg(p_{a,i} \wedge p_{b,i})$. So the C_is give a valid k-coloring of G.

Let $X_0 \subseteq X$ be finite, then let us define $G_0 = (V_0, E_0)$ where V_0 is the set of vertices appearing in formulas in X_0, and E_0 be the edges connecting them. By assumption, G_0 is k-colorable since it is finite. Now we define the valuation w_0 such that $w_0 \vDash p_{a,i}$ if and only if a is in the ith color class for $a \in V_0$. This must model X_0 since X_0 includes only statements saying that G_0 can be k-colored. So by the compactness theorem, X is satisfiable, as required. ■

There are more examples of applications of the compactness theorem. For example, the ultrafilter theorem, which we will visit later on.

1.3 Hilbert Calculi

In this subsection we will define another form of sequent calculus.

1.3.1 Definition

A define $\Lambda \subseteq \mathcal{F}$ to be a set of axioms, called the **logical axiom scheme**. Now, let Γ be a set of **rules of inference**, predicates of the form $R \in \Lambda^n \times \Lambda$ for $n > 0$, where $R((\varphi_1, \dots, \varphi_n), \varphi)$ which is to be understood as "*if* $\varphi_1, \dots, \varphi_n$ *then* φ".

If $X \subseteq \mathcal{F}$ is a set of formulas, then a **proof** is a sequence $\Phi = (\varphi_0, \dots, \varphi_n)$ where for every i, φ_i is either in $X \cup \Lambda$ or there exists a rule of inference $R \in \Gamma$ and indexes $i_1, \dots, i_n < i$ such that $R((\varphi_{i_1}, \dots, \varphi_{i_n}), \varphi)$. In such a case, φ_n is termed **derivable** (or **provable**) from X, and is written $X \vdash \varphi_n$ ($\vdash$ to differentiate it from the derivability relation $\vdash$ from the previous subsection).

Hilbert-style calculi will use the following axiom scheme Λ:

$\Lambda 1$	$(\alpha \to \beta \to \gamma) \to (\alpha \to \beta) \to \alpha \to \gamma$	$\Lambda 2$	$\alpha \to \beta \to \alpha \wedge \beta$
$\Lambda 3$	$\alpha \wedge \beta \to \alpha, \quad \alpha \wedge \beta \to \beta$	$\Lambda 4$	$(\alpha \to \neg\beta) \to \beta \to \neg\alpha$

And there is only a single rule of inference: $R((\alpha, \alpha \to \beta), \beta)$ called *modus ponens*, abbreviated MP. Essentially if α and $\alpha \to \beta$ then β.

The finiteness theorem for $\vdash$ is immediate, since $X \vdash \alpha$ requires a *finite* proof from X. And notice that

$$X \vdash \alpha, \alpha \to \beta \implies X \vdash \beta$$

Since if $\Phi_1 = (\varphi_0, \dots, \varphi_n)$ is a proof of α, and $\Phi_2 = (\varphi'_0, \dots, \varphi'_m)$ is a proof of $\alpha \to \beta$, then

$$\Phi = (\varphi_0, \dots, \varphi_n, \varphi'_0, \dots, \varphi'_m, \beta)$$

is a proof of $\alpha \to \beta$.

1.3.2 Proposition (Principle of Induction for $\vdash$)

Let X be a set of formulas and $\mathcal{E}$ a property of formulas. Then if

(1) $\mathcal{E}\alpha$ is true for all $\alpha \in X \cup \Lambda$, and

(2) $\mathcal{E}\alpha$ and $\mathcal{E}\alpha \to \beta$ implies $\mathcal{E}\beta$ for all formulas α, β.

Then $X \vdash \alpha$ implies $\mathcal{E}\alpha$.

We will prove this by induction on n, the length of the proof of α. If $n = 1$ then α is in $X \cup \Lambda$ and so by assumption $\mathcal{E}\alpha$. Now suppose $\Phi = (\varphi_0, \dots, \varphi_n)$ is a proof of $\alpha = \varphi_n$. If $\alpha \in X \cup \Lambda$ then by assumption $\mathcal{E}\alpha$. Otherwise Φ must contain formulas of the form α_i and $\alpha_i \to \alpha$. Since initial segments of proofs are themselves proofs, by our inductive hypothesis $\mathcal{E}\varphi_i$ for $i < n$. And thus $\mathcal{E}\alpha_i$ and $\mathcal{E}\alpha_i \to \alpha$ and so $\mathcal{E}\alpha$ as required. ■

This can obviously be generalized to a principle of induction for general rules of inferences, where the second condition is replaced with a general notion of closure under rules of inference.

Now we can show that $\vdash \subseteq \vDash$, meaning if $X \vdash \alpha$ then $X \vDash \alpha$ by defining the property $\mathcal{E}\alpha := X \vDash \alpha$. Since Λ contains only tautologies, for every $\alpha \in X \cup \Lambda$, $X \vDash \alpha$ meaning $\mathcal{E}\alpha$ for all $\alpha \in X \cup \Lambda$. And if $X \vDash \alpha$ and $X \vDash \alpha \to \beta$ then we know $X \vDash \beta$. So $\mathcal{E}$ satisfies the inductive properties stated above, meaning $X \vdash \alpha$ implies $X \vDash \alpha$ as required.

Now, obviously $\vdash$ is reflexive, monotonic, and transitive. Reflexivity follows directly from its definition. Monotonicity follows because a proof in X is also a proof in $X \subseteq X'$. And transitivity follows because if $X \vdash Y$ and $Y \vdash \alpha$, then by concatenating the proofs of $\varphi \in Y$ in X together with the proof of α in Y gives a proof of α in X.

Our goal for the remainder of this subsection is showing that $\vdash = \vDash$, we will do this by showing that $\vdash \subseteq \vdash$. As explained above, $\vdash$ is reflexive and monotonic, meaning it satisfies (IS) and (MR) of the Gentzen-style calculus $\vdash$.

1.3.3 Lemma

(1) If $X \vdash \alpha \to \neg\beta$ then $X \vdash \beta \to \neg\alpha$

(2) $\vdash \alpha \to \beta \to \alpha$

(3) $\vdash \alpha \to \alpha$

(4) $\vdash \alpha \to \neg\neg\alpha$

(5) $\vdash \beta \to \neg\beta \to \alpha$

(1) By $\Lambda 4$, we have $X \vdash (\alpha \to \neg\beta) \to \beta \to \neg\alpha$, and since $X \vdash \alpha \to \neg\beta$ by modus ponens we get $X \vdash \beta \to \neg\alpha$.

(2) By $\Lambda 3$, $\vdash \beta \wedge \neg\alpha \to \neg\alpha$, and so by (1) we have $\vdash \alpha \to \neg(\beta \wedge \neg\alpha) = \alpha \to \beta \to \alpha$.

(3) Let $\gamma = \alpha$ and $\beta = \alpha \to \alpha$, then $\Lambda 1$ gives

$$\vdash (\alpha \to (\alpha \to \alpha) \to \alpha) \to (\alpha \to \alpha \to \alpha) \to \alpha \to \alpha$$

We know by (2), $\vdash \alpha \to (\alpha \to \alpha) \to \alpha$ and $\vdash \alpha \to \alpha \to \alpha$, so by applying modus ponens twice we get $\vdash \alpha \to \alpha$.

(4) Since $\vdash \neg\alpha \to \neg\alpha$ by (3), and applying (1) gives $\vdash \alpha \to \neg\neg\alpha$.

(5) By $\Lambda 3$, $\vdash \neg\beta \wedge \neg\alpha \to \neg\beta$, applying (1) gives $\vdash \beta \to \neg(\neg\beta \to \neg\alpha) = \beta \to \neg\beta \to \alpha$. ■

Since $\Lambda 3$ gives $\alpha \wedge \beta \to \alpha, \beta$, $\vdash$ satisfies ($\wedge 2$) of $\vdash$. $\Lambda 2$ gives $\alpha \to \beta \to (\alpha \wedge \beta)$ and so by applying MP twice, we get $\alpha, \beta \vdash \alpha \wedge \beta$ and so $\vdash$ satisfies ($\wedge 1$) of $\vdash$. Now by (5) of the above lemma, since $\vdash \alpha \to \neg\alpha \to \beta$, by applying MP twice we get that $X, \alpha, \neg\alpha \vdash \beta$ for all formulas β. By transitivity, this means that $X \vdash \alpha, \neg\alpha$ implies $X \vdash \beta$. Thus $\vdash$ satisfies (IS), (MR), ($\wedge 1$), ($\wedge 2$), and ($\neg 1$) of $\vdash$. We will now do a bit more work to show that it also satisfies ($\neg 2$).

1.3.4 Lemma (The Deduction Theorem)

$X, \alpha \vdash \gamma$ implies $X \vdash \alpha \to \gamma$.

We will prove this using the principle of induction for $\vdash$. Let $\mathcal{E}\gamma$ mean $X \vdash \alpha \to \gamma$, we will show that $X, \alpha \vdash \gamma$ implies $\mathcal{E}\gamma$ by showing $\mathcal{E}$ is closed under the inductive properties stated in the Principle of Induction for $\vdash$. If $\gamma \in \Lambda \cup X \cup \{\alpha\}$, if $\gamma = \alpha$ then we showed above that $X \vdash \alpha \to \alpha$. Otherwise if $\gamma \in X \cup \Lambda$ then $X \vdash \gamma$ and $X \vdash \gamma \to \alpha \to \gamma$, so by MP $X \vdash \alpha \to \gamma$ meaning $\mathcal{E}\gamma$ as required.
Now, if $\mathcal{E}\beta$ and $\mathcal{E}\beta \to \gamma$, meaning $X \vdash \alpha \to \beta$ and $X \vdash \alpha \to \beta \to \gamma$. Then by $\Lambda 1$, applying MP twice gives $X \vdash \alpha \to \gamma$ as required. ■

1.3.5 Lemma

$\vdash \neg\neg\alpha \to \alpha$

By $\Lambda 3$ and MP, we have $\neg\neg\alpha \wedge \neg\alpha \vdash \neg\alpha, \neg\neg\alpha$. Let τ be any formula where $\vdash \tau$, then since we have already verified rule ($\neg 1$), $\neg\neg\alpha \wedge \neg\alpha \vdash \neg\tau$. And so by the deduction theorem, $\vdash \neg\neg\alpha \wedge \neg\alpha \to \neg\tau$. We showed above that this means $\vdash \tau \to \neg(\neg\neg\alpha \wedge \neg\alpha)$, and since $\vdash \tau$ by MP we get $\vdash \neg(\neg\neg\alpha \wedge \neg\alpha) = \neg\neg\alpha \to \alpha$ as required. ■

1.3.6 Lemma

$\vdash$ also satisfies rule ($\neg 2$) of the Gentzen-style calculus $\vdash$.

This is the rule that $X, \alpha \vdash \beta$ and $X, \neg\alpha \vdash \beta$ implies $X \vdash \beta$. If $X, \alpha \vdash \beta$ and $X, \neg\alpha \vdash \beta$ then $X, \alpha \vdash \neg\neg\beta$ and $X, \neg\alpha \vdash \neg\neg\beta$. By the deduction theorem, this means $X \vdash \alpha \to \neg\neg\beta, \neg\alpha \to \neg\neg\beta$. And thus $X \vdash \neg\beta \to \neg\alpha, \neg\beta \to \neg\neg\alpha$. Thus MP yields $X, \neg\beta \vdash \neg\alpha, \neg\neg\alpha$. So let $\vdash \tau$ and so by ($\neg 1$), we get $X, \neg\beta \vdash \neg\tau$, and so again by the deduction theorem, $X \vdash \neg\beta \to \neg\tau$, meaning $X \vdash \tau \to \neg\neg\beta$. Since $\vdash \tau$ by MP we get $X \vdash \neg\neg\beta$ and since $X \vdash \neg\neg\beta \to \beta$, by MP we get $X \vdash \beta$ as required. ■

1.3.7 Theorem (The Completeness Theorem)

$X \mathrel{|\!\!\sim} \alpha$ if and only if $X \vDash \alpha$. More suggestively,

$$\mathrel{|\!\!\sim} \, = \, \vDash$$

We have already shown $\mathrel{|\!\!\sim} \subseteq \vDash$. Since $\mathrel{|\!\!\sim}$ satisfies all the basic rules of $\vdash$, $\vdash \subseteq \mathrel{|\!\!\sim}$ (by $\vdash$'s principle of induction). Now since $\vdash = \vDash$, we get that $\vDash \subseteq \mathrel{|\!\!\sim} \subseteq \vDash$, and so $\mathrel{|\!\!\sim} = \vDash$. ■

It is important to note that Λ is sufficient to obtain all tautologies only because $\rightarrow$ was defined via $\neg$ and $\wedge$. Had it been taken as just another connective, we would've needed to add axioms to Λ stating the relation between $\rightarrow$ and $\neg$ and $\wedge$.

2 First Order Logic

2.1 Mathematical Structures

Our first step in studying first order logic (which will be defined later) is defining the general notion of a *mathematical structure*. Mathematical structures (also known as first order structures) give a useful generalization of many of the algebraic and relational objects mathematicians study.

2.1.1 Definition

An **extralogical signature** is a set σ of symbols of three types: function symbols, relational symbols, and constant symbols. Function symbols and relational symbols are also given an **arity**, a positive integer.

Formally, we can view σ as a tuple: $\sigma = (\sigma_f, \sigma_r, \sigma_c, \mathsf{ar})$, where σ_f is a set of function symbols, σ_r is a set of relational symbols, and σ_c is a set of constant symbols (meaning that they are all just sets of symbols). Further assume that σ_f, σ_r, and σ_c are all disjoint. ar is a function mapping symbols in σ_f and σ_r to positive integers.

2.1.2 Definition

Let σ be an extralogical signature (for short, a signature), **mathematical structure** over σ (for short, a σ-structure) is a pair $\mathcal{A} = (A, \sigma^{\mathcal{A}})$ where A is some non-empty set, called the **domain** of the structure, and $\sigma^{\mathcal{A}}$ is an **interpretation** of σ. This means that for every function symbol $f \in \sigma$, $\sigma^{\mathcal{A}}$ consists of an operation $f^{\mathcal{A}}\colon A^{\mathsf{ar}(f)} \longrightarrow A$, for every relational symbol $r \in \sigma$, $\sigma^{\mathcal{A}}$ contains a relation $r^{\mathcal{A}} \subseteq A^{\mathsf{ar}(r)}$, and for every constant symbol $c \in \sigma$, $\sigma^{\mathcal{A}}$ contains a constant $c^{\mathcal{A}} \in A$.

Constants may be viewed as 0-ary operations.

The domain of a mathematical structure $\mathcal{A}$ will always be denoted by A.

We now define some general notions relating to structures.

Suppose $A \subseteq B$, and f is an n-ary operation on B. Then A is *closed under* f if $f(A^n) \subseteq A$, meaning that for every $\vec{a} \in A^n$, $f\vec{a} \in A$. If $n = 0$, ie. if f is a constant c, then this simply means that $c \in A$. It is obvious that the intersection of a family of sets closed under f is itself closed under f, and thus we can discuss the smallest set closed under f. For example, $\mathbb{N}$ is closed under $+$ (when viewed as a binary operation of $\mathbb{N}$, $\mathbb{Q}$, etc.), but not under $-$.

Suppose $A \subseteq B$ again, and r^B is an n-ary relation on B. Then the *restriction* of r^B to A is the n-ary relation $r^A = r^B \cap A^n$. For example the restriction of $<^{\mathbb{Z}}$, the standard order of $\mathbb{Z}$, to $\mathbb{N}$ is $<^{\mathbb{N}}$, the standard order of $\mathbb{N}$. If f^B is an n-ary operation on B and $A \subseteq B$ is closed under f^B, then we define f^B's restriction to A to be the operation $f^A\vec{a} = f^B\vec{a}$.

So if $\mathcal{B}$ is a σ-structure and $A \subseteq B$ is closed under all operations (including constants), then A can be given the structure of a σ-structure naturally: define $\mathcal{A} = (A, \sigma^{\mathcal{A}})$ where for $f \in \sigma_f$ take $f^{\mathcal{A}} = f^A$ the restriction of $f^{\mathcal{B}}$ to A, for $r \in \sigma_r$ take $r^{\mathcal{A}} = r^A$ the restriction of $r^{\mathcal{B}}$ to A, and for $c \in \sigma_c$ take $c^{\mathcal{A}} = c^{\mathcal{B}}$. $\mathcal{A}$ is called a *substructure* of $\mathcal{B}$, denoted $\mathcal{A} \subseteq \mathcal{B}$.

Note that not every subset $A \subseteq B$ can be extended to a substructure of $\mathcal{B}$. For example, $\{1\} \subseteq \mathbb{Z}$ but if the signature σ is taken to include the constant 0, then since $\{1\}$ does not contain $0^{\mathbb{Z}} = 0$ it cannot be extended to a substructure. And similarly if σ includes $+$, then since $\{+\}$ is not closed under $+$, it cannot be extended to a substructure.

Suppose $\mathcal{A}$ is a σ-structure and $\sigma_0 \subseteq \sigma$ is another extralogical signature (meaning $\sigma_{0_x} \subseteq \sigma_x$ for $x = f, r, c$ and $\mathsf{ar}_0(\mathsf{s}) = \mathsf{ar}(\mathsf{s})$ for all relational and function symbols $\mathsf{s} \in \sigma_0$). Then we define the σ_0-structure $\mathcal{A}_0$ where the interpretation of each symbol $\mathsf{s} \in \sigma_0$ is $\mathsf{s}^{\mathcal{A}_0} = \mathsf{s}^{\mathcal{A}}$. $\mathcal{A}_0$ is called the σ_0*-reduct* of $\mathcal{A}$, and conversely $\mathcal{A}$ is called the σ*-expansion* of $\mathcal{A}_0$.

Many times, if σ is a signature consisting of the symbols $\mathsf{s}_1, \mathsf{s}_2, \ldots,$, we will write a σ-structure as $(A, \mathsf{s}_1, \mathsf{s}_2, \ldots)$ instead of writing out the signature. And further, we will often write the signature as a set instead of as a tuple of sets and an arity function. What symbols are functions, relational, and constants, and their arities are to be understood from context.

Mathematical structures defined over a signature without relational symbols are termed *algebraic structures*, while structures defined over a signature without function or constant symbols are termed *relational structures*.

For example, mathematical structures of the form $\mathcal{A} = (A, \circ)$ where $\circ$ is a binary operation are called *magmas*. If $\circ$ is associative, $\mathcal{A}$ is a *semigroup*, if it is invertible in each argument then it is a *group*, etc. These are examples of very common algebraic structures. Another common algebraic structure are *rings* and *fields*: both are structures of the form $\mathcal{A} = (A, +, \cdot, 0, 1)$ which satisfy certain axioms. Notice that a structure of this form is not necessarily a ring, but all rings are structures of this form.

A *semilattice* is another type of algebraic structure, and is a special case of a magma where $\circ$ is associative, commutative, and idempotent (meaning $a \circ a = a$ for all $a \in A$). For example $(\{0,1\}, \wedge)$ is a semilattice. We can define the partial order $\leq$ by $a \leq b \iff a \circ b = a$. This is reflexive since $\circ$ is, anticommutative since $\circ$ is commutative, and if $a \leq b$ and $b \leq c$ then $a = a \circ b = a \circ (b \circ c) = (a \circ b) \circ c = a \circ c$ so $a \leq c$. And a *lattice* is an algebraic structure of the form $\mathcal{A} = (A, \cap, \cup)$ where $(A, \cap)$ and $(A, \cup)$ are both semilattices and the following absorption laws hold: $a \cap (a \cup b) = a$ and $a \cup (a \cap b) = a$. A *distributive lattice* is a lattice which satisfies the distributive properties: $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$ and $x \cup (y \cap z) = (x \cup y) \cap (z \cup z)$. For example if M is a set, then $(\mathcal{P}(M), \cap, \cup)$ is a lattice.

A *boolean algebra* is an algebraic structure $\mathcal{A} = (A, \cap, \cup, \neg)$ where the reduct $(A, \cap, \cup)$ is a distributive lattice and

$$\neg\neg x = x, \quad \neg(x \cap y) = \neg x \cup \neg y, \quad x \cap \neg x = y \cap \neg y$$

The standard example is the boolean algebra $\mathit{2} = (\{0,1\}, \wedge, \vee, \neg)$.

A relational structure $\mathcal{A} = (A, \lhd)$ where $\lhd$ is a binary relation is often called a *graph* (this coincides with the definition of a directed graph). If $\lhd$ is irreflexive and transitive, this is a *(strict) partially ordered set*, or poset for short, and we generally write $<$ for $\lhd$. A *partially ordered set* is when $\lhd$ is reflexive, transitive, and antisymmetric, then we usually write $\leq$ for $\lhd$. Each partially ordered set gives rise to a strict partially ordered set and vice versa, by defining $a \leq b \iff a < b \vee a = b$

2.1.3 Definition

Let σ be some signature, and $\mathcal{A}$ and $\mathcal{B}$ be σ-structures. Then a map $h\colon A \longrightarrow B$ (though we will generally write $h\colon \mathcal{A} \longrightarrow \mathcal{B}$) is called a **homomorphism** provided that for every function symbol f, relational symbol r, and constant symbol c in σ, and $\vec{a} \in A^n$:

$$h(f^{\mathcal{A}}(\vec{a})) = f^{\mathcal{B}}(h(\vec{a})), \qquad h(c^{\mathcal{A}}) = c^{\mathcal{B}}, \qquad r^{\mathcal{A}}(\vec{a}) \implies r^{\mathcal{B}}(h(\vec{a}))$$

where $h(\vec{a}) = (h(a_1), \ldots, h(a_n))$.

A **strong homomorphism** is a homomorphism where the third condition on relations is replaced by the stronger $r^{\mathcal{B}}(h(\vec{a}))$ if and only if there exists a $\vec{b} \in A^n$ such that $h(\vec{a}) = h(\vec{b})$ and $r^{\mathcal{A}}(\vec{b})$ (thus we need not require that every $\vec{b}$ with the same image as $\vec{a}$ under h satisfy $r^{\mathcal{A}}$, only that one does). In other words, the condition is replaced with

$$r^{\mathcal{B}}(h(\vec{a})) \iff (\exists \vec{b} \in A^n)\big(h(\vec{a}) = h(\vec{b}) \wedge r^{\mathcal{A}}(\vec{b})\big)$$

An injective strong homomorphism $\mathcal{A} \longrightarrow \mathcal{B}$ is called an **embedding** of $\mathcal{A}$ into $\mathcal{B}$. If further the embedding is surjective, it is termed a **isomorphism**. If there exists an isomorphism between $\mathcal{A}$ and $\mathcal{B}$, the two structures are called **isomorphic**, and this is denoted $\mathcal{A} \cong \mathcal{B}$. Similarly if $\mathcal{A} = \mathcal{B}$ then an isomorphism is called a **automorphism**.

We will sometimes dispense of parentheses and write $f\vec{a}$ instead of $f(\vec{a})$.

Notice that for algebraic structures, strong and "weak" homomorphisms are one and the same. Furthermore, if $h\colon \mathcal{A} \longrightarrow \mathcal{B}$ is an embedding, the condition on h being a strong isomorphism is simply

$$r^{\mathcal{A}}\vec{a} \iff r^{\mathcal{B}}h\vec{a}$$

as $(\exists \vec{b} \in A^n)(h\vec{a} = h\vec{b} \wedge r^{\mathcal{A}}\vec{a})$ is equivalent to $r^{\mathcal{A}}\vec{a}$ as $h\vec{a} = h\vec{b}$ implies $\vec{a} = \vec{b}$.

The composition of homomorphisms is itself a homomorphism: if $h_1\colon \mathcal{A} \longrightarrow \mathcal{B}$ and $h_2\colon \mathcal{B} \longrightarrow \mathcal{C}$ are homomorphisms then

$$\begin{gathered} h_2 \circ h_1(f^{\mathcal{A}}\vec{a}) = h_2(f^{\mathcal{B}}h_1\vec{a}) = f^{\mathcal{C}}h_2 \circ h_1(\vec{a}) \\ h_2 \circ h_1(c^{\mathcal{A}}) = h_2 c^{\mathcal{B}} = c^{\mathcal{C}} \\ r^{\mathcal{A}}\vec{a} \implies r^{\mathcal{B}}h_1\vec{a} \implies r^{\mathcal{C}}h_2h_1\vec{a} \end{gathered}$$

And if h_1 and h_2 are strong homomorphisms, and h_1 is surjective, then $h_2 \circ h_1$ is also a strong homomorphism:

$$r^{\mathcal{C}}h_2 \circ h_1\vec{a} \iff (\exists \vec{b} \in B^n)\big(h_2\vec{b} = h_2h_1\vec{a} \wedge r^{\mathcal{B}}\vec{b}\big)$$

Since h_1 is surjective, suppose $h_1\vec{a}_0 = \vec{b}$ then

$$\iff (\exists\vec{a}_0 \in A^n)\big(h_2h_1\vec{a}_0 = h_2h_1\vec{a} \wedge r^{\mathcal{B}}h_1\vec{a}_0\big)$$

Since $r^{\mathcal{B}}h_1\vec{a}_0$ if and only if there exists an a_1 such that $h_1\vec{a}_0 = h_1\vec{a}_1$ and $r^{\mathcal{A}}\vec{a}_1$, so this is equivalent to

$$\iff (\exists\vec{a}_1 \in A^n)\big(h_2h_1\vec{a}_1 = h_2h_1\vec{a} \wedge r^{\mathcal{A}}\vec{a}_1\big)$$

As required.

2.1.4 Definition

Let σ be a signature and $\mathcal{A}$ be a σ-structure. Then a **congruence** on $\mathcal{A}$ is an equivalence relation on A, $\approx$, such that for all function symbols $f \in \sigma$ with arity $n > 0$,

$$\vec{a} \approx \vec{b} \implies f^{\mathcal{A}}\vec{a} \approx f^{\mathcal{A}}\vec{b}$$

where $\vec{a} \approx \vec{b}$ means $a_i \approx b_i$ for $i = 1, \dots, n$ where $\vec{a} = (a_1, \dots, a_n)$ and $\vec{b} = (b_1, \dots, b_n)$.

Let us denote $a/{\approx}$ to be the equivalence class of a under $\approx$, and $\vec{a}/{\approx} = (a_1/{\approx}, \dots, a_n/{\approx})$ for $\vec{a} \in A^n$. Let $f \in \sigma$ be a function symbol, $r \in \sigma$ be a relational symbol, and $c \in \sigma$ be a constant symbol, then let us define the σ-structure $\mathcal{A}'$ over the domain partition $A/{\approx}$ by

$$f^{\mathcal{A}'}(\vec{a}/{\approx}) := (f^{\mathcal{A}}(\vec{a}))/{\approx}, \qquad r^{\mathcal{A}'}(\vec{a}/{\approx}) \iff (\exists\vec{b} \approx \vec{a})(r^{\mathcal{A}}\vec{b}), \qquad c^{\mathcal{A}'} = (c^{\mathcal{A}})/{\approx}$$

These are well-defined as they are independent of the choice of representative from an equivalence class (only the first definition, for $f^{\mathcal{A}'}$, is not true for general equivalence relations). $\mathcal{A}'$ is the **quotient structure** of $\mathcal{A}$ modulo $\approx$, also denoted by $\mathcal{A}/{\approx}$ (the use of $\mathcal{A}'$ was to make it more readable in superscripts).

Let G be a group with the identity e and $\approx$ be a congruence on G. Then let us define $N = \{g \in G \mid g \approx e\}$, and N is a normal subgroup: if $g \in N$ and $h \in G$ then $hgh^{-1} \approx heh^{-1} = e$, and so $hgh^{-1} \in N$. And if N is a normal subgroup, let us define $a \approx_N b$ if and only if $ab^{-1} \in N$, then if $a_1 \approx_N a_2$ and $b_1 \approx_N b_2$ then

$$a_1b_1 \approx_N a_2b_2 \iff a_1b_1b_2^{-1}a_2^{-1} \in N \iff a_1(b_1b_2^{-1}a_2^{-1}a_1)a_1^{-1} \in N$$

since $b_1b_2^{-1} \in N$ and $a_2^{-1}a_1 \in N$, and since N is normal, this is indeed correct. So $\approx_N$ is a congruence on G. This relation is deeper: recall that normal groups are simply kernels of group homomorphisms. So we can define the kernel of general homomorphisms:

2.1.5 Definition

Let $h\colon \mathcal{A} \longrightarrow \mathcal{B}$ be a homomorphism of σ-structures. Then h's **kernel** is the congruence on $\mathcal{A}$ defined by

$$a \approx_h b \iff h(a) = h(b)$$

This is indeed a congruence on $\mathcal{A}$: if $\vec{a} \approx_h \vec{b}$ and $f \in \sigma$ then

$$f^{\mathcal{A}}\vec{a} \approx_h f^{\mathcal{A}}\vec{b} \iff hf^{\mathcal{A}}\vec{a} = hf^{\mathcal{A}}\vec{b} \iff f^{\mathcal{B}}h\vec{a} = f^{\mathcal{B}}h\vec{b}$$

which is true since $h\vec{a} = h\vec{b}$ as $\vec{a} \approx_h \vec{b}$.

Let h be a group homomorphism, and K be its kernel (viewed as a normal subgroup) then $\approx_h = \approx_K$ where $\approx_K$ is defined for groups as previously: $h(a) = h(b)$ if and only if $h(ab^{-1}) = e$ if and only if $ab^{-1} \in K$ if and only if $a \approx_K b$. So this definition of a kernel is natural, and generalizes much nicer than the group-theoretic definition.

2.1.6 Theorem (The Isomorphism Theorem)

(1) Let $\mathcal{A}$ be a σ-structure, and $\approx$ a congruence on $\mathcal{A}$. Then $k\colon a \mapsto a/{\approx}$ is a strong homomorphism from

$\mathcal{A}$ onto $\mathcal{A}/\approx$.

(2) Conversely, if $h\colon \mathcal{A} \longrightarrow \mathcal{B}$ is a surjective strong homomorphism of σ-structures, then $\iota\colon a/\approx_h \mapsto h(a)$ is an isomorphism between $\mathcal{A}/\approx_h$ and $\mathcal{B}$. Furthermore, $h = \iota \circ k$.

Let $f, r, c \in \sigma$ be function, relational, and constant symbols respectively. For readability, we will ignore superscripts.

(1) We do this directly:

$$k(f\vec{a}) = (f\vec{a})/\approx = f\big(\vec{a}/\approx\big) = f(k\vec{a})$$
$$(\exists \vec{b} \in A^n)\big(k\vec{a} = k\vec{b} \wedge r\vec{b}\big) \iff (\exists \vec{b} \approx \vec{a})(r\vec{b}) \iff r(\vec{a}/\approx) \iff rk\vec{a}$$
$$k(c) = c/\approx = c^{\mathcal{A}/\approx}$$

So k is indeed a strong homomorphism.

(2) The definition of ι is obviously sound (ie. it is well-defined) and injective by the definition of $\approx_h$:

$$\iota(a/\approx_h) = \iota(b/\approx_h) \iff h(a) = h(b) \iff a \approx_h b \iff a/\approx_h = b/\approx_h$$

It is surjective since if $b \in \mathcal{B}$, since h is surhjective there exists an $a \in \mathcal{A}$ such that $h(a) = b$ and so $\iota(a/\approx_h) = h(a) = b$. Now, ι is a strong homomorphism:

$$\iota f(\vec{a}/\approx_h) = \iota(f\vec{a})/\approx_h = h(f\vec{a}) = f(h\vec{a}) = f\iota(a/\approx_h)$$
$$r\iota(\vec{a}/\approx_h) \iff rh(\vec{a}) \iff (\exists \vec{b} \approx_h \vec{a})(r(\vec{b})) \iff r(\vec{a}/\approx_h)$$
$$\iota c/\approx_h = h(c) = c$$

By the definitions of ι and k, $h = \iota \circ k$. ■

We need not require h be surjective: instead we alter the claim and ι becomes an isomorphism between $\mathcal{A}$ and the image of $\mathcal{A}$ under h (denoted $h\mathcal{A}$), which is a substructure of $\mathcal{B}$ (this is easy to verify). This corollary is a direct result of the above theorem, as h is a strong homomorphism from $\mathcal{A}$ to $h\mathcal{A}$.

2.1.7 Definition

Let $\{A_i\}_{i\in I}$ be a family of sets, then we define their **direct product** to be the set of function $I \longrightarrow \bigcup_{i\in I} A_i$ such that for every $i \in I$, $i \mapsto a_i$ where $a_i \in A_i$. Such a function is denoted $(a_i)_{i\in I}$ (similar to how a sequence is denoted $(a_n)_{n=1}^{\infty}$ as it represents a function $\mathbb{N} \longrightarrow \mathbb{R}$ which maps $n \mapsto a_n$). So the direct product is defined as, in set-theoretic terms:

$$\prod_{i\in I} A_i = \left\{ f\colon I \longrightarrow \bigcup_{i\in I} A_i \;\middle|\; (\forall i \in I)(f(i) \in A_i) \right\}$$

Where the function f is written as $(f(i))_{i\in I}$ (this is generally more readable).

If $\{\mathcal{A}_i\}_{i\in I}$ is a family of σ-structures, we define their **direct product** to be a σ-structure $\mathcal{B} = \prod_{i\in I} \mathcal{A}_i$ whose domain is the direct product of the domains of $\mathcal{A}_i$ (so if A_i is the domain of $\mathcal{A}_i$, the domain is $B = \prod_{i\in I} A_i$) and for every function symbol f, relational symbol r, and constant symbol c in σ we define

$$f^{\mathcal{B}}\vec{a} = (f^{\mathcal{A}_i}\vec{a}_i)_{i\in I}, \qquad r^{\mathcal{B}}\vec{a} \iff r^{\mathcal{A}_i}\vec{a}_i \text{ for all } i \in I, \qquad c^{\mathcal{B}} = (c^{\mathcal{A}_i})_{i\in I}$$

Where $\vec{a} = \big((a_i^1)_{i\in I}, \dots, (a_i^n)_{i\in I}\big) \in B^n$ and $\vec{a}_i = (a_i^1, \dots, a_i^n) \in A_i^n$ is obtained by looking at the components of $\vec{a}$ at a specific $i \in I$ (take care, this is not the ith component of $\vec{a}$).

If all the structures are the same, $\mathcal{A}_i = \mathcal{A}$ for all $i \in I$, then $\prod_{i\in I} \mathcal{A}_i$ is called the **direct power** of $\mathcal{A}$ and is denoted $\mathcal{A}^I$. If $I = \{1, \dots, n\}$ then $\prod_{i\in I} \mathcal{A}_i$ is also written $\mathcal{A}_1 \times \cdots \times \mathcal{A}_n$ and $\prod_{i\in I} \mathcal{A}$ is written $\mathcal{A}^n$.

Notice that our concept of $\mathbb{R}^n$ as an abelian group corresponds with the above definition. But here we have also defined the coordinate-wise product of vectors in $\mathbb{R}^n$.

We can define the *projection homomorphism* from a direct product to one of its components:

$$\pi_j\colon \prod_{i\in I}\mathcal{A}_i \longrightarrow \mathcal{A}_j, \qquad (a_i)_{i\in I}\mapsto a_j$$

where $j\in I$. This is indeed a homomorphism, let $\vec{a}=\big((a_i^1)_{i\in I},\dots,(a_i^n)_{i\in I}\big)$, then $f\vec{a}=\big((fa_i^1)_{i\in I},\dots,(fa_i^n)_{i\in I}\big)$ and so

$$\pi_j f\vec{a}=(fa_j^1,\dots,fa_j^n)=f(a_j^1,\dots,a_j^n)=f\pi_j\vec{a}$$

As required, and $r\vec{a}$ if and only if $r(a_i^1,\dots,a_i^n)$ for all $i\in I$, which implies $r(a_j^1,\dots,a_j^n)=r\pi_j\vec{a}$. The case for constants is implied by the proof for functions.

But it is not necessarily strong: the condition for strongness is that $r\pi_j\vec{a}$ must be equivalent to

$$(\exists\vec{b})(\pi_j\vec{b}=\pi_j\vec{a}\wedge r\vec{b}) \iff (\exists\vec{b})\big(\pi_j\vec{b}=\pi_i\vec{a}\wedge(\forall i)(r\pi_i\vec{b})\big)$$

Since the definition of $r^{\mathcal{B}}\vec{a}$ is literally $r^{\mathcal{A}_i}\pi_i\vec{a}$ for all $i\in I$. So this is clearly stronger than $r\pi_j\vec{a}$, and unless we know that for every i, $r^{\mathcal{A}_i}$ can be satisfied, it is strictly stronger. But if we know that for all $i\in I$ (except for potentially j), there exists a $\vec{a}_i\in\mathcal{A}_i$ such that $r^{\mathcal{A}_i}\vec{a}_i$, then this is equivalent.

2.2 Syntax of First-Order Languages

First-order logic allow us to discuss precise concepts relating to mathematical structures. Unlike propositional logic, first-order logic has the ability to discuss individual variables within a mathematical structure, and it can *quantify* them as well. Like propositional logic, we must first discuss the syntax of first-order logic, which is more involved.

Let us define a set of variables, which is taken to be a countably infinite set of distinct symbols: $\mathit{Var}=\{\boldsymbol{v}_1,\boldsymbol{v}_2,\dots\}$. Like any language, we must first define the *alphabet* over which we define the language of first-order logic. First-order logic over an extralogical signature σ has the alphabet consisting of: the extralogical symbols of σ; the variables in Var; the logical connectives $\wedge$ and $\neg$; the quantifier $\forall$ (*for all*); the equality sign $\boldsymbol{=}$ (in boldface to distinguish it from the metalogical symbol $=$); and parentheses (and). Other logical connectives, like $\vee$, $\leftrightarrow$, and $\to$ can be defined via $\wedge$ and $\neg$, as discussed before. Similarly other quantifiers like $\exists$ (*there exists*) and $\exists!$ (*there exists a unique*) can be defined as well, which will be discussed later.

From the set of all strings over this alphabet are many meaningless ones, for example $)\forall\wedge$ would be a string over this alphabet, but it has no useful meaning. Like what we did in the previous section, we will recursively define meaningful strings from this alphabet.

2.2.1 Definition

We first define **terms** in this language. Terms are defined recursively as:

(1) Variables and constant symbols in σ, are **prime terms**.

(2) If $f\in\sigma$ is an n-ary function symbol, and $t_1,\dots,t_n$ are terms, then $ft_1\cdots t_n$ is a term as well.

The set of all terms (ie. all strings constructed in this matter) is denoted $\mathcal{T}$.

Notice that we do not use parentheses with terms, as this simplifies syntax and parentheses turn out to be unnecessary. Despite this, when we actually need to write terms (ie. outside of proofs about terms), we may add parentheses for readability. Also note that all of these definitions (and all the coming definitions) are dependent on the choice of extralogical signature σ, so we may speak of *terms over* σ, or σ-*terms*.

Notice that we can view $\mathcal{T}$ as a σ'-structure where σ' is the signature obtained from σ after removing all relational symbols. This is as for $f\in\sigma'$ we define $f^{\mathcal{T}}(t_1,\dots,t_n)=ft_1\cdots t_n$ (the right hand side is a term, a string, and is to be read literally) and for $c\in\sigma'$ we define $c^{\mathcal{T}}=c$. So $\mathcal{T}$ is also sometimes called the *term algebra*.

2.2.2 Proposition (Principle of Term Induction)

Let $\mathcal{E}$ be a property of strings (over this language) such that $\mathcal{E}$ is true for all prime terms, and for all $n>0$ and each n-ary function symbol $f\in\sigma$, $\mathcal{E}t_1,\dots,\mathcal{E}t_n$ implies $\mathcal{E}ft_1\cdots t_n$. Then $\mathcal{E}$ holds for all terms.

This is true since $\mathcal{T}$ is taken as the smallest set obtained by the two rules (that it contains all prime terms, and if $t_1,\dots,t_n$ are terms then so is $ft_1\cdots t_n$). So $\mathcal{E}$ must then contain $\mathcal{T}$.

2.2.3 Lemma

Let t be a term, then no proper intial segment of t is a term.

This follows from the principle of term induction: let $\mathcal{E}t$ be the property "*no proper initial segment of t is a term, and t is not a proper initial segment of some other term*". Then $\mathcal{E}$ holds for prime terms, as these are atomic characters from the alphabet and thus have no proper initial segments. And let p be a prime term since all other terms are either prime, or of the form $ft_1 \cdots t_n$, since $p \neq f$ are distinct symbols, p cannot be a proper initial segment of another term. Now if $\mathcal{E}t_1, \dots, \mathcal{E}t_n$ and $f \in \sigma$ is n-ary then any proper initial segment of $ft_1 \cdots t_n$ is of the form $ft_1 \cdots t_k \xi$, where ξ is a proper initial segment of t_{k+1} (it may also be empty). But in order for $ft_1 \cdots t_k \xi$ to be a term, it must be equal to $fs_1 \cdots s_n$ for other terms s_i, but by $\mathcal{E}t_1$, s_1 can not be an initial segment of t_1 nor can t_1 be an initial segment of s_1, so $t_1 = s_1$. Continuing inductively, we have that $t_i = s_i$ for $i \leq k$, and so we get that $\xi = s_{k+1} \cdots s_n$, but this implies s_{k+1} is an initial segment of ξ, but then s_{k+1} is a proper initial segment of t_{k+1}, and so it cannot be a term by $\mathcal{E}t_{k+1}$ in contradiction.

So no proper initial segment of $ft_1 \cdots t_n$ is a term. And if $ft_1 \cdots t_n$ is the proper initial segment of some other term $fs_1 \cdots s_n$, then by induction we see that $t_i = s_i$ (since t_1 is either an initial segment of s_1 or vice versa) contradicting it being proper. ■

2.2.4 Proposition (Unique Term Concatenation Property)

Suppose t_i and s_j are terms, then if $t_1 \cdots t_n = s_1 \cdots s_m$ then $n = m$ and $t_i = s_i$ for all $1 \leq i \leq n$.

If $t_1 \cdots t_n = s_1 \cdots s_m$ then t_1 is either a initial segment of s_1 or vice versa, but by the lemma above, this cannot be proper, so $t_1 = s_1$. Thus $t_2 \cdots t_n = s_2 \cdots s_m$ and so inducting on n, we get $t_i = s_i$ and $n = m$ as required. ■

Using the unique term concatenation property, we can recursively define functions on terms without worrying about them being well-defined:

2.2.5 Definition

Let t be a term, then we define its **set of variables** recursively as follows:

$$\mathit{var}c = \varnothing \text{ for constant symbols } c, \qquad \mathit{var}x = \{x\} \text{ for } x \in \mathit{Var}, \qquad \mathit{var}ft_1 \cdots t_n = \mathit{var}t_1 \cup \cdots \cup \mathit{var}t_n$$

Alternatively we could simply define it as the set of all symbols in Var which occur in t.

2.2.6 Definition

We now define **formulas** in our language (again, these are defined with respect to a specific signature, and may be called *formulas over σ* or *σ-formulas*). These are strings defined recursively by the rules

(1) If s and t are terms, then $s = t$ is a formula, called an **equation**.

(2) If $t_1, \dots, t_n$ are terms and $r \in \sigma$ is an n-ary relational symbol, then $rt_1 \cdots t_n$ is a formula.

(3) If α and β are formulas and x is a variable, then $(\alpha \wedge \beta)$, $\neg\alpha$, and $\forall x \alpha$ are formulas.

Formulas defined by the first two rules are **prime formulas**. Formulas which do not contain any quantifiers (no occurences of $\forall$, and since other quantifiers like $\exists$ are defined using $\forall$, this includes all other quantifiers) are called **quantifier-free**.

The set of all formulas over a signature σ is denoted $\mathcal{L}\sigma$. In the case that σ contains only a single symbol, $\sigma = \{\mathsf{s}\}$, we may write $\mathcal{L}_{\mathsf{s}}$ instead. And the set of all formulas over the signature $\varnothing$ is denoted $\mathcal{L}_{=}$ and is called the **language of pure identity**.

If $\circ$ is a binary operation, we will often write $t \circ s$ instead of $\circ ts$ as dictated by the definition of compound terms. Similarly if $\triangleleft$ is a binary relation, we will often write $t \triangleleft s$ instead of $\triangleleft ts$. Formally these are abbreviations which refer to the correct form of writing the terms and formulas.

We define the following abbreviations:

$$(\alpha \vee \beta) := \neg(\neg\alpha \wedge \neg\beta), \qquad (\alpha \to \beta) := \neg(\alpha \wedge \neg\beta), \qquad (\alpha \leftrightarrow \beta) := ((\alpha \to \beta) \wedge (\beta \to \alpha))$$
$$\exists x\alpha := \neg\forall x\neg\alpha$$

The first line of definitions should be familiar; they are the same as those defined in the previous section. The definition on the second line should make sense intuitively: "*there exists an x such that α*" if and only if not all xs don't satisfy α. The symbols $\forall$ and $\exists$ are called *quantifiers*. $\forall$ is also called the *universal quantifier*, and $\exists$ is the *existential quantifier*.

2.2.7 Proposition (Principle of Formula Induction)

If $\mathcal{E}$ is a property of strings such that $\mathcal{E}$ holds for all prime formulas and $\mathcal{E}\alpha$ and $\mathcal{E}\beta$ implies $\mathcal{E}(\alpha \wedge \beta)$, $\mathcal{E}\neg\alpha$, and $\mathcal{E}\forall x\alpha$, then $\mathcal{E}$ holds for all formulas in $\mathcal{L}$.

Again, this is directly due to the definition of $\mathcal{L}$.

2.2.8 Proposition (Unique Formula Reconstruction Property)

Every formula $\alpha \in \mathcal{L}$ is either prime or can be written uniquely as $(\alpha \wedge \beta)$, $\neg\alpha$, or $\forall x\alpha$ for $\alpha, \beta \in \mathcal{L}$ and $x \in Var$.

The proof of this is similar to all similar previous propositions: first show that no proper initial segment of a formula is itself a formula, then this follows immediately.

Now, instead of discussing σ-structures and σ-terms, we will refer to them as $\mathcal{L}$-structures and $\mathcal{L}$-terms where $\mathcal{L}$ is a first-order language (which is itself defined over a signature σ. This is simply accepted terminology.)

2.2.9 Definition

We define the set of variables in a formula φ recursively. First we define it on equations: $vars = t = vars \cup vart$, then we define it on prime formulas which are not equations: $vart_1 \cdots t_n = vart_1 \cup \cdots vart_n$. Now for the recursive par:

$$var(\alpha \wedge \beta) = var\alpha \cup var\beta, \quad var\neg\alpha = var\alpha, \quad var\forall x\alpha = var\alpha \cup \{x\}$$

Alternatively we could define it as all the variables which occur in φ.

As before, we define the **rank** of a formula recursively as follows:

$$rank\pi = 0 \text{ for prime formulas } \pi, \quad rank(\alpha \wedge \beta) = \max\{rank\alpha, rank\beta\} + 1, \quad rank\neg\alpha = rank\alpha + 1,$$
$$rank\forall x\alpha = rank\alpha + 1$$

And we similarly define the **quantifier rank** of a formula, which measures the maximum nesting depth of a quantifier in the formula:

$$qrank\pi = 0 \text{ for prime formulas } \pi, \quad qrank(\alpha \wedge \beta) = \max\{qrank\alpha, qrank\beta\}, \quad qrank\neg\alpha = qrank\alpha,$$
$$qrank\forall x\alpha = qrank\alpha + 1$$

The set of **subformulas** of a formula is defined similar to before:

$$Sf\pi = \{\pi\} \text{ for prime formulas } \pi, \quad Sf(\alpha \wedge \beta) = Sf\alpha \cup Sf\beta \cup \{(\alpha \wedge \beta)\}, \quad Sf\neg\alpha = Sf\alpha \cup \{\alpha\},$$
$$Sf\forall x\alpha = Sf\alpha \cup \{\forall x\alpha\}$$

2.2.10 Definition

A string of the form $\forall x$ (and by extension $\exists x$) is called a **prefix**. And given a subformula of the form $\forall x\alpha$,

α is called the **scope** of $\forall x$. Occurrences of x within the scope of an occurrence of $\forall x$ are termed **bound occurrences** of x, all other occurrences of x are termed **free occurrences** of x. In general we say that a variable x **occurs bound** in a formula φ if the prefix $\forall x$ occurs in φ.

We define $bnd\varphi$ to be the set of all variables which occur bound in φ, and $free\varphi$ to be the set of all variables which have free occurrences in φ.

$bnd\varphi$ and $free\varphi$ can also be defined recursively:

$$bnd\pi = \varnothing \text{ for prime formulas } \pi, \quad bnd(\alpha \wedge \beta) = bnd\alpha \cup bnd\beta, \quad bnd\neg\alpha = bnd\alpha,$$
$$bnd\forall x\alpha = bnd\alpha \cup \{x\}$$
$$free\pi = var\pi \text{ for prime formulas } \pi, \quad free(\alpha \wedge \beta) = free\alpha \cup free\beta, \quad free\neg\alpha = free\alpha,$$
$$free\forall x\alpha = free\alpha \setminus \{x\}$$

Notice that a variable can occur both free and bound in a formula, for example in the below formula x occurs both free and bound

$$\forall x(x = y) \wedge (x = y)$$

This will generally be avoided, but it can happen. We could strengthen our definitions of formulas to ensure that this does not occur, but there is no need to do so.

2.2.11 Definition

Let us define $\mathcal{L}^k$ to be the set of all formulas φ such that $free\varphi \subseteq \{\boldsymbol{v}_0, \dots, \boldsymbol{v}_{k-1}\}$. Thus $\mathcal{L}^0 \subseteq \mathcal{L}^1 \subseteq \cdots$ and $\mathcal{L} = \bigcup_{k=0}^{\infty} \mathcal{L}^k$. $\mathcal{L}^0$ is the set of all formulas which contain no free variables, formulas belonging to $\mathcal{L}^0$ are called **sentences** or **closed formulas**.

Note

If φ is a formula, we write $\varphi(\vec{x})$ to mean that $free\varphi \subseteq \{x_1, \dots, x_n\}$, where $\vec{x} = (x_1, \dots, x_n)$ and x_i are all arbitrary and distinct. Similarly if t is a term, $t(\vec{x})$ means $vart \subseteq \{x_1, \dots, x_n\}$.

And we write $f\vec{t}$ to mean the compound term $ft_1 \cdots t_n$ where $\vec{t} = (t_1, \dots, t_n)$ where t_i are terms. Similarly we write $r\vec{t}$ to mean the prime formula $rt_1 \cdots t_n$.

Now we would like to define a notion of substitution, which is a natural concept to have. But importantly we'd only like to substitute variables at their free occurrences, why? Suppose you have the formula $\varphi(y) = \exists x(x+x = y)$ (in the context of integers, this means y is even). We could substitute y for 2 and get $\exists x(x + x = 2)$. But now say we wanted to substitute x for 2, then should we get $\exists 2(2 + 2 = y)$ (which is not a valid formula), or $\exists x(2 + 2 = y)$? Well, neither, because neither really makes sense. This is since bound occurrences already have meaning associated with them by their quantifier, so it makes little sense to substitute them.

2.2.12 Definition

A **substitution** (also called a *global* substitution) is a function which assigns to every variable a term, meaning it is a function $\sigma\colon Var \longrightarrow \mathcal{T}$, where $x \mapsto x^\sigma$. We first extend it to a substitution of terms, ie. a function $\sigma\colon \mathcal{T} \longrightarrow \mathcal{T}$ as follows:

$$c^\sigma = c \text{ for constant symbols } c, \quad (ft_1 \cdots t_n)^\sigma = ft_1^\sigma \cdots t_n^\sigma$$

and now we extend it to a substitution of formulas, ie. a function $\sigma\colon \mathcal{L} \longrightarrow \mathcal{L}$ as follows:

$$(s = t)^\sigma = s^\sigma = t^\sigma, \quad (rt_1 \cdots t_n)^\sigma = rt_1^\sigma \cdots t_n^\sigma, \quad (\alpha \wedge \beta)^\sigma = \alpha^\sigma \wedge \beta^\sigma, \quad (\neg\alpha)^\sigma = \neg\alpha^\sigma, \quad (\forall x\alpha)^\sigma = \forall x\alpha^\tau$$

where τ is the substitution $y^\tau = y^\sigma$ for variables y distinct from x, and $x^\tau = x$ as we'd like to substitute only at free occurrences of variables.

A **simultaneous substitution** is a substitution σ such that there exist variables $x_1, \ldots, x_n \in Var$ and terms $t_1, \ldots, t_n \in \mathcal{T}$ such that

$$x^\sigma = \begin{cases} t_i & x = x_i \\ x & \text{else} \end{cases}$$

So we substitute only x_is with t_is, and leave all other variables the same. Instead of φ^σ we write instead $\varphi\frac{t_1\cdots t_n}{x_1\cdots x_n}$. In the case that $n = 1$ (we only substitute a single variable), this is called a **simple substitution**.

Notice that while by definition there is no significance in the order of writing the variables and their substitutions in a simultaneous substitution (meaning there is no difference between $\varphi\frac{t_1\cdots t_n}{x_1\cdots x_n}$ and $\varphi\frac{t_{\sigma 1}\cdots t_{\sigma n}}{x_{\sigma 1}\cdots x_{\sigma n}}$ where σ is a permutation), it is not true in general that

$$\varphi\frac{t_1 t_2}{x_1 x_2} = \varphi\frac{t_1}{x_1}\frac{t_2}{x_2} \left(= \left(\varphi\frac{t_1}{x_1}\right)\frac{t_2}{x_2}\right)$$

For example, let $\varphi = x_1 < x_2$, then $\varphi\frac{x_2 x_1}{x_1 x_2} = x_2 < x_1$, while $\varphi\frac{x_2}{x_1}\frac{x_1}{x_2} = x_2 < x_2\frac{x_1}{x_2} = x_1 < x_1$. Though it is the case that

$$\varphi\frac{\vec{t}}{\vec{x}} = \varphi\frac{y}{x_n}\frac{t_1\cdots t_{n-1}}{x_1\cdots x_{n-1}}\frac{t_n}{y}$$

where y is a variable not in $var\varphi \cup var\vec{x} \cup var\vec{t}$. This should make sense, as we substitute x_n first with y, which remains unchanged by the next simultaneous substitution, and then substitute y for t_n. Thus inductively we see that every simultaneous substitution can be written as a composition of simple substitutions.

2.3 Semantics of First-Order Languages

Similar to how in propositional logic we defined models to give meaning to propositional formulas, we do the same for first-order logic.

2.3.1 Definition

Suppose $\mathcal{L}$ is a first-order language, then an **$\mathcal{L}$-model** (or an **$\mathcal{L}$-interpretation**) is a pair $\mathcal{M} = (\mathcal{A}, w)$ where $\mathcal{A}$ is an $\mathcal{L}$-structure and w is a **valuation function**, $w\colon Var \longrightarrow A$, $x \mapsto x^w$. We denote $f^\mathcal{A}$, $r^\mathcal{A}$, $c^\mathcal{A}$, and x^w also by $f^\mathcal{M}$, $r^\mathcal{M}$, $c^\mathcal{M}$, and $x^\mathcal{M}$ respectively.

We can extend valuations to $\mathcal{T}$ in an obvious manner:

$$c^\mathcal{M} = c \text{ for constant symbols } c, \qquad (ft_1\cdots t_n)^\mathcal{M} = f^\mathcal{M} t_1^\mathcal{M}\cdots t_n^\mathcal{M}$$

In place of $t^\mathcal{M}$ we may write $t^{\mathcal{A},w}$ or simply t^w if the structure is understood. But we will usually stick with $t^\mathcal{M}$. Notice that the valuation of a term t depends only on the valuation of the variables and extralogical symbols occurring in t:

2.3.2 Proposition

Suppose t is an $\mathcal{L}$-term, and $\mathcal{M}$ and $\mathcal{M}'$ are two $\mathcal{L}$-models. Let V be a set of variables where $var t \subseteq V$. Now suppose that $\mathcal{M}$ and $\mathcal{M}'$ agree on their valuations of V and extralogical symbols in t: for every $x \in V$, $x^\mathcal{M} = x^{\mathcal{M}'}$ and for every extralogical symbol s occuring in t, $\mathsf{s}^\mathcal{M} = \mathsf{s}^{\mathcal{M}'}$. Then $t^\mathcal{M} = t^{\mathcal{M}'}$.

This is done by term induction. If t is a prime term, then $t = c$ for some constant or $t = x$ for some variable. In either case the proposition is satisfied by its assumption (that $\mathcal{M}$ and $\mathcal{M}'$ agree on variables and extralogical symbols occurring in t). Now suppose $t = ft_1\cdots t_n$ then by the assumption of the proposition, $f^\mathcal{M} = f^{\mathcal{M}'}$ and by the induction hypothesis $t_i^\mathcal{M} = t_i^{\mathcal{M}'}$. Thus

$$t^\mathcal{M} = f^\mathcal{M} t_1^\mathcal{M}\cdots t_n^\mathcal{M} = f^{\mathcal{M}'} t_1^{\mathcal{M}'}\cdots t_n^{\mathcal{M}'} = t^{\mathcal{M}'}$$

as required. ∎

2.3.3 Definition

We now define the **satisfiability relation** for first-order models. Let $\mathcal{M} = (\mathcal{A}, w)$ be a model. For every $a \in \mathcal{A}$ and $x \in \mathit{Var}$ let us define the model $\mathcal{M}^a_x = (\mathcal{A}, w')$ where $y^{w'} = y^w$ for variables y distinct from x, and $x^{w'} = a$. Meaning

$$y^{\mathcal{M}^a_x} = \begin{cases} a & y = x \\ y^w & \text{else} \end{cases}$$

So now we define the satisfiability relation $\vDash$ recursively as follows:

$$\begin{aligned} \mathcal{M} \vDash s = t &\iff s^{\mathcal{M}} = t^{\mathcal{M}}, & \mathcal{M} \vDash r\vec{t} &\iff r^{\mathcal{M}}\vec{t}^{\mathcal{M}}, \\ \mathcal{M} \vDash (\alpha \wedge \beta) &\iff \mathcal{M} \vDash \alpha \text{ and } \mathcal{M} \vDash \beta, & \mathcal{M} \vDash \neg\alpha &\iff \mathcal{M} \nvDash \alpha, \end{aligned}$$
$$\mathcal{M} \vDash \forall x\alpha \iff \mathcal{M}^a_x \vDash \alpha \text{ for all } a \in \mathcal{A}$$

If $\mathcal{M} \vDash \varphi$, then $\mathcal{M}$ is said to model φ. And if $X \subseteq \mathcal{L}$ is a set of formulas, we write $\mathcal{M} \vDash X$ if for all $\varphi \in X$, $\mathcal{M} \vDash \varphi$, and we similarly say $\mathcal{M}$ models X.

We can generalize $\mathcal{M}^a_x$ to $\mathcal{M}^{\vec{a}}_{\vec{x}}$ where the underlying structure remains the same and

$$y^{\mathcal{M}^{\vec{a}}_{\vec{x}}} = \begin{cases} a_i & y = x_i \\ y^{\mathcal{M}} & \text{else} \end{cases}$$

Notice that $\mathcal{M}^{\vec{a}}_{\vec{x}} = (\mathcal{M}^{a_1}_{x_1})^{a_2}_{x_2} \dots$. It follows immediately that if we use $\forall\vec{x}$ as an abbreviation for $\forall x_1 \forall x_2 \cdots \forall x_n$, then we get

$$\mathcal{M} \vDash \forall\vec{x}\alpha \iff \mathcal{M}^{\vec{a}}_{\vec{x}} \vDash \alpha \text{ for all } \vec{a} \in \mathcal{A}^n$$

It is easily verifiable that

$$\begin{aligned} \mathcal{M} \vDash (\alpha \vee \beta) &\iff \mathcal{M} \vDash \alpha \text{ or } \mathcal{M} \vDash \beta & \mathcal{M} \vDash (\alpha \to \beta) &\iff \text{if } \mathcal{M} \vDash \alpha \text{ then } \mathcal{M} \vDash \beta \\ \mathcal{M} \vDash (\alpha \leftrightarrow \beta) &\iff \mathcal{M} \vDash \alpha \text{ if and only if } \mathcal{M} \vDash \beta \end{aligned}$$

And also $\mathcal{M} \vDash \exists x\alpha = \neg\forall x\neg\alpha$ if and only if $\mathcal{M} \nvDash \forall x\neg\alpha$ so there exists an $a \in \mathcal{A}$ such that $\mathcal{M}^a_x \nvDash \neg\alpha$, meaning there exists an $a \in \mathcal{A}$ such that $\mathcal{M}^a_x \vDash \alpha$. This chain of reasoning is readily seen to be reversible. So we have shown

$$\mathcal{M} \vDash \exists x\alpha \iff \text{there exists an } a \in \mathcal{A} \text{ such that } \mathcal{M}^a_x \vDash \alpha$$

2.3.4 Definition

A formula or set of formulas is said to be **satisfiable** if it has a model. $\varphi \in \mathcal{L}$ is called a **tautology** (or **generally/logically valid**), denoted $\vDash \varphi$, if $\mathcal{M} \vDash \varphi$ for every model $\mathcal{M}$. Two formulas α and β are said to be **logically equivalent**, denoted $\alpha \equiv \beta$, if for every model $\mathcal{M}$,

$$\mathcal{M} \vDash \alpha \iff \mathcal{M} \vDash \beta$$

Now, say $\mathcal{A}$ is an $\mathcal{L}$-structure, then we write $\mathcal{A} \vDash \varphi$ for a formula φ if $(\mathcal{A}, w) \vDash \varphi$ for all valuations $w\colon \mathit{Var} \longrightarrow A$. Similarly one writes $\mathcal{A} \vDash X$ for a set of formulas X if $\mathcal{A} \vDash \varphi$ for all $\varphi \in X$.

2.3.5 Definition

Finally we define the **consequence relation** for first-order logic. Suppose X is a set of formulas and φ is a formula, then we write $X \vDash \varphi$ if every model of X models φ. Meaning $\mathcal{M} \vDash X \implies \mathcal{M} \vDash \varphi$.

Again, $\vDash$ is used to denote both the satisfaction and consequence relations. The meaning of the notation is to be understood from context. Moreso, $\vDash$ is also used for the satisfaction relation of structures. And again we write $\varphi_1, \dots, \varphi_n \vDash \varphi$ in place of $\{\varphi_1, \dots, \varphi_n\} \vDash \varphi$ and all the usual shorthands.

Notice that while by definition if $\mathcal{M}$ is a model then $\mathcal{M} \vDash \varphi$ or $\mathcal{M} \vDash \neg\varphi$ for all formulas φ. But if $\mathcal{A}$ is a structure, then it is possible for $\mathcal{A}$ to satisfy neither φ nor $\neg\varphi$ (but if it does satisfy one, it cannot satisfy the other obviously). Take for example the formula $x = y$, then if $\mathcal{A}$ is a structure with at least two elements, suppose $a \neq b \in \mathcal{A}$, then we can define a valuation which satisfies $x = y$ and one which does not. And so $\mathcal{A}$ satisfies neither $x = y$ nor $\neg x = y = x \neq y$.

Now suppose φ is a formula and let $x_1, \dots, x_n$ be an enumeration of *free*φ (according to some accepted total order of *Var*, for example by index), then we define the *generalized of* φ or its *universal closure* to be the sentence

$$\varphi^g := \forall x_1 \cdots \forall x_n \varphi$$

From the definitions provided above, it is immediate that if $\mathcal{A}$ is a structure then

$$\mathcal{A} \vDash \varphi \iff \mathcal{A} \vDash \varphi^g$$

And in general $\mathcal{A} \vDash X \iff \mathcal{A} \vDash X^g := \{\varphi^g \mid \varphi \in X\}$.

2.3.6 Theorem (The Coincidence Theorem)

Let φ be a formula, and V be a set of variables such that *free*$\varphi \subseteq V$. Let $\mathcal{M}$ and $\mathcal{M}'$ be two models over the same domain A such that $x^{\mathcal{M}} = x^{\mathcal{M}'}$ for all variables $x \in V$, and $\mathsf{s}^{\mathcal{M}} = \mathsf{s}^{\mathcal{M}'}$ for all extralogical symbols s occurring in φ. Then $\mathcal{M} \vDash \varphi$ if and only if $\mathcal{M}' \vDash \varphi$.

We prove this by induction on φ. If φ is a prime formula of the form $rt_1 \cdots t_n$, by the assumptions of the theorem and proposition 2.3.2, $t_i^{\mathcal{M}} = t_i^{\mathcal{M}'}$ for all $1 \leq i \leq n$, and $r^{\mathcal{M}} = r^{\mathcal{M}'}$, so $r^{\mathcal{M}}\vec{t}^{\mathcal{M}} \iff r^{\mathcal{M}'}\vec{t}^{\mathcal{M}'}$ as required. This proof holds for equations as well. Now by the inductive hypothesis we get

$$\mathcal{M} \vDash \alpha \wedge \beta \iff \mathcal{M} \vDash \alpha \text{ and } \mathcal{M} \vDash \beta \iff \mathcal{M}' \vDash \alpha \text{ and } \mathcal{M}' \vDash \beta \iff \mathcal{M}' \vDash \alpha \wedge \beta$$

Similar for formulas of the form $\neg\alpha$.
Now, let $a \in \mathcal{A}$ and suppose $\mathcal{M}_x^a \vDash \varphi$. Then let $V' = V \cup \{x\}$ then *free*$\varphi \subseteq V'$ (since *free*$\varphi \subseteq$ *free*$\forall x\varphi \cup \{x\} \subseteq V \cup \{x\}$) and $\mathcal{M}_x^a$ and $\mathcal{M}'^a_x$ coincide for all $y \in V'$ (though it is possible that $x^{\mathcal{M}} \neq x^{\mathcal{M}'}$). Thus by our inductive hypothesis $\mathcal{M}_x^a \vDash \varphi$ if and only if $\mathcal{M}'^a_x \vDash \varphi$. Thus

$$\mathcal{M} \vDash \forall x\varphi \iff \mathcal{M}_x^a \vDash \varphi \text{ for all } a \in \mathcal{A} \iff \mathcal{M}'^a_x \vDash \varphi \text{ for all } a \in \mathcal{A} \iff \mathcal{M}' \vDash \forall x\varphi$$

as required. ∎

Let $\sigma \subseteq \sigma'$ be two signatures, and $\mathcal{L} \subseteq \mathcal{L}'$ be their respective first-order languages. Now, if $\mathcal{M} = (\mathcal{A}, w)$ is an $\mathcal{L}$-model, it can be arbitrarily extended to an $\mathcal{L}'$-model $\mathcal{M}' = (\mathcal{A}', w)$, where $\mathcal{A}'$ is the σ'-expansion of $\mathcal{A}$, by arbitrarily setting $\mathsf{s}^{\mathcal{M}'}$ for $\mathsf{s} \in \sigma' \setminus \sigma$. Now, let us set $V = Var$ and by the coincidence theorem we get that for every $\varphi \in \mathcal{L}$ since $\mathcal{M}$ and $\mathcal{M}'$ agree on the extralogical symbols (as $\mathcal{A}'$ is an expansion of $\mathcal{A}$) and variables n V (since the valuation remains the same), we get that

$$\mathcal{M} \vDash \varphi \iff \mathcal{M}' \vDash \varphi$$

If we denote the consequence relation of $\mathcal{L}$ by $\vDash_{\mathcal{L}}$, then it follows that if $\mathcal{L} \subseteq \mathcal{L}'$, $\vDash_{\mathcal{L}'}$ is a *conservative* extension of $\vDash_{\mathcal{L}}$: for every $\varphi \in \mathcal{L}$ and $X \subseteq \mathcal{L}$, $X \vDash_{\mathcal{L}'} \varphi$ if and only if $X \vDash_{\mathcal{L}} \varphi$. Indeed: if $\mathcal{M}'$ is an $\mathcal{L}'$-model then let $\mathcal{M}$ be the $\mathcal{L}$-reduct of $\mathcal{M}'$ and so $\mathcal{M} \vDash_{\mathcal{L}} X$ if and only if $\mathcal{M}' \vDash_{\mathcal{L}'} X$, and same for φ.
So the satisfiability of φ depends only on the symbols occurring in φ, we need not the subscripts in $\vDash$.
Another consequence of the coincidence theorem is the *omission of superfluous quantifiers*:

$$\forall x\varphi \equiv \varphi \equiv \exists x\varphi \text{ if } x \notin \mathit{free}\varphi$$

To see this, let $\mathcal{M}$ be a model and $a \in \mathcal{A}$ be arbitrary. Then let $V = $ *free*φ and $\mathcal{M}' = \mathcal{M}_x^a$, and by the coincidence theorem since $y^{\mathcal{M}} = y^{\mathcal{M}'}$ for all $y \in V$ (since $x \notin$ *free*φ) we have that $\mathcal{M} \vDash \varphi$ if and only if $\mathcal{M}_x^a \vDash \varphi$. So $\mathcal{M} \vDash \forall x\varphi$ if and only if $\mathcal{M}_x^a \vDash \varphi$ for all $a \in \mathcal{A}$, which is if and only if $\mathcal{M} \vDash \varphi$, which is if and only if $\mathcal{M}_x^a \vDash \varphi$ for some $a \in \mathcal{A}$, which is by definition $\mathcal{M} \vDash \exists x\varphi$.
This fact should be intuitive, for example $\forall x\exists x(x > 0)$ is the same as $\exists x(x > 0)$ and $\exists x\exists x(x > 0)$ since the outermost quantifier is superfluous.
Another thing which is simple to show is that if $\mathcal{A}$ is a substructure of $\mathcal{B}$, and $\mathcal{M} = (\mathcal{A}, w)$ and $\mathcal{M}' = (\mathcal{B}, w)$ are models with the same valuation $w\colon Var \longrightarrow A$, then for every term t, $t^{\mathcal{M}} = t^{\mathcal{M}'}$. For prime terms this is obvious (for terms of the form x, this is since the same valuation is used, and for constant terms this is by definition of substructures). Now inductively, if $t_i^{\mathcal{M}} = t_i^{\mathcal{M}'}$ then since $t_i^{\mathcal{M}} \in \mathcal{A}$, and $f^{\mathcal{M}'}\vec{a} = f^{\mathcal{M}}\vec{a}$ for values $\vec{a} \in \mathcal{A}^n$, we have that

$$(f\vec{t})^{\mathcal{M}'} = f^{\mathcal{M}'}\vec{t}^{\mathcal{M}'} = f^{\mathcal{M}}\vec{t}^{\mathcal{M}} = (f\vec{t})^{\mathcal{M}}$$

Now by the coincidence theorem, we know that the satisfaction of a formula φ is dependent only on the valuations of variables occurring in φ. So let $\varphi = \varphi(\vec{x})$ (meaning *free*$\varphi \subseteq \{x_1, \dots, x_n\}$), then saying

$$(\mathcal{A}, w) \vDash \varphi \text{ for a valuation } w \text{ where } x_1^w = a_1, \dots, x_n^w = a_n$$

can be written more suggestively by

$$(\mathcal{A}, \vec{a}) \vDash \varphi, \text{ or } \mathcal{A} \vDash \varphi[a_1, \dots, a_n], \text{ or } \mathcal{A} \vDash \varphi[\vec{a}]$$

So there is no need to mention the global valuation w.

2.3.7 Theorem (The Substructure Theorem)

For $\mathcal{L}$-structures $\mathcal{A}$ and $\mathcal{B}$ such that $A \subseteq B$ (where A represents the domain of $\mathcal{A}$, and so on), the following conditions are equivalent:

(1) $\mathcal{A}$ is a substructure of $\mathcal{B}$: $\mathcal{A} \subseteq \mathcal{B}$,

(2) $\mathcal{A} \vDash \varphi[\vec{a}]$ if and only if $\mathcal{B} \vDash \varphi[\vec{a}]$ for all quantifier-free $\varphi = \varphi(\vec{x})$ and $\vec{a} \in A^n$,

(3) $\mathcal{A} \vDash \varphi[\vec{a}]$ if and only if $\mathcal{B} \vDash \varphi[\vec{a}]$ for all prime $\varphi = \varphi(\vec{x})$ and $\vec{a} \in A^n$.

First we prove (1) $\implies$ (2): so we must show that for models $\mathcal{M} = (\mathcal{A}, w)$ and $\mathcal{M}' = (\mathcal{B}, w)$, $\mathcal{M} \vDash \varphi$ if and only if $\mathcal{M}' \vDash \varphi$. For prime formulas this is obvious as we showed that $\vec{t}^{\mathcal{M}} = \vec{t}^{\mathcal{M}'}$ for all vectors of terms $\vec{t}$. The inductive step for $\wedge$ and $\neg$ are simple. (2) $\implies$ (3) is trivial as prime formulas are quantifier-free. Now to show that (3) $\implies$ (1) we must show that $f^{\mathcal{A}}\vec{a} = f^{\mathcal{B}}\vec{a}$ and $r^{\mathcal{A}}\vec{a} \iff r^{\mathcal{B}}\vec{a}$ for $\vec{a} \in A^n$. By (3), we know that

$$r^{\mathcal{A}}\vec{a} \iff \mathcal{A} \vDash r\vec{x}[\vec{a}] \iff \mathcal{B} \vDash r\vec{x}[\vec{a}] \iff r^{\mathcal{B}}\vec{a}$$

($\vec{x} = (x_1, \dots, x_n)$ is any vector of n distinct variables.) Similarly

$$f^{\mathcal{A}}\vec{a} = b \iff \mathcal{A} \vDash f\vec{x} = y[\vec{a}, b] \iff \mathcal{B} \vDash f\vec{x} = y[\vec{a}, b] \iff f^{\mathcal{B}}\vec{a} = b$$

and so $f^{\mathcal{A}}\vec{a} = f^{\mathcal{B}}\vec{a}$, meaning $\mathcal{A} \subseteq \mathcal{B}$ as required (constants are viewed as 0-ary functions; the proof is the same). ∎

2.3.8 Definition

A **universal formula** (or a **$\forall$-formula**) is a formula α of the form $\forall\vec{x}\beta$ where β is quantifier-free. If α is a sentence, it is also called a **universal sentence** or **$\forall$-sentence**. Similarly, an **existential formula** (or a **$\exists$-formula**) is a formula α of the form $\exists\vec{x}\beta$ where β is quantifier-free. And again if α is a sentence, it is also called an **existential sentence** or **$\exists$-sentence**.

For example, we can define the following existential sentences:

$$\exists_1 := \exists \boldsymbol{v}_0 \boldsymbol{v}_0 = \boldsymbol{v}_0, \qquad \exists_n := \exists \boldsymbol{v}_0 \cdots \exists \boldsymbol{v}_{n-1} \bigwedge_{0 \le i < j < n} \boldsymbol{v}_i \neq \boldsymbol{v}_j$$

So $\exists_n$ states "there exists at least n (distinct) elements". This means that $\neg\exists_{n+1}$ means "there exists at most n elements", and so we can define $\exists_{=n} := \exists_n \wedge \neg\exists_{n+1}$ which means "there exists exactly n elements".

Since structures are non-empty by definition, $\exists_1$ is a tautology, and so we define $\top := \exists_1$, and $\bot := \exists_0 := \neg\top$. From the substructure theorem we get the following corollary:

2.3.9 Corollary

Let $\mathcal{A} \subseteq \mathcal{B}$, then every $\forall$-sentence which is valid in $\mathcal{B}$ is also valid in $\mathcal{A}$. Conversely, every $\exists$-sentence which is valid in $\mathcal{A}$ is also valid in $\mathcal{B}$.

Suppose $\mathcal{B} \vDash \forall\vec{x}\beta$ where β is quantifier-free. Then let $\vec{a} \in A^n$, and so $\mathcal{B} \vDash \beta[\vec{a}]$, and by the substructure theorem since β is quantifier-free, we get $\mathcal{A} \vDash \beta[\vec{a}]$. Since $\vec{a}$ is arbitrary, this means $\mathcal{A} \vDash \forall\vec{x}\beta$ as required. And similarly, if $\mathcal{A} \vDash \exists\vec{x}\beta$ then there exists a $\vec{a} \in A^n$ such that $\mathcal{A} \vDash \beta[\vec{a}]$ and so again by the substructure theorem we get $\mathcal{B} \vDash \beta[\vec{a}]$, so $\mathcal{B} \vDash \exists\vec{x}\beta$. ∎

2.3.10 Theorem (The Invariance Theorem)

Suppose $\mathcal{A}$ and $\mathcal{B}$ are isomorphic $\mathcal{L}$-structures with an isomorphism $\iota\colon \mathcal{A} \longrightarrow \mathcal{B}$. Then for every formula

$\varphi = \varphi(\vec{x})$, and all $\vec{a} \in A^n$:

$$\mathcal{A} \vDash \varphi[\vec{a}] \iff \mathcal{B} \vDash \varphi[\iota\vec{a}]$$

where again, $\iota\vec{a} = (\iota a_1, \ldots, \iota a_n)$.

It is sufficient to show that for models $\mathcal{M} = (\mathcal{A}, w)$ and $\mathcal{M}' = (\mathcal{B}, w')$ where $w' = \iota \circ w$, $\mathcal{M} \vDash \varphi \iff \mathcal{M}' \vDash \varphi$. We will first prove by term induction that for every term t, $\iota(t^{\mathcal{M}}) = t^{\mathcal{M}'}$. For prime terms, this is obvious:

$$x^{\mathcal{M}'} = x^{w'} = \iota x^w = \iota(x^{\mathcal{M}}), \qquad c^{\mathcal{M}'} = c^{\mathcal{B}} = \iota(c^{\mathcal{A}}) = \iota(c^{\mathcal{M}})$$

For compound terms, suppose $\iota(t_i^{\mathcal{M}}) = t_i^{\mathcal{M}'}$ for $1 \leq i \leq n$, then we get $\vec{t}^{\mathcal{M}'} = \iota\vec{t}^{\mathcal{M}}$ and so

$$(f\vec{t})^{\mathcal{M}'} = f^{\mathcal{M}'}\vec{t}^{\mathcal{M}} = f^{\mathcal{B}}\iota\vec{t}^{\mathcal{M}} = \iota(f^{\mathcal{A}}\vec{t}^{\mathcal{M}}) = \iota(f\vec{t})^{\mathcal{M}}$$

the third equality is due to ι being a homomorphism.

Now we will prove the goal: $\mathcal{M} \vDash \varphi \iff \mathcal{M}' \vDash \varphi$, by formula induction on φ. For prime formulas φ:

$$\mathcal{M}' \vDash t = s \iff t^{\mathcal{M}'} = s^{\mathcal{M}'} \iff \iota t^{\mathcal{M}} = \iota s^{\mathcal{M}} \iff t^{\mathcal{M}} = s^{\mathcal{M}} \iff \mathcal{M} \vDash t = s$$
$$\mathcal{M}' \vDash r\vec{t} \iff r^{\mathcal{B}}\vec{t}^{\mathcal{M}'} \iff r^{\mathcal{B}}\iota\vec{t}^{\mathcal{M}} \iff r^{\mathcal{A}}\vec{t}^{\mathcal{M}} \iff \mathcal{M} \vDash r\vec{t}$$

For compound formulas we use the inductive hypothesis:

$$\mathcal{M}' \vDash \alpha \wedge \beta \iff \mathcal{M}' \vDash \alpha \text{ and } \mathcal{M}' \vDash \beta \iff \mathcal{M} \vDash \alpha \text{ and } \mathcal{M} \vDash \beta \iff \mathcal{M} \vDash \alpha \wedge \beta$$
$$\mathcal{M}' \vDash \neg\alpha \iff \mathcal{M}' \nvDash \alpha \iff \mathcal{M} \nvDash \alpha \iff \mathcal{M} \vDash \neg\alpha$$

And for quantified formulas, $\mathcal{M}' \vDash \forall x\varphi$ if and only if $\mathcal{M}'^{b}_{x} \vDash \varphi$ for all $b \in \mathcal{B}$. Since ι is surjective, $b = \iota a$ for some $a \in A$ and so if $\mathcal{M}^a_x = (\mathcal{A}, w_0)$ then $\mathcal{M}'^{b}_{x} = (\mathcal{B}, \iota \circ w_0)$ and so by our inductive hypothesis (as $\mathcal{M}$ and $\mathcal{M}'$ are arbitrary), this is if and only if $\mathcal{M}^a_x \vDash \varphi$. Since b is arbitrary, and therefore so is a, $\mathcal{M} \vDash \forall x\varphi$. This chain of logic is reversible, which can also be seen since ι^{-1} is also an isomorphism. So we have indeed shown that for all formulas, $\mathcal{M} \vDash \varphi \iff \mathcal{M}' \vDash \varphi$ as required. ■

This proof obviously covers sentences $\varphi \in \mathcal{L}^0$. So for example, if G is a group and ι is an isomorphism to some other $\circ$-structure, then $\iota(G)$ is also a group. This is as we can have φ run over the axioms of group theory (which are sentences) and $G \vDash \varphi \implies \iota(G) \vDash \varphi$.

2.3.11 Definition

Two $\mathcal{L}$-structures are **elementarily equivalent** if $\mathcal{A} \vDash \varphi \iff \mathcal{B} \vDash \varphi$ for all sentences $\varphi \in \mathcal{L}^0$. This is denoted $\mathcal{A} \equiv \mathcal{B}$.

So by the invariance theorem isomorphic structures are elementarily equivalent.

Notice that substitutions do not pay attention to *collision of variables*, for example if $\varphi = \exists y(x \neq y)$ then certainly $\forall x\varphi$ is true when the domain has at least two elements. But $\varphi\frac{y}{x} = \exists y(y \neq y)$ is false. So we have shown that $\forall x\varphi$ does not necessarily imply $\varphi\frac{t}{x}$, as one may expect. But the issue is that we have changed the context of the variable x from a free to a bound variable, and so the meaning of $\varphi\frac{y}{x}$ is different from φ.

2.3.12 Definition

A simple substitution $\varphi, \frac{t}{x}$ is called **collision-free** if every $y \in \mathit{var}t$ distinct from x does not occur bound in φ (meaning $\mathit{bnd}\varphi \cap \mathit{var}t \subseteq \{x\}$, or $y \in \mathit{var}t \setminus \{x\} \implies y \notin \mathit{bnd}\varphi$). This means that a variable being inserted into φ does not get potentially inserted into a scope of its own quantifier.

A global substitution σ is collision-free if $\varphi, \frac{x^\sigma}{x}$ is collision-free for every variable $x \in \mathit{Var}$. So for a simultaneous substitution $\varphi, \frac{\vec{t}}{\vec{x}}$ it is only necessary to verify that $\varphi, \frac{t_i}{x_i}$ is collision-free.

The reason for not requiring $x \notin \mathit{bnd}\varphi$ in a substitution $\frac{t}{x}$ is since t is only substituted at free occurrences of x, so whether or not x occurs bound in φ is immaterial. (Though such a restriction wouldn't matter practically: it is bad practice to have a formula where the same variable occurs both free and bound. In some texts, different symbols are used for free and bound variables even.)

This is a crude definition, for example if $\varphi = \forall x(x = x) \wedge \forall y(y = y)$, then $\varphi\frac{y}{x}$ is not collision-free since y occurs bound in φ. But such a substitution would not alter φ. We could refine the definition of collision-free substitutions to mean that x does not occur within the scope of any variable in *vart*, other than potentially x itself. Some texts refer to this definition as *t is free for x in φ*. Such a refinement is unnecessary for our purposes though.

For a model $\mathcal{M} = (\mathcal{A}, w)$ and a global substitution σ, we define $\mathcal{M}^\sigma = (\mathcal{A}, w^\sigma)$ where $x^{w^\sigma} = (x^\sigma)^{\mathcal{M}}$ (first substitute x to get a term, and then valuate it via $\mathcal{M}$) for every $x \in Var$. In other words $x^{\mathcal{M}^\sigma} = x^{\sigma\mathcal{M}}$ $(= (x^\sigma)^{\mathcal{M}})$. This extends to terms: $t^{\mathcal{M}^\sigma} = t^{\sigma\mathcal{M}}$. This is obvious for prime terms (variables and constants), and for compound terms:

$$(ft_1 \cdots t_n)^{\mathcal{M}^\sigma} = f^{\mathcal{M}} t_1^{\mathcal{M}^\sigma} \cdots t_n^{\mathcal{M}^\sigma} = f^{\mathcal{M}} t_1^{\sigma\mathcal{M}} \cdots t_n^{\sigma\mathcal{M}} = (ft_1 \cdots t_n)^{\sigma\mathcal{M}}$$

Notice that in the case $\sigma = \frac{\vec{t}}{\vec{x}}$,

$$x^{\mathcal{M}^\sigma} = x^{\sigma\mathcal{M}} = \left(x\frac{\vec{t}}{\vec{x}}\right)^{\mathcal{M}}$$

and so $\mathcal{M}^\sigma$ coincides with $\mathcal{M}_{\vec{x}}^{\vec{t}^{\mathcal{M}}}$ ($\vec{t}^{\mathcal{M}}$ since the terms must be interpreted, $\mathcal{M}_{\vec{x}}^{\vec{t}}$ makes no sense sine $\vec{t}$ is not a vector in the domain).

2.3.13 Theorem (The Substitution Theorem)

Let $\mathcal{M}$ be a model and σ be a global substitution. If φ, σ are collision-free then

$$\mathcal{M} \vDash \varphi^\sigma \iff \mathcal{M}^\sigma \vDash \varphi$$

In particular, $\mathcal{M} \vDash \varphi\frac{\vec{t}}{\vec{x}}$ if and only if $\mathcal{M}_{\vec{x}}^{\vec{t}^{\mathcal{M}}} \vDash \varphi$, provided $\varphi, \frac{\vec{t}}{\vec{x}}$ is collision-free.

This will be proven by formula induction. In the case that φ is an equation $t_1 = t_2$,

$$\mathcal{M} \vDash (t_1 = t_2)^\sigma \iff t_1^{\sigma\mathcal{M}} = t_2^{\sigma\mathcal{M}} \iff t_1^{\mathcal{M}^\sigma} = t_2^{\mathcal{M}^\sigma} \iff \mathcal{M}^\sigma \vDash t_1 = t_2$$

Similarly, for prime formulas of the form $r\vec{t}$:

$$\mathcal{M} \vDash (r\vec{t})^\sigma \iff r^{\mathcal{M}}\vec{t}^{\sigma\mathcal{M}} \iff r^{\mathcal{M}^\sigma}\vec{t}^{\mathcal{M}^\sigma} \iff \mathcal{M}^\sigma \vDash r\vec{t}$$

The induction steps for $\wedge$ and $\neg$ are obvious and simple. Now, recall that $\mathcal{M} \vDash \forall x\varphi$ if and only if $\mathcal{M} \vDash \forall x\varphi^\tau$ where $x^\tau = x$ and $y^\tau = y^\sigma$ for $x \neq y$,

$$\iff \mathcal{M}_x^a \vDash \varphi^\tau \text{ for every } a \in A \iff (\mathcal{M}_x^a)^\tau \vDash \varphi$$

now we claim that $(\mathcal{M}_x^a)^\tau = (\mathcal{M}^\sigma)_x^a$. Since $\forall x\varphi, \sigma$ is collision-free, meaning $\forall x\varphi, \frac{y^\sigma}{y}$ for every y is collision-free so $x \notin vary^\sigma$ for every $y \neq x$. And since $y^\tau = y^\sigma$, we get $y^{(\mathcal{M}_x^a)^\tau} = y^{\tau\mathcal{M}_x^a} = y^{\sigma\mathcal{M}_x^a} = y^{\sigma\mathcal{M}} = y^{\mathcal{M}^\sigma} = y^{(\mathcal{M}^\sigma)_x^a}$. And $x^{(\mathcal{M}_x^a)^\tau} = x^{\mathcal{M}_x^a} = a = x^{(\mathcal{M}^\sigma)_x^a}$. Thus

$$\iff (\mathcal{M}^\sigma)_x^a \vDash \varphi \text{ for every } a \in A \iff \mathcal{M}^\sigma \vDash \forall x\varphi$$

as required. ■

This proof also obviously works for the refinement of the concept of collision-free substitutions (what we termed "…free for x in φ").

2.3.14 Corollary

For all formulas φ and $\frac{\vec{t}}{\vec{x}}$ such that $\varphi, \frac{\vec{t}}{\vec{x}}$ is collision-free, the following are true:

(1) $\forall\vec{x}\varphi \vDash \varphi\frac{\vec{t}}{\vec{x}}$,

(2) $\varphi\frac{\vec{t}}{\vec{x}} \vDash \exists\vec{x}\varphi$,

(3) $\varphi\frac{s}{x}, s = t \vDash \varphi\frac{t}{x}$, provided $\varphi, \frac{s}{x}$ and $\varphi, \frac{t}{x}$ are collision-free.

(1) Let $\mathcal{M} \vDash \forall\vec{\varphi}$, so $\mathcal{M}_{\vec{x}}^{\vec{a}} \vDash \varphi$ for all $\vec{a} \in A^n$. This means that $\mathcal{M}_{\vec{x}}^{\vec{t}^{\mathcal{M}}} \vDash \varphi$, and so by the substitution theorem this is equivalent to $\mathcal{M} \vDash \varphi\frac{\vec{t}}{\vec{x}}$.

(2) Since $\neg\exists\vec{x}\varphi = \forall\vec{x}\neg\varphi$, by (1) $\neg\exists\vec{x}\varphi \vDash \neg\varphi\frac{\vec{t}}{\vec{x}}$. And so $\varphi\frac{\vec{t}}{\vec{x}} \vDash \exists\vec{\varphi}$ (since $\neg X \vDash \neg Y$ if and only if $Y \vDash X$).

(3) Suppose $\mathcal{M} \vDash \varphi\frac{s}{x}$, $s = t$ so $s^{\mathcal{M}} = t^{\mathcal{M}}$ and $\mathcal{M}_x^{s^{\mathcal{M}}} \vDash \varphi$ by the substitution theorem, and so $\mathcal{M}_x^{t^{\mathcal{M}}} \vDash \varphi$, and by applying the substitution theorem again we get $\mathcal{M} \vDash \varphi\frac{t}{x}$. ■

This shows that collision-free substitutions act as substitutions should: if φ is true for all $\vec{x}$, then we should be able to substitute any terms $\vec{t}$ in place of $\vec{x}$ and φ should hold. What we need to be careful of is that this substitution doesn't mess with the syntax of φ, which is why it must be collision-free.

Let us define the following quantifier, "there exists exactly one":

$$\exists!x\varphi := \exists x\varphi \wedge \forall x\forall y(\varphi \wedge \varphi\tfrac{y}{x} \rightarrow y = x)$$

where y is any variable not in $\mathit{var}\varphi$. So if $\mathcal{M} \vDash \forall x\forall y(\varphi \wedge \varphi\frac{y}{x} \rightarrow y = x)$ means that $\mathcal{M}_x^a \vDash \varphi$ and $\mathcal{M}_x^b \vDash \varphi\frac{y}{x}$ implies $a = b$ (since $y \notin \mathit{var}\varphi$). Putting it all together, $\mathcal{M} \vDash \exists!x\varphi$ if and only if there exists a single a such that $\mathcal{M}_x^a \vDash \varphi$. An example of a tautology is $\exists!xx = t$ for any term t.

An equivalent definition of $\exists!$ is

$$\exists!x\varphi := \exists x\forall y(\varphi\tfrac{y}{x} \leftrightarrow y = x)$$

Exercise

Suppose $X \vDash \varphi$ and $x \notin \mathit{free}\varphi$, then show that $X \vDash \forall x\varphi$.

Suppose $\mathcal{M} \vDash X$ and let $a \in A$, then notice that $\mathcal{M}_x^a$ and $\mathcal{M}$ agree on all variables in $\mathit{free}\varphi$ and extralogical symbols (as they are defined over the same structure). So by The Substitution Theorem since $\mathcal{M} \vDash \varphi$, $\mathcal{M}_x^a \vDash \varphi$. Since a is arbitrary, this means $\mathcal{M} \vDash \forall x\varphi$.

Exercise

Show that $\forall x(\alpha \rightarrow \beta) \vDash \forall x\alpha \rightarrow \forall x\beta$.

Let $\mathcal{M} \vDash \forall x(\alpha \rightarrow \beta)$. Then if $\mathcal{M} \vDash \forall x\alpha$ then let $a \in A$, so $\mathcal{M}_x^a \vDash \alpha, \alpha \rightarrow \beta$ so $\mathcal{M}_x^a \vDash \beta$ by modus ponens. Since a is arbitrary, this means $\mathcal{M} \vDash \forall x\beta$. Thus $\mathcal{M} \vDash \forall x\alpha \rightarrow \forall x\beta$.

Exercise

Suppose $\mathcal{A}$ is an $\mathcal{L}$-structure and $\mathcal{A}'$ is obtained by adjoining a constant symbol $\boldsymbol{a}$ for some $a \in A$ (meaning $\boldsymbol{a}^{\mathcal{A}'} = a$). Show that $\mathcal{A} \vDash \alpha[a] \iff \mathcal{A}' \vDash \alpha\frac{\boldsymbol{a}}{x}$ for a formula $\alpha = \alpha(x)$.

We will first show that $t(x)^{\mathcal{A},a} = (t\frac{\boldsymbol{a}}{x})^{\mathcal{A}'}$. For prime terms this is obvious. For compound terms this is simple by induction.

It is sufficient to show that for arbitrary $\mathcal{M} = (\mathcal{A}, w)$ and $\mathcal{M}' = (\mathcal{A}', w)$ such that $x^w = a$, $\mathcal{M} \vDash \alpha \iff \mathcal{M}' \vDash \alpha\frac{\boldsymbol{a}}{x}$. By formula induction on α: in the case that $\alpha = t_1 = t_2$,

$$\mathcal{M}' \vDash t_1\tfrac{\boldsymbol{a}}{x} = t_2\tfrac{\boldsymbol{a}}{x} \iff \left(t_1\tfrac{\boldsymbol{a}}{x}\right)^{\mathcal{M}'} = \left(t_2\tfrac{\boldsymbol{a}}{x}\right)^{\mathcal{M}'} \iff t_1^{\mathcal{M}} = t_2^{\mathcal{M}} \iff \mathcal{M} \vDash t_1 = t_2$$

Similar for prime formulas of the form $r\vec{t}$ (there is of course a similarity between equations and prime formulas of the form $r\vec{t}$, $=$ is a "primitive" binary relation; it represents the identity relation). For compound formulas using connectives $\wedge$ and $\neg$ the inductive step is clear. And for formulas $\alpha = \forall y\beta$, if $y \neq x$ then this is clear. If $y = x$ then this is clear by the substitution theorem, as $\mathcal{A}$ and $\mathcal{A}'$ act the same on α as it is a sentence.

Exercise

Prove the following:

(1) A conjunction of formulas $\exists_i$ and their negations is logically equivalent to a formula of the form $\exists_n \wedge \neg\exists_m$ for suitable n and m.

(2) A boolean combination of $\exists_i$ is equivalent to a formula of the form $\bigvee_{i\leq n} \exists_{=k_i}$ or $\exists_k \vee \bigvee_{i\leq n} \exists_{=k_i}$ where $k_0 < \cdots < k_n < k$.

(1) Every such conjunction is equivalent, by commutativity, to a formula of the form

$$\bigwedge_{1\leq i\leq n} \exists_{k_i} \wedge \bigwedge_{1\leq j\leq m} \neg\exists_{\ell_j}$$

Where $k_1 < \cdots < k_n$ and $\ell_1 < \cdots < \ell_m$. Since $\exists_i$ represents there existing at least i elements, and $\neg\exists_i$ means there exists $< i$ elements, this is equivalent to $\equiv \exists_{k_n} \wedge \neg\exists_{\ell_1}$. If $n = 0$ then take $k_0 = 1$ (since $\exists_1$ is a tautology) and if $m = 0$ then take $\ell_1 = 0$ (since $\neg\exists_0$ is a tautology).

(2) Let us prove this by induction on the rank of the boolean combination, r. If $r = 0$ then the formula is $\exists_n$, which is of the second form. Otherwise, there are three cases:

(1) The formula is of the form $\neg\alpha$ where $\alpha \equiv \bigvee_{i\leq n} \exists_{=k_i}$ or $\exists_k \vee \bigvee_{i\leq n} \exists_{=k_i}$. In the first case,

$$\neg\alpha \equiv \bigwedge_{i\leq n} \neg\exists_{=k_i} \equiv \exists_{k_n+1} \vee \bigwedge_{i\leq k_n, i\neq k_j} \exists_{=i}$$

This is as the negation of α refers to there not existing precisely k_i elements, so this means there is either $\geq k_n + 1$ elements or precisely i elements for $i \leq k_n$ where i differs from any k_j. In the second case,

$$\neg\alpha \equiv \neg\exists_k \wedge \bigwedge_{i\leq n} \neg\exists_{=k_i}$$

By above, both parts of the conjunction is equivalent to one of the forms required by the exercise, so this case is covered by case (3).

(2) The formula is of the form $\alpha \wedge \beta$ where α and β are boolean combinations. It is obvious how this formula is equivalent to a formula of one of the forms required.

(3) If the formula is of the form $\alpha \vee \beta$, then it is already (after α and β are converted to the correct form) in a required form.

2.4 General Validity and Logical Equivalence

2.4.1 Example

You may recall from set theory that the so-called "Russellian set" of sets not containing themselves is a proper class (not a set). This is reflective of the more general logical result, $\vDash \neg\exists u\forall x(x \in u \leftrightarrow x \notin x)$. This sentence states that there does not exist a u such that $x \in u$ if and only if $x \notin x$ for all x. But note that we require no conditions on the binary relation $\in$, so this result is purely logical one.

We will prove this in steps. Firstly, $\forall x(x \in u \leftrightarrow x \notin x) \vDash u \in u \leftrightarrow u \notin u$ which is due to corollary 2.3.14, using the substitution $\frac{u}{x}$. Clearly $u \in u \leftrightarrow u \notin u$ is not satisfiable, and thus neither is $\forall x(x \in u \leftrightarrow x \notin x)$, and therefore neither is $\exists u\forall x(x \in u \leftrightarrow x \notin x)$. Therfore $\vDash \neg\exists u\forall x(x \in u \leftrightarrow x \notin x)$.

This proof probably parallels the proof you are familiar with from an introductory set theory/discrete math course: one considers the Russelian set u itself.

It is obvious that $X \vDash \alpha \to \beta$ if and only if $X, \alpha \vDash \beta$. Suppose $X \vDash \alpha \to \beta$, and so suppose $\mathcal{M} \vDash X, \alpha$ then $\mathcal{M} \vDash \alpha \to \beta, \alpha$ and so $\mathcal{M} \vDash \beta$. Thus $X, \alpha \vDash \beta$. And conversely if $X, \alpha \vDash \beta$, suppose $\mathcal{M} \vDash X$. If $\mathcal{M} \vDash \alpha$ then $\mathcal{M} \vDash X \vDash \beta$, thus $X \vDash \alpha \to \beta$ as required.

Thus we immediately get, for example

$$\forall\vec{x}\alpha \to \alpha\frac{\vec{t}}{\vec{x}}$$

for collision-free $\alpha, \frac{\vec{t}}{\vec{x}}$.

Now, $\mathcal{L}$ is a structure over the signature $\{\wedge, \neg\} \cup \{\forall x \mid x \in \mathit{Var}\}$. And so we can talk about *congruences in* $\mathcal{L}$. These are equivalence relations $\approx$ such that for all formulas $\alpha, \alpha', \beta, \beta' \in \mathcal{L}$ and $x \in \mathit{Var}$:

$$\alpha \approx \alpha',\, \beta \approx \beta' \implies \alpha \wedge \beta \approx \alpha' \wedge \beta',\, \neg\alpha \approx \neg\alpha',\, \forall x\alpha \approx \forall x\alpha'$$

(As a technicality the symbol $\wedge$ represents the function $(\alpha, \beta) \mapsto (\alpha \wedge \beta)$ and not $(\alpha, \beta) \mapsto \alpha \wedge \beta$ but I digress.)

2.4.2 Theorem (The Replacement Theorem)

Let $\approx$ be a congruence in $\mathcal{L}$ and $\alpha \approx \alpha'$. Let φ be an arbitrary formula. If φ' is obtained by replacing some of the occurrences (meaning not necessarily all the occurrences are replaced) of α with α', then $\varphi \approx \varphi'$.

We will prove this by induction on φ. If φ is a prime formula, then either $\varphi = \alpha$ or $\varphi \neq \alpha'$, but in any case $\varphi' = \alpha'$ or $\varphi' = \varphi = \alpha$ (since prime formulas are the smallest units of formulas). And so $\varphi \approx \varphi$ since $\alpha \approx \alpha$ and $\alpha \approx \alpha'$. In general, the case where $\varphi = \alpha$ is trivial.

Now, if $\varphi = \varphi_1 \wedge \varphi_2$ then $\varphi' = \varphi_1' \wedge \varphi_2'$ (since we don't consider $\varphi = \alpha$) and inductively $\varphi_i \approx \varphi_i'$. Since $\approx$ is a congruence, $\varphi' \approx \varphi$. The steps for $\neg$ and $\forall x$ are analogous. ∎

So for example, since $\equiv$ is a congruence in $\mathcal{L}$, the replacement theorem tells us that if we replace subformulas of φ with equivalent subformulas, the resulting formula remains equivalent to φ. This is not surprising, but provides a rigorous proof of our intuition.

By this theorem, the "congruence-ness" holds for all defined connectives and quantifiers (meaning $\alpha \approx \alpha', \beta \approx \beta'$ implies $\alpha \to \alpha' \equiv \beta \to \beta'$ and $\exists x\alpha \approx \exists x\alpha'$ and so on).

2.4.3 Definition

Suppose $\mathcal{A}$ is an $\mathcal{L}$-structure. Two $\mathcal{L}$-formulas α and β are **equivalent in** $\mathcal{A}$ (or **equivalent modulo** $\mathcal{A}$) if $\mathcal{A}, w \vDash \alpha$ if and only if $\mathcal{A}, w \vDash \beta$ for all valuations $w\colon \mathcal{A} \longrightarrow \mathit{Var}$. This is denoted $\alpha \equiv_{\mathcal{A}} \beta$.

Obviously $\alpha \equiv_{\mathcal{A}} \beta$ is equivalent to $\mathcal{A} \vDash \alpha \leftrightarrow \beta$. And if $\alpha \equiv \beta$ then $\alpha \equiv_{\mathcal{A}} \beta$, or more suggestively: $\equiv \subseteq \equiv_{\mathcal{A}}$.

2.4.4 Proposition

Suppose $\{\approx_i\}_{i\in I}$ is a non-empty family of congruences in $\mathcal{L}$, then so is their intersection, $\approx = \bigcap_{i\in I} \approx_i$.

This is trivial. ∎

Thus if $\boldsymbol{K} \neq \varnothing$ is a class of $\mathcal{L}$-structures, $\cong_{\boldsymbol{K}} := \bigcap_{\mathcal{A}\in\boldsymbol{K}} \equiv_{\mathcal{A}}$ is also a congruence in $\mathcal{L}$. If $\boldsymbol{K}$ is the class of *all* $\mathcal{L}$-structures, then $\equiv_{\boldsymbol{K}}$ is equal to $\equiv$.

We will state some basic properties of $\equiv$ below:

$$\begin{array}{llll} (1) & \forall x(\alpha \wedge \beta) \equiv \forall x\alpha \wedge \forall x\beta & (2) & \exists x(\alpha \vee \beta) \equiv \exists x\alpha \vee \exists x\beta \\ (3) & \forall x\forall y\alpha \equiv \forall y\forall x\alpha & (4) & \exists x\exists y\alpha \equiv \exists y\exists x\alpha \end{array}$$

If x does not occur free in the formula β, then

$$\begin{array}{llll} (5) & \forall x(\alpha \vee \beta) \equiv \forall x\alpha \vee \beta & (6) & \exists x(\alpha \wedge \beta) \equiv \exists x\alpha \wedge \beta \\ (7) & \forall x\beta \equiv \beta & (8) & \exists x\beta \equiv \beta \\ (9) & \forall x(\alpha \to \beta) \equiv \exists x\alpha \to \beta & (10) & \exists x(\alpha \to \beta) \equiv \forall x\alpha \to \beta \end{array}$$

These proofs are relatively simple. Proofs of (5) through (10) will of course rely on The Substitution Theorem.

Another important property is the *renaming of bound variables* (*bound renaming* for short): let α be a formula and $x \neq y$ be distinct variables where $y \notin \mathit{var}\alpha$. Then

$$\forall x\alpha \equiv \forall y(\alpha\tfrac{y}{x}), \qquad \exists x\alpha \equiv \exists y(\alpha\tfrac{y}{x})$$

The second equivalence results from the first due to the replacement theorem (its necessity can easily be directly verified). So we prove the first equivalence ($\mathcal{M}_x^y$ means $\mathcal{M}_x^{y^{\mathcal{M}}}$):

$$\begin{array}{rlll} \mathcal{M} \vDash \forall x\alpha & \iff \mathcal{M}_x^a \vDash \alpha & \text{for all } a & \text{(definition)} \\ & \iff (\mathcal{M}_y^a)_x^a \vDash \alpha & \text{for all } a & \text{(by The Coincidence Theorem)} \end{array}$$

$$\begin{aligned}
&\iff (\mathcal{M}^a_y)^y_x \vDash \alpha \quad \text{for all } a && ((\mathcal{M}^a_y)^y_x = (\mathcal{M}^a_y)^a_x)\\
&\iff \mathcal{M}^a_y \vDash \alpha\tfrac{y}{x} \quad \text{for all } a && (\text{by The Substitution Theorem})\\
&\iff \mathcal{M} \vDash \forall y(\alpha\tfrac{y}{x}). && (\text{definition})
\end{aligned}$$

We also have the following properties: (assuming $\alpha, \frac{t}{x}$ are collision-free)

$$\begin{aligned}
&\forall x(x = t \to \alpha) \equiv \alpha\tfrac{t}{x} \equiv \exists x(x = t \wedge \alpha) && (x \notin \mathrm{var}t)\\
&\forall y(y = t \to \alpha\tfrac{y}{x}) \equiv \alpha\tfrac{t}{x} \equiv \exists y(y = t \wedge \alpha\tfrac{y}{x}) && (y \notin \mathrm{var}\alpha, \mathrm{var}t)
\end{aligned}$$

We will prove the first line of equivalences. $\forall x(x = t \to \alpha) \vDash (x = t \to \alpha)\frac{t}{x} = t = t \to \alpha\frac{t}{x} \vDash \alpha\frac{t}{x}$, by corollary 2.3.14. Conversely, let $\mathcal{M} \vDash \alpha\frac{t}{x}$, so by The Substitution Theorem $\mathcal{M}^{t^\mathcal{M}}_x \vDash \alpha$. If $\mathcal{M}^a_x \vDash x = t$, then obviously $a = t^\mathcal{M}$, and so $\mathcal{M}^a_x \vDash \alpha$. So for all $a \in A$, $\mathcal{M}^a_x \vDash x = t \to \alpha$, thus $\mathcal{M} \vDash \forall x(x = t \to \alpha)$ as required. Now, for the rightmost formula by what we just proved,

$$\exists x(x = t \wedge \alpha) = \neg\forall x\neg(x = t \wedge \alpha) \equiv \neg\forall x(x = t \to \neg\alpha) \equiv \neg\neg\alpha\tfrac{t}{x} \equiv \alpha\tfrac{t}{x}$$

The second line is proven similarly.

2.4.5 Definition

A formula of the form $\alpha = \mathtt{Q}_1 x_1 \cdots \mathtt{Q}_n x_n \beta$ where $\mathtt{Q}_i$ are either universal or existential quantifiers and β is quantifier-free is a **prenex normal form**, for short PNF.

In PNFs, we may assume that all the variables x_i are distinct, as we can drop superfluous quantifiers as stated earlier. For example, $\forall$- and $\exists$-formulas are prenex normal forms.

2.4.6 Theorem

Every formula is equivalent to a prenex normal form.

Let φ be a formula. Without loss of generality, we can assume that φ contains only $\neg, \wedge, \forall, \exists$ and not any other connectives or quantifiers (formally φ would not have $\exists$, but I digress). For each prefix $\mathtt{Q}x$ in φ, consider the number of occurrences of the symbols $\neg$ and $\wedge$ to the left of the prefix, and let $s\varphi$ be the sum of these numbers, summed over all prefixes in φ. Obviously φ is a PNF if and only if $s\varphi = 0$. So we will induct on $s\varphi$. In the case that $s\varphi = 0$, then φ is already a PNF. Otherwise, there must be some prefix $\mathtt{Q}x$ such that immediately to its left there is an occurrance of $\wedge$ or $\neg$. Applying the appropriate equivalence from the following

$$\neg\forall x\alpha \equiv \exists x\neg\alpha, \quad \neg\exists x\alpha \equiv \forall x\neg\alpha, \quad \beta \wedge \mathtt{Q}x\alpha \equiv \mathtt{Q}y(\beta \wedge \alpha\tfrac{y}{x}) \quad (y \notin \mathrm{var}\alpha, \mathrm{var}\beta)$$

(the third equivalence is the result of applying the appropriate basic properties (1), (2), (5), (6), (7), (8) as well as renaming bound variables) will reduce $s\varphi$, and the proof (and construction of a PNF for φ) proceeds inductively. ■

As a shorthand, if $\triangleleft$ is a binary relation and t is a term, then we will write $(\forall x \triangleleft t)\alpha$ in place of $\forall x(x \triangleleft t \to \alpha)$. Similarly we will write $(\exists x \triangleleft t)\alpha$ in place of $\exists x(x \triangleleft t \wedge \alpha)$. So for example, in fields we may write $(\forall x \neq 0)\exists y\, x \cdot y = 1$.

Exercise

Suppose $\alpha \equiv \beta$ are equivalent formulas, show that $\alpha\frac{\vec{t}}{\vec{x}} \equiv \beta\frac{\vec{t}}{\vec{x}}$ where the substitutions are collision-free.

We will use the notation $\mathcal{M}^{\vec{t}}_{\vec{x}}$ as an abbreviation for $\mathcal{M}^{\vec{t}^\mathcal{M}}_{\vec{x}}$. Then we have by The Substitution Theorem:

$$\mathcal{M} \vDash \alpha\tfrac{\vec{t}}{\vec{x}} \iff \mathcal{M}^{\vec{t}}_{\vec{x}} \vDash \alpha \iff \mathcal{M}^{\vec{t}}_{\vec{x}} \vDash \beta \iff \mathcal{M} \vDash \beta\tfrac{\vec{t}}{\vec{x}}$$

as required.

Exercise

Show that $\neg(\forall x \triangleleft t)\alpha \equiv (\exists x \triangleleft t)\neg\alpha$ and $\neg(\exists x \triangleleft t)\alpha \equiv (\forall x \triangleleft t)\neg\alpha$.

We know that
$$\neg(\forall x \triangleleft t)\alpha = \neg\forall x(x \triangleleft t \to \alpha) \equiv \exists x(x \triangleleft t \wedge \neg\alpha) = (\exists x \triangleleft t)\neg\alpha$$
This means that
$$\neg(\exists x \triangleleft t)\alpha \equiv \neg\neg(\forall x \triangleleft t)\neg\alpha \equiv (\forall x \triangleleft t)\neg\alpha$$

Exercise

Show that the conjunction and disjunction of $\forall$-formulas is equivalent to a $\forall$-formula. Show the same for $\exists$-formulas.

Suppose $\forall\vec{x}\alpha$ and $\forall\vec{y}\beta$ are $\forall$-formulas (α, β are quantifier-free). By bound renaming, we can assume that $\vec{x}$ is distinct from $free\beta$ and $\vec{y}$, and $\vec{y}$ is distinct from $free\alpha$ and $\vec{x}$. Thus by $\forall\vec{x}\alpha \wedge \beta \equiv \forall\vec{x}(\alpha \wedge \beta)$,

$$\forall\vec{x}\alpha \wedge \forall\vec{y}\beta \equiv \forall\vec{x}(\alpha \wedge \forall\vec{y}\beta) \equiv \forall\vec{x}\forall\vec{y}(\alpha \wedge \beta)$$

similar for disjunction and $\exists$-formulas.

Exercise

Show that every formula $\varphi \in \mathcal{L}$ is equivalent to some formula $\varphi' \in \mathcal{L}$ built up from literals using only $\wedge$, $\vee$, and $\exists$.

We will prove this by induction on φ. If φ is a prime formula, this is trivial. Now suppose $\varphi = \neg\varphi_1$, then by induction φ_1 is equivalent to some formula built up from literals using only conjunction, disjunction, and existential quantifiers. Let us assume for simplicity that φ_1 is (replace φ_1 with its equivalent formula). Then φ_1 is either a literal, a conjunction, a disjunction, or a combination quantified using an existential quantifier. If φ_1 is a literal, then $\neg\varphi_1$ is equivalent to a literal. If $\varphi_1 = \varphi_2 \vee \varphi_3$, then $\varphi \equiv \neg\varphi_2 \wedge \neg\varphi_3$ and by induction $\neg\varphi_2$ and $\neg\varphi_3$ are equivalent to formulas of the required form, and so therefore so is their conjunction, φ. Similar for $\varphi_1 = \varphi_2 \wedge \varphi_3$. If $\varphi_1 = \exists x\varphi_2$ then by induction φ_2 is equivalent to a formula of the required form and therefore so is φ_1.

Exercise

Let P be a unary relational symbol, show that $\exists x(Px \to \forall yPy)$ is a tautology.

The negation of this is equivalent to $\forall x(Px \wedge \neg\forall yPy) \equiv \forall x(Px \wedge \exists y(\neg Py) \equiv \forall x\, Px \wedge \exists y(\neg Py)$ (since $\forall x(\alpha \wedge \beta) \equiv \forall x\alpha \wedge \forall x\beta$, and $\forall x\exists y(\neg Py) \equiv \exists y(\neg Py)$ by removing superfluous quantifiers). This is is obviously not satisfiable.

2.5 Logical Consequence and Theories

Similar to the expansion and reduct of structures, if $\mathcal{L} \subseteq \mathcal{L}'$, $\mathcal{L}$ is called the *reduct* of $\mathcal{L}'$ and $\mathcal{L}'$ is called the *expansion* of $\mathcal{L}$. Since the satisfaction requirements for $\wedge$ and $\neg$ are the same as in propositional logic, the basic rules for $\vDash$ in propositional logic carry over to first order logic. These are

$$\text{(IS)}\quad \frac{}{\alpha \vDash \alpha} \qquad\qquad \text{(MR)}\quad \frac{X \vDash \alpha}{X' \vDash \alpha}\quad (X \subseteq X')$$
$$(\wedge 1)\quad \frac{X \vDash \alpha, \beta}{X \vDash \alpha \wedge \beta} \qquad\qquad (\wedge 2)\quad \frac{X \vDash \alpha \wedge \beta}{X \vDash \alpha, \beta}$$
$$(\neg 1)\quad \frac{X \vDash \alpha, \neg\alpha}{X \vDash \beta} \qquad\qquad (\neg 2)\quad \frac{X, \alpha \vDash \beta \mid X, \neg\alpha \vDash \beta}{X \vDash \beta}$$

But we also have new properties:

$$(1)\quad \frac{X \vDash \forall x\alpha}{X \vDash \alpha\frac{t}{x}} \qquad\qquad (\alpha, \tfrac{t}{x} \text{ collision-free})$$

$$(2)\quad \frac{X \vDash \alpha\frac{s}{x},\ s = t}{X \vDash \alpha\frac{t}{x}} \qquad (\alpha, \tfrac{s}{x} \text{ and } \alpha, \tfrac{t}{x} \text{ collision-free})$$

$$(3)\quad \frac{X, \beta \vDash \alpha}{X, \forall x\beta \vDash \alpha} \qquad (\text{anterior generalization})$$

$$(4)\quad \frac{X \vDash \alpha}{X \vDash \forall x\alpha} \qquad (x \notin \mathit{free}X, \text{ posterior generalization})$$

$$(5)\quad \frac{X, \beta \vDash \alpha}{X, \exists x\beta \vDash \alpha} \qquad (x \notin \mathit{free}X, \mathit{free}\alpha, \text{ anterior particularization})$$

$$(6)\quad \frac{X \vDash \alpha\frac{t}{x}}{X \vDash \exists x\alpha} \qquad (\alpha, \tfrac{t}{x} \text{ collision-free, posterior particularization})$$

(1) is due to $\forall x\alpha \vDash \alpha\frac{t}{x}$ and the transititvity of $\vDash$. Similarly (2) is due to $\alpha\frac{s}{x},\ s = t \vDash \alpha\frac{t}{x}$. And (3) is due to $\forall x\beta \vDash \beta$. We now prove (4). Suppose $X \vDash \alpha$, $\mathcal{M} \vDash X$, and $x \notin \mathit{free}X$. Then by The Coincidence Theorem, for every $a \in A$, $\mathcal{M}^a_x \vDash X$ and so $\mathcal{M}^a_x \vDash \alpha$ meaning $\mathcal{M} \vDash \forall x\alpha$ as required. For (5), notice that by (4):

$$X, \beta \vDash \alpha \implies X, \neg\alpha \vDash \neg\beta \implies X, \neg\alpha \vDash \forall x\neg\beta \implies X, \neg\forall x\neg\beta \vDash \alpha \implies X, \exists x\beta \vDash \alpha$$

And by corollary 2.3.14, $\alpha\frac{t}{x} \vDash \exists x\alpha$ proving (6).

Some texts define a stricter consequence relation, which we call the *global consequence relation*, denoted here by $\overset{g}{\vDash}$. For a set of formulas $X \subseteq \mathcal{L}$ and a formula $\varphi \in \mathcal{L}$, we define $X \overset{g}{\vDash} \varphi$ if and only if $\mathcal{A} \vDash X$ implies $\mathcal{A} \vDash \varphi$ for all *structures* $\mathcal{A}$. (The difference here is subtle: the "local" consequence relation deals in models, while the global consequence relation deals in structures.)

Obviously $X \vDash \varphi$ implies $X \overset{g}{\vDash} \varphi$: let $\mathcal{A} \vDash X$ then since for every valuation $\mathcal{A}, w \vDash X \implies \mathcal{A}, w \vDash \varphi$ so $\mathcal{A} \vDash \varphi$. The converse is not true though: for example $x = y \overset{g}{\vDash} \forall x\forall y\, x = y$ since if $\mathcal{A} \vDash x = y$ then $\mathcal{A}$ has a single element. But $x = y \nvDash \forall x\forall y\, x = y$, since any model over a structure with two or more elements whose valuations of x and y are equal is a counterexample. By posterior generalization, $X \vDash \varphi \implies X \vDash \varphi^g$ is true in general only if the free variables of φ do not occur free in X. But on the other hand $X \overset{g}{\vDash} \varphi \implies X \overset{g}{\vDash} \varphi^g$ is always true (since $\mathcal{A} \vDash \varphi \iff \mathcal{A} \vDash \varphi^g$).

Now, since for every structure $\mathcal{A}$, valuation w, and set of formulas X: $\mathcal{A}, w \vDash X^g \iff \mathcal{A} \vDash X^g$ by the coincidence theorem ($\mathit{free}X^g = \varnothing$). Thus

$$X \overset{g}{\vDash} \varphi \iff X^g \vDash \varphi$$

If $X \overset{g}{\vDash} \varphi$, let $\mathcal{A}, w \vDash X^g$ then $\mathcal{A} \vDash X^g$ (explained above) and so $\mathcal{A} \vDash X$ meaning $\mathcal{A} \vDash \varphi$ and in particular $\mathcal{A}, w \vDash \varphi$. Conversely, if $X^g \vDash \varphi$ suppose $\mathcal{A} \vDash X$ then $\mathcal{A} \vDash X^g$ and so $\mathcal{A} \vDash \varphi$. If S is a set of sentences, then $S^g = S$ and so we get that

$$S \overset{g}{\vDash} \varphi \iff S \vDash \varphi$$

so for sets of sentences, there is no difference. But otherwise the two consequence relations act differently, for example neither the rule of case distinction nor the deduction theorem hold. These are the rules:

$$\frac{X, \alpha \overset{g}{\vDash} \beta \;\big|\; X, \neg\alpha \overset{g}{\vDash} \beta}{X \overset{g}{\vDash} \beta}\ (\text{rule of case distinction}), \qquad \frac{X, \alpha \overset{g}{\vDash} \beta}{X \overset{g}{\vDash} \alpha \to \beta}\ (\text{deduction theorem})$$

For example, $x = y \overset{g}{\vDash} \forall x\forall y\, x = y$ but it is not necessarily true that $\overset{g}{\vDash} x = y \to \forall x\forall y\, x = y$.

2.5.1 Definition

A **first-order theory** over a language $\mathcal{L}$ (also a **$\mathcal{L}$-theory**), is a set of sentences $T \subseteq \mathcal{L}^0$ which is **deductively closed** in $\mathcal{L}$, meaning $T \vDash \varphi \iff \varphi \in T$ for all sentences $\varphi \in \mathcal{L}^0$. If $\varphi \in T$ then we say that φ is a **theorem** of T (or holds or is true in T). If $T \subseteq T'$ are both $\mathcal{L}$-theories, then T is called a **subtheory** of T' and T' is called an **extension** of T. An $\mathcal{L}$-structure $\mathcal{A}$ such that $\mathcal{A} \vDash T$ is called an **model** of T, we denote the class of all models of T by $\mathrm{Md}\, T$.

Different authors define theories slightly differently, for example some may not require the theory be deductively closed, etc.

For example, if $X \subseteq \mathcal{L}$ is a set of formulas we define its *deductive closure* to be $\{\alpha \in \mathcal{L}^0 \mid X \vDash \alpha\}$. Since $\vDash$ is transitive, the deductive closure of a set of formulas is indeed a theory. When we discuss the *theory of S*, where S is a set of sentences, we mean the deductive closure of S. A set $X \subseteq \mathcal{L}$ is called an *axiom system* for a theory T if $T = \{\alpha \in \mathcal{L}^0 \mid X^g \vDash \alpha\}$. Notice how we tacitly generalize the free variables in axioms.

Furthermore, if T is a theory then $T \vDash \varphi \iff \mathcal{A} \vDash \varphi$ for all $\mathcal{A} \vDash T$ since T is a set of sentences (in general we require that $\mathcal{M} \vDash \varphi$ for all $\mathcal{M} \vDash T$ but as discussed previously, there is no distinction between global and local consequence). And $T \vDash \varphi \iff T \vDash \varphi^g$.

The smallest theory in $\mathcal{L}$ is the theory containing only tautologies, which we denote $Taut_{\mathcal{L}}$. An axiom system for $Taut_{\mathcal{L}}$ is the empty set. The set of all sentences, $\mathcal{L}^0$, is the largest theory in $\mathcal{L}$ which is also called the *inconsistent theory*, and all other theories are called or *satisfiable*.

Notice that if $\{T_i\}_{i\in I}$ is a family of $\mathcal{L}$-theories, $T = \bigcap_{i\in I} T_i$ is also an $\mathcal{L}$-theory: if $T \vDash \alpha$ then $T_i \vDash \alpha$ for all $i \in I$ (by (MR)) and so $\alpha \in T$.

2.5.2 Definition

Let T be an $\mathcal{L}$-theory, and $\varphi \in \mathcal{L}^0$ be a sentence. Then we define $T + \varphi$ to be the smallest theory which extends T and contains φ. Similarly if $S \subseteq \mathcal{L}^0$ a set of sentences, then $T + S$ is the smallest theory which extends T and contains S. Alternatively,

$$T + S = \bigcap \{T' \mid T' \text{ is a theory which extends } T \text{ and contains } S\}$$

If $T + \varphi$ is satisfiable, then φ is said to be **compatible with** T. If $\neg\varphi$ is compatible with T, then φ is said to be **refutable in** T. If φ is both compatible and refutable (meaning $\neg\varphi$ is also compatible) with T, then φ is said to be **independent of** T.

2.5.3 Definition

If T is a $\mathcal{L}$-theory and $\alpha, \beta \in \mathcal{L}$ are two formulas, we say α and β are **equivalent modulo** T, denoted $\alpha \equiv_T \beta$, if $\alpha \equiv_{\mathcal{A}} \beta$ for all $\mathcal{A} \vDash T$. Meaning

$$\equiv_T := \bigcap_{\mathcal{A}\in \operatorname{Md} T} \equiv_{\mathcal{A}}$$

So $\equiv_T$ is an intersection of congruences, and is therefore also a congruence. Obviously

$$\alpha \equiv_T \beta \iff T \vDash \alpha \leftrightarrow \beta \iff T \vDash (\alpha \leftrightarrow \beta)^g$$

Similarly if t and s are $\mathcal{L}$-terms, we call them **equivalent modulo** T, denoted $t \approx_T s$, if $T \vDash s = t$ (meaning for all $\mathcal{A} \vDash T$ and valuations $w\colon Var \longrightarrow A$, $\mathcal{A}, w \vDash s = t$).

For example, let $T_G^=$ be the theory of groups in the signature $\{\circ, e, ^{-1}\}$ (an equivalent formulation can be defined over the signature $\{\circ, e\}$ and gives us the equivalent theory T_G), then $(x \circ y)^{-1} \approx_{T_G^=} y^{-1} \circ x^{-1}$.

Let $\mathcal{A}$ be an $\mathcal{L}$-structure, then we define its *theory* to be:

$$Th\mathcal{A} := \{\alpha \in \mathcal{L}^0 \mid \mathcal{A} \vDash \alpha\}$$

this is indeed a theory, as if $Th\mathcal{A} \vDash \alpha$ then since $\mathcal{A} \vDash Th\mathcal{A}$ by definition, we have $\mathcal{A} \vDash \alpha$ so $\alpha \in Th\mathcal{A}$. And if $\boldsymbol{K}$ is a class of $\mathcal{L}$-structures, then its theory is defined to be

$$Th\boldsymbol{K} := \bigcap_{\mathcal{A}\in\boldsymbol{K}} Th\mathcal{A}$$

which is a theory as the intersection of theories.

Exercise

Suppose $x \notin freeX$ and c is not in X or α. Prove

$$X \vDash \alpha \iff X \vDash \forall x\alpha \iff X \vDash \alpha\tfrac{c}{x}$$

Since $x \notin freeX$, we have $X \vDash \alpha \implies \forall x\alpha$. And similarly $X \vDash \forall x\alpha \implies X \vDash \alpha$ since $\alpha, \frac{x}{x}$ is collision-free. And since $\alpha, \frac{c}{x}$ is collision-free, we have $X \vDash \forall x\alpha \implies X \vDash \alpha\frac{c}{x}$.

Now suppose $X \vDash \alpha\frac{c}{x}$, then suppose $\mathcal{M} \vDash X$ then $\mathcal{M}^c_x \vDash \alpha$ by the substitution theorem. Let $a \in A$, since c is not in X, if we set $c^{\mathcal{M}} = a$ then $\mathcal{M}$ still satisfies X, and so $\mathcal{M}^a_x \vDash \alpha$, thus $\mathcal{M} \vDash \forall x\alpha$. Thus $X \vDash \forall x\alpha$ as required.

Exercise

Let S be a set of sentences, α and β be formulas, $x \notin \mathit{free}\beta$, and c be a constant not occurring in S, α, or β. Show that

$$S \vDash \alpha\frac{c}{x} \to \beta \iff S \vDash \exists x\alpha \to \beta$$

If $S \vDash \alpha\frac{c}{x} \to \beta$, then if $\mathcal{M} \vDash S$, if $\mathcal{M} \vDash \exists x\alpha$ then there exists an $a \in A$ such that $\mathcal{M}^a_x \vDash \alpha$. Since we can set $c^{\mathcal{M}} = a$ as c does not occur in S, α, β (again, the coincidence theorem), so $\mathcal{M} \vDash \alpha\frac{c}{x}$ by the substitution theorem. Thus $\mathcal{M} \vDash \beta$, meaning $\mathcal{M} \vDash \exists x\alpha \to \beta$ so $S \vDash \exists x\alpha \to \beta$.

And if $S \vDash \exists x\alpha \to \beta$ then since $\vDash \alpha\frac{c}{x} \to \exists x\alpha$ we have $S \vDash \alpha\frac{c}{x} \to \beta$ as required.

Exercise

Show that for all sentences $\alpha, \beta \in \mathcal{L}^0$, if T is an $\mathcal{L}$-theory then

$$\beta \in T + \alpha \iff \alpha \to \beta \in T$$

If $\beta \in T + \alpha$ then $T, \alpha \vDash \beta$ so $T \vDash \alpha \to \beta$ by the deduction theorem, thus $\alpha \to \beta \in T$ since T is deductively closed. Conversely if $\alpha \to \beta \in T$ then $T, \alpha \vDash \alpha \to \beta, \alpha$ and so $T, \alpha \vDash \beta$ thus $\beta \in T + \alpha$.

Exercise

Let $T \subseteq \mathcal{L}$ be an $\mathcal{L}$-theory and $\mathcal{L}_0 \subseteq \mathcal{L}$ a reduction of the language $\mathcal{L}$. Show that $T_0 := T \cap \mathcal{L}_0$ is also a ($\mathcal{L}_0$-)theory.

Suppose $T_0 \vDash \alpha$ for a $\mathcal{L}_0$-sentence α, then $T \vDash \alpha$ and so $\alpha \in T$ and $\alpha \in \mathcal{L}_0$ so $\alpha \in T_0$. (Note that T_0 is not necessarily a $\mathcal{L}$-theory. For example if c is a constant symbol not occurring in $\mathcal{L}_0$ then $T_0 \vDash c = c$ but $c = c$ is a sentence not in $\notin \mathcal{L}_0$, so $c = c \notin T_0$).

2.6 Explicit Definitions—Language Expansions

It is often very useful to define notions within a theory on top of the language of the theory. For example, in rings the concept of divisibility is often useful, where divisibility is defined by $x|y \iff \exists z\, x \cdot z = y$.

In this section, if $\varphi(\vec{x})$ is a formula (meaning $\mathit{free}\varphi \subseteq \{x_1, \dots, x_n\}$), we use the notation $\varphi(\vec{t})$ in place of $\varphi\frac{\vec{t}}{\vec{x}}$. Substitutions here are tacitly assumed to be collision-free.

2.6.1 Definition

Let r be an n-ary relational symbol not occurring in $\mathcal{L}$. An **explicit definition of** r **in** $\mathcal{L}$ is a formula of the form

$$\eta_r: \qquad r\vec{x} \leftrightarrow \delta(\vec{x})$$

where $\delta(\vec{x}) \in \mathcal{L}$ and $\vec{x}$ is a vector of distinct variables. δ is called the **defining formula of** r. If T is an $\mathcal{L}$-theory, then $T_r := T + \eta_r^g$ is the **definitorial extension** of T by r.

Similarly if f is an n-ary ($n \geq 0$) function (or constant, in the case $n = 0$) symbol not occurring in $\mathcal{L}$ its explicit definition is of the form

$$\eta_f: \qquad y = f\vec{x} \leftrightarrow \delta(\vec{x}, y)$$

where $\delta \in \mathcal{L}$ is the defining formula, and $\vec{x}, y$ are distinct. η_f is called **legitimate** in a theory T if $T \vDash \forall\vec{x}\exists!y\delta$ (so f is indeed a function), and if η_f is legitimate then $T_f := T + \eta_f^g$ is the definitorial extension of T by f.

2.6.2 Definition

Let s be a (function, relational, or constant) symbol and $\mathcal{L}$ be a first-order language. Then we define the

language extension $\mathcal{L}[s]$ to be the first-order language obtained by adding s to the signature.

If r is an n-ary relational symbol not occuring in $\mathcal{L}$ defined by η_r, then for every $\varphi \in \mathcal{L}[r]$ we define the **reduced formula** φ^{rd} by replacing every prime formula $r\vec{t}$ in φ by $\delta_r(\vec{t})$.

Now, if f is an n-ary function symbol legitimately defined by η_f in T, then for every $\varphi \in \mathcal{L}[f]$ define φ^{rd} in steps: for the first occurrence of f in φ on the left, we may write $\varphi = \varphi_0\frac{f\vec{t}}{y}$ for appropriate φ_0, $\vec{t}$, and $y \notin \mathrm{var}\varphi$. So $\varphi \equiv_{T_f} \exists y(\varphi_0 \wedge y = f\vec{t}) \equiv_{T_f} \varphi_1$ where $\varphi_1 = \exists y(\varphi_0 \wedge \delta_f(\vec{t}, y))$. Continue inductively on φ_1 until getting a formula φ_m in which f does not occur, set $\varphi^{rd} = \varphi_m$.

Obviously if $\varphi \in \mathcal{L}$ (ie. they do not include the defined symbols), $\varphi^{rd} = \varphi$. For example, in T_G we can define $^{-1}$ by $y = x^{-1} \leftrightarrow x \circ y = e$, and this is legitimate in T_G.

2.6.3 Theorem (The Elimination Theorem)

Let T be an $\mathcal{L}$-theory, and T_{s} be the definitorial extension of T by some η_{s}. Then for all $\varphi \in \mathcal{L}[\mathsf{s}]$,

$$T_{\mathsf{s}} \vDash \varphi \iff T \vDash \varphi^{rd}$$

We will prove this in the case that $\mathsf{s} = r$ is an n-ary relational symbol. If $\mathcal{A} \vDash T$ then $\mathcal{A}$ can be expanded to a model $\mathcal{A}' \vDash T_r$ by setting $r^{\mathcal{A}'}\vec{a} \iff \mathcal{A} \vDash \delta[\vec{a}]$. Since $r\vec{t} \equiv_{T_r} \delta(\vec{t})$ for any $\vec{t}$, by The Replacement Theorem $\varphi \equiv_{T_r} \varphi^{rd}$ for all formulas $\varphi \in \mathcal{L}[r]$. Thus

$$\begin{aligned}
T_r \vDash \varphi &\iff \mathcal{A}' \vDash \varphi && \text{for all } \mathcal{A} \vDash T && (\mathrm{Md}\, T_r = \{\mathcal{A}' \mid \mathcal{A} \vDash T\}) \\
&\iff \mathcal{A}' \vDash \varphi^{rd} && \text{for all } \mathcal{A} \vDash T && (\varphi \equiv_{T_r} \varphi^{rd}) \\
&\iff \mathcal{A} \vDash \varphi^{rd} && \text{for all } \mathcal{A} \vDash T && (\text{The Coincidence Theorem}) \\
&\iff T \vDash \varphi^{rd}
\end{aligned}$$

■

Sometimes function symbols are defined similar to

$$f\vec{x} := t(\vec{x})$$

where t is a term. This is simply equivalent to

$$\eta_f: \qquad y = f\vec{x} \leftrightarrow y = t(\vec{x})$$

and so it is a special case of explicit definitions. And all definitions of this form are legitimate: $\forall\vec{x}\exists!y\, y = t(\vec{x})$ is a tautology.

2.6.4 Definition

Suppose T is an $\mathcal{L}$-theory, and Δ is a set of explicit definitions, then $T' = T + \Delta$ is a **definitorial extension** of T. Let $\mathcal{L}'$ be the language expansion of $\mathcal{L}$ obtained by adjoining all the new symbols defined in Δ to $\mathcal{L}$, then for $\varphi \in \mathcal{L}'$, we define φ^{rd} as above where the procedure is done stepwise for every new symbol in φ.

From The Elimination Theorem it is immediate

2.6.5 Theorem (The General Elimination Theorem)

Let T' be a definitorial extension of T, then $\alpha \in T' \iff \alpha^{rd} \in T$ for all $\alpha \in \mathcal{L}'$. In particular if $\alpha \in \mathcal{L}$, $\alpha \in T' \iff \alpha \in T$, thus T' is a conservative extension of T.

2.6.6 Definition

Let T be an $\mathcal{L}$-theory, and s be a symbol in $\mathcal{L}$. Then s is **explicitly definable** in T if T contains an explicit definition of s where s's defining formula is in $\mathcal{L}_0$, the language obtained by removing s from $\mathcal{L}$.

For example, the constant e is explicitly definable in the theory of groups T_G, since $x = e \leftrightarrow x \circ x = x$ is an explicit definition of e which is in T_G (more specifically, its generalized is).

2.6.7 Proposition

Suppose s is explicitly definable in $Th\mathcal{A}$ and legitimate, suppose we extend $\mathcal{A}$ to $\mathcal{A}'$ by adding the symbol s to its signature. Then automorphisms of $\mathcal{A}$ are automorphisms of $\mathcal{A}'$ (the converse is trivial).

Let σ be an automorphism of $\mathcal{A}$. If $\mathsf{s} = r$ is a relational symbol, we must show that $r\vec{x} \iff r\sigma\vec{x}$. Now suppose that δ is the defining formula of r, so this is equivalent to showing $\mathcal{A} \vDash \delta[\vec{a}]$ if and only if $\mathcal{A} \vDash \delta[\sigma\vec{a}]$ for all $\vec{a} \in A^n$. This is true by The Invariance Theorem. Now suppose that $\mathsf{s} = f$ is a function symbol, we must show that $f\sigma\vec{x} = \sigma f\vec{x}$. Since $y = f\vec{x}$ if and only if $\delta[\vec{x}, y]$, we must show that $\delta[\vec{x}, y]$ if and only if $\delta[\sigma\vec{x}, \sigma y]$. This is as then if $y = f\vec{x} \implies \sigma y = f\sigma\vec{x}$ so $\sigma f\vec{x} = f\sigma\vec{x}$. But this is again true by The Invariance Theorem. ■

Let us now turn our attention to another type of normal forms. We call two formulas α and β *satisfiably equivalent* if they are both satisfiable (not necessarily by the same model) or they are both not satisfiable. Now suppose α is a formula, we assume without loss of generality that it is a PNF $\alpha = \mathtt{Q}_1 x_1 \cdots \mathtt{Q}_n x_n \beta$ (β quantifier-free). We will construct a satisfiably equivalent $\forall$-formula $\hat{\alpha}$ such that α and $\hat{\alpha}$ are satisfiably equivalent. We do this inductively/stepwise: let $\alpha_0 = \alpha$ and suppose we have already constructed α_i. If α_i is already a $\forall$-formula, set $\hat{\alpha} = \alpha_i$, otherwise α_i has the form $\forall x_1 \cdots \forall x_k \exists y \beta_i$ for $k \geq 0$. Let f be some k-ary (if $k = 0$ this is a constant) function symbol not yet used (not in $\mathcal{L}$, not used in the previous i steps) and let $\alpha_{i+1} = \forall x_1 \cdots \forall x_k \beta_i \frac{f\vec{x}}{y}$. $\hat{\alpha}$ is called the *Skolem Normal Form* (SNF for short) of α.

Essentially what is done here is that variables which are existentially quantified are viewed instead as functions of the universally quantified variables (which preceed them on the left).

If m is the number of $\exists$ quantifiers in $\mathtt{Q}_1, \dots, \mathtt{Q}_n$, then after m steps we will have $\hat{\alpha} = \alpha_m$ and $\mathit{free}\hat{\alpha} = \mathit{free}\alpha$ (since $\mathit{free}\alpha_i = \mathit{free}\alpha$ for all i). And $\hat{\alpha}$ is a $\forall$-formula.

For example, let α be the formula $\forall x \exists y\, x < y$ then $\hat{\alpha}$ is $\forall x\, x < fx$. And for $\alpha = \exists x \forall y\, x \cdot y = y$, we have $\hat{\alpha} = \forall y\, c \cdot y = y$. And for $\alpha = \forall x \forall y \exists z (x < z \land y < z)$ then $\hat{\alpha} = \forall x \forall y (x < fxy \land y < fxy)$.

2.6.8 Theorem

Let $\hat{\alpha}$ be the Skolem normal form of the formula α. Then

(1) $\hat{\alpha} \vDash \alpha$ (2) $\hat{\alpha}$ and α are satisfiably equivalent

To prove (1), it is sufficient to show that $\alpha_{i+1} \vDash \alpha_i$ for each of the construction steps of $\hat{\alpha}$ (this is sufficient due to the transitivity of $\vDash$). Since $\beta_i, \frac{f\vec{x}}{y}$ is collision-free (β_i is a PNF not containing quantifiers of $\vec{x}$) we have that $\beta_i \frac{f\vec{x}}{y} \vDash \exists y \beta_i$. Then by anterior and posterior generalization, $\alpha_{i+1} = \forall \vec{x} \beta_i \frac{f\vec{x}}{y} \vDash \forall \vec{x} \exists y \beta_i = \alpha_i$ as required.

By (1), if $\hat{\alpha}$ is satisfiable, so is α. Conversely, if $\mathcal{A} \vDash \alpha_i[\vec{c}] = \forall \vec{x} \exists y \beta_i(\vec{x}, y, \vec{z})[\vec{c}]$. For each $\vec{a} \in A^n$, choose a $b \in A$ such that $\mathcal{A} \vDash \beta_i[\vec{a}, b, \vec{c}]$ (*choosing* such a b is possible due to the axiom of choice), and expand $\mathcal{A}$ to $\mathcal{A}'$ by setting $f^{\mathcal{A}'}\vec{a} = b$ where f is the new function symbol. Then $\mathcal{A}' \vDash \alpha_{i+1}[\vec{c}]$. So if α is satisfiable, so is every α_i, and in particular $\hat{\alpha}$ (and the model which satisfies $\hat{\alpha}$ is an expansion of the one which satisfies α). ■

Notice that if α is a formula then $\neg\alpha$ has a SNF, and so its negation is an $\exists$-formula and is satisfiably equivalent to α. We will denote this $\exists$-formula by $\check{\alpha}$: if $\beta = \neg\alpha$ then $\check{\alpha} = \neg\hat{\alpha}$.

For example, if $\alpha = \exists x \forall y (ry \to rx)$ then $\neg\alpha \equiv \beta = \forall x \exists y (ry \land \neg rx)$ and then $\hat{\beta} = \forall x (rfx \land \neg rx)$ and so $\check{\alpha} = \exists x (rfx \to rx)$. This is actually a tautology. Skolem normal forms are used in model theory and logic programming.

Exercise

Let T_f result from adjoining an explicit definition η_f for f to a theory T. Show that T_f is a conservative extension of T if and only if η_f is legitimate.

Suppose δ_f is the defining formula for f, meaning $\eta_f = y = f\vec{x} \leftrightarrow \delta(\vec{x}, y)$. We have already shown that if η_f is legitimate then T_f is a conservative extension of T, we will now show that if T_f is conservative then η_f is legitimate. We know that $T_f \vDash \forall\vec{x}\exists! y\, y = f\vec{x}$ and since $y = f\vec{x} \equiv_{T_f} \delta(\vec{x}, y)$ we have that $T_f \vDash \forall\vec{x}\exists! y\, \delta(\vec{x}, y)$ and since T_f is conservative this means $T \vDash \forall\vec{x}\exists! y\delta(\vec{x}, y)$ so η_f is legitimate.

Exercise

Let $\mathtt{S}\colon n \mapsto n+1$ be the successor function in $\mathcal{N} = (\mathbb{N}, 0, \mathtt{S}, +, \cdot)$. Show that $Th\mathcal{N}$ is a definitorial extension of $Th(\mathbb{N}, \mathtt{S}, \cdot)$; ie. 0 and $+$ are explicitly definable by $\mathtt{S}$ and $\cdot$ in $\mathbb{N}$.

0 is the only number which is not the successor of some other number, so we could define 0 by $\delta_0(x) = \forall y\, \mathtt{S}y \neq x$. Alternatively 0 is the only number whose product with every number is itself, so we could also define $\delta_0(x) = \forall y\, x \cdot y = x$. Addition requires more consideration.

Exercise

$<$ is explicitly definable in $(\mathbb{N}, 0, +)$ by $x < y \leftrightarrow (\exists z \neq 0)x + z = y$. Show that $<$ is not explicitly definable in $(\mathbb{Z}, 0, +)$.

We showed that explicitly definable relations are invariant under automorphisms. But $n \mapsto -n$ is an automorphism of $(\mathbb{Z}, 0, +)$ (since $-0 = 0$ and $(-x) + (-y) = -(x + y)$) but $<$ is certainly not invariant under it.

3 Complete Logical Calculi

3.1 A Calculus of Natural Deduction

Let $\mathcal{L}$ be an arbitrary first-order language in the logical signature $\neg, \wedge, \forall, =$. We define a calculus $\vdash$ using the following basic rules (the definitions for sequents and proofs remain the same as they were in propositional logic):

$$\text{(IR)} \quad \frac{}{X \vdash \alpha} \quad (\alpha \in X \cup \{t = t\}) \qquad \text{(MR)} \quad \frac{X \vdash \alpha}{X' \vDash \alpha} \quad (X \subseteq X')$$

$$(\wedge 1) \quad \frac{X \vdash \alpha, \beta}{X \vdash \alpha \wedge \beta} \qquad (\wedge 2) \quad \frac{X \vdash \alpha \wedge \beta}{X \vdash \alpha, \beta}$$

$$(\neg 1) \quad \frac{X \vdash \alpha, \neg\alpha}{X \vdash \beta} \qquad (\neg 2) \quad \frac{X, \alpha \vdash \beta \mid X, \neg\alpha \vdash \beta}{X \vdash \beta}$$

$$(\forall 1) \quad \frac{X \vdash \forall x \alpha}{X \vdash \alpha \frac{t}{x}} \quad (\alpha, \tfrac{t}{x} \text{ collision-free}) \qquad (\forall 2) \quad \frac{X \vdash \alpha \frac{y}{x}}{X \vdash \forall x \alpha} \quad (y \notin \mathit{free} X \cup \mathit{var} \alpha)$$

$$(=) \quad \frac{X \vdash s = t, \alpha \frac{s}{x}}{X \vdash \alpha \frac{t}{x}} \quad (\alpha \text{ prime})$$

Note that $(\forall 2)$ is the result of posterior generalization: if $X \vDash \alpha$ then $X \vDash \forall x \alpha$ for $x \notin \mathit{free} X$. Set $\alpha = \alpha \frac{y}{x}$ then then $X \vDash \alpha \frac{y}{x}$ implies $X \vDash \forall y \alpha \frac{y}{x} \equiv \forall x \alpha$ for $y \notin \mathit{free} X \cup \mathit{var} \alpha$ by renaming bound variables. Thus all of these rules are sound in the sense that they are true for $\vDash$. The rule $(=)$ could be strengthened to require $\alpha, \frac{t}{x}, \frac{s}{x}$ be collision-free, but requiring α be prime instead is sufficient.

Since this calculus is an extension of the Gentzen-style calculus of propositional logic, all of the rules we proved in propositional logic carry over here.

Again, let R be a rule of the form

$$R\colon \quad \frac{X_1 \vdash \alpha_1 \mid \cdots \mid X_n \vdash \alpha_n}{X \vdash \alpha}$$

Then we say that a property of sequents (which are again pairs (X, α) where $X \subseteq \mathcal{L}$ and $\alpha \in \mathcal{L}$) $\mathcal{E}$ is *closed under* R if $\mathcal{E}(X_1, \alpha_1), \ldots, \mathcal{E}(X_n, \alpha_n)$ implies $\mathcal{E}(X, \alpha)$. The same principle from propositional calculus holds here, and the proof is the same.

3.1.1 Proposition (Principle of Rule Induction)

Let $\mathcal{E}$ be a property of sequents (X, α) such that $\mathcal{E}$ is closed under all of the basic rules defined above, then $X \vdash \alpha$ implies $\mathcal{E}(X, \alpha)$.

Thus if we let $\mathcal{E}(X, \alpha)$ be the property that $X \vDash \alpha$ then we get that $\mathcal{E}$ is closed under all of the basic rules, and so $X \vdash \alpha \implies X \vDash \alpha$, or more suggestively $\vdash \subseteq \vDash$. Similarly if we have $\mathcal{L} \subseteq \mathcal{L}'$ and we let $\mathcal{E}(X, \alpha)$ be the property $X \vdash_{\mathcal{L}'} \alpha$ then we get $X \vdash_{\mathcal{L}} \alpha \implies X \vdash_{\mathcal{L}'} \alpha$, or $\vdash_{\mathcal{L}} \subseteq \vdash_{\mathcal{L}'}$.

Similarly let $\mathcal{E}(X, \alpha)$ be the property that there exists a finite $X_0 \subseteq X$ and a finite $\mathcal{L}_0 \subseteq \mathcal{L}$ (meaning the signature of $\mathcal{L}_0$ is finite, not $\mathcal{L}_0$ itself, which is always infinite) where $X_0 \vdash_{\mathcal{L}_0} \alpha$. Notice that this requires that $X_0 \cup \{\alpha\} \subseteq \mathcal{L}_0$. We will first show that (IR) is closed under $\mathcal{E}$, so if $\alpha \in X \cup \{t = t\}$ then let $X_0 = \{\alpha\}$ and let $\mathcal{L}_0$ be constructed from the symbols in α which is finite. And then we get $X_0 \vdash_{\mathcal{L}_0} \alpha$. For (MR), it is trivial. For $(\wedge 1)$ suppose $X_1 \vdash_{\mathcal{L}_1} \alpha_1$ and $X_2 \vdash_{\mathcal{L}_2} \alpha_2$ where $X_i \subseteq X$ and the signatures of $\mathcal{L}_i$ are finite. Then let $\mathcal{L}_0$ be constructed by taking the union of the signatures of $\mathcal{L}_1$ and $\mathcal{L}_2$, and $X_0 = X_1 \cup X_2$ then we get that $X_0 \vdash_{\mathcal{L}} \alpha, \beta$ and so $X_0 \vdash_{\mathcal{L}} \alpha \wedge \beta$ and $X_0 \subseteq X$ is finite and $\mathcal{L}$ has a finite signature, as required. The proofs for the rest of the rules are similar.

3.1.2 Theorem

If $X \vdash_{\mathcal{L}} \alpha$ then there exists a finite $X_0 \subseteq X$ and a language $\mathcal{L}_0 \subseteq \mathcal{L}$ with finitely many symbols such that $X_0 \vdash_{\mathcal{L}_0} \alpha$.

Let us now prove the following

$$(1) \quad \frac{X \vdash s = t, s = t'}{X \vdash t = t'} \qquad (2) \quad \frac{X \vdash s = t}{X \vdash t = s} \qquad (3) \quad \frac{X \vdash t = s, s = t'}{X \vdash t = t'}$$

To show (1), let $x \notin \text{var}t'$ and let $\alpha = x = t'$ then the premise of (1) can be rewritten as $X \vdash s = t, \alpha\frac{s}{x}$. Rule $(=)$ yields $X \vdash \alpha\frac{t}{x} = t = t'$ as required. For (2), by (IR) $X \vdash s = t, s = s$ so by (1) we get $X \vdash t = s$. And so follows (3), as the premise gives $X \vdash s = t, s = t'$ by (2) which gives $X \vdash t = t'$ by (1).

Now we show (f and r are function and relation symbols of arity n),

$$(1)\ \frac{X \vdash t_i = t}{X \vdash f\vec{t} = ft_1\cdots t_{i-1}tt_{i+1}\cdots t_n} \qquad (2)\ \frac{X \vdash \vec{t} = \vec{t'}}{X \vdash f\vec{t} = f\vec{t'}}$$

$$(3)\ \frac{X \vdash t_i = t, r\vec{t}}{X \vdash rt_1\cdots t_{i-1}tt_{i+1}\cdots t_n} \qquad (4)\ \frac{X \vdash \vec{t} = \vec{t'}, r\vec{t}}{X \vdash r\vec{t'}}$$

We prove (1): let us define $\alpha = f\vec{t} = ft_1\cdots x\cdots t_n$ where x does not occur in any t_j. Then if $X \vdash t_i = t$ we also have $X \vdash \alpha\frac{t_i}{x}$ by (IR), as $\alpha\frac{t_i}{x} = f\vec{t} = f\vec{t}$. So by $(=)$ we get $X \vdash \alpha\frac{t}{x} = f\vec{t} = ft_1\cdots t\cdots t_n$, as required. for (2), apply (1) n times (since $\vec{t'} = \vec{t}$ means $t'_i = t_i$ for each i). Rule (3) is proven similarly to rule (1), and (4) is proven similar to (2).

We will also show

$$(1)\ \vdash \exists x\, t = x\ (x \notin \text{var}t) \qquad (2)\ \vdash \exists x\, x = x$$

These are trivial to show semantically, but require a bit more work to show derive them using our calculus. ($\forall$1) gives $\forall x\, t \neq x \vdash (t \neq x)\frac{t}{x} = t \neq t$. And by (IR), $\forall x\, t \neq x \vdash t = t$, and so by ($\neg$1) we get $\forall x\, t \neq x \vdash \exists x\, t = x$. And so $\neg\forall x\, t \neq x \vdash \exists x\, t = x$ (as by definition these are equal, so this is by (IR)). So by ($\neg$2), we get $\vdash \exists x\, t = x$ as required. (2) is verified similarly, starting with $\forall x\, x \neq x \vdash x \neq x, x = x$.

3.1.3 Definition

A set $X \subseteq \mathcal{L}$ is called **inconsistent** if $X \vdash \alpha$ for all formulas $\alpha \in \mathcal{L}$. A non-inconsistent set is called **consistent**.

If a set is satisfiable, it is obviously consistent (since $\vdash \subseteq \vDash$). Also by ($\neg$1), the inconsistency of X is equivalent to $X \vdash \alpha, \neg\alpha$ for any α, which is equivalent to $X \vdash \bot$ (since $\bot = \neg\top = \exists \boldsymbol{v}_0\, \boldsymbol{v}_0 = \boldsymbol{v}_0$ and $X \vdash \top$ as we showed above).

And as before (the proof remains valid), we have the following two equivalences:

$$\mathtt{C}^+\colon\quad X \vdash \alpha \iff X, \neg\alpha \vdash \bot \qquad \mathtt{C}^-\colon\quad X \vdash \neg\alpha \iff X, \alpha \vdash \bot$$

Similarly again to propositional logic, a set X is *maximally consistent* if X is consistent but every $X \subset X'$ is inconsistent.

Exercise

Derive the rule $\dfrac{X \vdash \alpha\frac{t}{x}}{X \vdash \exists x\alpha}$ where $\alpha, \frac{t}{x}$ are collision-free.

By ($\forall$1) we have $X, \forall x\neg\alpha \vdash \neg\alpha\frac{t}{x}$, then by the premise and (MR) we have $X, \forall x\neg\alpha \vdash \alpha\frac{t}{x}$. Thus by ($\neg$1) we have $X, \forall x\neg\alpha \vdash \neg\forall x\neg\alpha$. And by (IR) and (MR) we also have $X, \neg\forall x\neg\alpha \vdash \neg\forall x\neg\alpha$, and so $X \vdash \neg\forall x\neg\alpha = \exists x\alpha$ as required.

Exercise

Prove $\forall x\alpha \vdash \forall y\big(\alpha\frac{t}{x}\big)$ and $\forall y\big(\alpha\frac{y}{x}\big) \vdash \forall x\alpha$ provided $y \notin \text{var}\alpha$.

Since let $z \notin \text{var}\alpha$ and $z \neq y$ then $\alpha\frac{z}{x} = \alpha\frac{y}{x}\frac{z}{y}$, and so by ($\forall$1), $\forall x\alpha \vdash \alpha\frac{y}{x}\frac{z}{y}$. Then by ($\forall$2) we get $\forall x\alpha \vdash \forall y\big(\alpha\frac{y}{x}\big)$.

Exercise

Prove the rule $\dfrac{X \vdash \forall y\big(\alpha\frac{y}{x}\big)}{X \vdash \forall z\big(\alpha\frac{z}{x}\big)}$ where $y, z \notin \text{var}\alpha$.

We have that $\forall y\big(\alpha\frac{y}{x}\big) \vdash \forall z\Big(\alpha\frac{y}{x}\frac{z}{y}\Big) = \forall z\big(\alpha\frac{z}{x}\big)$ by the previous problem. So we have $X \vdash \forall y\big(\alpha\frac{y}{x}\big)$ and $X, \forall y\big(\alpha\frac{y}{x}\big) \vdash \forall z\big(\alpha\frac{z}{x}\big)$ so by the cut rule, we get $X \vdash \forall z\big(\alpha\frac{z}{x}\big)$, as required.

Exercise

Show that a consistent set $X \subseteq \mathcal{L}$ is maximally consistent if and only if for all $\varphi \in \mathcal{L}$, $\varphi \in X$ or $\neg\varphi \in X$. This means that maximally consistent sets are deductively closed, ie. $X \vdash \varphi \implies \varphi \in X$.

Let $\varphi \in \mathcal{L}$ then suppose $\varphi, \neg\varphi \notin X$ then $X, \varphi \vdash \bot$ and $X, \neg\varphi \vdash \bot$ since X is maximally consistent, so any extension is inconsistent. Thus by $(\neg 2)$ we have $X \vdash \bot$ contradicting X being consistent. Now, if $X \vdash \varphi$ then if $\varphi \notin X$ we have $\neg\varphi \in X$ and so $X \vdash \varphi, \neg\varphi$ contradicting X's consistency.

3.2 The Completeness Proof

We previously used the notation $\mathcal{L}[\mathsf{s}]$ for the language obtained by adjoining the symbol s to the signature of $\mathcal{L}$. For a set of symbols S we will similarly use the notation $\mathcal{L}[S]$ for the language obtained by adjoining the symbols in S to the signature of $\mathcal{L}$. We will also sometimes omit the brackets and instead write $\mathcal{L}\mathsf{s}$ or $\mathcal{L}S$.

If C is a set of constants, then $\mathcal{L}C$ is a *constant expansion of* $\mathcal{L}$. If $\alpha \in \mathcal{L}c$ and $z \in Var$ then we write $\alpha\frac{z}{c}$ to be the formula in $\mathcal{L}$ obtained by replacing every occurrence of c in α with the variable z. Similarly if $X \subseteq \mathcal{L}c$, then $X\frac{z}{c} := \{\alpha\frac{z}{c} \mid \alpha \in X\}$.

3.2.1 Lemma

Suppose $X \vdash_{\mathcal{L}c} \alpha$ then $X\frac{z}{c} \vdash_{\mathcal{L}} \alpha\frac{z}{c}$ for almost all (meaning all but a finite number) of variables z.

We do this by rule induction on $\vdash_{\mathcal{L}c}$. We start with (IR): if $\alpha \in X$ then $\alpha\frac{z}{c} \in X\frac{z}{c}$ as well, so $X\frac{c}{z} \vdash_{\mathcal{L}} \alpha\frac{z}{c}$ as required. If $\alpha = t = t$ then $\alpha\frac{z}{c}$ is also of the form $t' = t'$ and so $X\frac{z}{c} \vdash_{\mathcal{L}} \alpha\frac{z}{c}$ as well. (MR) through $(\neg 2)$ are similarly obvious and impose no restrictions on z; only the steps for $(\forall 1)$, $(\forall 2)$, and $(=)$ aren't immediately obvious. We will show only $(\forall 1)$ as the other two are proven in a similar manner. Let $\alpha, \frac{t}{x}$ be collision-free and $X \vdash_{\mathcal{L}c} \forall x\alpha$, then by induction we have $X\frac{z}{c} \vdash_{\mathcal{L}} (\forall x\alpha)\frac{z}{c}$ for almost all z. We may assume $z \notin var\forall x\alpha, vart$. We can verify through formula induction that $\alpha\frac{t}{x}\frac{z}{c} = \alpha'\frac{t'}{x}$ where $\alpha' = \alpha\frac{z}{c}$ and $t' = t\frac{z}{c}$. Then $\alpha', \frac{t'}{x}$ is collision-free (the bound variables of α' are the bound variables of α, and t' contains potentially only one more variable than t, z which does not occur in α). So we have that $X\frac{z}{c} \vdash_{\mathcal{L}} (\forall x\alpha)\frac{z}{c} = \forall x\alpha'$, thus by $(\forall 1)$ we have $X\frac{z}{c} \vdash_{\mathcal{L}} \alpha'\frac{t'}{x} = \left(\alpha\frac{t}{x}\right)\frac{z}{c}$, and this is true for almost all variables z, as required. ■

This lemma gives rise to the following rule:

$$(\forall 3) \quad \frac{X \vdash \alpha\frac{c}{x}}{X \vdash \forall x\alpha} \quad (c \text{ not in } X, \alpha)$$

Suppose $X \vdash \alpha\frac{c}{x}$, by the finiteness of $\vdash$ we may assume that X is finite. By the above lemma, where $\mathcal{L}c = \mathcal{L}$ (ie. we adjoin a constant already in $\mathcal{L}$), we can find a y not occurring in $X \cup \{\alpha\}$ (as X is finite) such that $X\frac{y}{c} \vdash \alpha\frac{c}{x}\frac{y}{c} = \alpha\frac{y}{x}$ (since c does not occur in α). Since c does not occur in X, we have $X\frac{y}{c} = X$ and thus we get $X \vdash \alpha\frac{y}{x}$ and so $X \vdash \forall x\alpha$ by $(\forall 2)$.

3.2.2 Lemma

Let C be a set of constant symbols and $\mathcal{L}' = \mathcal{L}C$. Then $X \vdash_{\mathcal{L}} \alpha \iff X \vdash_{\mathcal{L}'} \alpha$ for all $X \subseteq \mathcal{L}$ and $\alpha \in \mathcal{L}$. Thus $\vdash_{\mathcal{L}'}$ is a conservative expansion of $\vdash_{\mathcal{L}}$.

By the monotonicity of $\vdash$, we have $X \vdash_{\mathcal{L}} \alpha \implies X \vdash_{\mathcal{L}'} \alpha$. Now conversely, suppose $X \vdash_{\mathcal{L}'} \alpha$ then we may assume by finiteness that C is finite, and since adjoining finitely many constants can be done stepwise, we may assume that $\mathcal{L}' = \mathcal{L}c$ for a single constant c not occurring in $\mathcal{L}$ (if c occurs in $\mathcal{L}$ this is trivial). Then lemma 3.2.1 tells us that $X\frac{z}{c} \vdash_{\mathcal{L}} \alpha\frac{z}{c}$ for some variable z, but since $X \subseteq \mathcal{L}$ and $\alpha \in \mathcal{L}$ so c occurs in neither X nor α so this is equivalent to $X \vdash_{\mathcal{L}} \alpha$ as required. ■

By this lemma, we may write $\vdash$ for the derivability relation $\mathcal{L}$ and every constant expansion $\mathcal{L}'$ and there will be no misunderstandings as the expansions are conservative. This also means that X is consistent in $\mathcal{L}$ if and only if it is consistent in $\mathcal{L}'$ as $X \vdash_{\mathcal{L}} \bot \iff X \vdash_{\mathcal{L}'} \bot$.

For every variable x and formula $\alpha \in \mathcal{L}$ let us define a distinct constant symbol $c_{x,\alpha}$ which does not occur in $\mathcal{L}$. Then we define

$$\alpha^x := \neg\forall x\alpha \wedge \alpha\frac{c_{x,\alpha}}{x}$$

Notice that $\neg\alpha^x \equiv \exists x\neg\alpha \to \neg\alpha\frac{c}{x}$, meaning $\neg\alpha^x$ says that "if α is not true for all x, c provides a counterexample for α". Thus $\neg\alpha^x \equiv \top$ whenever $x \notin free\alpha$ (since then $\neg\alpha^x \equiv \neg\alpha \to \neg\alpha$).

3.2.3 Lemma

Let $\Gamma_{\mathcal{L}} := \{\neg\alpha^x \mid \alpha \in \mathcal{L},\, x \in \mathit{Var}\}$, and let $X \subseteq \mathcal{L}$ be consistent. Then $X \cup \Gamma_{\mathcal{L}}$ is also consistent.

Suppose $X \cup \Gamma_{\mathcal{L}} \vdash \bot$, then by the finiteness of $\vdash$, there exist $\neg\alpha_0^{x_0}, \dots, \neg\alpha_n^{x_n} \in \Gamma_{\mathcal{L}}$ such that $X \cup \{\neg\alpha_i^{x_i} \mid i \leq n\} \vdash \bot$. Since X is consistent and so $X \nvdash \bot$, we can assume that n is minimal so that $X' = X \cup \{\neg\alpha_i^{x_i} \mid i < n\} \nvdash \bot$. Let $x = x_n$, $\alpha = \alpha_n$, and $c = c_{x,\alpha}$ then $X' \cup \{\neg\alpha^x\} \vdash \bot$ and so by $\mathtt{C}^+$ we have $X' \vdash \alpha^x = \neg\forall x\alpha \wedge \alpha\frac{c}{x}$. Thus by ($\wedge$2) we have $X' \vdash \neg\forall x\alpha, \alpha\frac{c}{x}$. By ($\forall$3), we get that $X' \vdash \alpha\frac{c}{x}$ means $X' \vdash \forall x\alpha$, and this means that X' is inconsistent, in contradiction. ■

3.2.4 Definition

A set $X \subseteq \mathcal{L}$ is called a **Henkin set** if X satisfies the following two conditions:

(H1) $X \vdash \neg\alpha \iff X \nvdash \alpha$, (equivalently $X \vdash \alpha \iff X \nvdash \neg\alpha$)

(H2) $X \vdash \forall x\alpha \iff X \vdash \alpha\frac{c}{x}$ for all constants c in $\mathcal{L}$.

(H1) and (H2) imply a third property:

(H3) For each term t there is a constant c such that $X \vdash t = c$

We showed $X \vdash \neg\forall x\, t \neq x$ $(= \exists x\, t = x)$, and so $X \nvdash \forall x\, t \neq x$ by (H1) and so $X \nvdash t \neq c$ for some c by (H2) and thus $X \vdash t = c$ by (H1).

3.2.5 Lemma

Let $X \subseteq \mathcal{L}$ be consistent. Then there exists a Henkin set $Y \supseteq X$ in a suitable constant expansion $\mathcal{L}C$ of $\mathcal{L}$.

Let $\mathcal{L}_0 := \mathcal{L}$ and $X_0 := X$, and assume that $\mathcal{L}_n$ and X_n have already be defined. Then let $\mathcal{L}_{n+1}$ be obtained by adding new constants $c_{x,\alpha,n}$ for all $x \in \mathit{Var}$ and $\alpha \in \mathcal{L}_n$, meaning $\mathcal{L}_{n+1} = \mathcal{L}_n C_n$ where C_n is the set of constants $c_{x,\alpha,n}$. Furthermore, let $X_{n+1} = X_n \cup \Gamma_{\mathcal{L}_n}$ (where $\Gamma_{\mathcal{L}_n}$ is defined as in the previous lemma), and so $X_{n+1} \subseteq \mathcal{L}_{n+1}$. Inductively we see that X_n is consistent for each n by the previous lemma. Let $X' = \bigcup_{n=0}^{\infty} X_n$ and $\mathcal{L}' = \bigcup_{n=0}^{\infty} \mathcal{L}_n = \mathcal{L}C$ where $C = \bigcup_{n=0}^{\infty} C_n$. Thus $X' \subseteq \mathcal{L}'$, and X' must be consistent: if $X' \vdash \bot$ then by the finiteness theorem we get that some $\varphi_1, \dots, \varphi_n \vdash \bot$. Since X' is the union of a chain of sets, we have that $\varphi_1, \dots, \varphi_n \in X_m$ for some m and thus $X_m \vdash \bot$ contradicting its consistency.

Let $\alpha \in \mathcal{L}'$ and $x \in \mathit{Var}$, then say $\alpha \in \mathcal{L}_n$ where n is minimal, and take α^x defined with respect to $\mathcal{L}^n$, and so $\neg\alpha^x \in X_{n+1} \subseteq X'$. Let us define

$$H = \{Y \subseteq \mathcal{L}' \mid X' \subseteq Y \text{ and } Y \text{ is consistent}\}$$

the set of all consistent extensions of X' in $\mathcal{L}'$. This is partially ordered by $\subseteq$. Let K be a chain in H, then $\bigcup K$ is an upper bound of the chain since the union of a chain of consistent sets is consistent (we showed this earlier in the proof). And $H \neq \varnothing$ as $X \in H$, so by Zorn's lemma H contains a maximal element Y; Y is a maximally consistent extension of X. Since $\neg\alpha^x \in X' \subseteq Y$, we have that $Y \vdash \neg\alpha^x$ for all $\alpha \in \mathcal{L}'$.

Y is also Henkin: for (H1), $Y \vdash \neg\alpha$ implies $Y \nvdash \alpha$ by Y's consistency. And if $Y \nvdash \alpha$ then $\alpha \notin Y$ and so $Y, \alpha \vdash \bot$ since Y is maximally consistent and so by $\mathtt{C}^-$ we get $Y \vdash \neg\alpha$. For (H2), $Y \vdash \forall x\alpha \implies Y \vdash \alpha\frac{c}{x}$ is true in general by ($\forall$1). Now, if $Y \vdash \alpha\frac{c}{x}$ for all c in $\mathcal{L}'$ and thus also for $c = c_{x,\alpha,n}$ where n is minimal such that $\alpha \in \mathcal{L}_n$. Assume $Y \nvdash \forall x\alpha$ then $Y \vdash \neg\forall x\alpha$ by (H1) and so $Y \vdash \neg\forall x\alpha, \alpha\frac{c}{x}$ implies $Y \vdash \neg\forall x\alpha \wedge \alpha\frac{c}{x} = \alpha^x$ by ($\wedge$1). But $Y \vdash \neg\alpha^x$ and so this contradicts Y's consistency. Therefore $Y \vdash \forall x\alpha$ as required.

Thus Y is indeed a Henkin extension of X. ■

This lemma is not true if we require Y be Henkin in $\mathcal{L}$. For example, let $\mathcal{L}$ consist of constants c_i for $i \in I$ where I is infinite and let $X = \{\boldsymbol{v}_0 \neq c_i \mid i \in I\}$. Then X is consistent, but in no consistent extension of X in $\mathcal{L}$ can $\boldsymbol{v}_0 = c_i$ be derived, contradicting (H3).

3.2.6 Lemma

Every Henkin set $Y \subseteq \mathcal{L}$ possess a model.

We will construct a model for Y called a *term model*. Let us define $t \approx t'$ whenever $Y \vdash t = t'$, this is a congruence relation on the term algebra $\mathcal{T}$ of $\mathcal{L}$. This means that $\approx$ is an equivalence relation and $\vec{t} \approx \vec{t'}$ implies $f\vec{t} \approx f\vec{t'}$ for all vectors of terms $\vec{t}, \vec{t'}$ and function symbols f of $\mathcal{L}$ (these were both proven explicitly in the previous subsection). Let us denote the equivalence class of t under $\approx$ by $\bar{t}$, and let us denote the quotient algebra of $\mathcal{T}$ by A; meaning $A = \{\bar{t} \mid t \in \mathcal{T}\}$. A will be the domain of the model $\mathcal{M}$ of Y.

Let C be the set of constants in $\mathcal{L}$, then by (H3) for each term $t \in \mathcal{T}$ there exists a $c \in C$ such that $c \approx t$. This means that $A = \{\bar{c} \mid c \in C\}$. For every variable x, let us define $x^{\mathcal{M}} = \bar{x}$ and for constants $c \in C$, define $c^{\mathcal{M}} = \bar{c}$. For a function symbol f in $\mathcal{L}$ we define

$$f^{\mathcal{M}}(\bar{t}_1, \ldots, \bar{t}_n) = \overline{ft_1 \cdots t_n}$$

This is well-defined since $\approx$ is a congruence: so if $t_i' \approx t_i$ then $ft_1' \cdots t_n' \approx ft_1 \cdots t_n$. And if r is a relation symbol, then define

$$r^{\mathcal{M}}\bar{t}_1 \cdots \bar{t}_n \iff Y \vdash rt_1 \cdots t_n$$

this too is well-defined since we showed that $Y \vdash \vec{t} = \vec{t'}$, $Y \vdash r\vec{t}$ implies $Y \vdash r\vec{t'}$.

We will now show that

$$(1)\ t^{\mathcal{M}} = \bar{t}, \qquad (2)\ \mathcal{M} \vDash \alpha \iff Y \vdash \alpha$$

(2) is certainly sufficient to show that $\mathcal{M}$ models Y: if $\alpha \in Y$ then $Y \vdash \alpha$ and so $\mathcal{M} \vDash \alpha$, meaning $\mathcal{M} \vDash Y$. We prove (1) by term induction. This is obvious by definition for prime terms. For compound terms

$$(ft_1 \cdots t_n)^{\mathcal{M}} = f^{\mathcal{M}} t_1^{\mathcal{M}} \cdots t_n^{\mathcal{M}} = f^{\mathcal{M}} \bar{t}_1 \cdots \bar{t}_n = \overline{ft_1 \cdots t_n}$$

by definition, as required. We prove (2) by formula induction.

$$\begin{array}{llll}
\mathcal{M} \vDash t = s & \iff t^{\mathcal{M}} = s^{\mathcal{M}} & \iff \bar{t} = \bar{s} & (\text{since } t^{\mathcal{M}} = \bar{t}) \\
 & & \iff Y \vDash t = s & (\text{definition}) \\
\mathcal{M} \vDash r\vec{t} & \iff r^{\mathcal{M}}\vec{t}^{\mathcal{M}} & \iff r^{\mathcal{M}}\bar{\vec{t}} & \iff Y \vdash r\vec{t} \\
\mathcal{M} \vDash \alpha \wedge \beta & \iff \mathcal{M} \vDash \alpha, \beta & \iff Y \vdash \alpha, \beta & (\text{induction hypothesis}) \\
 & & \iff Y \vdash \alpha \wedge \beta & (\text{by } (\wedge 1) \text{ and } (\wedge 2)) \\
\mathcal{M} \vDash \neg\alpha & \iff \mathcal{M} \nvDash \alpha & \iff Y \nvdash \alpha & (\text{induction hypothesis}) \\
 & & \iff Y \vdash \neg\alpha & (\text{by (H1)}) \\
\mathcal{M} \vDash \forall x\alpha & \iff \mathcal{M}_x^{\bar{c}} \vDash \alpha \text{ for all } c \in C & & (\text{since } A = \{\bar{c} \mid c \in C\}) \\
 & \iff \mathcal{M}_x^{c^{\mathcal{M}}} \vDash \alpha \text{ for all } c \in C & & (\text{since } c^{\mathcal{M}} = \bar{c}) \\
 & \iff \mathcal{M} \vDash \alpha\frac{c}{x} \text{ for all } c \in C & & (\text{by The Substitution Theorem}) \\
 & \iff Y \vdash \alpha\frac{c}{x} \text{ for all } c \in C & & (\text{induction hypothesis}) \\
 & \iff Y \vdash \forall x\alpha & & (\text{by (H2)})
\end{array}$$

So we have shown (2), and as explained (2) is sufficient to show the premise of the lemma. ∎

3.2.7 Theorem (The Model Existence Theorem)

Every consistent set $X \subseteq \mathcal{L}$ has a model.

Let $Y \supseteq X$ be a Henkin expansion of X in a suitable constant expansion $\mathcal{L}C$ of $\mathcal{L}$. By the above lemma, Y has a model $\mathcal{M}'$ in $\mathcal{L}C$. Let $\mathcal{M}$ be the $\mathcal{L}$-reduct of $\mathcal{M}'$. Then since $\mathcal{M}' \vDash X$, by The Coincidence Theorem we get $\mathcal{M} \vDash X$. ∎

3.2.8 Theorem (The Completeness Theorem)

Let $\mathcal{L}$ be a first-order language, and $X \subseteq \mathcal{L}$ and $\alpha \in \mathcal{L}$. Then $X \vdash \alpha \iff X \vDash \alpha$. More suggestively, $\vdash = \vDash$.

The soundness of $\vdash$ states $X \vdash \alpha \implies X \vDash \alpha$. If $X \nvdash \alpha$ then $X, \neg\alpha$ is consistent and therefore $X \cup \{\neg\alpha\}$ has a model and therefore $X \nvDash \alpha$. ∎

So we can freely alternate between using $\vdash$ and $\vDash$, and we will often prove $X \vdash \alpha$ by showing $X \vDash \alpha$ in a semi-formal manner as is common in mathematics. In particular for theories T, $T \vDash \alpha$ is equivalent to $T \vdash \alpha$ and we will often write $\vdash_T \alpha$ in place of it. For sentences α, $\vdash_T \alpha$ means the same as $\alpha \in T$. More generally, we write $X \vdash_T \alpha$ to mean $X \cup T \vdash \alpha$. Notice that

$$\alpha \vdash_T \beta \iff \vdash_T \alpha \to \beta \iff \vdash_{T+\alpha} \beta, \qquad \vdash_T \alpha \iff \vdash_T \alpha^g$$

are all immediate due to $\vdash = \vDash$.

Exercise

Let $K \neq \varnothing$ be a chain of theories in $\mathcal{L}$. Show that $T_K = \bigcup K$ is a theory that is consistent if and only if every $T \in K$ is consistent.

If every $T \in K$ is consistent, then suppose $T_K \vdash \bot$ then by finiteness, we get $\varphi_1, \dots, \varphi_n \vdash \bot$ for $\varphi_i \in T_K$. Since T_K is the union of a chain, $\varphi_1, \dots, \varphi_n \in T$ for some $T \in K$ and thus $T \vdash \bot$, contradicting T's consistency. If T_K is consistent, since $T \subseteq T_K$ for every $T \in K$, T is consistent as well.

Exercise

Suppose T is consistent and $Y \subseteq \mathcal{L}$. Prove that the following are equivalent:

(1) $Y \vdash_T \bot$,

(2) $\vdash_T \neg\alpha$ for some conjunction α of formulas in Y.

$Y \vdash_T \bot$ if and only if $Y \cup T \vdash \bot$ and since T is consistent and by finiteness, this is equivalent to $T, \alpha_1, \dots, \alpha_n \vdash \bot$. By $\mathtt{C}^-$, this is if and only if $T, \alpha_1, \dots, \alpha_{n-1} \vdash \neg\alpha_n$, which is equivalent to

$$\vdash_T \alpha_1 \to \cdots \to \alpha_{n-1} \to \neg\alpha_n \equiv \alpha_1 \wedge \cdots \wedge \alpha_{n-1} \to \neg\alpha_n \equiv \neg(\alpha_1 \wedge \cdots \wedge \alpha_n)$$

so setting $\alpha = \alpha_1 \wedge \cdots \wedge \alpha_n$ gives the desired result. This chain is obviously reversible.

Exercise

Let $x \notin vart$ and $\alpha, \frac{t}{x}$ collision-free. Verify the following equivalence chain:

$$(1)\ \vdash_T \alpha\frac{t}{x} \iff (2)\ x = t \vdash_T \alpha \iff (3)\ \vdash_T x = t \to \alpha$$
$$\iff (4)\ \vdash_T \forall x(x = t \to \alpha) \iff (5)\ \vdash_T \exists x(x = t \wedge \alpha)$$

If $T \vdash \alpha\frac{t}{x}$ then $T, x = t \vdash x = t, \alpha\frac{t}{x}$ so by $(=)$ (but which is true, due to the completeness theorem for all formulas α where $\alpha, \frac{t}{x}$ is collision-free) we get $T, x = t \vdash \alpha$ as required. (2) $\implies$ (3) is due to the deduction theorem. As explained above (it is due to the completeness theorem), $\vdash_T \alpha \iff \vdash_T \alpha^g$ and in particular $\vdash_T \alpha \iff \vdash_T \forall x\alpha$, which proves (3) $\implies$ (4). (4) $\implies$ (5) is due to $(\forall x = t)\alpha \vDash \exists(x = t)\alpha$ which we showed in a previous exercise. For (5) $\implies$ (1): suppose $\mathcal{M} \vDash \exists x(x = t \wedge \alpha)$ then $\mathcal{M}_x^a \vDash x = t \wedge \alpha$ for some $a \in A$ and thus $t^{\mathcal{M}} = a$ and $\mathcal{M}_x^t \vDash \alpha$ thus $\mathcal{M} \vDash \alpha\frac{t}{x}$ by the substitution theorem, so $\exists x(x = t \wedge \alpha) \vDash \alpha\frac{t}{x}$ and this is sufficient by the completeness theorem.

3.3 First Applications: Nonstandard Models

By the completeness theorem, $\vdash = \vDash$ and so immediately we get

3.3.1 Theorem (The Finiteness Theorem)

$X \vDash \alpha$ implies $X_0 \vDash \alpha$ for a finite $X_0 \subseteq X$.

Let us define the sentence $\texttt{char}_p = 1 + \cdots + 1 = 0$ (1 is added p times) in the language of rings $\mathcal{L}\{+, \cdot, 0, 1\}$. A *field of characteristic* p is a field F such that $F \vDash \texttt{char}_p$ and $F \nvDash \texttt{char}_q$ for $q < p$ (it is simple to show that if a field is of characteristic p, then p is prime). We define the first-order theory of fields of characteristic 0 to be the theory axiomatized by X: the set of axioms for the theory of fields, as well as $\neg\texttt{char}_p$ for all prime p.

3.3.2 Proposition

A sentence α which is valid in all fields of characteristic 0 is also valid in all fields of sufficiently high prime characteristic p (p is dependent on α).

A sentence α is valid in all fields of characteristic 0 if and only if $X \vDash \alpha$. By the finiteness theorem this means $X_0 \vDash \alpha$ for some finite $X_0 \subseteq X$. Since X_0 is finite, this means it can only contain finitely many sentences of the form $\neg\texttt{char}_q$, thus there exists some p such that for all $\neg\texttt{char}_q \in X_0$, $q < p$. Thus for all fields F of characteristic $\geq p$, we have $F \vDash X_0$ and so $F \vDash \alpha$. Meaning α is valid in all fields whose characteristic is at least p, as required. ■

3.3.3 Definition

A set of strings Z is **decidable** if there is an algorithm (this will be formally defined later) which after finitely many steps provides an answer to whether a string of symbols ξ belongs to Z. Otherwise Z is **undecidable**. For example, it is intuitively clear that it is decidable whether ξ is a formula (or a sentence). A theory T is **recursively axiomatizable** (for short **axiomatizable**) if it possess a decidable axiom system. For example, T is recursively axiomatizable if it is finitely axiomatizable: if it has a finite axiom system.

So by the above proposition, the theory of fields of characteristic zero is not finitely axiomatizable. For if F were a finite set of axioms, then their conjunction $\alpha = \bigwedge F$ would be valid in fields of sufficiently large characteristic. But then these fields would be of characteristic zero, in contradiction.

Here is another example: an abelian group $\mathcal{G}$ is *n-divisible* if $G \vDash \vartheta_n$ where $\vartheta_n = \forall x \exists y\, x = ny$ where $ny = y + \cdots + y$ (n times). And $\mathcal{G}$ is *divisible* if it is n-divisible for all $n \geq 1$. The theory of divisible abelian groups is denoted DAG and is axiomatized by the axioms of abelian groups and the sentences ϑ_n for $n \geq 1$.

3.3.4 Proposition

A sentence $\alpha \in \mathcal{L}\{+, 0\}$ valid in all divisible abelian groups is also valid in at least one nondivisible abelian group.

If $\mathsf{DAG} \vDash \alpha$ then $X_0 \vDash \alpha$ for a finite subset of the axioms X_0. In particular there exists some m such that for every $\vartheta_n \in X_0$, $n < m$, let p be a prime larger than m. Now let $\mathbb{Z}_p$ be the cyclic group of order p, then $\mathbb{Z}_p \vDash \vartheta_n$ for $0 < n < p$ (since n is coprime with p, and thus is divisible in $\mathbb{Z}_p$ thus take $y = n^{-1}x$). Thus $\mathbb{Z}_p \vDash X_0$, but $\mathbb{Z}_p$ is not divisible, as $\mathbb{Z}_p \nvDash \vartheta_p$ (since $py = 0$). ■

So DAG is not finitely axiomatizable, for the same reason that the theory of fields of characteristic zero is not finitely axiomatizable.

3.3.5 Theorem (The Compactness Theorem)

A set of formulas X is satisfiable if and only if every finite subset $X_0 \subseteq X$ is satisfiable.

If X is satisfiable, then so is every $X_0 \subseteq X$. If X is not satisfiable, then by the model existence theorem this is equivalent to X being inconsistent so $X \vdash \bot$ and thus $X_0 \vdash \bot$ meaning $X_0 \vDash \bot$ for some finite subset $X_0 \subseteq X$. ■

3.3.6 Definition

An $\mathcal{L}$-theory T is **complete** if it is consistent and has no consistent proper extension in $\mathcal{L}^0$.

If T is complete and if $\nvdash_T \alpha, \neg\alpha$ then $T + \alpha$ and $T + \neg\alpha$ are consistent by $\mathtt{C}^+$ and $\mathtt{C}^-$ and are therefore consistent proper extensions in contradiction. So if T is complete then $\vdash_T \alpha$ or $\vdash_T \neg\alpha$ (but not both, for then T would not be consistent) for every $\alpha \in \mathcal{L}^0$. Conversely if this is true then T is consistent as if $\vdash_T \perp$ then $\vdash_T \alpha, \neg\alpha$ contradicting the "not both" part. And it is maximally consistent as then if $\alpha \notin T$, $\neg\alpha \in T$ and so $T + \alpha \vdash \alpha, \neg\alpha$ and is therefore inconsistent. So T is complete if and only if $\vdash_T \alpha$ or $\vdash_T \neg\alpha$ but not both, or equivalently $\vdash_T \alpha \iff \nvdash_T \neg\alpha$ for all sentences α.

So for example, if $\mathcal{A}$ is an $\mathcal{L}$-structure, then since for every sentence α, either $\mathcal{A} \vDash \alpha$ or $\mathcal{A} \nvDash \alpha$, which is equivalent to $\mathcal{A} \vDash \neg\alpha$ (since α this is true for specific models over $\mathcal{A}$, and since α is a sentence, all the models agree on the validity of α). Thus for every sentence α, $\mathcal{A} \vDash \alpha$ or $\mathcal{A} \nvDash \neg\alpha$, and it obviously cannot satisfy both. Thus $Th\mathcal{A}$ is a complete theory for every structure $\mathcal{A}$.

Now, a frequently occurring theory is the theory $Th\mathcal{N}$, which is the theory of the structure $\mathcal{N} = (\mathbb{N}, 0, \mathtt{S}, +, \cdot)$ where $\mathtt{S}$ is the *successor function*: $\mathtt{S}\colon n \mapsto n + 1$. $\mathcal{N}$ is the standard structure of the arithmetical language $\mathcal{L}_{ar} := \mathcal{L}\{0, \mathtt{S}, +, \cdot\}$. We can define the relationship $\leq$ in $\mathcal{L}_{ar}$ by $x \leq y \leftrightarrow \exists z\, x + z = y$ and the relationship $<$ by $x < y \leftrightarrow x \leq y \wedge x \neq y$.

Even more frequent than $Th\mathcal{N}$ is one of its subtheories: *Peano arithmetic* denoted PA. It is axiomatized by

$$\forall x\, \mathtt{S}x \neq 0 \qquad \forall x\, x + 0 = x \qquad \forall x\, x \cdot x = 0$$
$$\forall x \forall y\, (\mathtt{S}x = \mathtt{S}y \to x = y) \qquad \forall x \forall y\, x + \mathtt{S}y = \mathtt{S}(x + y) \qquad \forall x \forall y\, x \cdot \mathtt{S}y = x \cdot y + x$$
$$\text{IS: } \varphi\tfrac{0}{x} \wedge \forall x\big(\varphi \to \varphi\tfrac{\mathtt{S}x}{x}\big) \to \forall x\varphi$$

IS is the *induction schema* and is a *schema* of rules: it runs over all formulas φ in $\mathcal{L}_{ar}$. Since φ may not be a sentence, and axioms are, by our convention IS must be generalized, so the correct axiom is $\big(\varphi\frac{0}{x} \wedge \forall x\big(\varphi \to \varphi\frac{\mathtt{S}x}{x}\big) \to \forall x\varphi\big)^g$. The purpose of IS is to prove $\forall x\varphi$ *by induction on* x: first one proves φ for when $x = 0$ (hence showing $\varphi\frac{0}{x}$), and then one shows that if φ then $\varphi\frac{\mathtt{S}}{x}$. Proving $\vdash_{\mathsf{PA}} \forall x\big(\varphi \to \varphi\frac{\mathtt{S}x}{x}\big)$ is equivalent to showing $\varphi \vdash_{\mathsf{PA}} \varphi\frac{\mathtt{S}x}{x}$.

For example, let $\varphi = x = 0 \vee \exists v\, \mathtt{S}v = x$. We will prove $\vdash_{\mathsf{PA}} \forall x\varphi$, meaning every $x \neq 0$ has a predecessor. Obviously $\varphi\frac{0}{x}$ $(= 0 = 0 \vee \exists v\, \mathtt{S}v = x)$. Now we must show that $\varphi \vdash_{\mathsf{PA}} \varphi\frac{\mathtt{S}x}{x}$. In general we have $\mathtt{S}v = x \vdash_{\mathsf{PA}} \mathtt{SS}v = \mathtt{S}x$, and so by particularization we get $\exists v\, \mathtt{S}v = x \vdash_{\mathsf{PA}} \exists v\, \mathtt{SS}v = \mathtt{S}x$. And since $x = 0 \vdash_{\mathsf{PA}} \exists v\, \mathtt{S}v = \mathtt{S}x$, we get $\varphi \vdash_{\mathsf{PA}} \exists v\, \mathtt{S}v = \mathtt{S}x \vdash_{\mathsf{PA}} \mathtt{S}x = 0 \vee \exists v\, \mathtt{S}v = \mathtt{S}x = \varphi\frac{\mathtt{S}x}{x}$ as required.

We will now show that $Th\mathcal{N}$ and PA have what are called *nonstandard models*, models which are not isomorphic to the standard model of the theory, which is $\mathcal{N}$ in this case. Let us define $\underline{n} := \mathtt{S}^n 0 := \mathtt{S} \cdots \mathtt{S}0$ where $\mathtt{S}$ is composed with itself n times. So for example $\underline{1} = \mathtt{S}0$, $\underline{2} = \mathtt{S}1$ and in general $\underline{\mathtt{S}n} = \mathtt{S}\underline{n}$. Let $x \in \mathit{Var}$ and we define $X := Th\mathcal{N} \cup \{\underline{n} < x \mid n \in \mathbb{N}\}$.

Let $X_0 \subseteq X$ be finite, then X_0 is satisfiable: since X_0 is finite, there must be some m such that $X_0 \subseteq Th\mathcal{N} \cup \{\underline{n} < x \mid n < m\}$. Thus, we can simply assign x the value of m and then $(\mathcal{N}, m)$ (meaning the valuation function maps x to m) is a model of X_0. So by The Compactness Theorem, X has a model $(\mathcal{N}', c)$ where the domain of the model is $\mathbb{N}'$ and $c \in \mathbb{N}'$ is the valuation of x. Since $Th\mathcal{N}$ is a subset of X and is complete, $\mathbb{N}'$ satisfies precisely the same sentences as $\mathcal{N}$. In particular the following sentences are valid in $\mathcal{N}'$: $\mathtt{S}\underline{n} = \underline{\mathtt{S}n}$, $\underline{n+m} = \underline{n} + \underline{m}$, and $\underline{n \cdot m} = \underline{n} \cdot \underline{m}$. Thus $n \mapsto \underline{n}^{\mathcal{N}'}$ is an embedding of $\mathcal{N}$ into $\mathcal{N}'$. We can identify the image of this embedding with $\mathcal{N}$, meaning it is legitimate to assume $\underline{n}^{\mathcal{N}'} = n$ so $\mathcal{N} \subseteq \mathcal{N}'$.

Since $\mathcal{N}' \vDash X$, $\mathcal{N}$ and $\mathcal{N}'$ are elementarily equivalent as $Th\mathcal{N}$ is complete. But on the other hand, $n < a$ for any $a \in \mathbb{N}' \setminus \mathbb{N}$, since in $\mathcal{N}$ and thus in $\mathcal{N}'$ we have $(\forall x \leq \underline{n}) \bigvee_{i \leq n} x = \underline{i}$. So if $a \leq n$ then we'd have $a = \underline{i}$ and so $a \in \mathbb{N}$ in contradiction. And since $\mathbb{N}' \setminus \mathbb{N}$ is nonempty by the formulas added to X, $\mathbb{N}$ is a proper initial segment of $\mathbb{N}'$. The elements of $\mathbb{N}' \setminus \mathbb{N}$ are called *nonstandard numbers* (the existence of nonstandard numbers are precisely what make $\mathbb{N}'$ and $\mathbb{N}$ non-isomorphic). c is an example of a nonstandard number, as well as $c + c$ and so on.

Since $(\forall x \neq 0)\exists! y\, \mathtt{S}y = x$ is a theorem of $\mathcal{N}$ it too must be valid in $\mathcal{N}'$ and so c must have an immediate successor. But if we were to chase the successors of c, we would never find one which is in $\mathbb{N}$ (ie. there is no natural x and n such that $\mathtt{S}^n x = c$ as then c would be natural).

In no nonstandard model $\mathcal{N}'$ of $Th\mathcal{N}$ is $\mathbb{N}$ definable (for the same reason as above, $\mathbb{N}$ can be embedded in all models of $Th\mathcal{N}$, so we really mean that the image of this embedding is not definable). In fact it is not even *parameter definable*, meaning there is no $\alpha = \alpha(x, \vec{y})$ and $b_1, \ldots, b_n \in \mathbb{N}'$ such that $\mathbb{N} = \Big\{a \in \mathbb{N}' \,\Big|\, \mathcal{N}' \vDash \alpha[a, \vec{b}]\Big\}$.

Otherwise, we'd have $\mathcal{N}' \vDash \alpha\frac{0}{x}[\vec{b}]$ since $0 \in \mathbb{N}$ and $\mathcal{N}' \vDash \forall x\big(\alpha \to \alpha\frac{\mathtt{S}x}{x}\big)[\vec{b}]$ since $\mathbb{N}$ is closed under $\mathtt{S}$. Thus by IS we have $\mathcal{N}' \vDash \forall x\alpha[\vec{b}]$, which contradicts $\mathbb{N}' \setminus \mathbb{N}$ being nonempty.

Similar to $\mathcal{N}$ we can find nonstandard models for the theory of real numbers, $\mathcal{R} = (\mathbb{R}, +, \cdot, <, \{\boldsymbol{a} \mid a \in \mathbb{R}\})$ which contains a constant symbol $\boldsymbol{a}$ for every number $a \in \mathbb{R}$. Consider $X = Th\mathcal{R} \cup \{\boldsymbol{a} < x \mid a \in \mathbb{R}\}$, which is finitely satisfiable and therefore by The Compactness Theorem X has a model $\mathcal{R}^*$ which is a *nonstandard model of analysis*. $\mathcal{R}^*$ models $Th\mathcal{R}$ and therefore every theorem valid in $\mathcal{R}$ is valid in $\mathcal{R}^*$ and vice versa. We could also add functions like $\exp, \ln, \sin, \cos$ to the signature of $\mathcal{R}$ and obtain a model $\mathcal{R}^*$ where the properties of these functions which can be formulated in first-order language are preserved, like

$$\forall x, y \; \exp(x+y) = \exp x \cdot \exp y, \qquad (\forall x > 0) \exp \ln x = x, \qquad \forall x \; \sin^2 x + \cos^2 x = 1$$

Since continuity and differentiability can be formulated in first-order language (via ε-δ shenanigans), these functions remain continuous and repeatedly differentiable. Though topological properties like Bolzano-Weirestrass are not generally true, but can be replaced by infinitesimal arguments.

For a nonstandard model $\mathcal{R}^*$ of $Th\mathcal{R}$, with $\mathcal{R} \subseteq \mathcal{R}^*$ contains infinitely large numbers (meaning numbers c where $r < c$ for all $r \in \mathbb{R}$), but also infinitely small numbers (meaning numbers ε where $0 < \varepsilon < r$ for all $r \in \mathbb{R}$, as we can take $\varepsilon := \frac{1}{c}$). These infinitely small numbers are termed *infinitesimal numbers*, and utilizing their properties one can study *nonstandard analysis* which is a branch of analysis utilizing these nonstandard models of analysis and their infinitesimal numbers. This branch gives arguably more intuitive proofs for analytic results, which can more closely follow the arguments of the original analytical researchers like Leibniz.

While every bounded set in $\mathcal{R}$ possesses a supremum, this is not true in $\mathcal{R}^*$. Rather we can only say that every bounded definable set in $\mathcal{R}^*$ possesses a supremum, as this is first-order formulatable,

$$\exists x\varphi \land \exists y \forall x(\varphi \to x \leq y) \to \exists z \forall x\big((\varphi \to x \leq z) \land \forall y((\varphi \to x \leq y) \to z \leq y)\big)$$

(the set is defined by φ, ie. it is the set $\{x \mid \mathcal{R}^* \vDash \varphi\}$.)

Exercise

Prove in PA the associativity, commutativity, and distributivity of $+$ and $\cdot$.

We will first derive $\mathtt{S}x + y = x + \mathtt{S}y$ by induction on y, ie. we define $\varphi = \mathtt{S}x + y = x + \mathtt{S}y$. The base case $\varphi\frac{0}{y} = \mathtt{S}x + 0 = x + \mathtt{S}0$: since $\mathtt{S}x + 0 = \mathtt{S}x$ and $x + \mathtt{S}0 = \mathtt{S}(x+0) = \mathtt{S}x$, this is true. Now we show $\varphi \vdash \varphi\frac{\mathtt{S}y}{y} = \mathtt{S}x + \mathtt{S}y = x + \mathtt{SS}y$. Now, $\mathtt{S}x + \mathtt{S}y = \mathtt{S}(\mathtt{S}x + y) = \mathtt{S}(x + \mathtt{S}y) = x + \mathtt{SS}y$ as required. So we have shown $\forall y\varphi$ as required.

Now, we will also show that $0 + y = y$ by induction on y: let $\varphi = 0 + y = y$. Then $\varphi\frac{0}{y} = 0 + 0 = 0$ which is valid in PA. Now we must show $0 + y = y \vdash 0 + \mathtt{S}y = \mathtt{S}y$. Since $0 + \mathtt{S}y = \mathtt{S}(0 + y) = \mathtt{S}y$ we have finished.

Now we will show associativity, by inducting $\varphi = (x+y)+z = x+(y+z)$ on y. $\varphi\frac{0}{y} = (x+0)+z = x+(0+z)$ and since $x+0 = x$ and $0+z = z$ this is true. For the inductive step, $(x+\mathtt{S}y)+z = \mathtt{S}(x+y)+z = \mathtt{S}((x+y)+z) = \mathtt{S}(x+(y+z)) = x+\mathtt{S}(y+z) = x+(y+\mathtt{S}z) = x+(\mathtt{S}y+z)$ as required.

For commutativity, we induct $x+y = y+x$ on y (we could also on x, but why change?). This is obviously true when substituting 0 for y, as $x+0 = x = x+0$. Now $x + \mathtt{S}y = \mathtt{S}(x+y) = \mathtt{S}(y+x) = \mathtt{S}y + x$, as required. Proofs for $\cdot$ and distributivity are similar.

Exercise

Prove the antisymmetry of $\leq$ in PA.

So we must show that $x \leq y$ and $y \leq x$ implies $x = y$. Now,

$$x \leq y \leftrightarrow \exists z \, x + z = y, \qquad y \leq x \leftrightarrow \exists u \, y + u = x$$

This means that $x \leq y \land y \leq x \to \exists u, z \, y + u + z = y$. Now simple induction on x will yield $y + x = y \to x = 0$, and so $u + z = 0$. And induction on z will yield $u + z = 0 \to z = 0, u = 0$ (the base case is trivial, the inductive step holds vacuously). And so $u + z = 0$ and thus $y = x$.

Exercise

Prove $x < y \equiv_{\mathsf{PA}} \mathsf{S}x \leq y$. Use this to prove $\vdash_{\mathsf{PA}} x \leq y \vee y \leq x$ by induction on x.

We will first prove $x < y \vdash_{\mathsf{PA}} \mathsf{S}x \leq y$ By definition if $x < y$ then $\exists z\, x + z = y$ and since $x \neq y$ we have that $z \neq 0$ and so z has a predecessor, u. Thus $x < y$ implies $\exists u\, x + \mathsf{S}u = y$ and since $x + \mathsf{S}u = \mathsf{S}x + u$ this means $\mathsf{S} \leq y$ as required. Conversely, since $x < \mathsf{S}x$ by induction on x, we have that if $\mathsf{S}x \leq y$ by transitivity $x < y$.

Let $\varphi = x \leq y \vee y \leq x$, then substituting 0 for x, we get $0 \leq y \vee y \leq 0$ and since we know that $\forall y\, 0 \leq y$ (by induction on y), this is true. Now for the inductive step we must show $\mathsf{S}x \leq y \vee y \leq \mathsf{S}x \equiv x < y \vee y \leq \mathsf{S}x$. We know that $x \leq y \vee y \leq x$, so either $x < y$ in which case this is true, or $y \leq x$ in which case $y \leq \mathsf{S}x$.

Exercise

Let $\alpha, \beta, \gamma \in \mathcal{L}_{ar}$ and $y \notin \mathit{var}\{\alpha, \beta\}$ and $z \notin \mathit{var}\gamma$. Verify the following:

(1) $\vdash_{\mathsf{PA}} \forall x\big((\forall y < x)\alpha\frac{y}{x} \to \alpha\big) \to \forall x\alpha$, the *schema of <-induction* (or *strong induction*).

(2) $\vdash_{\mathsf{PA}} \exists x\beta \to \exists x\big(\beta \wedge (\forall y < x)\neg\beta\frac{y}{x}\big)$, the *minimum schema* (or *well-ordering principle*).

(3) $\vdash_{\mathsf{PA}} (\forall x < v)\exists y\gamma \to \exists z(\forall x < v)(\exists y < z)\gamma$, the *schema of collection*.

(1) Let $\varphi = (\forall y < x)\alpha\frac{y}{x}$, we will prove $\forall x(\varphi \to \alpha) \vdash_{\mathsf{PA}} \varphi\frac{0}{x}, \varphi \to \varphi\frac{\mathsf{S}x}{x}$ which proves $\forall x\varphi$ by IS. This means that in particular $\forall x\varphi\frac{\mathsf{S}x}{x}$, and so for every $y < \mathsf{S}x$, $\alpha\frac{y}{x}$ and since $x < \mathsf{S}x$ we get α as required. Since $\varphi\frac{0}{x} = (\forall y < 0)\alpha\frac{y}{x}$ which holds vacuously. And since $\forall x(\varphi \to \alpha) = \forall x\big((\forall y < x)\alpha\frac{y}{x} \to \alpha\big)$, we have that if $\varphi = (\forall y < x)\alpha\frac{y}{x}$ then α, meaning $(\forall y \leq x)\alpha\frac{y}{x}$, which is equivalent to $\varphi\frac{\mathsf{S}x}{x}$. Thus $\forall x(\varphi \to \alpha) \vdash_{\mathsf{PA}} \varphi \to \varphi\frac{\mathsf{S}x}{x}$, as required.

(2) Note that the contrapositive of this is $\forall x\big((\forall y < x)\neg\beta\frac{y}{x} \to \neg\beta\big) \to \forall x(\neg\beta)$, which follows from (1).

(3) Let $\varphi = (\forall x < v)\exists y\gamma \to \exists z(\forall x < v)(\exists y < z)\gamma$, we will prove $\forall v\varphi$ by induction on v. $\vdash_{\mathsf{PA}} \varphi\frac{0}{v}$ holds vacuously. Now we must prove $\varphi \vdash_{\mathsf{PA}} \varphi\frac{\mathsf{S}v}{v} \equiv_{\mathsf{PA}} (\forall x \leq v)\exists y\gamma \to \exists z(\forall x \leq v)(\exists y < z)\gamma$. By φ, if $(\forall x \leq v)\exists y\gamma$, then we have $\exists z(\forall x < v)(\exists y < z)\gamma$. So the only "issue" is when $x = v$, but in which case we know there exists a y_x such that γ, so let $z' = \max\{z, \mathsf{S}y_x\}$ and we get that for every $x < v$ there exists a $y < z \leq z'$ such that γ and for $x = v$ there exists a $y_x < \mathsf{S}y_x \leq z'$ such that γ. And so we get $\exists z(\forall x < v)(\exists y < z)\gamma$ (where the z is z') as required.

3.4 ZFC and Skolem's Paradox

We briefly review fundamentals of set theory. Recall that a set M is *countable* if there exists a surjection $f\colon \mathbb{N} \longrightarrow M$ (assuming the axiom of choice, this is equivalent to there being an injection $M \longrightarrow \mathbb{N}$). A set which is not countable is called *uncountable*. Two sets M and N are *equipotent* if there exists a bijection between M and N, we denote this as $M \sim N$. If $M \sim \mathbb{N}$, M is *countably infinite*. Recall that $\mathbb{R}$ is uncountable and equipotent to $\mathcal{P}(\mathbb{N})$. In general due to Cantor's theorem which states that the powerset of a set, $\mathcal{P}(M)$, is of a higher cardinality than M (meaning there is an injection $M \longrightarrow \mathcal{P}(M)$ but no surjection).

And if M and N are countable, than so is $M \times N$. And if $\{M_i\}_{i \in I}$ is a countable family of countable sets (meaning I is countable and M_i is countable for every $i \in I$), then their union $\bigcup_{i \in I} M_i$ is also countable. Notice then that if M is countable, the set of all finite sequences of symbols in M is $\bigcup_{n \in \mathbb{N}} M^n$, which is countable. Thus if $\mathcal{L}$ is a first-order language over a countable signature then $\mathcal{L}$ is countable as well.

A *countable theory* is a theory formalized in a countable first-order language.

3.4.1 Theorem (Löwenheim-Skolem Theorem)

A countable consistent theory T has a countable model.

By The Model Existence Theorem, T has a model whose domain A consists of the equivalence classes $\bar{c}$ for $c \in C$, where $C = \bigcup_{n \in \mathbb{N}} C_n$ is a new set of constants. By definition, C_0 is equipotent to $\mathcal{L} \times \mathit{Var}$ which is countable, and so $\mathcal{L}_1 := \mathcal{L}C_0$ is equipotent. And in general, C_n is equipotent to $\mathcal{L}_n \times \mathit{Var}$ so C_n and $\mathcal{L}_{n+1}$ are countable for every n, by induction. Thus C is also countable as the countable union of countable sets. The map $c \mapsto \bar{c}$ is a surjection from C to A, and therefore A is also countable, as required. ∎

This theorem will later be generalized, but this theorem already leads to interesting consequences. For example, there must be countable ordered fields which are nonstandard models of the first-order theory $Th(\mathbb{R}, 0, 1, +, \cdot, <, \exp, \sin, \ldots)$. Even more surprising (for reasons explained shortly) is that this theorem implies the existence of countable models of set theory.

To explain why, we must first define set theory. The most common and thus most important formalization of set theory is ZFC (short for Zermelo-Fraenkel+Choice, after the two mathematicians which first formalized it as well as the axiom of choice AC. Excluding AC from the axioms of the theory, gives rise to a proper (this is not a trivial fact) subtheory, ZF.).

The language upon which ZFC is defined is $\mathcal{L}\{\boldsymbol{\in}\}$, it contains only the extralogical symbols $\boldsymbol{\in}$ (and of course $=$). We denote this set-theoretic language by $\mathcal{L}_{\boldsymbol{\in}}$. We use a boldface $\boldsymbol{\in}$ symbol in order to differentiate it from the metalogical $\in$ symbol, much like $=$. Importantly, in ZFC all objects are sets, there are no other types of objects like in other formulations of set theory.

We define the relation $\subseteq$ by $x \subseteq y \leftrightarrow \forall z(z \boldsymbol{\in} x \to z \boldsymbol{\in} y)$.

We begin to list the axioms of ZFC (recall that these axioms must be generalized):

$$\begin{aligned} &\mathsf{AE}: \quad \forall z(x \boldsymbol{\in} x \leftrightarrow z \boldsymbol{\in} y) \to x = y \qquad &&\text{(axiom of extensionality)} \\ &\mathsf{AS}: \quad \exists y \forall z(z \boldsymbol{\in} y \leftrightarrow \varphi \wedge z \boldsymbol{\in} x) \qquad &&\text{(axiom of separation)} \end{aligned}$$

AS is an axiom schema: φ runs over all $\mathcal{L}_{\boldsymbol{\in}}$-formulas with $y \notin \mathit{free}\varphi$. Now, let $\varphi = \varphi(x, z, \vec{a})$, then we can derive $\forall x \exists! y \forall z(z \boldsymbol{\in} y \leftrightarrow \varphi \wedge z \boldsymbol{\in} x)$, meaning that $y = \{z \boldsymbol{\in} x \mid \varphi\} \leftrightarrow \forall z(z \boldsymbol{\in} y \leftrightarrow \varphi \wedge z \boldsymbol{\in} x)$ is a legitimate definition of the *set term* $\{z \boldsymbol{\in} x \mid \varphi\}$ (the set of all $z \in x$ such that $\varphi(x, z, \vec{a})$ holds). The set term $\{z \boldsymbol{\in} x \mid \varphi\}$ is simply a suggestive way of writing a function term $f_{\vec{a}}x$ (which depends on the "parameter" term $\vec{a}$). To derive the required formula for legitimacy, let $y, y' \notin \mathit{free}\varphi$ then it is obvious to see

$$(x \boldsymbol{\in} y \leftrightarrow \varphi \wedge z \boldsymbol{\in} x) \wedge (z \boldsymbol{\in} y' \leftrightarrow \varphi \wedge z \boldsymbol{\in} x) \to (z \boldsymbol{\in} y \leftrightarrow z \boldsymbol{\in} y')$$

This implies that $\forall z(x \boldsymbol{\in} y \leftrightarrow \varphi \wedge z \boldsymbol{\in} x) \wedge \forall z(z \boldsymbol{\in} y' \leftrightarrow \varphi \wedge z \boldsymbol{\in} x) \to y = y'$, and so we have proven the claim of legitimacy.

We can explicitly define the empty set by $y = \varnothing \leftrightarrow \forall z\, z \boldsymbol{\notin} y$. But we must show that this is legitimate. According to AS, $\exists y \forall z(z \boldsymbol{\in} y \leftrightarrow z \boldsymbol{\notin} x \wedge z \boldsymbol{\in} x)$. But since $z \boldsymbol{\in} y \leftrightarrow z \boldsymbol{\in} x \wedge z \boldsymbol{\in} x \equiv z \boldsymbol{\notin} y$, so this is equivalent to $\exists y \forall z(z \boldsymbol{\notin} y)$ (proving the existence part of legitimacy; we must still prove uniqueness). Now by AE, $\forall z\, z \boldsymbol{\notin} y \wedge \forall z\, z \boldsymbol{\notin} y' \to y = y'$, and since we just showed $\exists y \forall z(z \boldsymbol{\notin} y)$ gives the required $\exists! y \forall z(z \boldsymbol{\notin} y)$.

We also have

$$\mathsf{AU}: \quad \forall x \exists y \forall z(z \boldsymbol{\in} y \leftrightarrow (\exists u \boldsymbol{\in} x) z \boldsymbol{\in} u) \quad \text{(axiom of union)}$$

y here is the *union* of x, $z \boldsymbol{\in} y$ if and only if there exists a set $u \boldsymbol{\in} x$ such that $z \boldsymbol{\in} u$. By AE, the $\exists y$ may be replaced with $\exists! y$. So we can define an operator (we use the word operator instead of function as the word function has a special meaning in ZFC), the union operator $x \mapsto \bigcup x$. In view of AS, the axiom of union could be weakened to $\forall x \exists y \forall z((\exists u \boldsymbol{\in} x) z \boldsymbol{\in} u \to z \boldsymbol{\in} y)$ as then we can use AS to get the union from this set y. Similarly the following axiom could be weakened:

$$\mathsf{AP}: \quad \forall x \exists y \forall z(z \boldsymbol{\in} y \leftrightarrow z \subseteq x) \quad \text{(power set axiom)}$$

And again, $\exists y$ can be replaced with $\exists! y$ due to AE. We define the *powerset* of a set x to be this unique y, which we denote $\mathcal{P}x$, which contains all the subsets of x. Now, since $x \subseteq \varnothing \leftrightarrow \forall y \boldsymbol{\in} x\, y \boldsymbol{\in} \varnothing$, and $y \boldsymbol{\in} \varnothing \equiv_{\mathsf{ZFC}} \bot$, we have that $x \subseteq \varnothing \leftrightarrow \bot$, so $\forall x(x \in \mathcal{P}\varnothing \leftrightarrow x = \varnothing)$. Similarly $\forall x(x \in \mathcal{P}\mathcal{P}\varnothing \leftrightarrow x = \varnothing \vee x = \mathcal{P}\varnothing)$. Since $\mathcal{P}\varnothing \neq \varnothing$ as it is not empty, $\mathcal{P}\mathcal{P}\varnothing$ has precisely two elements.

The following axiom was added by Fraenkel to the axioms formulated by Zermelo:

$$\mathsf{AR}: \quad \forall x \exists! y \varphi \to \forall u \exists v \forall y(y \boldsymbol{\in} v \leftrightarrow (\exists x \in u)\varphi) \quad \text{(axiom of replacement)}$$

Here $\varphi = \varphi(x, y, \vec{a})$ is viewed as a function which maps x to y, where y is the unique value satisfying φ (hence the $\forall x \exists! y \varphi$). If φ does represent such a function, AR says that for every u we can define the image of u under this function. More suggestively, we can define an operator $x \mapsto Fx$ where $y = Fx \leftrightarrow \varphi$ (since F actually is dependent on $\vec{a}$, we should instead write $F_{\vec{a}}$ for F). Then AR states that for every u we can define its image under u which is denoted $\{Fx \mid x \in u\}$.

For example, let us define $\varphi = \varphi(x, y, a, b) := x = \varnothing \wedge y = a \vee x \neq \varnothing \wedge y = b$. $\forall x \exists! y \varphi$ is obvious, and so this defines an operator $F = F_{a,b}$ which satisfies $F\varnothing = a$ and $Fx = b$ for all $x \neq \varnothing$ (and such an x exists, for example $\mathcal{P}\varnothing$). We then define

$$\{a, b\} := \{F_{a,b}x \mid x \in \mathcal{P}\mathcal{P}\varnothing\}$$

(since $\mathcal{PP}\varnothing$ contains $\mathcal{P}\varnothing$ and $\varnothing$, two distinct elements.) This is called the *pair set* of a, b. Then we can define the union of two sets (which would not necessarily exist, as we have only defined the union of a set) to be $a \cup b := \bigcup\{a, b\}$. The intersection of two sets already exists by AS: $a \cap b := \{z \in a \mid z \in b\}$. But we can also define the intersection of a set as $\bigcap x := \{z \in \bigcup x \mid (\forall y \in x) z \in y\}$. Notice that $\bigcap\{a, b\} = \{z \in a \cup b \mid z \in a \wedge a \in b\}$ which by AE is equal to $a \cap b$. We further define $\{a\} := \{a, a\}$ and inductively $\{a_1, \ldots, a_{n+1}\} := \{a_1, \ldots, a_n\} \cup \{a_{n+1}\}$. So for example, we can prove $\mathcal{P}\varnothing = \{\varnothing\}$, $\mathcal{PP}\varnothing = \{\varnothing, \{\varnothing\}\}$. We define the *ordered pair* of a, b to be $(a, b) := \{\{a\}, \{a, b\}\}$. This definition is useful as it has the basic property $(a, b) = (c, d) \leftrightarrow a = c \wedge b = d$.

So far our theory has many features, but it lacks an important one: there does not necessarily exist an infinite set! Let us define the successor operator, defined by $\mathtt{S}\colon x \mapsto x \cup \{x\}$. So for example $\mathtt{S}\varnothing = \{\varnothing\}$, $\mathtt{SS}\varnothing = \{\varnothing, \{\varnothing\}\}, \ldots$. Notice that $\mathtt{S}x$ contains precisely one more element than x (assuming that $x \notin x$, which requires another axiom which will be stated next). So there is a natural correspondence between this successor operator and the one from PA (in fact, in some definitions of $\mathbb{N}$, these two operators are one and the same). Now we state the following axiom which asserts the existence of an infinite set

$$\mathsf{AI}: \quad \exists u(\varnothing \in u \wedge \forall(x \in u)\mathtt{S}x \in u) \quad \text{(axiom of infinity)}$$

This set u is a so-called *inductive set*: it is closed under successors and contains $\varnothing$.

We also have

$$\mathsf{AF}: \quad (\forall x \neq \varnothing)(\exists y \in x) x \cap y = \varnothing \quad \text{(axiom of foundation)}$$

This means that every nonempty set x has a $\in$-minimal element: an element y such that no other $y' \in x$ is in y. The axiom of foundation implies that there cannot be a circular chain of inclusion, ie. $x_0 \in \cdots \in x_n \in x_0$, as then $y = \{x_0, \ldots, x_n\}$ would have no $\in$-minimal element.

These previously mentioned axioms axiomatize the theory ZF. If we adjoin to ZF the *axiom of choice*, we get the theory ZFC.

$$\mathsf{AC}: \quad \forall u\big(\varnothing \notin u \wedge (\forall x \in u)(\forall y \in u)(x \neq y \to x \cap y = \varnothing) \to \exists z(\forall x \in u)\exists! y(y \in x \cap z)\big) \quad \text{(axiom of choice)}$$

The axiom of choice states that if u is a set of nonempty disjoint sets, then there exists a *choice set* z which contains precisely one element from each $x \in u$. Equivalent to AC (among many other important equivalences) is that $\prod_{i \in I} A_i = \varnothing$ for any index set I.

ZFC is often viewed as the purest of all first-order theories, as it is sophisticated and thought to be able to formalize (almost) all proofs in mathematics (the *almost* here is technical: ZFC is not complete, and so some mathematicians may study results independent of it. One such example is the *continuum hypothesis*). But whether or not this is true is not of too much significance, as most mathematics does not care whether or not it can be formalized in ZFC.

If ZFC is consistent (which no one really doubts, although we cannot prove), then by the Löwenheim-Skolem Theorem, it must contain a countable model $\mathcal{V} = (V, \in^{\mathcal{V}})$. But certainly the existence of an uncountable set is provable within ZFC: for every inductive set u, $\mathcal{P}u$ is uncountable. But at the same time, $(\mathcal{P}u)^{\mathcal{V}} \subseteq V$ and so $\mathcal{P}u$ must be countable. The fallacy here is that our notion of *countable* "outside" of $\mathcal{V}$ is different from the notion "within" $\mathcal{V}$. This "paradox" is known as *Skolem's paradox*.

The explanation of Skolem's paradox is that the countable model $\mathcal{V}$ contains fewer sets and functions than expected, and in particular there do not exist enough to satisfy a bijection between $u^{\mathcal{V}}$ and $(\mathcal{P}u)^{\mathcal{V}}$ although "outside" of $\mathcal{V}$ such a bijection exists.

Another "paradox" is that while V is by definition a set, but $\vdash_{\mathsf{ZFC}} \neg\exists u \forall z\, z \in u$, ie. there does not exist a universal set. This is derived as follows: $\exists u \forall z\, z \in u$ implies, using AE and AS, the existence of a Russelian set $v = \{x \in u \mid x \notin x\}$. Thus $\exists u \forall z\, z \in u \vdash_{\mathsf{ZFC}} \exists u \forall x(x \in u \leftrightarrow x \notin x)$. Though we have previously shown that the right-hand formula is logically invalid, and thus equivalent to $\perp$. So $\exists u \forall z\, z \in u \vdash_{\mathsf{ZFC}} \perp$ and so $\vdash_{\mathsf{ZFC}} \neg\exists u \forall z\, z \in u$. So similar to how there is no absolute concept of countability, there is no absolute concept of a set: a set "outside" a model of ZFC is not necessarily a set within ZFC.

Exercise

Let T be an elementary theory with arbitrarily large finite models. Prove that T also has an infinite model.

Since T has arbitrarily large finite models, $\exists_n$ does not contradict T for every n (since $T + \exists_n$ is not unsatisfiable). Thus let $X = \{\exists_n \mid n \in \mathbb{N}\}$, then $T + X$ is satisfiable: let $X_0 \subseteq T \cup X$ be finite then $X_0 \subseteq T \cup \{\exists_n \mid n \leq m\}$ for some m. Then X_0 is satisfiable as $T + \exists_m$ is satisfiable, and therefore so is $T \cup \{\exists_n \mid n \leq m\}$ (as $\exists_m \to \exists_n$ is a tautology) and thus so is X_0. So every finite subset of $T \cup X$ is satisfiable, and so by The Compactness Theorem, so is $T \cup X$.

Now, a model of $T + X$ must satisfy $\exists_n$ for every n, and so it must have at least n elements for every n. So it cannot be finite, and thus must be infinite. Furthermore, by definition a model of $T + X$ models T as well.

Exercise

Suppose $\mathcal{A} = (A, <)$ is an infinite well-ordered set. Show that there is a not well-ordered set elementarily equivalent to $\mathcal{A}$. Thus being well-ordered is not a first-order property (at least in the language $\mathcal{L}\{<\}$).

Let us define $X = Th\mathcal{A} \cup \{\boldsymbol{v}_{n+1} < \boldsymbol{v}_n \mid n \in \mathbb{N}\}$. X is finitely satisfiable (as finite subsets of X can be modeled over $\mathcal{A}$), but every model of X has an infinitely descending chain (and such a chain therefore has no minimum element) and therefore models of X are not well-ordered.

Exercise

Prove that a consistent theory T is equal to the intersection of all its complete extensions. Meaning $T = \bigcap\{T' \supseteq T \mid T' \text{ complete}\}$.

Firstly, T does have a complete extension as there exists a $\mathcal{A} \vDash T$ and $Th\mathcal{A}$ is a complete extension of T. Obviously then $T \subseteq \bigcap\{T' \supseteq T \mid T' \text{ complete}\}$. Now suppose $\varphi \in T'$ for every complete extension of T, then if $\varphi \notin T$ then $\neg\varphi$ doesn't contradict T since $T, \neg\varphi \nvdash \bot$ (in general if $X \nvdash \varphi$ then by $\mathtt{C}^+$, $X, \neg\varphi \nvdash \bot$). But then $T, \neg\varphi$ is consistent and therefore has a complete extension, which contradicts φ being in every complete extension of T.

Exercise

Derive AS from AR.

Let $\varphi = \varphi(x, z, \vec{a})$ be a formula and let $y, b \notin free\varphi$.

$$\psi = \left(\varphi\tfrac{b}{z} \wedge \left((y = z \wedge \varphi) \vee (y = b \wedge \neg\varphi)\right)\right) \vee \left(\neg\varphi\tfrac{b}{x} \wedge y = \varnothing\right)$$

Then $\forall z \exists! y \psi$, as either φ or $\neg\varphi$ is true. This operator defined by ψ maps z to z if φ and z to b otherwise. If $\{z \in x \mid \varphi\}$ is nonempty then there exists a b such that $\varphi\frac{b}{z}$ and for this b, the operator defined by ψ when parameterized with b and x gives this set. If φ is invalid, then $\{z \in x \mid \varphi\}$ is the empty set which exists by definition.

3.5 Enumerability and Decidability

3.5.1 Definition

A countable set M is *effectively enumerable* if there exists an algorithm which produces as output every element in M stepwise. This definition will be made more formal in a later section.

So for example, all provable finite sequents of countable first-order language (ie. (X, α) such that $X \vdash \alpha$ and X is finite) is effectively enumerable. To do so, we can enumerate all initial sequents in a reproducible sequence $S_0, S_1, \dots$. Then at every iteration we check if one of the basic rules are applicable, and if so we form a second sequence of sequents and so on. And this can also help effectively enumerate all tautologies of such a language: in this enumeration, only output sequents with $X = \varnothing$.

A similar argument can be applied to the theorems of an axiomatizable countable theory. One obtains an effective enumeration of the axioms (which exists as they are decidable) of the theory and adds those to the sequents (along with the initial sequents) of the machine described above. This is an intuitive explanation of the following theorem which will be formally proven in a later section.

3.5.2 Theorem

The theorems of an axiomatizable countable theory are effectively enumerable.

For example, ZFC and PA are axiomatizable as their axioms are decidable. Thus the theorems of these theories are effectively enumerable. But just because something is effectively enumerable does not mean that it is decidable: one has no method to determine whether an object will show up in the sequence. One just knows that if it belongs to the set then it will show up in the sequence, but if an object does not belong to the set they will have no way of knowing this.

The following theorem is also only proven at present in an informal manner, a formal proof is provided later.

3.5.3 Theorem

A complete axiomatizable theory T is decidable.

By the previous theorem, there exists an effective enumeration of all the theorems of T. Given a sentence φ we iterate over this enumeration and check for φ and $\neg\varphi$. Since T is complete, one of them must exist in the enumeration and so this takes finite time. If we find φ then φ belongs to T so we can accept, and if we find $\neg\varphi$ then φ does not belong to T (as T is consistent) and we can reject. ■

On the other hand, a complete decidable theory T is axiomatizable by T itself. Thus for complete theories, being decidable is equivalent to being axiomatizable.

Exercise

Let $T' = T + \alpha$ be a finite extension of T. Show that if T is decidable then so too is T'.

Notice that $T + \alpha \vdash \varphi$ if and only if $T \vdash \alpha \to \varphi$ and so we can check if $\varphi \in T + \alpha$ by checking if $\alpha \to \varphi \in T$ which is decidable as T is.

Exercise

Suppose T is consistent and has only finitely many completions. Prove that every completion of T is a finite extension.

Suppose T had a completion which is an infinite extension, of the form $T' = T + \{\alpha_i\}_{i\in\mathbb{N}}$. We can assume that $\bigwedge_{i=1}^n \alpha_i \nvdash_T \alpha_{n+1}$ as otherwise we could just remove α_{n+1}. But then $T + \bigwedge_{i=1}^n \alpha_i \wedge \neg\alpha_{n+1}$ is consistent, and therefore has a completion T_n. But for $n \neq m$ we have $T_n \neq T_m$ (since one contains $\neg\alpha_{n+1}$ and the other does not), and so T has infinite distinct completions in contradiction.

Exercise

Show that an axiomatizable theory with finitely many completions is decidable.

We have shown previously that if T is consistent then $T = \bigcap\{T' \supseteq T \mid T' \text{ complete}\}$. By above, this set is finite and each T' is a finite extension. Since T' is a finite extension, and T is axiomatizable, so is T'. By the above theorem, T' is therefore decidable, and so T is the finite intersection of decidable sets, which is decidable.

Exercise

Show that a decidable countable theory T has a decidable completion.

Let $(\alpha_n)_{n=1}^\infty$ be an effective enumeration of the sentences in $\mathcal{L}^0$. Let us set $T_0 := T$ and $T_{n+1} := T_n + \alpha_n$ if it is consistent, otherwise $T_{n+1} := T_n$. Then $T' = \bigcup_{n=0}^\infty T_n$ is a complete theory: it being a theory and consistent are immediate from T_n being a chain of consistent theories and the finiteness theorem. It is complete since if $\alpha, \neg\alpha \notin T'$ suppose $\alpha = \alpha_n$ and $\neg\alpha = \alpha_m$ and $n < m$ then $T_m + \neg\alpha_m$ was inconsistent and so $T_m, \neg\alpha \vdash \bot$ so $\alpha \in T_m$ but $T_n + \alpha \subseteq T_m + \alpha = T_m$ is also inconsistent, contradicting every T_i being consistent.

T' is also effectively enumerable as it can be axiomatized: for every sentence, check if it is in T or if $T_n + \alpha_n$ is consistent (which requires finding the n and checking consistency, which can both be done in finite time). Since T' is complete and axiomatizable, it is decidable.

3.6 Complete Hilbert Calculi

As with propositional logic, we define Hilbert calculi for first-order logic. These are calculi defined by a logical axiom system and rules of inference. Here our only rule of inference is *modus ponens* (MP): $\alpha, \alpha \to \beta / \beta$ (meaning if both α and $\alpha \to \beta$ are derivable then β). We will denote the derivation relation on a first-order language $\mathcal{L}$ by $\vdash$ (or $\vdash_{\mathcal{L}}$ if we are to specify the language). Again, $\alpha \to \beta$ is defined to mean $\neg(\alpha \wedge \neg\beta)$. The logical axiom scheme is defined to be all formulas of the form $\forall x_1 \cdots \forall x_n \varphi$ where x_i are arbitrary and φ is a formula of one of the following forms:

Λ1: $(\alpha \to \beta \to \gamma) \to (\alpha \to \beta) \to \alpha \to \gamma$	Λ2: $\alpha \to \beta \to \alpha \wedge \beta$
Λ3: $\alpha \wedge \beta \to \alpha, \quad \alpha \wedge \beta \to \beta$	Λ4: $(\alpha \to \neg\beta) \to \beta \to \neg\alpha$
Λ5: $\forall x\alpha \to \alpha\frac{t}{x}$ ($\alpha, \frac{t}{x}$ collision-free)	Λ6: $\alpha \to \forall x\alpha$ ($x \notin free\alpha$)
Λ7: $\forall x(\alpha \to \beta) \to \forall x\alpha \to \forall x\beta$	Λ8: $\forall y\alpha\frac{y}{x} \to \forall x\alpha$ ($y \notin var\alpha$)
Λ9: $t = t$	Λ10: $x = y \to \alpha \to \alpha\frac{y}{x}$ (α prime)

These are all easy to see as tautologies. Λ1 to Λ4 are simply the axioms from Hilbert calculi in propositional logic. Now we say that $X \vdash \alpha$ (α is derivable from X) if there exists a *proof* from X, $\Phi = (\varphi_0, \ldots, \varphi_n)$. This means that $\alpha = \varphi_n$ and for every $i < n$, $\varphi_i \in X \cup \Lambda$ or is derivable by MP from previous φ_js (meaning $\varphi_j \to \varphi_i$ and φ_j appear before φ_i in Φ). This is the same definition as used in propositional logic, and so the principle of induction holds here too: let $\mathcal{E}$ be a property of formulas such that $\mathcal{E}\varphi$ for all $\varphi \in X \cup \Lambda$ and $\mathcal{E}\alpha$, $\mathcal{E}(\alpha \to \beta)$ implies $\mathcal{E}\beta$, then $X \vdash \alpha$ implies $\mathcal{E}\alpha$.

Again, we can prove by induction that $\vdash \subseteq \vDash$, meaning $X \vdash \alpha \implies X \vDash \alpha$. The proof is identical to the one provided for propositional logic.

3.6.1 Theorem (The Completeness Theorem for $\vdash$)

$X \vdash \alpha$ if and only if $X \vDash \alpha$, ie. $\vdash = \vDash$.

We need only show that $X \vDash \alpha$ implies $X \vdash \alpha$. Since $\vDash = \vdash$, we will show that $\vdash$ satisfies all the basic rules of $\vdash$, which shows that $\vdash \subseteq \vdash$. The rules of (MR) through ($\neg$2) are handled as in propositional logic, (IR) requires only the addition of (Λ9). ($\forall$1) follows from Λ5 by use of MP. So we need now only show ($\forall$2) and (=).

For ($\forall$2), we must show that if $X \vdash \alpha\frac{y}{x}$ then $X \vdash \forall x\alpha$ where $y \notin freeX \cup var\alpha$. Let us first show by induction that if $x \notin freeX$ and $X \vdash \alpha$ then $X \vdash \forall x\alpha$. For the base case, if $\alpha \in X$ then $x \notin free\alpha$ and so Λ6 yields $\alpha \to \forall x\alpha$ and by MP we get $X \vdash \forall x\alpha$. If $\alpha \in \Lambda$ then $\forall x\alpha \in \Lambda$ and so $X \vdash \forall x\alpha$ as required. For the inductive step, if $X \vdash \forall x(\alpha \to \beta)$ and $X \vdash \forall x\alpha$ then by Λ7 we get $X \vdash \forall x\alpha \to \forall x\beta$ and thus by MP we get the desired $X \vdash \forall x\beta$. To verify ($\forall$2) if $X \vdash \alpha\frac{y}{x}$ and $y \notin freeX \cup var\alpha$, then by what we just proved we get $X \vdash \forall y\alpha\frac{y}{x}$. So by Λ8 we have $X \vdash \forall y\alpha\frac{y}{x} \to \forall x\alpha$ and so MP gives $X \vdash \forall x\alpha$, as required.

For (=), we must show that if $X \vdash s = t, \alpha\frac{s}{x}$ then $X \vdash \alpha\frac{t}{x}$, where α is prime. Let y be a variable distinct from x not occurring in s or α. By Λ10 we have $X \vdash \forall x\forall y(x = y \to \alpha \to \alpha\frac{y}{x})$. ($\forall$1) gives us

$$X \vdash \big(\forall y(x = y \to \alpha \to \alpha\tfrac{y}{x})\big)\tfrac{s}{x} = \forall y(s = y \to \alpha\tfrac{s}{x} \to \alpha\tfrac{y}{x})$$

Then applying ($\forall$1) again gives

$$X \vdash \big(s = y \to \alpha\tfrac{s}{x} \to \alpha\tfrac{y}{x}\big)\tfrac{t}{y} = s = t \to \alpha\tfrac{s}{x} \to \alpha\tfrac{t}{x}$$

since y does not occur in α or s, $\alpha\frac{s}{x}\frac{t}{y} = \alpha\frac{s}{x}$ and $s\frac{t}{y} = s$. So if $X \vdash s = t, \alpha\frac{s}{x}$ applying MP twice gives $\alpha\frac{t}{x}$. ■

Another rule of inference is MQ: $\alpha/\forall x\alpha$. This is not a rule of inference in this calculus, and it is not provable. But Λ is, by definition, closed under MQ.

3.6.2 Corollary

For any $\alpha \in \mathcal{L}$, the following are equivalent

(1) $\vdash \alpha$, meaning α is derivable from α be means of MP only,

(2) α is derivable from Λ1 through Λ10 by means of MP and MQ only,

(3) $\vDash \alpha$, meaning α is a tautology.

(1) $\iff$ (2) is obvious, Λ is simply the closure of Λ1 – Λ10 by means of MQ. And (1) $\iff$ (3) is a direct result of the completeness theorem. ■

We define the Hilbert calculus $\overset{g}{\vdash}$ which has the same logical axiom system as $\vdash$, but it has MQ as well as MP as its rules of inference. The definition of proofs and derivability in $\overset{g}{\vdash}$ must take MQ into account. It is obvious to see that (like for every Hilbert calculus), $X \overset{g}{\vdash} Y$ and $Y \overset{g}{\vdash} \alpha$ implies $X \overset{g}{\vdash} \alpha$.

3.6.3 Theorem (The Completeness Theorem for $\overset{g}{\vdash}$)

$X \overset{g}{\vdash} \alpha$ if and only if $X \overset{g}{\vDash} \alpha$, ie. $\overset{g}{\vdash} = \overset{g}{\vDash}$.

By induction on $\overset{g}{\vdash}$, it is easy to see that $\overset{g}{\vdash} \subseteq \overset{g}{\vDash}$. If $X \overset{g}{\vDash} \alpha$ then $X^g \vDash \alpha$, and so by completeness $X^g \vdash \alpha$. Since certainly by definition $\vdash \subseteq \overset{g}{\vdash}$ we get $X^g \overset{g}{\vdash} \alpha$. By MQ we also have $X \overset{g}{\vdash} X^g$ and thus $X \overset{g}{\vdash} \alpha$. ■

3.6.4 Definition

A sentence $\alpha \in \mathcal{L}^0$ is called **generally valid in the finite** if $\mathcal{A} \vDash \alpha$ for all finite $\mathcal{L}$-structures $\mathcal{A}$. Let $\mathsf{Tautfin}_{\mathcal{L}}$ be the set of all generally valid in the finite sentences in $\mathcal{L}^0$. Obviously $\mathsf{Taut}_{\mathcal{L}} \subseteq \mathsf{Tautfin}_{\mathcal{L}}$.

For example, let f be a unary function symbol. In finite sets, an injective endomorphism (meaning an operator from a structure to itself) is also surjective. Thus $\forall x \forall y(fx = fy \to x = y) \to \forall y \exists x\, y = fx$ (f is injective implies f is surjective) is generally valid in the finite, but not so in infinite structures (for example $n \mapsto 2n$ in $\mathbb{N}$).
A theory T has the *finite model property* if every compatible $\alpha \in \mathcal{L}^0$ has a finite T-model. In other words, if $T + \alpha$ is satisfiable then it can be satisfied by a finite model. For example if $\boldsymbol{K}$ is a class of finite $\mathcal{L}$-structures then $T = Th\boldsymbol{K}$ has the finite model property: if $T + \alpha$ is consistent, meaning $\neg\alpha \notin T$ and so $\mathcal{A} \nvDash \neg\alpha$ for some $\alpha \in \boldsymbol{K}$, meaning $\mathcal{A} \vDash \alpha$ so $\mathcal{A}$ is a finite model of $T + \alpha$.
So for example $\mathsf{FSG} = Th\{S \mid S \text{ is a finite semigroup}\}$ and $\mathsf{FG} = Th\{G \mid G \text{ is a finite group}\}$ in $\mathcal{L}_\circ$. Both of these theories are undecidable (the tools to prove this will be discussed later).

3.6.5 Lemma

Suppose T has the finite model property, and T's finite models are effectively enumerable (more precisely, it has an effectively enumerable family of representatives of isomorphism classes of models of T). Then the set of sentences refutable in T are effectively enumerable and if T is axiomatizable then it is decidable.

To check if a sentence α is refutable with T, we need only to check that $T + \neg\alpha$ is consistent. Since T has the finite model property, we need only check that $\neg\alpha$ has a finite T-model. So we can iterate over the enumeration of finite T-models, and for each model we iterate over all sentences α and check if the model models $\neg\alpha$. If so, we can add α to the sequence, then to ensure we go over every sentence we can go to the beginning of the sequence and begin checking from the next sentence and repeat.
Notice that $\alpha \notin T$ is equivalent to $\nvdash_T \alpha$ which is equivalent to $T, \neg\alpha \nvdash \bot$ by C^+ meaning $\neg\alpha$ is consistent with T, ie. α is refutable in T. So for every sentence α, either $\alpha \in T$ or α is refutable in T. So we can simply check for every sentence if it is provable (which can be done since T is axiomatizable: iterate over every possible proof), or if α is refutable in T (which can be done by what we just proved). ■

3.6.6 Theorem (Trachtenbrot's Theorem)

$\mathsf{Tautfin}_{\mathcal{L}}$ is not axiomatizable for any first-order language $\mathcal{L}$ containing at least one binary operation or binary relation symbol.

We will prove this for $\mathcal{L}$ containing a binary operation symbol. Notice how $\mathsf{Tautfin}_{\mathcal{L}}$'s finite models are effectively enumerable, as they are simply all the finite models. So by the above lemma if $\mathsf{Tautfin}_{\mathcal{L}}$ were axiomatizable, it would be decidable. Now, in the previous subsection we proved in the first exercise that if T is decidable, then so is $T + \alpha$. If $\mathsf{Tautfin}_{\mathcal{L}}$ were decidable, then so too must $\mathsf{Tautfin}_{\mathcal{L}_\circ}$ be, and since FSG is simply the extension of this theory with the law of associativity, this would mean that FSG is decidable. But as said above, it is not. ■

Exercise

Show that MQ is unprovable in $\vdash$, meaning $X \vdash \alpha$ does not mean $X \vdash \forall x\alpha$.

We know that $x = y \vdash x = y$, but $x = y \vdash \forall x\, x = y$ does not hold.

Exercise

Suppose T is a finitely axiomatizable theory with the finite model property, show that T is decidable.

Since T is finitely axiomatizable, its finite models are effectively enumerable: enumerate all finite models and check that all the axioms hold (there are only finitely many). Since T also has the finite model property, we showed that this means T is decidable.

Exercise

Show that $\forall x \exists y\, y + y = x \to \forall x(x + x = 0 \to x = 0)$ holds in all finite abelian groups. Show that it does not hold in all infinite abelian groups.

Suppose $x + x = 0$ in a finite abelian group, or $2x = 0$. Let us denote $x = x_1$, then since there exists $2x_2 = x_1$, we have $4x_2 = 0$, and so on we set $2x_{k+1} = x_k$ and so $2^k x_k = 0$ for every k. Now since the group is finite, we must have $x_{k+\ell} = x_k$ for some $k, \ell > 0$. But since $2^\ell x_{k+\ell} = x_k$ we get $2^\ell x_k = x_k$. Now let us assume that $x \neq 0$ and so we must have $x_k \neq 0$ as well and thus $2^\ell x_k \neq 0$ and so $\ell < k$ (as otherwise we'd have that $2^\ell x_k = 0$ since $2^k x_k = 0$). Thus $0 = 2^k x_k = 2^{k-\ell} x_k$, but since $2^{k-\ell} x_k = x_\ell$ we get that $x_\ell = 0$ contradicting $x_k \neq 0$ for every k.

The group $(\mathbb{C} \setminus \{0\}, \cdot)$ is a counterexample: every x has a square root (two in fact), but $x^2 = 1$ (since 1 is the unit) does not imply $x = 1$. For example, take $x = -1$.

3.7 First-Order Fragments

A first-order fragment is a language which reduces the expressive power of first-order logic, for example by omitting some logical connectives. Fragments are useful as they can be more easily simulated by a computer, whose computational power is limited. We will deal with the *language of equations* closely here, whose formulas are equations over an algebraic signature (an extralogical signature with no relation symbols). The variables in these formulas are tacitly generalized, and such generalizations are called *identities*. Theories with an axiom system consisting of only identities are called *equational theories*, and their model classes *varieties*.

Suppose Γ is a set of identities which form the axiom system of some equational theory, and suppose γ is some equation. If $\Gamma^g \vDash \gamma$ then by the completeness theorem there is some proof of γ from Γ^g. But one might expect that to verify $\Gamma^g \vdash \gamma$, we need not utilize every formula, and indeed we can restrict ourselves only to other identities.

The following are the *Birkhoff rules* which are the rules of inference of the Hilbert-style calculus $\overset{B}{\vDash}$,

(B0) $/t = t$	(B1) $s = t/t = s$	(B2) $t = s, s = t'/t = t'$
(B3) $\vec{t} = \vec{t'}/f\vec{t} = f\vec{t'}$	(B4) $s = t/s^\sigma = t^\sigma$	

These rules are all stated with respect to ungeneralized equations, but in a derivation sequence (a proof) we must generalize all equations as (B4) does not hold in general if its premise is not generalized, we have in general only $(s = t)^g \vDash s^\sigma = t^\sigma$. (B0) has no premise, meaning $t = t$ is derivable from every set Γ. Now we claim that for equations γ, $\Gamma \overset{B}{\vDash} \gamma \iff \Gamma^g \vDash \gamma$ (or $\Gamma \overset{g}{\vDash} \gamma$). One direction is simpler than the other: to show $\Gamma \overset{B}{\vDash} \gamma \implies \Gamma^g \vDash \gamma$, we need only prove that $\overset{g}{\vDash}$ is closed under (B0) through (B4). This is simple, and we have basically done it already.

Let us define the congruence $\approx$ on $\mathcal{T}$ by $s \approx t \iff \Gamma \overset{B}{\vDash} s = t$. By (B4), $\approx$ is also substitution invariant: $s \approx t \implies s^\sigma \approx t^\sigma$. Let $\mathcal{F}$ be the quotient algebra of $\mathcal{T}$ with respect to $\approx$, and let $\bar{t}$ be the congruence class of t with respect to $\approx$. This means $\overline{t_1} = \overline{t_2} \iff \Gamma \overset{B}{\vDash} t_1 = t_2$. Now let $w\colon \mathit{Var} \longrightarrow \mathcal{F}$, and denote $x^w = \overline{t_x}$ (t_x is an arbitrary representative of x^w). Any choice of representative defines a global substitution $\sigma\colon x \mapsto t_x$. Term induction gives

$$t^{\mathcal{F},w} = \overline{t^{\sigma_w}}$$

3.7.1 Lemma

$\Gamma \overset{B}{\vDash} t_1 = t_2$ if and only if $\mathcal{F} \vDash t_1 = t_2$.

We prove the first direction: suppose $\Gamma \overset{B}{\vDash} t_1 = t_2$, let $w\colon \mathit{Var} \longrightarrow \mathcal{F}$ and $\sigma = \sigma_w$. By (B4) this means $\Gamma \overset{B}{\vDash} t_1^\sigma = t_2^\sigma$, meaning $\overline{t_1^\sigma} = \overline{t_2^\sigma}$, and so this means $t_1^{\mathcal{F},w} = t_2^{\mathcal{F},w}$. Since w was arbitrary, this means $\mathcal{F} \vDash t_1 = t_2$. Now suppose the other direction, $\mathcal{F} \vDash t_1 = t_2$, and let $\varkappa$ be the *canonical evaluation* $x \mapsto \bar{x}$, and we choose $\sigma_\varkappa = \iota$ (the identity substitution). Thus $t_i^{\mathcal{F},\varkappa} = \overline{t_i}$, and $\mathcal{F} \vDash t_1 = t_2$ implies $t_1^{\mathcal{F},\varkappa} = t_2^{\mathcal{F},\varkappa}$ meaning $\overline{t_1} = \overline{t_2}$, so $\Gamma \overset{B}{\vdash} t_1 = t_2$ as required. ■

Notice that if Γ is a set of equations, then by this lemma $\mathcal{F} \vDash \Gamma$ and so $\mathcal{F} \vDash \Gamma^g$.

3.7.2 Theorem (Birkhoff's Completeness Theorem)

Let Γ be a set of identities and $t_1 = t_2$ an equation, then $\Gamma \overset{B}{\vdash} t_1 = t_2 \iff \Gamma^g \vDash t_1 = t_2$.

We have already shown $\Longrightarrow$, we now show the other direction. Now if $\Gamma^g \vDash t_1 = t_2$, then $\mathcal{F} \vDash t_1 = t_2$ since $\mathcal{F} \vDash \Gamma^g$ and therefore by the above lemma $\Gamma \overset{B}{\vdash} t_1 = t_2$. ■

This theorem has many variations, for example for sentences of the form $\forall \vec{x}\pi$ where π is an arbitrary prime formula. In this case we must add $\vec{t} = \vec{t}', r\vec{t}/r\vec{t}'$ to the Birkhoff rules.

Exercise

Let $\boldsymbol{K}$ be a variety. Show that it is closed with respect to homomorphic images, taking subalgebras, and forming arbitrary direct products of members of $\boldsymbol{K}$.

Essentially all we must prove is that if $\mathcal{A} \vDash t_1 = t_2$, then (a) if $h\colon \mathcal{A} \longrightarrow \mathcal{B}$ is a homomorphism then $h\mathcal{A} \vDash t_1 = t_2$, (b) if $\mathcal{B} \subseteq \mathcal{A}$ then $\mathcal{B} \vDash t_1 = t_2$, and (c) if $\{\mathcal{A}_i\}_{i\in I} \vDash t_1 = t_2$ then $\prod_{i\in I} \mathcal{A}_i \vDash t_1 = t_2$. (a) Let $w\colon \mathit{Var} \longrightarrow h\mathcal{A}$, then suppose $x^w = h(a_x)$ (a_x is arbitrary) then each choice of a_x defines the valuation $w'\colon \mathit{Var} \longrightarrow \mathcal{A}$ by $x^{w'} = a_x$. By term induction we get $t^w = h(t^{w'})$ and so since $\mathcal{A}, w' \vDash t_1 = t_2$ we must have $t_1^{w'} = t_2^{w'}$ and thus $h(t_1^{w'}) = h(t_2^{w'})$ meaning $t_1^w = t_2^w$. Since w is arbitrary this means $h\mathcal{A} \vDash t_1 = t_2$ as required.

(b) By The Substructure Theorem, for every $\vec{b} \in \mathcal{B}^n$, since $\mathcal{A} \vDash t_1 = t_2[\vec{b}]$, we get $\mathcal{B} \vDash t_1 = t_2[\vec{b}]$. Since $\vec{b}$ is arbitrary this means that $\mathcal{B} \vDash t_1 = t_2$.

(c) Let $\mathcal{B} = \prod_i \mathcal{A}_i$. Then let $w\colon \mathit{Var} \longrightarrow \mathcal{B}$ be a valuation, so $x^w = (a_{i,x})_{i\in I}$ and this defines valuations $w_i\colon \mathit{Var} \longrightarrow \mathcal{A}_i$ by $x^w = a_{i,x}$. Again by term induction $t^w = (t^{w_i})_{i\in I}$, and so $t_1^w = t_2^2$ if and only if $t_1^{w_i} = t_2^{w_i}$ for every $i \in I$. Thus if $\mathcal{A}_i \vDash t_1 = t_2$ for every $i \in I$ we get $\mathcal{B} \vDash t_1 = t_2$ as required.

4 Foundations of Logic Programming

4.1 Term Models and Herbrand's Theorem

In previous sections we proved various lemmas utilizing models whose domains are equivalence classes of terms of a first-order language. This method can and will be extended upon in this subsection.

4.1.1 Definition

A **term model** over a first-order language $\mathcal{L}$ is an $\mathcal{L}$-model $\mathcal{F}$ whose domain is the quotient algebra $\mathcal{T}/\approx$, where $\approx$ is a congruence of the term algebra $\mathcal{T}$. We denote $\bar{t} = t/\approx$ to be the equivalence class of the term t. In term models function and constant symbols are interpreted canonically: $f^{\mathcal{F}}\overline{t_1}\cdots\overline{t_n} = \overline{ft_1\cdots t_n}$ and $c^{\mathcal{F}} = \bar{c}$, but no condition is required of the interpretations of relation symbols.

Let $\varkappa$ be the canonical valuation $x \mapsto \bar{x}$, and let $\mathcal{F} = (\mathcal{T}/\approx, \varkappa)$ be a term model (when discussing a term model, if its valuation is not explicitly defined then it is implicitly taken to be the canonical valuation). Then we claim that

$$t^{\mathcal{F}} = \bar{t} \text{ for all terms } t, \qquad \mathcal{F} \vDash \forall\vec{x}\alpha \iff \mathcal{F} \vDash \alpha\tfrac{\vec{t}}{\vec{x}} \text{ for all } \vec{t} \in \mathcal{T}^n \text{ and } \alpha \text{ quantifier-free}$$

The first claim is proven via term induction: it is true by definition for variables as $x^{\varkappa} = \bar{x}$, and $(f\vec{t})^{\mathcal{F}} = f^{\mathcal{F}}\vec{\bar{t}} = \overline{f\vec{t}}$. For the second claim, the left-to-right direction is true in general for all formulas. The converse can be proven as follows: $\mathcal{F} \vDash \alpha\frac{\vec{t}}{\vec{x}}$ if and only if $\mathcal{F}_{\vec{x}}^{\vec{t}^{\mathcal{F}}} = \mathcal{F}_{\vec{x}}^{\vec{\bar{t}}} \vDash \alpha$ by The Substitution Theorem and the first claim. Since $\vec{\bar{t}}$ is arbitrary and encompasses all the elements of the domain, this means $\mathcal{F} \vDash \forall\vec{x}\alpha$ as required.

4.1.2 Definition

Let $X \subseteq \mathcal{L}$ be a set of formulas, then its associated term model $\mathcal{F} = \mathcal{F}X$ is the term model defined by the congruence $\approx_{\mathcal{F}X}$ and whose relations are defined by:

$$s \approx_{\mathcal{F}X} t \iff X \vdash s = t, \qquad r^{\mathcal{F}X}\vec{\bar{t}} \iff X \vdash r\vec{t}$$

(Oftentimes this **term structure** is denoted $\mathfrak{F}$.)

It is readily verifiable that $\approx_{\mathcal{F}X}$ is indeed a congruence and that the definition of $r^{\mathcal{F}X}$ is independent on the choice of representatives. If X is the axiom system of some theory T then we may write $\mathcal{F}T$ and $\approx_T$ in place of $\mathcal{F}X$ and $\approx_{\mathcal{F}X}$ respectively.

As proven above, $\mathcal{F}X \vDash s = t \iff s^{\mathcal{F}X} = t^{\mathcal{F}X} \iff \bar{s} = \bar{t} \iff X \vdash s = t$. And similarly $\mathcal{F}X \vDash r\vec{t} \iff X \vdash r\vec{t}$. Thus we get that

$$\mathcal{F}X \vDash \pi \iff X \vdash \pi \quad (\pi \text{ prime})$$

But $\mathcal{F}X$ is not necessarily a model of X.

As before, let $Var_k = \{v_0, \dots, v_{k-1}\}$ and $\mathcal{L}^k$ be the set of all formulas where $free\varphi \subseteq Var_k$. Pairs $(\mathcal{A}, w)$ where $\mathcal{A}$ is an $\mathcal{L}$-structure and the domain of w contains Var_k (so it can be viewed as a valuation $w\colon Var_k \longrightarrow \mathcal{A}$) are called *$\mathcal{L}^k$-models*. Thus if $k = 0$ then w is the empty function (ie. $w = \varnothing$) and so $\mathcal{L}^0$-models can be identified with $\mathcal{L}$-structures. Let us define $\mathcal{T}_k = \{t \in \mathcal{T} \mid var t \subseteq Var_k\}$. We tacitly assume that the set of ground terms $\mathcal{T}_0$ is nonempty by assuming that $\mathcal{L}$ contains at least one constant. $\mathcal{T}_k$ is obviously a subalgebra of $\mathcal{T}$ since if $t_1, \dots, t_n \in \mathcal{T}_k$ then $f\vec{t} \in \mathcal{T}_k$.

We can extend the definition of term models to $\mathcal{L}^k$: let $\approx$ be a congruence in $\mathcal{T}^k$ then we define the quotient structure by $\mathfrak{F}_k$ (the interpretation of function and constant symbols are the same as for term models). We can canonically extend $\mathfrak{F}_k$ to an $\mathcal{L}^k$-model by the (partial) valuation $x \mapsto \bar{x}$ for $x \in Var_k$ (for $k = 0$ it is empty), and this defines the model $\mathcal{F}_k$. For $X \subseteq \mathcal{L}^k$ we define the $\mathcal{L}^k$-model $\mathcal{F}_kX$ analogously to as its term model was defined previously. Then as before we get the following

$$t^{\mathcal{F}_k} = \bar{t} \text{ for } t \in \mathcal{T}_k, \qquad \mathcal{F}_k \vDash \forall\vec{x}\alpha \iff \mathcal{F}_k \vDash \alpha\tfrac{\vec{t}}{\vec{x}} \text{ for all } \vec{t} \in \mathcal{T}_k^n \text{ and } \alpha \text{ quantifier-free,}$$
$$\mathcal{F}_kX \vDash \pi \iff X \vdash \pi \ (\pi \text{ prime})$$

Let $\varphi = \forall\vec{x}\alpha$ be a universal formula, then we call $\alpha\frac{\vec{t}}{\vec{x}}$ an *instance* of φ. If $\vec{t} \in \mathcal{T}_k^n$ then $\alpha\frac{\vec{t}}{\vec{x}}$ is a *$\mathcal{T}_k$-instance* of φ, if $k = 0$ also a *ground instance*. If U is a set of universal formulas then we define $\mathrm{GI}(U)$ to be the set of all ground instances of $\varphi \in U$. If $k = 0$ and $U \neq \varnothing$ then $\mathrm{GI}(U) \neq \varnothing$ if $\mathcal{L}$ contains constants.

4.1.3 Theorem

Let U be a set of universal formulas and $\tilde{U}$ be the set of all instances of formulas in U. Then the following are equivalent:

(1) U is consistent,

(2) $\tilde{U}$ is consistent,

(3) U has a term model in $\mathcal{L}$.

If $U \subseteq \mathcal{L}^k$ and $\tilde{U}$ is the set of all $\mathcal{T}_k$-instances of formulas in U then this holds as well.

Since $U \vdash \tilde{U}$ by particularization, we immediately get (1) $\Longrightarrow$ (2). For (2) $\Longrightarrow$ (3): let $X \supseteq \tilde{U}$ be maximally consistent, then $\mathcal{F}X \vDash \pi \iff X \vdash \pi$ for prime π. By induction on $\wedge$ and $\neg$ this immediately yields $\mathcal{F}X \vDash \alpha \iff X \vdash \alpha$ for quantifier-free α. Since $\tilde{U}$ contains only quantifier-free formulas and $X \vdash \tilde{U}$, we get that $\mathcal{F}X \vDash \tilde{U}$, and so U has a term model ($\mathcal{F}X$). (3) $\Longrightarrow$ (1) is trivial. For $U \subseteq \mathcal{L}^k$ the proof runs analogously but with $\mathcal{F}_k X$. ■

Notice then that if U is a consistent set of universal sentences, it has a term model whose domain is a quotient algebra of $\mathcal{T}_0$, the set of all ground terms. If U is =-free then there is no need to take a quotient, and so the domain is $\mathcal{T}_0$ itself. A model of U whose domain is the set of ground terms $\mathcal{T}_0$ is called a *Herbrand model* of U, and $\mathcal{T}_0$ is also called its *Herbrand universe*. In a Herbrand model $\mathfrak{T}$ the interpretations of functions and constants are canonical: $c^{\mathfrak{T}} = c$ and $f^{\mathfrak{T}}\vec{t} = f\vec{t}$ for $\vec{t} \in \mathcal{T}_0^n$, the interpretations of relations may vary though.

4.1.4 Example

Let $U \subseteq \mathcal{L}\{0, \mathtt{S}, <\}$ contain the following two universal sentences:

$$\forall x\, x < \mathtt{S}x, \qquad \forall x, y, z(x < y \wedge y < z \to x < z)$$

Then $\mathcal{N} = (\mathbb{N}, 0, \mathtt{S}, <)$ is a model of U. Since for every ground term t (meaning just for 0), $\mathtt{S}^{\mathcal{N}}t = \mathtt{S}t$, $\mathcal{N}$ is a Herbrand model of U. There are many other Herbrand models for U as $<$ may be interpreted in numerous ways.

4.1.5 Lemma

If $X \cup \{\neg\alpha \mid \alpha \in Y\}$ is inconsistent and Y is nonempty, then there exist formulas $\alpha_0, \ldots, \alpha_m \in Y$ such that $X \vdash \alpha_0 \vee \cdots \vee \alpha_m$.

We have that $X \cup \{\neg\alpha \mid \alpha \in Y\} \vdash \bot$ since it is inconsistent, and so by the compactness theorem we have that there exist $\alpha_0, \ldots, \alpha_m \in Y$ such that $X, \neg\alpha_0, \ldots, \neg\alpha_m \vdash \bot$ and so $X, \neg\alpha_0, \ldots, \neg\alpha_{m-1} \vdash \alpha_m$. By the deduction theorem we get that $X \vdash \neg\alpha_0 \to \cdots \to \neg\alpha_{m-1} \to \alpha_m = \alpha_0 \vee \cdots \vee \alpha_m$ as required. ■

4.1.6 Theorem (Herbrand's Theorem)

Let $U \subseteq \mathcal{L}$ be a set of universal formulas, $\exists\vec{x}\alpha \in \mathcal{L}$ with α quantifier-free, and let $\tilde{U}$ be the set of all instances of U. Then the following are equivalent:

(1) $U \vdash \exists\vec{x}\alpha$,

(2) $U \vdash \bigvee_{i \leq m} \alpha\frac{\vec{t}_i}{\vec{x}}$ for some m and $\vec{t}_0, \ldots, \vec{t}_m \in \mathcal{T}^n$,

(3) $\tilde{U} \vdash \bigvee_{i \leq m} \alpha\frac{\vec{t}_i}{\vec{x}}$ for some m and $\vec{t}_0, \ldots, \vec{t}_m \in \mathcal{T}^n$,

The same holds if $\mathcal{L}$ is replaced with $\mathcal{L}^k$, $\mathcal{T}$ by $\mathcal{T}_k$, and $\tilde{U}$ is the set of all $\mathcal{T}_k$-instances of U.

Since $U \vdash \tilde{U}$, we get (3) $\Longrightarrow$ (2) $\Longrightarrow$ (1). So we will show (1) $\Longrightarrow$ (3): by (1) we get that $X = U \cup \{\forall\vec{x}\neg\alpha\}$ is inconsistent and therefore by the previous theorem so is $\tilde{X} = \tilde{U} \cup \left\{\neg\alpha\frac{\vec{t}}{\vec{x}} \,\middle|\, \vec{t} \in \mathcal{T}^n\right\}$. So by the above lemma we get that there exist $\vec{t}_0, \ldots, \vec{t}_m \in \mathcal{T}^n$ such that $\tilde{U} \vdash \bigvee_{i \leq m} \alpha \frac{\vec{t}_i}{\vec{x}}$. The proof in the case of $\mathcal{L}^k$ is analogous. ■

Notice that in the case $\exists x\alpha = \exists x \forall y (ry \to rx)$ and $U = \varnothing$, then $\vdash \exists x\alpha$ as it is a tautology. But there aren't necessarily terms (variables) such that $\vdash \bigvee_{i \leq m} \alpha \frac{t_i}{x}$: if there are $m+2$ elements in the domain then we can have it not satisfy rx_i for $i \leq m$ but have it satisfy rx_{m+1}. So in Herbrand's theorem, the assumption that α is quantifier-free is necessary.

4.2 Horn Formulas

4.2.1 Definition

A **Horn formula** in a first-order language $\mathcal{L}$ is defined recursively as follows:

(1) Literals (prime formulas and their negations) are *basic Horn formulas.* If α is prime and β is a basic Horn formula, then $\alpha \to \beta$ is also a basic Horn formula.

(2) Basic Horn formulas are Horn formulas, and if α and β are Horn formulas so too are $(\alpha \wedge \beta), \forall x\alpha, \exists x\alpha$.

If a formula is equivalent to a (respectively, basic) Horn formula, it is also called a (respectively, basic) Horn formula.

Thus the general form of a basic Horn formula is $\alpha_1 \to \cdots \to \alpha_n \to \beta$ where α_i are prime and β is a literal ($n \geq 0$). This is logically equivalent to $\beta \vee \neg\alpha_1 \vee \cdots \vee \neg\alpha_n$. If β is prime, we write $\beta = \alpha_0$ and if β is the negation of a prime formula we write $\beta = \neg\alpha_0$. So up to logical equivalence, all basic Horn formulas are equivalent to one of the following (α_i prime):

$$\mathrm{I} : \alpha_0 \vee \neg\alpha_1 \vee \cdots \vee \neg\alpha_n \qquad \text{or} \qquad \mathrm{II} : \neg\alpha_0 \vee \neg\alpha_1 \vee \cdots \vee \neg\alpha_n$$

Basic Horn formulas of the first type are termed *positive* and basic Horn formulas of the second type are *negative.* Now recall that $\alpha_1 \to \cdots \to \alpha_n \to \beta \equiv \alpha_1 \wedge \cdots \wedge \alpha_n \to \beta$ and so such a formula is also a basic Horn formula.

Reviewing our proof on the existence of prenex normal forms, it is apparent that all Horn formulas are equivalent to prenex Horn formulas. If the prefix of its prenex form contains only $\forall$-quantifiers, it is called a *universal Horn formula.* If its kernel of its prenex form is the conjunction of positive basic Horn formulas it is called a *positive Horn formula.* A Horn formula without free variables is called a *Horn sentence.*

For example the following are positive Horn sentences which describe the congruence-ness of $=$:

$$(x = x)^g, \quad (x = y \wedge x = z \to y = z)^g, \quad (\vec{x} = \vec{y} \to r\vec{x} \to r\vec{y})^g, \quad (\vec{x} = \vec{y} \to f\vec{x} = f\vec{y})^g$$

A *Horn theory* is a theory T with an axiom system of Horn sentences. If these axioms are all universal Horn sentences, then T is a *universal Horn theory.* For example, the theory of groups and rings (in the appropriate languages) are universal Horn theories.

4.2.2 Theorem

Let U be a consistent set of universal Horn formulas in a language $\mathcal{L}$. Then $\mathcal{F} = \mathcal{F}U$ is a model for U. If $U \subseteq \mathcal{L}^k$ then $\mathcal{F} = \mathcal{F}_k U$ is a model for U as well.

We will show that $U \vdash \alpha \implies \mathcal{F} \vDash \alpha$ for all universal Horn formulas α. We will prove this by induction on the recursive definition of Horn formulas. For prime formulas this holds in general ($U \vdash \alpha \iff \mathcal{F} \vDash \alpha$) and thus this holds for literals as well. Now let α be prime and β be a basic Horn formula such that $U \vdash \alpha \to \beta$, then if $\mathcal{F} \vDash \alpha$ we have $U \vdash \alpha$ (again these are equivalent for prime formulas) and so $U \vdash \beta$ and thus $\mathcal{F} \vDash \beta$, meaning $\mathcal{F} \vDash \alpha \to \beta$ as required. Induction on $\wedge$ is clear. Now if $U \vdash \forall \vec{x}\alpha$ where α is quantifier-free, then let $t \in \mathcal{T}^n$ and we get $U \vdash \alpha \frac{\vec{t}}{\vec{x}}$ and so $\mathcal{F} \vDash \alpha \frac{\vec{t}}{\vec{x}}$ by induction. Since $\vec{t}$ is arbitrary and exhausts the domain of $\mathcal{F}$, we get $\mathcal{F} \vDash \forall \vec{x}\alpha$ as required. The proof for the case of $\mathcal{L}^k$ is analogous. ■

Then if U is the axiom system of a universal Horn theory T, we get that $\mathcal{F}U \vDash T$ and since $U \subseteq \mathcal{L}^k$ for all k, also $\mathcal{F}_k U \vDash T$.

Let $U \subseteq \mathcal{L}^0$ be a consistent set of $=$-free universal Horn sentences, and let T be axiomatized by U. We assume $\mathcal{L}$ contains at least one constant symbol, so $\mathcal{F}_0 U$ is well-defined and is a Herbrand model for T, called the *minimal Herbrand model for T* and is denoted by $\mathcal{C}_U$ or $\mathcal{C}_T$. The domain of $\mathcal{C}_U$ is the set of ground terms (notice that denoting it by $\mathcal{C}_T$ will not cause any confusion as all term models agree in their interpretation of function and constant symbols, and their relations are defined by $r^{\mathcal{C}_T}\vec{t} \iff U \vdash r\vec{t} \iff T \vdash r\vec{t}$).

4.2.3 Example

Let $\mathcal{N} = (\mathbb{N}, 0, \mathtt{S}, <)$ and U be the set of universal Horn sentences:

$$(1)\ \forall x\, x < \mathtt{S}x, \qquad (2)\ \forall x, y, z(x < y \land y < z \to x < z)$$

We will show that $\mathcal{N} \cong \mathcal{C}_U$ with the isomorphism $n \mapsto \underline{n}$. Since $\mathcal{C}_U \vDash \underline{m} < \underline{k} \iff U \vdash \underline{m} < \underline{k}$, to show that this is an isomorphism we must show only that $m < k \iff U \vdash \underline{m} < \underline{k}$. We prove $\implies$ by induction on k, starting with $k = \mathtt{S}m$ which is clear in lieu of (1). Now if $m < \mathtt{S}k$, then by the induction hypothesis we get that $m < k$ or $m = k$, and in both cases we get $U \vdash \underline{m} < \underline{\mathtt{S}k}$ by (1) and (2).

So this mapping preserves $<$, and it also preserves 0 and $\mathtt{S}$ trivially. It is also injective as $\underline{m} = \underline{k}$ if and only if $m = k$ as the domain of $\mathcal{C}_U$ is the set of all ground terms (and not a quotient of them). It is also surjective as the set of ground terms of this language is the set of all terms of the form $\underline{n}$, so it is indeed an isomorphism.

4.2.4 Theorem

Let $U \subseteq \mathcal{L}^k$ be a consistent set of universal Horn formulas, and $\gamma = \gamma_0 \land \cdots \land \gamma_m$ be a conjunction of prime formulas, and $\exists\vec{x}\gamma \in \mathcal{L}^k$. Then the following are equivalent:

(1) $\mathcal{F}_k U \vDash \exists\vec{x}\gamma$,

(2) $U \vdash \gamma\frac{\vec{t}}{\vec{x}}$ for some $\vec{t} \in \mathcal{T}_k^n$,

(3) $U \vdash \exists\vec{x}\gamma$.

This theorem, and proof, holds if we replace $\mathcal{L}^k$ with $\mathcal{L}$ and $\mathcal{T}_k$ with $\mathcal{T}$.

(1) $\implies$ (2): if $\mathcal{F}_k U \vdash \exists\vec{x}\gamma$ then if $\mathcal{F}_k U \nvDash \gamma\frac{\vec{t}}{\vec{x}}$ for all $\vec{t} \in \mathcal{T}_k^n$, then we get that $\mathcal{F}_k U \vDash \forall\vec{x}\neg\gamma$ which contradicts (1). So $\mathcal{F}_k U \vdash \gamma_i\frac{\vec{t}}{\vec{x}}$ for all i, and since this is prime we get that $U \vdash \gamma_i\frac{\vec{t}}{\vec{x}}$ and thus $U \vdash \gamma\frac{\vec{t}}{\vec{x}}$. (2) $\implies$ (3) holds in general. (3) implies $\mathcal{F}_k U \vDash \exists\vec{x}\gamma$ since $\mathcal{F}_k U$ is a model for U. ■

Notice then that if T is a consistent universal Horn theory in a $=$-free language with constants, $\mathcal{C}_T \vDash \exists\vec{x}\gamma$ is equivalent to $\vdash_T \exists\vec{x}\gamma$. This is since we can take $k = 0$ and so $\mathcal{C}_T = \mathcal{F}_0 U \vDash \exists\vec{x}\gamma \iff U \vdash \exists\vec{x}\gamma \iff \vdash_T \exists\vec{x}\gamma$ by (1) $\iff$ (3).

Exercise

Let T be a Horn theory, show that $\operatorname{Md} T$ is closed under direct products. And if T is a universal Horn theory, it is also closed under substructures.

To show this we need to show that if φ is a Horn formula, then if $\mathcal{A}_i \vDash \varphi$ for all $i \in I$ then $\mathcal{B} = \prod_{i\in I} \mathcal{A}_i \vDash \varphi$. We will show this by induction on the construction of the Horn formula φ. For prime formulas, this is a direct result of the definition of the direct product of structures: $\mathcal{B} \vDash r\vec{t}$ if and only if $r^{\mathcal{A}_i}\vec{t}^{\mathcal{A}_i}$ ie. if and only if $\mathcal{A}_i \vDash r\vec{t}$. So here the converse holds as well, meaning $\mathcal{B} \vDash \pi$ if and only if $\mathcal{A}_i \vDash \pi$ for all $i \in I$. This holds for prime formulas of the form $t = s$ as well. Now if $\mathcal{A}_i \vDash \alpha \to \beta$ where α is prime and β is a basic Horn formula, suppose $\mathcal{B} \vDash \beta$ then $\mathcal{A}_i \vDash \beta$ for all $i \in I$ by above and so $\mathcal{A}_i \vDash \beta$ for all $i \in I$ and by induction $\mathcal{B} \vDash \beta$, thus $\mathcal{B} \vDash \alpha \to \beta$ as required.

The induction step for $\land$ is obvious. Now if φ is a Horn formula and $\mathcal{A}_i \vDash \forall x\varphi$, then let us add a constant symbol $\boldsymbol{a}$ to the language and so for every $a_i \in \mathcal{A}_i$ if we interpret $\boldsymbol{a}$ by a_i, then in $\mathcal{B}$ it is interpreted as $(a_i)_{i\in I}$. Now we also have that $\mathcal{A}_i \vDash \varphi\frac{\boldsymbol{a}}{x}$ and so by induction we have $\mathcal{B} \vDash \varphi\frac{\boldsymbol{a}}{x}$. Since the interpretation of $\boldsymbol{a}$ is arbitrary, this means that $\mathcal{B} \vDash \forall x\varphi$ as required. The induction step for $\exists$ proceeds analogously.

For a universal Horn theory, we proceed by induction as well. The step for basic Horn formulas is due to The Substructure Theorem. Now if $\mathcal{A} \vDash \forall\vec{x}\alpha$ where α is the conjunction of basic Horn formulas, and $\mathcal{A}' \subseteq \mathcal{A}$ then by corollary 2.3.9 we get that $\mathcal{A}' \vDash \forall\vec{x}\alpha$ as required.

Exercise

Prove that a set of positive Horn formulas is always consistent.

A positive Horn formula is equivalent to a formula of the form $\mathtt{Q}_1 \cdots \mathtt{Q}_n(\alpha_1 \wedge \cdots \wedge \alpha_m)$ where α_i are all positive (basic) Horn formulas. We can simply take the structure of one element, and so this formula is equivalent to $\alpha_1 \wedge \cdots \wedge \alpha_m$, and so we must show that the one element structure satisfies any positive basic Horn formula and that is sufficient. A positive basic Horn formula is of the form $\alpha_0 \vee \neg\alpha_1 \vee \cdots \vee \neg\alpha_k$ where α_i are prime. In the structure of one element, we can interpret relations as always being true, and so α_0 is always satisfied, as required.

5 Elements of Model Theory

5.1 Elementary Extensions

Suppose $\mathcal{L}$ is a first-order language and A is a set, then let us denote $\mathcal{L}A$ the language obtained by adjoining constant symbols $\boldsymbol{a}$ to $\mathcal{L}$ for every $a \in A$. Having $\boldsymbol{a}$ be boldface is in order to distinguish it from a, but we will later remove the boldface. Let $\mathcal{B}$ be an $\mathcal{L}$-structure, and $A \subseteq B$ be a subset of the domain of $\mathcal{B}$, then let us define the $\mathcal{L}A$-expansion of $\mathcal{B}$ by interpreting $\boldsymbol{a}$ as $a \in A$, and this structure will be denoted $\mathcal{B}_A$. We previously showed in an exercise that for every $\alpha(\vec{x}) \in \mathcal{L}$ and $\vec{a} \in A^n$,

$$\mathcal{B} \vDash \alpha[\vec{a}] \iff \mathcal{B}_A \vDash \alpha(\vec{\boldsymbol{a}}) \qquad (\alpha(\vec{\boldsymbol{a}}) := \alpha\tfrac{\boldsymbol{a}_1 \cdots \boldsymbol{a}_n}{x_1 \cdots x_n})$$

Notice that every sentence in $\mathcal{L}A$ is of the form $\alpha(\vec{\boldsymbol{a}})$ for suitable $\alpha(\vec{x}) \in \mathcal{L}$ and $\vec{a} \in A^n$. So instead of $\mathcal{B}_A \vDash \alpha(\vec{\boldsymbol{a}})$ we may just write $\mathcal{B}_A \vDash \alpha(\vec{a})$ or $\mathcal{B} \vDash \alpha(\vec{a})$. Meaning we may sometimes disregard the distinction between $\mathcal{B}$ and its extension $\mathcal{B}_A$ if there can be no misunderstandings.

5.1.1 Definition

Let $\mathcal{A}$ be an $\mathcal{L}$-structure, then its **diagram** is the set of all variable-free literals $\lambda \in \mathcal{L}A$ such that $\mathcal{A}_A \vDash \lambda$. $\mathcal{A}$'s diagram is denoted by $D\mathcal{A}$.

So for example, $D(\mathbb{R}, <)$ contains for every $a, b \in \mathbb{R}$ one of the following, depending on the relationship between a and b: $\boldsymbol{a} = \boldsymbol{b}$, $\boldsymbol{a} \neq \boldsymbol{b}$, $\boldsymbol{a} < \boldsymbol{b}$, or $\boldsymbol{b} < \boldsymbol{a}$.

In general if $\mathcal{A}$ can be embedded into $\mathcal{B}$ then this means that $\mathcal{A}$ is isomorphic to a substructure of $\mathcal{B}$, and so we can view $\mathcal{A}$ as a substructure of $\mathcal{B}$. Suppose $\mathcal{L}_0 \subseteq \mathcal{L}$, then we may say that an $\mathcal{L}_0$-structure $\mathcal{A}$ can be embedded into an $\mathcal{L}$-structure $\mathcal{B}$ to mean that $\mathcal{A}$ can be embedded into the $\mathcal{L}_0$-reduct of $\mathcal{B}$. And so we may also write $\mathcal{A} \subseteq \mathcal{B}$ in this case. So for example the ring $\mathbb{Z}$ is embeddable into the field $\mathbb{Q}$ in this sense.

5.1.2 Proposition

Let $\mathcal{L}_0 \subseteq \mathcal{L}$, $\mathcal{A}$ be an $\mathcal{L}_0$-structure, and $\mathcal{B}$ an $\mathcal{L}A$-structure. Then $\mathcal{B} \vDash D\mathcal{A}$ if and only if $\iota\colon a \mapsto \boldsymbol{a}^{\mathcal{B}}$ is an embedding of $\mathcal{A}$ into $\mathcal{B}$.

If $\mathcal{B} \vDash D\mathcal{A}$ then for every $a \neq b$, we have that $\mathcal{B} \vDash \boldsymbol{a} \neq \boldsymbol{b}$ and so $\iota a \neq \iota b$, meaning ι is injective. And for relational symbols r,

$$r^{\mathcal{A}}\vec{a} \iff r\vec{\boldsymbol{a}} \in D\mathcal{A} \iff \mathcal{B} \vDash r\vec{\boldsymbol{a}} \iff r^{\mathcal{B}}\iota\vec{a}$$

Similarly for function symbols f,

$$f^{\mathcal{A}}\vec{a} = b \iff f\vec{\boldsymbol{a}} = \boldsymbol{b} \in D\mathcal{A} \iff \mathcal{B} \vDash f\vec{\boldsymbol{a}} = \boldsymbol{b} \iff f^{\mathcal{B}}\iota\vec{a} = \iota b = \iota f\vec{a}$$

Thus ι is an injective mapping which preserves relational and function symbols, meaning it is an embedding of $\mathcal{A}$ into $\mathcal{B}$. Now conversely, suppose ι is an embedding. For variable-free terms t in $\mathcal{L}_0A$, it is easy to verify that $\iota t^{\mathcal{A}} = t^{\mathcal{B}}$ by term induction ($t^{\mathcal{A}}$ here is to be read as $t^{\mathcal{A}_A}$). So for variable-free equations $t_1 = t_2$ in $\mathcal{L}_0A$,

$$t_1 = t_2 \in D\mathcal{A} \iff t_1^{\mathcal{A}} = t_2^{\mathcal{A}} \iff \iota t_1^{\mathcal{B}} = \iota t_2^{\mathcal{B}} \iff t_1^{\mathcal{B}} = t_2^{\mathcal{B}} \iff B \vDash t_1 = t_2$$

we can similarly show this for inequalities and literals of the form $r\vec{\boldsymbol{a}}$ and its negation. So $B \vDash D\mathcal{A}$ as required. ∎

5.1.3 Corollary

Let $\mathcal{A}$ and $\mathcal{B}$ be $\mathcal{L}$-structures. Then $\mathcal{A}$ is embeddable into $\mathcal{B}$ if and only if $\mathcal{B}$ has an $\mathcal{L}A$-extension $\mathcal{B}'$ such that $\mathcal{B}' \vDash D\mathcal{A}$. Furthermore, if $A \subseteq B$, then $\mathcal{B}_A \vDash D\mathcal{A}$ if and only if $\mathcal{A} \subseteq \mathcal{B}$.

Using the above proposition with $\mathcal{L} = \mathcal{L}_0$, we get that if $\mathcal{B}' \vDash D\mathcal{A}$ then $\iota\colon a \mapsto a^{\mathcal{B}'}$ embeds $\mathcal{A}$ into $\mathcal{B}$. And if ι is an embedding, then define $\boldsymbol{a}^{\mathcal{B}'} = \iota a$, and then $\mathcal{B}' \vDash D\mathcal{A}$. If $A \subseteq B$ then ι is the identity mapping, meaning $\mathcal{B}' = \mathcal{B}_A$. ∎

5.1.4 Definition

A **prime model** of a theory T is a T-model $\mathcal{A}$ which can be embedded into every other T model.

This corollary means that $\mathcal{A}_A$ is a prime model of the theory defined by $D\mathcal{A}$.

5.1.5 Definition

Let $\mathcal{L}$ be a first-order language and $\mathcal{A}, \mathcal{B}$ be $\mathcal{L}$-structures. Then $\mathcal{A}$ is an **elementary substructure** of $\mathcal{B}$ (and $\mathcal{B}$ is an **elementary extension** of $\mathcal{A}$), denoted $\mathcal{A} \preceq \mathcal{B}$, if $A \subseteq B$ and

$$\mathcal{A} \vDash \alpha[\vec{a}] \iff \mathcal{B} \vDash \alpha[\vec{a}], \text{ for all } \alpha = \alpha(\vec{x}) \in \mathcal{L} \text{ and } \vec{a} \in A^n$$

Further let us define $\mathcal{A}$'s **elementary diagram** to be the set of $\mathcal{L}A$-sentences valid in $\mathcal{A}$:

$$D_{el}\mathcal{A} := \left\{\alpha \in \mathcal{L}A^0 \mid \mathcal{A}_A \vDash \alpha\right\}$$

By The Substructure Theorem, if $\mathcal{A} \preceq \mathcal{B}$ then $\mathcal{A} \subseteq \mathcal{B}$, but the converse is not generally true. In fact being an elementary substructure is a very strong condition. It is obvious that $\mathcal{A} \preceq \mathcal{B}$ is equivalent to $A \subseteq B$ and $\mathcal{B} \vDash D_{el}\mathcal{A}$, as $\mathcal{A} \vDash \alpha[\vec{a}] \iff \mathcal{A}_A \vDash \alpha(\vec{\boldsymbol{a}})$. Notice that this is also equivalent to $A \subseteq B$ and $\mathcal{A}_A \equiv_{\mathcal{L}A} \mathcal{B}_A$. So being an elementary substructure implies elementary equivalence, but it is also stronger: take for example $\mathcal{A} = (\mathbb{N}_+, <)$ and $\mathcal{B} = (\mathbb{N}, <)$ then $A \subseteq B$ and since $\mathcal{A} \cong \mathcal{B}$, so we have that $\mathcal{A} \equiv \mathcal{B}$. But they are not equivalent modulo $\mathcal{L}A$; $\mathcal{A}$ is not an elementary substructure of $\mathcal{B}$ as $\exists x\, x < \mathbf{1}$ is true in $\mathcal{B}$ but not $\mathcal{A}$.

5.1.6 Theorem (Tarski's Criterion)

Let $\mathcal{A}, \mathcal{B}$ be $\mathcal{L}$-structures with $A \subseteq B$, then the following are equivalent:

(1) $\mathcal{A} \preceq \mathcal{B}$,

(2) For all $\varphi(\vec{x}, y) \in \mathcal{L}$ and $\vec{a} \in A^n$, one has $\mathcal{B} \vDash \exists y \varphi(\vec{a}, y)$ implies $\mathcal{B} \vDash \varphi(\vec{a}, a)$ for some $a \in A$. In other words every existential formula, if witnessed in $\mathcal{B}$, is witnessed in $\mathcal{A}$.

If $\mathcal{A} \preceq \mathcal{B}$ and $\mathcal{B} \vDash \exists y \varphi(\vec{a}, y)$ then $\mathcal{A} \vDash \exists y \varphi(\vec{a}, y)$ as a substructure. And so $\mathcal{A} \vDash \varphi(\vec{a}, a)$ for some $a \in A$ and so $\mathcal{B} \vDash \varphi(\vec{a}, a)$ since $\mathcal{B}$ is an elementary extension. Conversely, notice that the condition for elementary extensions holds for quantifier-free formulas (and in particular prime formulas). The induction step for $\neg$ and $\wedge$ are obvious, so all we must show is

$$\begin{aligned} \mathcal{A} \vDash \forall y \varphi(\vec{a}, y) &\iff \mathcal{A} \vDash \varphi(\vec{a}, a) \text{ for all } a \in A \\ &\iff \mathcal{B} \vDash \varphi(\vec{a}, a) \text{ for all } a \in A \quad \text{(induction hypothesis)} \\ &\iff \mathcal{B} \vDash \forall y \varphi(\vec{a}, y) \quad \text{(see below)} \end{aligned}$$

One direction of the final equivalence is trivial, we will prove the converse ($\Longrightarrow$) by contrapositive: if $\mathcal{B} \nvDash \forall y \varphi(\vec{a}, y)$ then $\mathcal{B} \vDash \exists y \neg\varphi(\vec{a}, y)$ and therefore $\mathcal{B} \vDash \neg\varphi(\vec{a}, a)$ for some $a \in A$ and so we cannot have $\mathcal{B} \vDash \varphi(\vec{a}, a)$ for all $a \in A$. ■

Tarski's Criterion allows us to prove nontrivial elementary extensions, for example we can utilize the following:

5.1.7 Theorem

Let $\mathcal{A} \subseteq \mathcal{B}$, and suppose that for all n and $\vec{a} \in A^n$ and $b \in B$, there exists an automorphism $\iota\colon \mathcal{B} \longrightarrow \mathcal{B}$ such that $\iota\vec{a} = \vec{a}$ and $\iota b \in A$. Then $\mathcal{A} \preceq \mathcal{B}$.

We will prove that Tarski's criterion holds: suppose $\mathcal{B} \vDash \exists y \varphi(\vec{a}, y)$ and so $\mathcal{B} \vDash \varphi(\vec{a}, b)$ for some $b \in B$. Then let ι be an automorphism of $\mathcal{B}$ where $\iota\vec{a} = \vec{a}$ and $\iota b \in B$, then $\mathcal{B} \vDash \varphi(\iota\vec{a}, \iota b) = \varphi(\vec{a}, a)$ with $a = \iota b \in A$. And so we have proven Tarski's criterion, meaning $\mathcal{A} \preceq \mathcal{B}$ as required. ■

Notice that we need consider only if $b \notin A$, as otherwise we can just take the identity mapping as the automorphism.

5.1.8 Example

We claim that $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$. Indeed, if $a_1 < \cdots < a_n \in \mathbb{Q}$ and $b \in \mathbb{R}$ then we consider three cases: $b < a_1$, $a_i < b < a_{i+1}$, and $a_n < b$.

(1) If $b < a_1$, then by the density of $\mathbb{Q}$ in $\mathbb{R}$ there must exist a rational $b < q < a_1$, and so let us define ι to map $x \mapsto x$ for $x \geq a_1$. And for $x < a_1$ let us define ι to be the linear function connecting (b, q) to (a_1, a_1), which is increasing and therefore an automorphism.

(2) If $a_i < b < a_{i+1}$ we can similarly take a rational $b < q < a_{i+1}$, and have ι be the identity outside the interval $[a_i, a_{i+1}]$ and within this interval we split it up into the linear function connecting (a_i, a_i) with (b, q) and the linear function connecting (b, q) with (a_{i+1}, a_{i+1}). This is again an automorphism.

(3) The case $b > a_n$ can be treated analogously to the case for $b < a_1$.

So we have satisfied the condition for the above theorem, proving that $(\mathbb{Q}, <)$ is indeed an elementary substructure of $(\mathbb{R}, <)$.

5.1.9 Theorem (Downward Löwenheim-Skolem Theorem)

Suppose $\mathcal{B}$ is an $\mathcal{L}$-structure such that $|\mathcal{L}| \leq |\mathcal{B}|$, and let $A_0 \subseteq B$ be arbitrary. Then $\mathcal{B}$ has an elementary substructure $\mathcal{A}$ of cardinality $\leq \max\{|A_0|, |\mathcal{L}|\}$ such that $A_0 \subseteq A$.

We will inductively define a sequence $A_0 \subseteq A_1 \subseteq \cdots \subseteq B$ as follows: assuming we have constructed A_k, for every $\alpha = \alpha(\vec{x}, y)$ and $\vec{a} \in A_k^n$ such that $\mathcal{B} \vDash \exists y \alpha(\vec{a}, y)$, then arbitrarily choose a $b \in B$ such that $\mathcal{B} \vDash \alpha(\vec{a}, b)$ and add b to A_k to obtain A_{k+1}. In particular, if α is $f\vec{x} = y$ then certainly $\mathcal{B} \vDash \exists! y\, f\vec{a} = y$ and so $f^{\mathcal{B}}\vec{a} \in A_{k+1}$. Thus $A = \bigcup_{k=0}^{\infty} A_k$ is closed under the operations of $\mathcal{B}$, and therefore defines a substructure $\mathcal{A} \subseteq \mathcal{B}$. And we will show that $\mathcal{A} \preceq \mathcal{B}$ by Tarski's criterion: if $\mathcal{B} \vDash \exists y \varphi(\vec{a}, y)$ for $\vec{a} \in A^n$, then there must be some k such that $\vec{a} \in A_k^n$ and so $\mathcal{B} \vDash \varphi(\vec{a}, b)$ for some $b \in \mathcal{B}$ and by definition one of these bs is in $A_{k+1} \subseteq A$. Thus $\mathcal{A}$ is indeed an elementary substructure.

Now all that remains to be shown is that $|A| \leq \kappa := \max\{|A_0|, |\mathcal{L}|\}$. Notice that there are at most κ formulas and κ finite sequences of elements in A_0, and so we adjoin at most κ new elements to A_0 in order to obtain A_1, meaning $|A_1| \leq \kappa$. And inductively we have that $|A_n| \leq \kappa$, meaning $|A| = |\bigcup A_k| \leq \kappa$ as required. ∎

5.1.10 Theorem (Upward Löwenheim-Skolem Theorem)

Let $\mathcal{C}$ be any infinite $\mathcal{L}$-structure and $\kappa \geq \max\{|\mathcal{C}|, |\mathcal{L}|\}$ then there exists an elementary extension $\mathcal{A} \succeq \mathcal{C}$ with $|\mathcal{A}| = \kappa$.

Let $D \supseteq C$ with a cardinality of κ. Since the alphabet of $\mathcal{L}D$ has a cardinality of κ, $|\mathcal{L}D| = \kappa$, and since $|\mathcal{C}| \geq \aleph_0$. Now, $D_{el}\mathcal{C} \cup \{c \neq d \mid c \neq d \in D\}$ is finitely satisfiable: every finite subset contains only finitely many rules for $c \neq d \in D$ and so we can interpret these constants as elements of $\mathcal{C}$, and so it is satisfiable by the compactness theorem. Let this model be $\mathcal{B}$. And since $d \mapsto d^{\mathcal{B}}$ is injective, we can assume $d^{\mathcal{B}} = d$ for all $d \in D$, ie. $D \subseteq B$. By the downward Löwenheim-Skolem theorem with $\mathcal{L} = \mathcal{L}D$ and $A_0 = D$, there exists some $\mathcal{A} \preceq \mathcal{B}$ with $D \subseteq A$ and $\kappa = |D| \leq |A| \leq \max\{|D|, |\mathcal{L}D|\} = \kappa$, thus $|A| = \kappa$. And since $C \subseteq D$ and $\mathcal{A} \equiv_{\mathcal{L}D} \mathcal{B} \vDash D_{el}\mathcal{C}$, we have that $\mathcal{A} \vDash D_{el}\mathcal{C}$ and since $C \subseteq A$, this means that $\mathcal{C} \preceq \mathcal{A}$, as required. ∎

In particular, if T is a countable satisfiable theory then the downward Löwenheim-Skolem theorem tells us that it must have a countable model (and so the previous Löwenheim-Skolem theorem is just a special case of the downward version), and then the upward version tells us that it must have models of every infinite cardinality.

Exercise

An embedding $\iota\colon \mathcal{A} \longrightarrow \mathcal{B}$ is **elementary** if $\iota\mathcal{A} \preceq \mathcal{B}$, where $\iota\mathcal{A}$ is the image of $\mathcal{A}$ under ι. Show that an $\mathcal{L}A$-structure $\mathcal{B}$ is a model of $D_{el}\mathcal{A}$ if and only if $\mathcal{A}$ is elementary embeddable into $\mathcal{B}$.

If $\mathcal{B} \vDash D_{el}\mathcal{A}$ then define $\iota\colon a \mapsto a^{\mathcal{B}}$, this is an embedding since $\mathcal{B} \vDash D\mathcal{A}$. So now it remains to be shown that

$\iota\mathcal{A} \preceq \mathcal{B}$, so let $\varphi(\vec{x}) \in \mathcal{L}$ and $\iota\vec{a} \in \iota A^n$. Then

$$\iota\mathcal{A} \vDash \varphi(\iota\vec{a}) \iff \mathcal{A} \vDash \varphi(\vec{a}) \iff \varphi(\vec{a}) \in D_{el}\mathcal{A} \iff \mathcal{B} \vDash \varphi(\vec{a}^{\mathcal{B}}) \iff \mathcal{B} \vDash \varphi(\iota\vec{a})$$

so we indeed have that $\iota\mathcal{A} \preceq \mathcal{B}$.

Since ι preserves constants, we have that $\iota a = a^{\mathcal{B}}$, and so for $\varphi(\vec{a}) \in D_{el}\mathcal{A}$, since $\mathcal{A} \vDash \varphi(\vec{a})$, we must have that $\mathcal{B} \vDash \varphi(\iota\vec{a}) = \varphi(\vec{a}^{\mathcal{B}})$, thus $\mathcal{B} \vDash \varphi(\vec{a})$ as required.

Exercise

Suppose $\mathcal{A} \equiv \mathcal{B}$, show that there exists a $\mathcal{C}$ in which both $\mathcal{A}$ and $\mathcal{B}$ can be embedded elementarily.

We simply must show that $\mathcal{D}_{el}\mathcal{A} \cup \mathcal{D}_{el}\mathcal{B}$ is consistent; otherwise by the compactness theorem there must be a $\gamma(\vec{b}) \in \mathcal{D}_{el}\mathcal{B}$ such that $\mathcal{D}_{el}\mathcal{A}, \gamma(\vec{b}) \vdash \bot$ (since elementary diagrams are closed under conjunctions). Thus $\mathcal{D}_{el}\mathcal{A} \vdash \neg\gamma(\vec{b})$. We can assume without loss of generality that $A \cap B = \varnothing$ in which case by ($\forall$2) we have that $\mathcal{D}_{el}\mathcal{A} \vdash \forall\vec{x}\neg\gamma$. And so $\mathcal{A} \vDash \forall\vec{x}\neg\gamma$ and since $\mathcal{A}$ and $\mathcal{B}$ are equivalent, we get $\mathcal{B} \vDash \forall\vec{x}\neg\gamma$ and in particular $\mathcal{B} \vDash \neg\gamma(\vec{b})$. But this contradicts $\mathcal{B} \vDash \gamma(\vec{b}) \in \mathcal{D}_{el}\mathcal{B}$.

Exercise

Let $\mathcal{A}$ be an $\mathcal{L}$-structure generated by $G \subseteq A$, and $\mathcal{T}_G$ the set of ground terms in $\mathcal{L}G$. Prove that

(1) For every $a \in A$ there exists some $t \in \mathcal{T}_G$ such that $a = t^{\mathcal{A}}$,

(2) if $\mathcal{A} \vDash T$ and $\mathcal{D}A \vdash_T \alpha \in \mathcal{L}G$ then $\mathcal{D}_G\mathcal{A} \vdash_T \alpha$ where $\mathcal{D}_G\mathcal{A} = \mathcal{D}A \cap \mathcal{L}G$.

(1) Let us define $\mathcal{B} = \{t^{\mathcal{A}} \mid t \in \mathcal{T}_G\}$, this is a substructure of $\mathcal{A}$: if $\vec{t} \in \mathcal{T}_G^n$ and f is a function symbol, then $f\vec{t} \in \mathcal{T}_G$. And so for $\vec{t}^{\mathcal{A}} \in \mathcal{B}^n$, $f^{\mathcal{A}}\vec{t}^{\mathcal{A}} = (f\vec{t})^{\mathcal{A}} \in \mathcal{B}$ so $\mathcal{B}$ is closed under the operations of $\mathcal{A}$ as required. And furthermore, $G \subseteq B$ since for every $g \in G$, $\boldsymbol{g} \in \mathcal{T}_G$ and so $\boldsymbol{g}^{\mathcal{A}} = g \in \mathcal{B}$. But since G generates $\mathcal{A}$, $\mathcal{A}$ must be the smallest structure containing G, meaning $\mathcal{A} = \mathcal{B}$ as required.

(2) Let $\mathcal{B}$ be an $\mathcal{L}G$-structure where $\mathcal{B} \vDash \mathcal{D}_G\mathcal{A}, T$ then let us define $\iota\colon \mathcal{A} \longrightarrow \mathcal{B}$ as follows: for every $a \in A$, there exists a $t_a \in \mathcal{T}_G$ such that $t_a^{\mathcal{A}} = a$, choose any of these terms and set $\iota a = t_a^{\mathcal{B}}$. We claim then that ι is an embedding: firstly it is well-defined since any choice of t_a will give the same function, if $t^{\mathcal{A}} = s^{\mathcal{A}}$ then $t = s \in \mathcal{D}_G\mathcal{A}$ and so $\mathcal{B} \vDash t = s$ and so $t^{\mathcal{B}} = s^{\mathcal{B}}$. Now if f is a function and $\vec{a} \in A^n$, let $b = f\vec{a}$ then suppose $\vec{t}^{\mathcal{A}} = \vec{a}$ and $s^{\mathcal{A}} = b$, then $f\vec{t} = s \in \mathcal{D}_G\mathcal{A}$, and so $f\vec{t}^{\mathcal{B}} = s^{\mathcal{B}}$ meaning $f\iota\vec{a} = \iota b = \iota f\vec{a}$ as required. And similar for relations.

Let us now define an $\mathcal{L}A$-structure, $\mathcal{B}'$, as follows: for every $a \in A$ let $a^{\mathcal{B}'} = \iota a$. $\mathcal{B}'$ is an extension of $\mathcal{B}$, and so by the coincidence theorem $\mathcal{B} \equiv_{\mathcal{L}G} \mathcal{B}'$. And ι is an embedding of $\mathcal{A}$ into $\mathcal{B}'$ so $\mathcal{B}' \vDash \mathcal{D}A, T$ and thus $\mathcal{B}' \vDash \alpha$ and since $\mathcal{B}$ is equivalent to $\mathcal{B}'$ we get that $\mathcal{B} \vDash \alpha$. Since this is true for every model of $\mathcal{D}_G\mathcal{A}, T$, we get that $\mathcal{D}_G\mathcal{A} \vdash_T \alpha$ as required.

5.2 Complete and κ-Categorical Theories

Recall that a theory $T \subseteq \mathcal{L}^0$ is consistent by definition if it has no consistent extensions $T \subsetneq T' \subseteq \mathcal{L}^0$. Note that a complete theory T may have a consistent extension in $\mathcal{L}$, as in general neither $\vdash_T x = y$ nor $\vdash_T x \neq y$, so $T \cup \{x = y\}$ for example may be consistent. But this of course is not interesting: we want to deal with theories, not arbitrary sets of formulae.

We can equivalently characterize completeness as follows:

5.2.1 Theorem

For a consistent theory T, the following are equivalent:

(1) T is complete,

(2) $T = Th\mathcal{A}$ for every $\mathcal{A} \vDash T$,

(3) $\mathcal{A} \equiv \mathcal{B}$ for every $\mathcal{A}, \mathcal{B} \vDash T$,

(4) $\vdash_T \alpha \vee \beta$ implies $\vdash_T \alpha$ or $\vdash_T \beta$ for any $\alpha, \beta \in \mathcal{L}^0$,

(5) $\vdash_T \alpha$ or $\vdash_T \neg\alpha$ for every $\alpha \in \mathcal{L}^0$.

(1) $\Longrightarrow$ (2): obviously if $\mathcal{A} \vDash T$, $T \subseteq Th\mathcal{A}$ and since $Th\mathcal{A}$ is consistent and an extension of T, there must be equality by definition of completeness. (2) $\Longrightarrow$ (3): if $\mathcal{A}, \mathcal{B} \vDash T$ then $Th\mathcal{A} = Th\mathcal{B} = T$ and so $\mathcal{A} \vDash \mathcal{B}$. (3) $\Longrightarrow$ (4): let $\mathcal{A} \vDash T$ then $\mathcal{A} \vDash \alpha \vee \beta$ and so $\mathcal{A} \vDash \alpha$ or $\mathcal{A} \vDash \beta$. Suppose that $\mathcal{A} \vDash \alpha$, then if $\mathcal{B} \vDash T$ as well we have $Th\mathcal{A} = Th\mathcal{B}$ and so $\mathcal{B} \vDash \alpha$, meaning that α is modeled by every model of T, and so $\vdash_T \alpha$ as required. (4) $\Longrightarrow$ (5): this is a specific case of (4) for $\beta = \neg\alpha$. (5) $\Longrightarrow$ (1): let $T \subset T'$ and let $\alpha \in T' \setminus T$, then $\nvdash_T \alpha$ so $\vdash_T \neg\alpha$ and therefore $\vdash_{T'} \neg\alpha, \alpha$ meaning T' is inconsistent. Thus T is complete as all of its extensions are inconsistent. ∎

Notice that if we take the inconsistent theory $T = \mathcal{L}^0$ to be complete, this theorem still holds even if we remove the condition that T be consistent.

5.2.2 Definition

A theory T is **κ-categorical** for a cardinal κ if all of its models of cardinality κ are isomorphic.

5.2.3 Example

A trivial example is the theory $\mathit{Taut}_=$ of tautologies in $\mathcal{L}_=$. This theory is κ-categorical for all cardinalities κ as all models of $\mathit{Taut}_=$ are simply sets, and so if they have the same cardinality they are by definition isomorphic.

5.2.4 Example

The theory DO is the theory of densely ordered sets: it is obtained by adjoining the following two axioms:

$$\exists x \exists y (x \neq y), \qquad \forall x \forall y \exists z (x < y \to x < z \wedge z < y)$$

Let us further define $\mathsf{L} = \exists x \forall y (x \leq y)$ and $\mathsf{R} = \exists x \forall y (y \leq x)$. And so we define extensions of DO as follows: $\mathsf{DO}_{ij} = \mathsf{DO} + i\mathsf{L} + j\mathsf{R}$ where $i, j \in \{0, 1\}$ and $i\varphi$ is to mean φ for $i = 1$ and $\neg\varphi$ for $i = 0$. Thus for example $\mathsf{DO}_{00} = \mathsf{DO} + \neg\mathsf{L} + \neg\mathsf{R}$ is the theory of densely ordered sets without endpoints.

It turns out that all DO_{ij} are $\aleph_0$-categorical, we will show this for DO_{00} (this proof is due to Cantor). Let us call a function f with $\mathit{dom} f \subseteq M$ and $\mathit{ran} f \subseteq N$ **partial function** from M to N. Now let $A = \{a_0, a_1, \ldots\}$ and $B = \{b_0, b_1, \ldots\}$ be countable models of DO_{00}. We will define a sequence of partial functions from A to B, $\{f_i\}_{i=0}^{\infty}$, by induction. We begin with $f_0 a_0 = b_0$. Now assuming that f_n has been constructed such that $a < a' \iff f_n a < f_n a'$ for all $a, a' \in \mathit{dom} f$ and $\{a_0, \ldots, a_n\} \subseteq \mathit{dom} f_n$ and $\{b_0, \ldots, b_n\} \subseteq \mathit{ran} f_n$ (such a function is caled a **partial isomorphism**), we construct f_n (note that for $n = 0$ these conditions have been satisfied trivially).

Let m be the minimum index such that $a_m \in A \setminus \mathit{dom} f_n$, choose $b \in B \setminus \mathit{ran} f_n$ such that $g_n := f_n \cup \{(a_m, b)\}$ is a partial isomorphism. Such a b exists since B is dense. Now similarly let m be the minimum index such that $b_m \in B \setminus \mathit{ran} g_n$, and choose $a \in A \setminus \mathit{dom} f_n$ such that $f_{n+1} := g_n \cup \{(a, b_m)\}$ is a partial isomorphism.

Notice that by this construction $a_n \in \mathit{dom} f_n$ and $b_n \in \mathit{ran} f_n$, so let us define $f = \bigcup_{n=0}^{\infty} f_n$. f is certainly a bijection between A and B, and it is an isomorphism since for every $x, y \in A$, $x, y \in \mathit{dom} f_n$ for some n and so $x < y \iff f_n x < f_n y \iff fx < fy$ as f_n is a partial isomorphism.

For $\mathcal{A}, \mathcal{B} \vDash \mathsf{DO}_{ij}$ of cardinality $\aleph_0$, let $\mathcal{A}', \mathcal{B}'$ be the DO_{00}-models obtained by removing the endpoints from $\mathcal{A}$ and $\mathcal{B}$ respectively. Then as shown above $\mathcal{A}' \cong \mathcal{B}'$. We can map the endpoints of $\mathcal{A}$ to their respective endpoints in $\mathcal{B}$ and adjoin this to this isomorphism between $\mathcal{A}'$ and $\mathcal{B}'$, giving us an isomorphism between $\mathcal{A}$ and $\mathcal{B}$. Thus we have shown that every DO_{ij} is $\aleph_0$-categorical.

An interesting (and important!) result which we will not prove is the following:

5.2.5 Theorem (Morley's Theorem)

If T is a κ-categorical theory for some uncountably infinite cardinal κ, T is κ-categorical for all uncountably infinite cardinals κ.

5.2.6 Theorem (Vaught's Test)

A countable consistent theory T without finite models which is also κ-categorical for some cardinal κ is complete.

Since T has no finite models κ must be infinite. And so if we assume that T is incomplete, there must be an α such that $\nvdash \alpha, \neg\alpha$. Thus $T + \alpha$ and $T + \neg\alpha$ are both consistent, and since T is countable they must have countably infinite models. By the Upward Löwenheim-Skolem Theorem, $T + \alpha$ and $T + \neg\alpha$ must have models of cardinality κ. So let $\mathcal{A} \vDash T + \alpha$ and $\mathcal{B} \vDash T + \neg\alpha$ be models of cardinality κ, then $\mathcal{A}, \mathcal{B} \vDash T$ and since T is κ-categorical, this means that they are isomorphic and in particular elementarily equivalent. But $\mathcal{A} \vDash \alpha$ and $\mathcal{B} \vDash \neg\alpha$ in contradiction. ■

This means that DO_{ij} is complete for every i, j. Since $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are both densely ordered sets without endpoints, we have once again confirmed that $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$.

5.2.7 Definition

Model classes of first-order sentences are **elementary classes**. An (arbitrary) intersection of elementary classes is called a **Δ-elementary class**. In particular, if T is a theory then $\operatorname{Md} T = \bigcap_{\alpha \in T} \operatorname{Md} \alpha$ is a Δ-elementary class. If T is a complete theory, then $\operatorname{Md} T$ is also termed an **elementary type**.

Notice that the elementary class of a theory is equal to the union of its completion's elementary types (this is a mouthful, yet a trivial result).

5.2.8 Definition

Let $X \subseteq \mathcal{L}$ be an arbitrary nonempty set of formulas and T an $\mathcal{L}$-theory. We define $\langle X \rangle_T$ (we will remove the subscript if the theory is understood) to be the set of all formulas equivalent modulo T to a boolean combination of formulas in X. Notice then that $\top \in \langle X \rangle_T$ since $\top \equiv_T \alpha \vee \neg\alpha$ for $\alpha \in X$, and therefore for every $\alpha \in T$ since $\alpha \equiv_T \top$ so we have that $T \subseteq \langle X \rangle_T$. And we say that X is a **boolean basis for $\mathcal{L}^0$ in** T if every sentence $\alpha \in \mathcal{L}^0$ belongs to $\langle X \rangle_T$.

We write $\mathcal{A} \equiv_X \mathcal{B}$ to mean that $\mathcal{A} \vDash \alpha \iff \mathcal{B} \vDash \alpha$ for every $\alpha \in X$.

5.2.9 Theorem (Basis Theorem for Sentences)

Let T be a theory and $X \subseteq \mathcal{L}^0$ a set of sentences with $\mathcal{A} \equiv_X \mathcal{B} \implies \mathcal{A} \equiv \mathcal{B}$ for all models $\mathcal{A}, \mathcal{B} \vDash T$. Then X is a boolean basis for $\mathcal{L}^0$.

Let $\alpha \in \mathcal{L}^0$, and let us define $Y_\alpha := \{\beta \in \langle X \rangle \mid \alpha \vdash_T \beta\}$. Then we claim that $Y_\alpha \vdash_T \alpha$. Otherwise there must be a model $\mathcal{A} \vDash Y_\alpha, T, \neg\alpha$. So let us define $T_X\mathcal{A} := \{\gamma \in \langle X \rangle \mid \mathcal{A} \vDash \gamma\}$, and we have that $T_X\mathcal{A} \vdash \neg\alpha$ since if $\mathcal{B} \vDash T_X\mathcal{A}$ then for every $\beta \in X$ if $\mathcal{A} \vDash \beta$ then $\beta \in T_X\mathcal{A}$ and so $\mathcal{B} \vDash \alpha$, and conversely if $\mathcal{A} \nvDash \beta$ then $\neg\beta \in T_X\mathcal{A}$ and so $\mathcal{B} \nvDash \beta$. Thus $\mathcal{A} \equiv_X \mathcal{B}$ so $\mathcal{A} \equiv \mathcal{B}$ and thus $\mathcal{B} \vDash \neg\alpha$. So by the compactness theorem, and since $T_X\mathcal{A}$ is closed under conjunctions, there exists a $\gamma \in T_X\mathcal{A}$ such that $\gamma \vdash \neg\alpha$ and so $\alpha \vdash \neg\gamma$, so $\neg\gamma \in Y_\alpha$ and thus $\mathcal{A} \vDash \neg\gamma$. But we know that $\gamma \in T_X\mathcal{A}$ so $\mathcal{A} \vDash \gamma$ in contradiction.

Thus $Y_\alpha \vdash_T \alpha$ and therefore again by the compactness theorem there exist $\beta_1, \dots, \beta_m \in Y_\alpha$ such that $\beta = \bigwedge_{i=1}^m \beta_i \vdash_T \alpha$. And since $\alpha \vdash_T \beta_i$, we have that $\alpha \vdash_T \beta$ as well. Therefore $\alpha \equiv_T \beta \in \langle X \rangle$ and therefore $\alpha \in \langle X \rangle$. So indeed $\langle X \rangle = \mathcal{L}^0$ as required. ■

5.2.10 Example

For $T = \mathsf{DO}$ and $X = \{\mathsf{L}, \mathsf{R}\}$, we have that $\mathcal{A} \equiv_X \mathcal{B}$ implies $\mathcal{A} \equiv \mathcal{B}$ for all $\mathcal{A}, \mathcal{B} \vDash T$. This is as if $\mathcal{A}$ and $\mathcal{B}$ are densely ordered sets which agree on X, meaning they have the same endpoint configuration, they belong to the same DO_{ij} which is complete. Therefore X is a boolean basis for $\mathcal{L}^0$ in T.

5.2.11 Corollary

Let $T \subseteq \mathcal{L}^0$ be a theory with arbitrarily large finite models, such that all finite models with the same number of elements and all infinite models are elementarily equivalent. Then

(1) the sentences $\exists_n$ form a Boolean basis for $\mathcal{L}^0$ in T, and

(2) T is decidable provided it is finitely axiomatizable.

Let us define $X := \{\exists_n \mid n \in \mathbb{N}\}$, then by our assumption $\mathcal{A} \equiv_X \mathcal{B}$ implies that $\mathcal{A} \equiv \mathcal{B}$ for all $\mathcal{A}, \mathcal{B} \vDash T$. Thus X is a boolean basis for $\mathcal{L}^0$ in T, so we have proven (1). For (2) we claim that T has the finite model property, and as shown earlier a theory which is finitely axiomatizable and has the finite model property is decidable. Since X is a boolean basis, every $\alpha \in \mathcal{L}^0$ is equivalent modulo T to a boolean combination of $\exists_n$s. As shown in a previous exercise, every boolean combination of $\exists_n$ is equivalent to one of the following forms: $\bigvee_{i=1}^n \exists_{=k_i}$ with $k_1 < \cdots < k_n$ or $\exists_k \vee \bigvee_{i=1}^n \exists_{=k_i}$ for some k. Since T has arbitrarily large models, $T + \alpha$ if satisfied must be so by a finite model as required. ■

Notice that if $\{T + X_i\}_{i \in I}$ is the set of all completions of a theory T where $X_i \subseteq \mathcal{L}^0$, then $X = \bigcup_{i \in I} X_i$ forms a boolean basis for $\mathcal{L}^0$ in T. This is as for T-models $\mathcal{A}$ and $\mathcal{B}$, the models must also model some completions of T suppose $\mathcal{A} \vDash T + X_i$ and $\mathcal{B} \vDash T + X_j$. So if $\mathcal{A} \equiv_X \mathcal{B}$ we must have $i = j$ and so $\mathcal{A} \equiv \mathcal{B}$ since $T + X_i$ is complete. Though finding this X is not always trivial.

Suppose though that T has only finitely many completions, then all of its extensions must be finite (of the form $T + \alpha$). As otherwise if $T + \{\alpha_i\}_{i \in \mathbb{N}}$ is a nonfinite extension, we can assume that $\bigwedge_{i<n} \alpha_i \nvDash \alpha_n$ which would imply that T has infinitely many completions (as each $T + \bigwedge_{i<n} \alpha_i + \neg\alpha_n$ is consistent and thus has a distinct completion), in contradiction. So now suppose that the completions of T are $T + \alpha_1, \ldots, T + \alpha_n$ then all of its consistent extensions are of the form $T + \bigvee_{i=1}^m \alpha_{k_i}$ for some $m \leq n$ and $k_1 < \cdots < k_m$.

On one hand all such extensions are consistent:

$$\mathcal{A} \vDash T + \bigvee_{i=1}^m \alpha_{k_i} \iff \mathcal{A} \vDash T \text{ and there exists an } 1 \leq i \leq m \text{ such that } \mathcal{A} \vDash \alpha_{k_i}$$

this means that

$$\mathrm{Md}\left(T + \bigvee_{i=1}^m \alpha_{k_i}\right) = \bigcup_{i=1}^m \mathrm{Md}(T + \alpha_{k_i})$$

and so surely $T + \bigvee \alpha_{k_i}$ is consistent. Conversely, if $T \subseteq T'$ then let $T + \alpha_{k_1}, \ldots, T + \alpha_{k_m}$ be all the completions of T' (a completion of T' is also a completion of T). We then claim that $T' = T + \bigvee_{i=1}^m \alpha_{k_i}$. Now, as a general lemma we have that $T = \bigcap_{\mathcal{A} \in \mathrm{Md}\, T} Th\mathcal{A}$: obviously $T \subseteq Th\mathcal{A}$ for every $\mathcal{A} \in \mathrm{Md}\, T$ so one inclusion is trivial. In the other direction suppose that there existed an α such that every T-model $\mathcal{A} \vDash \alpha$ but $\alpha \notin T$. This would mean $T + \neg\alpha$ is consistent and thus has a T-model, but this T-model would not satisfy α in contradiction.

Thus a theory depends only on its models, so we need only show that $\mathrm{Md}\, T' = \mathrm{Md}(T + \bigvee_{i=1}^m \alpha_{k_i}) = \bigcup_{i=1}^m \mathrm{Md}(T + \alpha_{k_i})$. Since $T' \subseteq T + \alpha_{k_i}$ we must have that $\mathrm{Md}(T + \alpha_{k_i}) \subseteq \mathrm{Md}\, T'$ so one direction is trivial. Now suppose that $\mathcal{A} \vDash T'$ then $Th\mathcal{A}$ is a completion of T' and thus is equal to some $T + \alpha_{k_i}$ and so $\mathcal{A} \vDash T + \alpha_{k_i}$. So we have proven the equality and so $T' = \bigvee_{i=1}^m T + \alpha_{k_i}$.

Notice that $T + \bigvee_{i=1}^n \alpha_i = T$, and so T has exactly $2^n - 1$ consistent extensions (as we have 2^n choices for $k_1 < \cdots < k_m$, and one choice gives T). And these extensions are all distinct since the extensions are determined by their completions. So we have proven the following:

5.2.12 Proposition

A theory with n completions has $2^n - 1$ consistent extensions.

Exercise

Show that the theory T of torsion-free divisible abelian groups is $\aleph_1$-categorical and therefore complete.

A torsion-free divisible abelian group is simply a $\mathbb{Q}$-vector space where scalar multiplication is defined by for $g \in G$ and $\frac{a}{b} \in \mathbb{Q}$, $\frac{a}{b}g = \frac{ag}{b}$. So if G is a torsion-free divisible abelian group of cardinality $\aleph_1$ it is a vector space and thus must have a basis B, meaning $G \cong \bigoplus_{b\in B} \mathbb{Q}$. Notice that this cardinality is dependent only on the cardinality of B, and we can swap B with any set of the same cardinality and get the same direct product. So all torsion-free divisible abelian groups are isomorphic to $\bigoplus_{b\in B} \mathbb{Q}$ and so T is indeed $\aleph_1$-categorical. Since a torsion-free group cannot be finite, by Vaught's test T is complete.

Exercise

Show that a countable $\aleph_0$-categorical theory T with no finite models has an elementary prime model.

Let $\mathcal{P}$ be a T-model with cardinality $\aleph_0$. Let $\mathcal{B} \vDash T$ and since T has no finite models it must be infinite. By the Downward Löwenheim-Skolem Theorem $\mathcal{B}$ must have an elementary substructure $\mathcal{A} \preceq \mathcal{B}$ of cardinality $\aleph_0$. Since T is $\aleph_0$-categorical, $\mathcal{P} \cong \mathcal{A}$ and so $\mathcal{P}$ can be elementarily embedded into $\mathcal{B}$ (using its isomorphism with $\mathcal{A}$), and is thus an elementary prime model.

5.3 Ehrenfeucht-Fraïssé Games

Vaught's test, while simple and useful in specific circumstances, is quite limited in application as its conditions are quite restrictive: many complete theories are not categorical for any infinite cardinality. One example of this is the theory SO of *discretely ordered sets*, where every element which is not the right edge element has an immediate successor and every element which is not the left edge element has an immediate predecessor. Then we define SO_{ij} for $i, j \in \{0, 1\}$ analogously to DO_{ij}. Clearly in SO_{10} and SO_{00} we can define $\mathtt{S}$ the successor function, and so $(\mathbb{N}, <)$ is a prime model of SO_{10}. Models of SO_{10} can be viewed as looking something like $\mathbb{N}\ \mathbb{Z}\ \mathbb{Z}\ \cdots$, ie. the initial segment of such a model is $(\mathbb{N}, <)$ then followed by an arbitrary amount of segments $(\mathbb{Z}, <)$. This demonstrates that SO_{10} cannot be categorical for any cardinality, but it is indeed complete as we will demonstrate.

We will prove this using an *Ehrenfeucht-Fraïssé game* (for short an EF-game). Let $\mathcal{L}$ be a first-order relational language, $\mathcal{A}$ and $\mathcal{B}$ be two $\mathcal{L}$-structures and $k \geq 0$. Then the EF-game played in k rounds on $\mathcal{A}$ and $\mathcal{B}$, denoted $\Gamma_k(\mathcal{A}, \mathcal{B})$ is a game played by two players: player I and player II. The game proceeds as follows: player I starts by picking any element in either $\mathcal{A}$ or $\mathcal{B}$, suppose he choose $a \in A$, then player II must respond with an element $b \in B$. Conversely if player I chooses an element $b \in B$, player II must respond with an element $a \in A$. After k such rounds suppose the elements chosen are $a_1, \ldots, a_k \in A$ and $b_1, \ldots, b_k \in B$ (where a_i and b_i are the elements chosen in the ith round), then player II wins if $a_i \mapsto b_i$ is a partial isomorphism, meaning that the substructures $\{a_1, \ldots, a_k\} \subseteq \mathcal{A}$ and $\{b_1, \ldots, b_k\} \subseteq \mathcal{B}$ are isomorphic. Player I wins otherwise.

We say that player II has a winning strategy in the game $\Gamma_k(\mathcal{A}, \mathcal{B})$ if no matter what moves player I plays, player II can always find a way to win (this will be formalized more later on this section). We write this as $\mathcal{A} \sim_k \mathcal{B}$. For the zero-round game, we just define $\mathcal{A} \sim_0 \mathcal{B}$.

5.3.1 Example

We will show that for every $\mathcal{B} \vDash \mathsf{SO}_{00}$, $\mathcal{A} = (\mathbb{Z}, <) \sim_k \mathcal{B}$ for every $k \geq 0$. Let us first define the distance function $d(x, y)$ on $\mathcal{B}$: if $x = y$ then $d(x, y) = 0$ and otherwise it is one plus the number of elements between x and y if finite, otherwise $d(x, y) = \infty$. Since we can embed $\mathcal{A}$ into $\mathcal{B}$, we will just assume that $\mathcal{A} \subseteq \mathcal{B}$.

Our goal is to try and ensure that after m rounds of the game, if $a_1 < a_2 < \cdots < a_m$ are the elements of $\mathcal{A}$ which have been played, and $b_1 < b_2 < \cdots < b_m$ are the elements in $\mathcal{B}$ which have been played, then the mapping $a_i \mapsto b_i$ is the partial embedding corresponding to the play of the game. Furthermore, if $d(a_i, a_{i+1}) > 3^{n-m}$ then $d(b_i, b_{i+1}) > 3^{n-m}$, and if $d(a_i, a_{i+1}) \leq 3^{n-m}$ then $d(b_i, b_{i+1}) = d(a_i, a_{i+1})$, for $i = 1, \ldots, m-1$.

Obviously since $a_i < a_{i+1}$ and $b_i < b_{i+1}$, the function will preserve the relations of the theory and thus be a partial embedding.

We claim that player II can always make a move to preserve this condition. In round 1, player II can choose any arbitrary element and the condition will hold. Now suppose we have played m rounds and $a_1 < \cdots < a_m$ and $b_1 < \cdots < b_m$ be defined as above. Now suppose player I plays $b \in B$, then there are several cases

(1) If $b < b_1$ then if $d(b, b_1) = k < \infty$ then player II plays $a_1 - k$. If $d(b, b_1) = \infty$ then player II plays $a_1 - 3^n$, but in any case the condition holds.

(2) If $b_i < b < b_{i+1}$ and $d(b_i, b_{i+1}) \leq 3^{n-m}$ then $d(a_i, a_{i+1}) = d(b_i, b_{i+1})$. Play $a = a_i + d(b, b_i)$, then $d(a, a_{i+1}) = d(b, b_{i+1})$ as required.

(3) If $b_i < b < b_{i+1}$ and $d(b_i, b_{i+1}) > 3^{n-m}$ and $d(b, b_i) < 3^{n-m-1}$ then $d(a_i, a_{i+1}) > 3^{n-m}$. Play $a = a_i + d(b, b_i)$, then $d(a, a_{i+1})$ and $d(a_i, a)$ are greater than 3^{n-m-1} as required.

(4) If $b_i < b < b_{i+1}$ and $d(b_i, b_{i+1}) > 3^{n-m}$ and $d(b, b_{i+1}) < 3^{n-m-1}$, play $a = a_{i+1} - d(b, b_{i+1})$.

(5) If $b_i < b < b_{i+1}$ and $d(b_i, b_{i+1}) > 3^{n-m}$, $d(b, b_i) > 3^{n-m-1}$, and $d(b, b_{i+1}) < 3^{n-m-1}$. Then $d(a_i, a_{i+1}) > 3^{n-m}$ and so choose an a such that $a_i < a < a_{i+1}$ and the distance of a between them both is greater than 3^{n-m-1}. Playing a satisfies the condition.

(6) If $b > b_m$, this is similar to the first condition.

Let us write $\mathcal{A} \equiv_k \mathcal{B}$ to mean that $\mathcal{A} \vDash \alpha \iff \mathcal{B} \vDash \alpha$ for all $\alpha \in \mathcal{L}^0$ of quantifier rank $\leq k$. In a relational language there are no sentences of quantifier rank zero so $\mathcal{A} \equiv_0 \mathcal{B}$ vacuously. It is our goal this section to prove the remarkable

5.3.2 Theorem

$\mathcal{A} \sim_k \mathcal{B}$ implies $\mathcal{A} \equiv_k \mathcal{B}$, thus $\mathcal{A} \equiv \mathcal{B}$ provided $\mathcal{A} \sim_k \mathcal{B}$ for all k.

In finite signatures the converse is also true, but we will not prove this. Notice that this proves that SO_{00} is complete, since every $\mathcal{B} \vDash \mathsf{SO}_{00}$ has that $\mathcal{B} \sim_k (\mathbb{Z}, <)$ for all k and thus $\mathcal{B} \equiv (\mathbb{Z}, <)$, meaning all models of SO_{00} have the same theory and therefore SO_{00} is complete. We can similarly show that SO_{10} is complete (as all its models are equivalent to $(\mathbb{N}, <)$) and so is SO_{01} (using $(\mathbb{N}, >)$). SO_{11} is not complete as it has the finite model property.

Let us generalize our concept of EF-games as follows: we define the game $\Gamma_k(\mathcal{A}, \mathcal{B}, \vec{a}, \vec{b})$ to be the EF-game with prior moves $\vec{a} \in A^n$ and $\vec{b} \in B^n$. In this game, in the first round player I plays $a_{n+1} \in A$ or $b_{n+1} \in B$ and player II responds with $b_{n+1} \in B$ or $a_{n+1} \in A$, etc. Then at the end we have $(a_1, \ldots, a_n, a_{n+1}, \ldots, a_{n+k}) \in A^{n+k}$ and $(b_1, \ldots, b_n, b_{n+1}, \ldots, b_{n+k}) \in B^{n+k}$ and player II wins if $a_i \mapsto b_i$ is a partial isomorphism. For $n = 0$ this just coincides with our previous definition of an EF-game.

Now we can formalize the concept of a winning strategy:

5.3.3 Definition

Player II has a **winning strategy** in $\Gamma_0(\mathcal{A}, \mathcal{B}, \vec{a}, \vec{b})$ if $a_i \mapsto b_i$ for $1 \leq i \leq n$ is a partial isomorphism. Inductively, player II has a winning strategy in $\Gamma_{k+1}(\mathcal{A}, \mathcal{B}, \vec{a}, \vec{b})$ provided that for any $a \in A$ there exists a $b \in B$ and for any $b \in B$ there exists an $a \in A$ such that player II has a winning strategy in $\Gamma_k(\mathcal{A}, \mathcal{B}, \vec{a}\#(a), \vec{b}\#(b))$ where $\vec{v}\#\vec{w}$ is the concatenation of two vectors.

Now we write $(\mathcal{A}, \vec{a}) \sim_k (\mathcal{B}, \vec{b})$ if player II has a winning strategy in $\Gamma_k(\mathcal{A}, \mathcal{B}, \vec{a}, \vec{b})$. In particular $\mathcal{A} \sim_k \mathcal{B}$ if this holds for $\vec{a}, \vec{b} = \varnothing$.

5.3.4 Lemma

Let $(\mathcal{A}, \vec{a}) \sim_k (\mathcal{B}, \vec{b})$ where $\vec{a} \in A^n$ and $\vec{b} \in B^n$, then for every $\varphi = \varphi(\vec{x})$ with quantifier rank $\leq k$, $\mathcal{A} \vDash \varphi(\vec{a}) \iff \mathcal{B} \vDash \varphi(\vec{b})$.

We prove this by induction on k: for $k = 0$ since $a_i \mapsto b_i$ is a partial isomorphism we have that $\{a_1, \ldots, a_n\} \vDash \varphi(\vec{a}) \iff \{b_1, \ldots, b_n\} \vDash \varphi(\vec{b})$ for all quantifier-free φ (ie. formulas with quantifier rank zero). We have that by

The Substructure Theorem this is equivalent to $\mathcal{A} \vDash \varphi(\vec{a}) \iff \mathcal{B} \vDash \varphi(\vec{b})$. Now suppose that $(\mathcal{A}, \vec{a}) \sim_{k+1} (\mathcal{B}, \vec{b})$ then let $\varphi = \forall y \alpha(\vec{x}, y)$ where $qrank\alpha \leq k$. Let $\mathcal{A} \vDash \forall y \alpha(\vec{a}, y)$ and $b \in B$, then player II can choose an $a \in A$ such that $(\mathcal{A}, \vec{a}\#(a)) \sim_k (\mathcal{B}, \vec{b}\#(b))$ and so by our inductive hypothesis we have that since $\mathcal{A} \vDash \alpha(\vec{a}, a)$ we get $\mathcal{B} \vDash \alpha(\vec{b}, b)$. Since $b \in B$ is arbitrary we have $\mathcal{B} \vDash \forall y \alpha(\vec{b}, x) = \varphi(\vec{b})$. Arguing similarly we obtain the converse. All formulas of quantifier rank $\leq k+1$ are boolean combinations of formulas of the form $\forall y \alpha(\vec{x}, y)$ and so we obtain the result for arbitrary formulas by inducting on $\wedge$ and $\neg$ which is simple. ■

Theorem 5.3.2 results simply in applying this lemma for all k. This is actually quite a practical method proving the completeness of certain theories, like $\mathsf{SO}_{10}, \mathsf{SO}_{01}, \mathsf{SO}_{00}$. SO_{11} is obviously not complete as each $\exists_n$ is independent of it. Since two finite discrete orders are isomorphic, and it can be shown by use of an EF-game that infinite models of SO_{11} are elementarily equivalent we get that $X = \{\mathsf{L}, \mathsf{R}, \exists_1, \exists_2, \dots\}$ is a boolean basis of $\mathcal{L}^0$ in SO (by the Basis Theorem for Sentences).

5.4 Embedding and Characterization Theorems

An *universal theory* is a theory with an axiom system of $\forall$-sentences. For example, the theory of groups in the signature $\{\cdot, e, ^{-1}\}$ and rings in $\{+, \cdot, 0, 1\}$ are universal theories. We know that universal theories are **S**-*invariant* (substructure invariant), meaning if $\mathcal{A} \subseteq \mathcal{B} \vDash T$ then $\mathcal{A} \vDash T$ (as the satisaction of $\forall$-sentences is preserved by substructures).

DO does not have this property (for example $(\mathbb{Z}, <) \subseteq (\mathbb{Q}, <)$) and therefore has no axiom system consisting of $\forall$-sentences.

Interestingly, universal theories are characterized precisely by **S**-invariance, as we will show.

Let $T^\forall = \{\alpha \in T \mid \alpha \text{ is a } \forall\text{-sentence}\}$ be the *universal part* of the theory T. Notice that $T^\forall$ is not a theory as the theory generated by it contains more than just $\forall$-sentences. For a sublanguage $\mathcal{L}_0 \subseteq \mathcal{L}$ let $T_0^\forall = \mathcal{L}_0 \cap T^\forall = (\mathcal{L}_0 \cap T)^\forall$. And as before if $\mathcal{A}$ is an $\mathcal{L}_0$-structure and $\mathcal{B}$ an $\mathcal{L}$-structure, then $\mathcal{A} \subseteq \mathcal{B}$ is to be taken as meaning that $\mathcal{A}$ is a substructure of the $\mathcal{L}_0$-reduct of $\mathcal{B}$.

5.4.1 Lemma

Every $T_0^\forall$-model $\mathcal{A}$ is embeddable into some T-model.

It is sufficient to show that $T + \mathcal{D}A$ is consistent, because for every $\mathcal{B} \vDash T + \mathcal{D}A$, we must have that $\mathcal{A}$ is embeddable into $\mathcal{B}$ as it models its diagram. So assume the opposite, then there must be a conjunction of sentences from $D\mathcal{A}$, $\varkappa(\vec{a})$, such that $\varkappa(\vec{a}) \vdash_T \bot$, equivalently $\vdash_T \neg\varkappa(\vec{a})$. Since $\vec{a}$ do not appear in T, by ($\forall$3) we have that $\vdash_T \forall\vec{x}\neg\varkappa$ and thus $\forall\vec{x}\neg\varkappa \in T_0^\forall$, so $\mathcal{A} \vDash \forall\vec{x}\neg\varkappa$. But this contradicts $\mathcal{A} \vDash \varkappa(\vec{a})$. ■

5.4.2 Lemma

$\mathrm{Md}\, T^\forall$ consists of precisely all substructures of T-models.

Since a substructure of a T-model is a $T^\forall$ model (since $T^\forall$ is **S**-invariant as a set of $\forall$-sentences), we have shown one direction. All that remains to show is that a $T^\forall$-model is a substructure of a T-model. By applying the above lemma for $\mathcal{L}_0 = \mathcal{L}$ we get that every $T^\forall$-model is embeddable into some T-model. Since T is closed under isomorphic images, we have that a $T^\forall$-model is a substructure of a T-model. ■

5.4.3 Example

Let AG be the theory of abelian groups in $\mathcal{L}\{\circ\}$. A substructure of an abelian group in this signature is a commutative cancellative semigroup (a cancellative semigroup is a semigroup with the left and right cancellation properties: $ab = ac \implies b = c$, $ba = ca \implies b = c$). Conversely it can be shown that a cancellative semigroup can be embedded into an abelian group, and is thus a substructure of an abelian group. Thus the theory (generated by) $\mathsf{AG}^\forall$ coincides with the theory of commutative cancellative semigroups (since $\mathrm{Md}\, T = \mathrm{Md}\, T' \implies T = T'$ as discussed before).

5.4.4 Theorem

T is a universal theory if and only if it is **S**-invariant.

If T is **S**-invariant then by the above lemma $\operatorname{Md} T^{\forall} = \operatorname{Md} T$, and so T can be axiomatized by its universal part $T^{\forall}$. ■

Recall that a universal Horn theory is a theory with an axiom system consiting of universal horn sentences: sentences of the form $\forall\vec{x}(\neg\alpha_0 \vee \neg\alpha_1 \cdots \vee \neg\alpha_n)$ or $\forall\vec{x}(\alpha_0 \vee \neg\alpha_1 \vee \cdots \vee \neg\alpha_n)$ where α_i are prime. We say that a theory T is **SP**-*invariant* (substructure and product invariant) if it is **S**-invariant and closed under direct products: if $\{\mathcal{A}_i\}_{i\in I}$ is a family of T-models, then $\prod_{i\in I}\mathcal{A}_i$ is also a T-model.

5.4.5 Theorem

T is a universal Horn theory if and only if T is **SP**-invariant.

$\Longrightarrow$: this was shown in a previous exercise. $\Longleftarrow$: if $\vdash_T \forall xy\, x = y$ then T can be axiomatized by $\forall xy\, x = y$ which is a universal Horn formula. Otherwise let U be the set of all universal Horn sentences in T, then we will show that $\operatorname{Md} T = \operatorname{Md} U$. Obviously $\operatorname{Md} T \subseteq \operatorname{Md} U$ so we will prove the other direction. Let $\mathcal{A} \vDash U$, then we claim that $T \cup D\mathcal{A}$ is consistent: since then there exists a $\mathcal{B} \vDash T + D\mathcal{A}$ and so $\mathcal{A}$ is embeddable into a T-structure so by **S**-invariance we get that $\mathcal{A} \vDash T$. Let us define $P := \{\pi \in D\mathcal{A} \mid \pi \text{ prime}\}$, and so $D\mathcal{A} = P \cup \{\neg\pi_i\}_{i\in I}$ for some nonempty I, π_i all prime.
We will now show that $P \nvdash_T \pi_i$ for all $i \in I$, as otherwise we'd have $\vdash_T \varkappa(\vec{a}) \to \pi_i(\vec{a})$ for some conjunction of sentences in P, $\varkappa(\vec{a})$. Since $\vec{a}$ are constants which do not appear in T, $\vdash_T \forall\alpha := \vec{x}(\varkappa(\vec{x}) \to \pi_i(\vec{x}))$. And since $\varkappa \to \pi_i$ is a Horn formula, we have that $\alpha \in U$ and so $\mathcal{A} \vDash \alpha$. But this contradicts $\mathcal{A} \vDash \varkappa(\vec{a}) \wedge \neg\pi_i(\vec{a})$.
So choose $\mathcal{A}_i \vDash T, P, \neg\pi_i$, then $\mathcal{B} := \prod_{i\in I}\mathcal{A}_i \vDash T \cup P \cup \{\neg\pi_i\}_{i\in I} = T \cup D\mathcal{A}$ by **P**-invariance and since for direct products and prime formulas π, $\prod \mathcal{A}_i \vDash \pi$ if and only if $\mathcal{A}_i \vDash \pi$ for all i. Thus $T + D\mathcal{A}$ is consistent as required. ■

5.4.6 Theorem

Let $\mathcal{L}_0 \subseteq \mathcal{L}$ and let $\mathcal{A}$ be an $\mathcal{L}_0$-structure. For $T \subseteq \mathcal{L}^0$, the following are equivalent:

(1) $\mathcal{A}$ is embeddable into some T-model,

(2) every $\mathcal{B} \subseteq \mathcal{A}$ is embeddable into a T-model,

(3) every finitely generated $\mathcal{B} \subseteq \mathcal{A}$ is embeddable into a T-model,

(4) $\mathcal{A} \vDash T_0^{\forall}$.

(1) $\Longrightarrow$ (2) and (2) $\Longrightarrow$ (3) are trivial. (3) $\Longrightarrow$ (4): let $\forall\vec{x}\alpha(\vec{x}) \in T_0^{\forall}$. Then let $\vec{a} \in A^n$ and $\mathcal{A}_0$ be the $\mathcal{A}$-substructure generated by $a_1, \ldots, a_n$. Thus $\mathcal{A}_0$ is embeddable into a T-model $\mathcal{B}$. Since $\mathcal{B} \vDash \forall\vec{x}\alpha(\vec{x})$, we have that $\mathcal{A}_0 \vDash \forall\vec{x}\alpha(\vec{x})$ as well and in particular $\mathcal{A}_0 \vDash \alpha(\vec{a})$. By The Substructure Theorem we get that $\mathcal{A} \vDash \alpha(\vec{a})$ and since $\vec{a} \in A^n$ is arbitrary, we get that $\mathcal{A} \vDash \forall\vec{x}\alpha(\vec{x})$ so $\mathcal{A} \vDash T_0^{\forall}$. (4) $\Longrightarrow$ (1) is simply lemma 5.4.1. ■

5.4.7 Example

Let T be the theory of ordered abelian groups in $\mathcal{L} = \mathcal{L}\{0, +, -, <\}$. Such a group must be torsion-free as if $a > 0$ then $0 < a < 2a < 3a < \cdots$ so $na \neq 0$ for all n, which can be expressed in $\mathcal{L}_0 = \mathcal{L}\{0, +, -\}$. The above theorem implies that a torsion-free abelian group is embeddable into an ordered abelian group (and is thus an ordered abelian group), as every finitely generated torsion-free abelian group is embeddable into an ordered abelian group. This is since a torsion-free abelian group G is isomorphic to $\mathbb{Z}^n$ for some $n > 0$, and $\mathbb{Z}^n$ can be ordered lexicographically. So we have shown that every abelian torsion-free group can be ordered.

We can also characterize $\forall$-formulas model-theoretically: say a formula $\alpha(\vec{x}) \in \mathcal{L}^0$ **S**-*persistent* or simply *persistent* in a theory T provided that for all T-models $\mathcal{A}, \mathcal{B}$ if for all $\vec{a} \in A^n$,

$$\mathcal{A} \subseteq \mathcal{B} \vDash \alpha(\vec{a}) \Longrightarrow \mathcal{A} \vDash \alpha(\vec{a})$$

5.4.8 Theorem

If $\alpha = \alpha(\vec{x})$ is persistent in T then α is equivalent to some $\forall$-formula α' in T, which can be chosen such that $free\alpha' \subseteq free\alpha$.

Let us define $Y = \{\forall\vec{y}\beta(\vec{x},\vec{y}) \mid \alpha \vdash_T \forall\vec{y}\beta(\vec{x},\vec{y})\}$ where β is quantifier-free. Then we claim that $Y \vdash_T \alpha(\vec{x})$ and this would be sufficient as by the compactness theorem there exists a conjunction of formulas in Y, $\varkappa(\vec{x})$, such that $\varkappa(\vec{x}) \vdash_T \alpha(\vec{x})$, and since $\alpha(\vec{x}) \vdash_T \varkappa(\vec{x})$ we'd have that they are equivalent in T. The conjunction of $\forall$-formulas is itself equivalent to a $\forall$-formula, and so α' can be chose to be $\varkappa$.

So now assume that $(\mathcal{A},\vec{a}) \vDash T, Y$ where $\vec{a} \in A^n$, and we must show that $(\mathcal{A},\vec{a}) \vDash \alpha$. To do so we claim that $T, \alpha(\vec{a}), D\mathcal{A}$ is consistent as then if $\mathcal{B} \vDash T, \alpha(\vec{a}), D\mathcal{A}$ we can assume that $\mathcal{A} \subseteq \mathcal{B}$ and so by the persistancy of α, since $\mathcal{B} \vDash \alpha(\vec{a})$ we have that $\mathcal{A} \vDash \alpha(\vec{a})$. So if $T, \alpha(\vec{a}), D\mathcal{A}$ was inconsistent then $\alpha(\vec{a}) \vdash_T \neg\varkappa(\vec{a},\vec{b})$ for a conjunction $\varkappa(\vec{a},\vec{b})$ of sentences from $D\mathcal{A}$. Since $\vec{b}$ does not occur in $\alpha(\vec{a})$ or T, we have that $\alpha(\vec{a}) \vdash_T \forall\vec{y}\neg\varkappa(\vec{a},\vec{y})$ and since $\vec{a}$ do not occur in T we have that $\alpha(\vec{x}) \vdash_T \forall\vec{y}\neg\varkappa(\vec{x},\vec{y})$ and so $\forall\vec{y}\neg\varkappa(\vec{x},\vec{y}) \in Y$, so $\mathcal{A} \vDash \forall\vec{y}\neg\varkappa(\vec{x},\vec{y})$. But this contradicts $\mathcal{A} \vDash \varkappa(\vec{a},\vec{b})$. ∎

Say that $\alpha = \alpha(\vec{x})$ is *kappa-persistent* for a cardinality κ if we have that for all T-models $\mathcal{A},\mathcal{B}$ of cardinality κ and $\vec{a} \in A^n$,

$$\mathcal{A} \subseteq \mathcal{B} \vdash \alpha(\vec{a}) \implies \mathcal{A} \vDash \alpha(\vec{a})$$

Then if T is countable this theorem holds even if we loosen the condition to α being κ-persistent for any $\kappa \geq \aleph_0$. This is as every T-model is elementarily equivalent (in their extended languages) to a T-model of cardinality κ by the Löwenheim-Skolem theorems.

A $\forall\exists$-*sentence* is a sentence of the form $\forall\vec{x}\exists\vec{y}\alpha(\vec{x},\vec{y})$ with α quantifier-free. A theory with an axiomatic system consisting of $\forall\exists$-sentences is called a $\forall\exists$-*theory*. Many theories are $\forall\exists$-theories, for example the theory of algebraically closed fields and divisible groups. We will now also characterize $\forall\exists$-theories semantically.

A *chain of structures* is a set K of $\mathcal{L}$-structures such that for every $\mathcal{A},\mathcal{B} \in K$ wither $\mathcal{A} \subseteq \mathcal{B}$ or $\mathcal{B} \subseteq \mathcal{A}$. Often times the structure will be countable and can thus be written as $\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \cdots$. In any case if K is a chain of structures, we can define the structure $\mathcal{C} := \bigcup K$ whose domain is $C = \bigcup\{A \mid \mathcal{A} \in K\}$ and for every relation symbol r and $\vec{a} \in C^n$, $r^{\mathcal{C}}\vec{a} \iff r^{\mathcal{A}}\vec{a}$ for any $\mathcal{A} \in K$ such that $\vec{a} \in A^n$. Notice that this is independent of the choice for $\mathcal{A}$ (and such a $\mathcal{A}$ must exist), as if $\vec{a} \in A_1^n, A_2^n$ then suppose $\mathcal{A}_1 \subseteq \mathcal{A}_2$ and since they are substructures we have $r^{\mathcal{A}_1}\vec{a} \iff r^{\mathcal{A}_2}\vec{a}$. And for function symbols f, let $f^{\mathcal{C}}\vec{a} = f^{\mathcal{A}}\vec{a}$ for any $\mathcal{A} \in K$ such that $\vec{a} \in A^n$, and for similar reasons as above this is well-defined. As such, $\mathcal{C}$ can be equivalently defined as the smallest $\mathcal{L}$-structure where every $\mathcal{A} \in K$ is a substructure.

5.4.9 Example

Define $\mathcal{D}_n$ to be the additive group of real numbers which have at most n decimals after the decimal point. Since $\mathcal{D}_n \subseteq \mathcal{D}_{n+1}$, they form a chain and we can take $\mathcal{D} := \bigcup_{n=0}^{\infty} \mathcal{D}_n$, which is just the additive group of reals with finitely many decimals after the point. Similarly if we take $\mathcal{D}_n$ to be ordered sets. But in this case we have $\mathcal{D}_n \vDash \mathsf{SO}$ while $\mathcal{D} \vDash \mathsf{DO}$, and therefore as we will see this means that SO is not a $\forall\exists$-theory.

Notice that if K is a chain of substructures and $\alpha = \forall\vec{x}\exists\vec{y}\beta(\vec{x},\vec{y})$ is a $\forall\exists$-sentence valid in all $\mathcal{A} \in K$, then it is also valid in $\mathcal{C} = \bigcup K$. This is as if we let $\vec{a} \in C^n$, then $\vec{a} \in A^n$ for some $\mathcal{A} \in K$ and so $\mathcal{A} \vDash \exists\vec{y}\beta(\vec{a},\vec{y})$ and since $\mathcal{A} \subseteq \mathcal{C}$ we have that $\mathcal{C} \vDash \exists\vec{y}\beta(\vec{a},\vec{y})$ and since $\vec{a} \in C^n$ is arbitrary we have $\mathcal{A} \vDash \forall\vec{x}\exists\vec{y}\beta(\vec{x},\vec{y})$. Suppose $\boldsymbol{K}$ is a class of $\mathcal{L}$-structures which is closed under chains (meaning if $K \subseteq \boldsymbol{K}$ is a chain then $\bigcup K \in \boldsymbol{K}$), then $\boldsymbol{K}$ is called *inductive*. So by this, if T is a $\forall\exists$-theory then $\mathrm{Md}\,T$ is inductive.

Let us say that a chain K is an *elementary chain* if for every $\mathcal{A},\mathcal{B} \in K$ either $\mathcal{A} \preceq \mathcal{B}$ or $\mathcal{B} \preceq \mathcal{A}$. Clearly an elementary chain is a chain in the normal sense as well.

5.4.10 Lemma (Tarski's Chain Lemma)

Let K be an elementary chain, and let $\mathcal{C} = \bigcup K$. Then $\mathcal{A} \preceq \mathcal{C}$ for every $\mathcal{A} \in K$.

We must show that for every $\vec{a} \in A^n$, $\mathcal{A} \vDash \alpha(\vec{a}) \iff \mathcal{C} \vDash \alpha(\vec{a})$. We will show this by induction on $\alpha(\vec{x})$. It is trivial for prime formulas as $\mathcal{A} \subseteq \mathcal{C}$, and the induction step on $\wedge$ and $\neg$ are obvious. So let $\mathcal{A} \vDash \forall y\alpha(y,\vec{a})$ then let $a_0 \in C$, then there must exist some $\mathcal{B} \in K$ such that $a_0, a_1, \ldots, a_n \in B$ and $\mathcal{A} \preceq \mathcal{B}$ so since $\mathcal{A} \vDash \alpha(a_0,\vec{a})$ we have that $\mathcal{B} \vDash \alpha(a_0,\vec{a})$ and by our inductive hypothesis $\mathcal{C} \vDash \alpha(a_0,\vec{a})$. Since $a_0 \in C$ is arbitrary we have that $\mathcal{C} \vDash \forall y\alpha(y,\vec{a})$. The converse $\mathcal{C} \vDash \forall y\alpha(y,\vec{a}) \implies \mathcal{A} \vDash \forall y\alpha(y,\vec{a})$ is due to $\mathcal{A} \subseteq \mathcal{C}$. ∎

We define another useful concept. Let $\mathcal{A} \subseteq \mathcal{B}$, then $\mathcal{A}$ is *existentially closed* in $\mathcal{B}$, in symbols $\mathcal{A} \subseteq_{ec} \mathcal{B}$ if

$$\mathcal{B} \vDash \exists\vec{x}\varphi(\vec{x},\vec{a}) \implies \mathcal{A} \vDash \exists\vec{x}\varphi(\vec{x},\vec{a}) \quad (\vec{a} \in A^n)$$

where $\varphi(\vec{x},\vec{a})$ is any conjunction of literals in $\mathcal{L}A$. Then this holds for all quantifier-free φ as we can convert φ to DNF and distributing $\exists\vec{x}$ over the disjunctions. We clearly have that $\mathcal{A} \preceq \mathcal{B} \implies \mathcal{A} \subseteq_{ec} \mathcal{B} \implies \mathcal{A} \subseteq \mathcal{B}$.

Call a chain of substructures K such that for every $\mathcal{A},\mathcal{B} \in K$ either $\mathcal{A} \subseteq_{ec} \mathcal{B}$ or $\mathcal{B} \subseteq_{ec} \mathcal{A}$, an *existentially closed chain*. Then we have the following lemma for existentially closed chains: if K is an existentially closed chain then for every $\mathcal{A} \in K$, $\mathcal{A} \subseteq_{ec} \bigcup K$. Suppose that $\mathcal{C} \vDash \exists\vec{x}\varphi(\vec{x},\vec{a})$ then there exists a $\vec{b} \in C^n$ such that $\mathcal{C} \vDash \varphi(\vec{b},\vec{a})$. Let $\mathcal{B} \in K$ such that $\vec{a},\vec{b} \in B$ and $\mathcal{A} \subseteq_{ec} \mathcal{B}$, so then we know that for any literal $\pi(\vec{a},\vec{b})$ we have that by definition $\mathcal{C} \vDash \pi(\vec{b},\vec{a}) \iff \mathcal{B} \vDash \pi(\vec{b},\vec{a})$. And so this must hold for φ in place of π as it is a conjunction of literals. So then we have that $\mathcal{B} \vDash \varphi(\vec{b},\vec{a})$ and thus $\mathcal{B} \vDash \exists\vec{x}\varphi(\vec{x},\vec{a})$ and since $\mathcal{A} \subseteq_{ec} \mathcal{B}$, we have the desired result.

Let us define the *universal diagram* of an $\mathcal{L}$-structure $\mathcal{A}$ to be the set of all $\forall$-sentences in $\mathcal{L}A$ valid in $\mathcal{A}$. We denote the universal diagram of $\mathcal{A}$ by $D_\forall\mathcal{A}$. Obviously $D_\forall\mathcal{A} \subseteq D_{el}\mathcal{A}$.

5.4.11 Lemma

Let $\mathcal{A},\mathcal{B}$ be $\mathcal{L}$-structures such that $\mathcal{A} \subseteq \mathcal{B}$. Then the followinga are equivalent:

(1) $\mathcal{A} \subseteq_{el} \mathcal{B}$,

(2) there exists an $\mathcal{A}' \supseteq \mathcal{B}$ such that $\mathcal{A} \preceq \mathcal{A}'$,

(3) $\mathcal{B} \vDash D_\forall\mathcal{A}$.

(1) $\implies$ (2): to do this it is sufficient to show that $D_{el}\mathcal{A} \cup D\mathcal{B}$ is consistent. Otherwise $D_{el}\mathcal{A} \vdash \neg\varkappa(\vec{b})$ for some conjunction of sentences in $D\mathcal{B}$, where $\vec{b}$ is an n-tuple of elements in $B \setminus A$. Then $D_{el}\mathcal{A} \vdash \forall\vec{x}\neg\varkappa(\vec{x})$, and so $\mathcal{A} \vDash \forall\vec{x}\neg\varkappa(\vec{x})$. But since $\mathcal{B} \vDash \exists\vec{x}\varkappa(\vec{x})$ we must have that $\mathcal{A} \vDash \exists\vec{x}\varkappa(\vec{x})$ in contradiction. (2) $\implies$ (3): since $\mathcal{A} \preceq \mathcal{A}'$ we have that $\mathcal{A}' \vDash D_{el}\mathcal{A} \supseteq D_\forall\mathcal{A}$. Since $\mathcal{B} \subseteq \mathcal{A}'$, we must have that $\mathcal{B} \vDash D_\forall\mathcal{A}$ as $D_\forall\mathcal{A}$ is a $\forall$-theory and is thus **S**-invariant. (3) $\implies$ (1): we have that $\mathcal{A} \vDash \alpha \implies \mathcal{B} \vDash \alpha$ for all $\forall$-sentences α of $\mathcal{L}A$. Since the negation of a $\forall$-sentence is an $\exists$-sentence, this is equivalent to $\mathcal{B} \vDash \alpha \implies \mathcal{A} \vDash \alpha$ for all $\exists$-sentences of $\mathcal{L}A$. ■

5.4.12 Theorem

A theory T is an $\forall\exists$-theory if and only if T is inductive.

We have already shown that a $\forall\exists$-theory is inductive, all which remains is to prove the converse. So let T be inductive, then we will show that $\operatorname{Md} T = \operatorname{Md} T^{\forall\exists}$ where $T^{\forall\exists}$ is the set of all $\forall\exists$-sentences in T. All that we must prove is that $\operatorname{Md} T^{\forall\exists} \subseteq \operatorname{Md} T$, so let $\mathcal{A} \vDash T^{\forall\exists}$. We claim that $T \cup D_\forall\mathcal{A}$ is consistent, as otherwise $\vdash_T \neg\varkappa(\vec{a})$ for a conjunction of sentences in $D_\forall\mathcal{A}$. Since $\vec{a}$ does not occur in T, we have that $\vdash_T \forall\vec{x}\neg\varkappa(\vec{x})$. Since $\varkappa(\vec{x})$ is the conjunction of $\forall$-formulas, it itself is equivalent to a $\forall$-formula and so $\forall\vec{x}\neg\varkappa(\vec{x})$ is equivalent to a $\forall\exists$-sentence, meaning $\forall\vec{x}\neg\varkappa(\vec{x})$ belongs to $T^{\forall\exists}$ up to equivalence. Thus $\mathcal{A} \vDash \forall\vec{x}\neg\varkappa(\vec{x})$, contradicting $\mathcal{A} \vDash \varkappa(\vec{a})$.

So let $\mathcal{A}_1 \vDash T, D_\forall\mathcal{A}$, then without loss of generality $\mathcal{A} \subseteq \mathcal{A}_1$ and in lieu of the above lemma we further have $\mathcal{A} \subseteq_{ec} \mathcal{A}_1$. By the same lemma we have a $\mathcal{A}_2$ such that $\mathcal{A} \preceq \mathcal{A}_2$, so $\mathcal{A}_2 \vDash T^{\forall\exists}$. If we repeat this construction with $\mathcal{A}_2$ in place of $\mathcal{A}$ and so on, we get a chain $\mathcal{A}_0 = \mathcal{A}, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \dots$ where $\mathcal{A}_{2n} \preceq \mathcal{A}_{2n+2}$ and $\mathcal{A}_{2n} \subseteq_{ec} \mathcal{A}_{2n+1} \vDash T$. So let us define $\mathcal{C} := \bigcup_{n=0}^\infty \mathcal{A}_n$, and so surely $\mathcal{C} := \bigcup_{n=0}^\infty \mathcal{A}_{2n} = \bigcup_{n=0}^\infty \mathcal{A}_{2n+1}$. Thus we have that $\mathcal{A} \preceq \mathcal{C}$ by Tarski's Chain Lemma. By the construction we have that $\mathcal{A}_{2n+1} \vDash T$ and since T is inductive we must have that $\mathcal{C} \vDash T$. Since $\mathcal{A} \preceq \mathcal{C}$ we have that $\mathcal{A} \vDash T$ as required. ■

One ready application of this theorem is that SO_{00}, SO_{10}, and SO_{10} are not $\forall\exists$-theories by example 5.4.9.

The construction utilized in the above theorem can be generalized, but we will not do so.

Let us say that two theories are T_0 and T_1 are *compatible* if $T_0 + T_1$ is consistent (this is not a new definition). We of course can not infer compatibility from the consistency of T_0 and T_1: DO and SO are consistent but $\mathsf{DO} + \mathsf{SO}$ is not. We further say that T_0 and T_1 are *model compatible* if every T_0-model is embeddable into some T_1-model and vice versa. By theorem 5.4.4 this is equivalent to $T_0^\forall = T_1^\forall$, and so model compatibility is therefore an equivalence relation and so partitions the set of consistent $\mathcal{L}$-theories.

Notice that model compatibility does not imply compatibility, as we see in the following example:

5.4.13 Example

SO and DO are both model compatible: let $\mathcal{A} \vDash \mathsf{SO}$ then we claim that $\mathcal{A} \vDash \mathsf{DO}^\forall$ (and is thus embeddable into a DO-model). Let $a_1 < \ldots < a_n \in A$ then the model generated by these elements is simply $\{a_1, \ldots, a_n\}$ as the theory has no function symbols. This can obviously be embedded into a DO-model, as it is isomorphic to $\{1, \ldots, n\}$ which we can embed into $\mathbb{Q}$. Thus every finitely generated $\mathcal{A}$-substructure can be embedded into a DO-model and as we showed this means that $\mathcal{A}$ can be embedded into a DO-model. The converse follows similarly.

Exercise

Let X be a set of *positive sentences*: sentences constructed from prime formulas utilizing only $\wedge, \vee, \forall, \exists$. Prove that if $\mathcal{A} \vDash X$ then $\mathcal{B} \vDash X$ where $\mathcal{B}$ is a homomorphic image of $\mathcal{A}$. Thus $\operatorname{Md} X$ is closed under homomorphic images (the converse is also true).

Let φ be a positive formula and h a homomorphism, then we claim that $\mathcal{A} \vDash \varphi(\vec{a}) \implies \mathcal{B} := h\mathcal{A} \vDash \varphi(h\vec{a})$ for all $\vec{a} \in A^n$. We prove this by induction on φ: for a prime formula $r\vec{t}$ then we have that $(r\vec{t})^{\mathcal{B},h\vec{a}} = r^{\mathcal{B}}\vec{t}^{\mathcal{B},h\vec{a}}$. By term induction we can see that $t^{\mathcal{B},h\vec{a}} = ht^{\mathcal{A},\vec{a}}$, and so this is if and only if $r^{\mathcal{A}}\vec{t}^{\mathcal{A},\vec{a}}$ since h is a homomorphism, which is equivalent to $\mathcal{A} \vDash r\vec{t}(\vec{a})$. The inductive step for $\wedge, \vee$ is simple. Now, $\mathcal{A} \vDash \forall x\varphi(x, \vec{a})$ means that for any arbitrary $a_0 \in A$, $\mathcal{A} \vDash \varphi(a_0, \vec{a})$ and so by induction $\mathcal{B} \vDash \varphi(ha_0, h\vec{a})$. Since a_0 is arbitrary and ha_0 exhausts B, this means $\mathcal{B} \vDash \forall x\varphi(h\vec{a})$. Similar for $\exists$. This proves the desired result for when φ has no free variables.

Exercise

Suppose that T_0 and T_1 are model compatible and inductive. Show then that $T_0 + T_1$ is also inductive and model compatible with T_0 and T_1.

$T_0 + T_1$ being inductive is trivial: any chain of $T_0 + T_1$-models is also a chain of T_0 and T_1-models and by their own inductiveness the union of the chain is a T_0 and T_1-model and is thus a $T_0 + T_1$-model. Alternatively, we have that the union of the $\forall\exists$-axiom systems of T_0 and T_1 forms a $\forall\exists$-axiom system for $T_0 + T_1$ so it is therefore a $\forall\exists$-theory.

Now we must show that $(T_0 + T_1)^\forall = T_0^\forall = T_1^\forall$. To do so we will show that $\operatorname{Md}(T_0 + T_1)^\forall = \operatorname{Md} T_0^\forall = \operatorname{Md} T_1^\forall$, the only nontrivial direction is $\operatorname{Md} T_0^\forall \subseteq \operatorname{Md}(T_0 + T_1)^\forall$. So let $\mathcal{A} \vDash T_0^\forall$, which means $\mathcal{A} \subseteq \mathcal{A}_0 \vDash T_0$. And since T_0 and T_1 are model compatible, this means there exists an $\mathcal{A}_1$ such that $\mathcal{A}_0 \subseteq \mathcal{A}_1 \vDash T_1$, and so on we can construct a chain $\mathcal{A}_/ \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots$ where $\mathcal{A}_{2n} \vDash T_0$ and $\mathcal{A}_{2n+1} \vDash T_1$. So let

$$\mathcal{C} := \bigcup_{n=0}^{\infty} \mathcal{A}_n = \bigcup_{n=0}^{\infty} \mathcal{A}_{2n} = \bigcup_{n=0}^{\infty} \mathcal{A}_{2n+1}$$

and since T_0 and T_1 are inductive, we get that $\mathcal{C} \vDash T_0, T_1$ and since $\mathcal{A} \subseteq \mathcal{C}$ we get that $\mathcal{A}$ is a substructure of a $T_0 + T_1$-model so $\mathcal{A} \in \operatorname{Md}(T_0 + T_1)^\forall$ as required.

Exercise

For an inductive theory T, show that of all its inductive extensions model compatible with T, there exists a largest one. This is called the *inductive completion* of T. For the theory of fields, for example, this is ACF.

Suppose $\{T_i\}_{i\in I}$ are of T's inductive extensions which are also model compatible with T, then let us define $T' = \bigcup_{i\in I} T_i$ (or rather T' is the theory generated by this). We claim that T' is inductive and model compatible with T. Suppose K is a chain of T'-models, then every $\mathcal{A} \in K$ is a T_i-model for every $i \in I$ and thus by the inductive nature of each T_i, $\mathcal{C} = \bigcup K$ is also a T_i-model, and therefore it is a T-model as required.

Now since $T \subseteq T'$ we must have that every T'-model can be trivially embedded into a T-model, so all that remains is to prove the converse. So let $\mathcal{A} \vDash T$, then we claim that $T' + D\mathcal{A}$ is consistent, and this is of course sufficient. Otherwise by the compactness theorem there exists $T_1, \ldots, T_n \in \{T_i\}_{i\in I}$ such that $T_1, \ldots, T_n, D\mathcal{A}$ is inconsistent. But by the above exercise $T_1 + \cdots + T_n$ is model compatible with T and so $\mathcal{A}$ can be embedded into a $T_1 + \cdots + T_n$-model $\mathcal{B}$, without loss of generality $\mathcal{A} \subseteq \mathcal{B}$. An so $\mathcal{B} \vDash D\mathcal{A}$, which means that $T_1, \ldots, T_n, D\mathcal{A}$ has a model in contradiction.

5.5 Model Completeness

5.5.1 Definition

A theory T is **model complete** if for every model $\mathcal{A} \vDash T$, then $\mathcal{L}\mathcal{A}$-theory $T + D\mathcal{A}$ is complete.

Notice that if $\mathcal{A}, \mathcal{B} \vDash T$ where $\mathcal{A} \subseteq \mathcal{B}$ then since $\mathcal{B}_A \vDash D\mathcal{A}$ we have that $\mathcal{A}_A \equiv \mathcal{B}_A$ by the completeness of $T + D\mathcal{A}$. And so if T is a model complete theory, then it has the property that for every $\mathcal{A}, \mathcal{B} \vDash T$, $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} \preceq \mathcal{B}$. Conversely if a theory has this property then it must be model complete: let $\mathcal{B} \vDash T + D\mathcal{A}$ and so we can assume $\mathcal{A} \subseteq \mathcal{B}$ and thus $\mathcal{A} \preceq \mathcal{B}$ so $\mathcal{B}$ is elementarily equivalent to $\mathcal{A}_A$ and this means that $T + D\mathcal{A}$ is complete.

Clearly if T is a model complete in $\mathcal{L}$ then so to is every one of its extensions (since $T + D\mathcal{A} \subseteq T' + D\mathcal{A}$). Furthermore a model complete theory is inductive: for if K is a chain of T-models then by above it is an elementary chain and so $\mathcal{A} \preceq \bigcup K$ for every $\mathcal{A} \in K$. In particular $\mathcal{A} \equiv \bigcup K$ so $\bigcup K \vDash T$. Thus only $\forall\exists$-theories can be model complete.

But not all $\forall\exists$-theories are model complete, for example DO is not model complete. Define $\mathbb{Q}_a := \{x \in \mathbb{Q} \mid a \leq x\}$ for every $a \in \mathbb{Q}$, then $(\mathbb{Q}_1, <) \subseteq (\mathbb{Q}_0, <)$ but not $(\mathbb{Q}_1, <) \preceq (\mathbb{Q}_1, <)$ since $\forall x\, 1 \leq x$ is valid only in $\mathbb{Q}_1$. These models also shows that the complete theory DO_{10} is not model complete. And similarly a model complete theory need not be complete: one such example is the theory of algebraically closed fields ACF which will be considered later on this section.

5.5.2 Theorem

A model complete theory T with a prime model is complete.

Let $\mathcal{P} \vDash T$ be a prime model, then for every $\mathcal{A} \vDash T$ we have that $\mathcal{P}$ is up to isomorphism a substructure of $\mathcal{A}$, and thus $\mathcal{P} \preceq \mathcal{A}$. In particular $\mathcal{A} \equiv \mathcal{P}$ and so all T-models are elementarily equivalent, meaning T is complete. ■

5.5.3 Theorem

For a theory T, the following are equivalent:

(1) T is model complete,

(2) $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} \subseteq_{ec} \mathcal{B}$ for all T-models $\mathcal{A} \subseteq \mathcal{B}$,

(3) every $\exists$-formula α is equivalent in T to a $\forall$-formula β with $\mathit{free}\beta \subseteq \mathit{free}\alpha$,

(4) every formula α is equivalent in T to a $\forall$-formula β with $\mathit{free}\beta \subseteq \mathit{free}\alpha$.

(1) $\implies$ (2) is trivial as $\mathcal{A} \preceq \mathcal{B} \implies \mathcal{A} \subseteq_{ec} \mathcal{B}$. (2) $\implies$ (3): by theorem 5.4.8, it is sufficient to show that α is persistent in T. So suppose $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{B} \vDash \alpha(\vec{a})$ for $\vec{a} \in A^n$, and since $\mathcal{A}$ is existentially closed in $\mathcal{B}$ this means precisely that $\mathcal{A} \vDash \alpha(\vec{a})$, meaning α is persistent in T. (3) $\implies$ (4): we induct on α. Prime formulas are $\exists$-formulas and are handled by assumption. The induction step for $\wedge$ and $\forall$ are obvious (since for $\forall x\alpha$, convert α to an equivalent $\forall$-formula and then you get a $\forall$-formula). For $\alpha = \neg\alpha'$ then α' is equivalent to a $\forall$-formula by induction, and so α is equivalent to an $\exists$-formula and thus by assumption a $\forall$-formula. (4) $\implies$ (1): let $\mathcal{A}, \mathcal{B} \vDash T$ and $\mathcal{B} \vDash \alpha(\vec{a})$ for $\vec{a} \in A^n$, then $\alpha(\vec{x})$ is equivalent to a $\forall$-formula and since $\forall$-formulas are preserved by substructures, we have $\mathcal{A} \vDash \alpha(\vec{a})$. And conversely if $\mathcal{A} \vDash \alpha(\vec{a})$, notice that $\alpha(\vec{x})$ is equivalent to an $\exists$-formula (since $\neg\alpha$ is equivalent to a $\forall$-formula) which are preserved by superstructures, so $\mathcal{B} \vDash \alpha(\vec{a})$. Thus $\mathcal{A} \preceq \mathcal{B}$. ■

The implication (2) $\implies$ (1) is called *Robinson's test* for model completeness.

If T is countable and has finite models only, then (2) can be restricted to models $\mathcal{A}, \mathcal{B}$ of any set cardinality $\kappa \geq \aleph_0$. Then it can be shown that every $\exists$-formula is κ-persistent, which is sufficient to show its equivalence to a $\forall$-formula as explained in the remark after theorem 5.4.8. So this restricted (2) still implies (3) and thus the rest of the chain of implications holds. This remark is important for the proof of Lindström's Criterion .

A simpler example of a model complete theory is the theory of nontrivial $\mathbb{Q}$-vector spaces $T_{V\mathbb{Q}}$ over the signature $\{+, 0, \mathbb{Q}\}$ where 0 is the zero vector and each $r \in \mathbb{Q}$ is a unary operation corresponding to scalar multiplication by r (so for example $\forall ab\big(r(a + b) = ra + rb\big)$ would be an axiom schema for all $r \in \mathbb{Q}$). Let $\mathcal{V}, \mathcal{V}' \vDash T_{V\mathbb{Q}}$ be two $\mathbb{Q}$-vector spaces with $\mathcal{V} \subseteq \mathcal{V}'$ then we claim $\mathcal{V} \subseteq_{ec} \mathcal{V}'$ (ie. we are trying to use Robinson's test). Suppose $\mathcal{V}' \vDash \exists \vec{x}\alpha(\vec{x}, \vec{a}, \vec{b})$ for $\vec{a} \in V^m, \vec{b} \in V^k$ where α is a conjunction of literals. Then α is equivalent to a system of the form

$$\begin{cases} r_{11}x_1 + \cdots + r_{1n}x_n = a_1 & s_{11}x_1 + \cdots + s_{1n} \neq b_1 \\ \vdots & \vdots \\ r_{m1}x_1 + \cdots + r_{mn}x_n = a_m & s_{k1}x_1 + \cdots + s_{kn}x_n \neq b_k \end{cases}$$

Now it can be shown that if this has a solution in $\mathcal{V}'$, it must have a solution in $\mathcal{V}$ as well.

The following concept generalizes the idea of closures in algebra, like the algebraic closure of a field, real closure of an ordered field, and the divisible closure of an abelian group.

5.5.4 Definition

Let T be a theory and $\mathcal{A} \vDash T^{\forall}$ (ie. $\mathcal{A}$ is a substructure of some T-model). Then the **closure of $\mathcal{A}$ in T** is the smallest T-model containing $\mathcal{A}$, ie. it is a T-model $\overline{\mathcal{A}} \supseteq \mathcal{A}$ such that if $\mathcal{A} \subseteq \mathcal{B} \vDash T$ then $\overline{\mathcal{A}} \subseteq \mathcal{B}$, if it exists. If every $\mathcal{A} \vDash T^{\forall}$ has a closure in T, then we say that T **permits a closure operation**.

Now suppose that T does indeed permit a closure operation, and $\mathcal{A}, \mathcal{B} \vDash T$ with $\mathcal{A} \subset \mathcal{B}$, then let $b \in B \setminus A$. Then let $\mathcal{A}(b)$ be the structure generated by $A \cup \{b\}$, which is a substructure of $\mathcal{B}$ and thus $\mathcal{A}(b) \vDash T^{\forall}$. We then denote the closure of $\mathcal{A}(b)$ in T by $\mathcal{A}^b$. As $\mathcal{A} \subset \mathcal{A}^b \subseteq \mathcal{B}$, $\mathcal{A}^b$ is called an *immediate extension* of $\mathcal{A}$ in T.

5.5.5 Example

Let $T = \mathsf{ACF}$ be the theory of algebraically closed fields. Then a $T^{\forall}$-model $\mathcal{A}$ is an integral domain (as every integral domain can be embedded into its field of fractions, which has an algebraic closure). $\overline{\mathcal{A}}$ is the *algebraic closure* of the field of fractions of $\mathcal{A}$, this is a well-known result from field theory. Now suppose $\mathcal{A}, \mathcal{B} \vDash T$ with $\mathcal{A} \subset \mathcal{B}$ and let $b \in \mathcal{B} \setminus \mathcal{A}$, then b is transcendental in $\mathcal{A}$ as it is algebraically closed, meaning that $a_0 + a_1 b + \cdots + a_n b^n \neq 0$ for $a_0, \ldots, a_n \in A$ with $a_n \neq 0$. Thus $\mathcal{A}(b)$ is isomorphic to the ring of polynomials $\mathcal{A}[x]$, and so we have that for every $\mathcal{A}, \mathcal{B}, \mathcal{C} \vDash T$ with $\mathcal{A} \subset \mathcal{B}, \mathcal{C}$ and $b \in B \setminus A$ and $c \in C \setminus A$, $\mathcal{A}(b) \cong \mathcal{A}[x] \cong \mathcal{A}(c)$. And so their closures $\mathcal{A}^b$ and $\mathcal{A}^c$ are also isomorphic. Thus every algebraically closed field has up to isomorphism a single immediate extension.

We can simplify Robinson's test in specific cases where the theory is inductive and permits a closure operation.

5.5.6 Lemma

Let T be an inductive theory which permits a closure operation. Then further assume that $\mathcal{A} \subseteq_{ec} \mathcal{A}'$ for all $\mathcal{A}, \mathcal{A}' \vDash T$ where $\mathcal{A}'$ is an immediate extension of $\mathcal{A}$ in T. Then T is model complete.

Let $\mathcal{A} \subseteq \mathcal{B}$ be two T-models, then by Robinson's test it is sufficient to show that $\mathcal{A} \subseteq_{ec} \mathcal{B}$. Let us define H to be the set of all $\mathcal{C} \subseteq \mathcal{B}$ such that $\mathcal{A} \subseteq_{ec} \mathcal{C} \vDash T$. We trivially have that $\mathcal{A} \in H$ and since T is inductive, for every chain $K \subseteq H$ we have $\bigcup K \vDash T$. Since a chain $K \subseteq H$ is an existentially closed chain (since for every $\mathcal{A}_1 \in K$ and $\mathcal{A}_2 \in K$ with $\mathcal{A}_1 \subseteq \mathcal{A}_2$ we have that $\mathcal{A} \subseteq_{ec} \mathcal{A}_2$ so every $\exists$-formula valid in $\mathcal{A}_2$ is valid in $\mathcal{A} \subseteq \mathcal{A}_1$ and is thus valid in $\mathcal{A}_1$). We can add $\mathcal{A}$ to the chain which does not affect its union, and then we have that $\mathcal{A} \subseteq_{ec} \bigcup K$ by the chain lemma for existentially closed chains. This means that $\bigcup K \in H$ and so H has a maximal element $\mathcal{A}_m$ by Zorn's lemma.

We now claim that $\mathcal{A}_m = \mathcal{B}$, which would mean $\mathcal{B} \in H$ and thus $\mathcal{A} \subseteq_{ec} \mathcal{B}$ as required. Assume that $\mathcal{A}_m \subset \mathcal{B}$ then there exists an immediate extension of $\mathcal{A}_m$ in T, $\mathcal{A}'_m \vDash T$ such that $\mathcal{A}_m \subset \mathcal{A}'_m \subseteq \mathcal{B}$. By assumption we have $\mathcal{A} \subseteq_{ec} \mathcal{A}_m \subseteq_{ec} \mathcal{A}'_m$, which means that $\mathcal{A}'_m \in H$ contradicting $\mathcal{A}_m$'s maximality. ■

5.5.7 Theorem

ACF is model complete and therefore so too is ACF_p, the theory of algebraically closed fields of characteristic p ($= 0$ or prime). Furthermore ACF_p is complete.

We will use the above lemma to prove ACF's model completeness. Let $\mathcal{A}, \mathcal{B} \vDash \mathsf{ACF}$ with $\mathcal{A} \subset \mathcal{B}$ and $b \in B \setminus A$, then by the above lemma all we must show is that $\mathcal{A} \subseteq_{ec} \mathcal{A}^b$. Let $\alpha := \exists \vec{x}\beta(\vec{x}, \vec{a}) \in \mathcal{L}A$, β quantifier-free, and suppose that $\mathcal{A}^b \vDash \alpha$. We must prove $\mathcal{A} \vDash \alpha$. Let us define

$$X := \mathsf{ACF} \cup D\mathcal{A} \cup \{p(x) \neq 0 \mid p(x) \text{ is a monic polynomial over } \mathcal{A}\}$$

We see that $(\mathcal{A}^b, b) \vDash X$ (with b for x, since $\mathit{free}X = \{x\}$) since b is transcendental over $\mathcal{A}$ so by definition $p(b) \neq 0$ for any monic polynomial over $\mathcal{A}$. Let $(\mathcal{C}, c) \vDash X$, since $\mathcal{C} \vDash D\mathcal{A}$ without loss of generality $\mathcal{A} \subseteq \mathcal{C}$. By the above example we have that $\mathcal{A}^c \cong \mathcal{A}^b$ and therefore $\mathcal{A}^c \vDash \alpha$ and as an $\exists$-formula, $\mathcal{C} \vDash \alpha$. Since $(\mathcal{C}, c)$ is arbitrary we have that $X \vdash \alpha$, and by the compactness theorem

$$D\mathcal{A}, \bigwedge_{i=1}^{k} p_i(x) \neq 0 \vdash_{\mathsf{ACF}} \alpha$$

for some k where $p_1(x), \dots, p_k(x)$ are monic polynomials over $\mathcal{A}$. Then by particularization and the deduction theorem we have $D\mathcal{A} \vdash_{\mathsf{ACF}} \exists x \bigwedge_{i=1}^{k} p_i(x) \neq 0 \to \alpha$. Therefore $\mathcal{A} \vDash \exists x \bigwedge_{i=1}^{k} p_i(x) \neq 0 \to \alpha$. Since $\mathcal{A}$ is an algebraically closed field it is infinite (as if it were finite $\prod_{a \in A}(x-a)+1$ would have no zeroes), and polynomials have finitely many zeroes in fields, meaning that $\mathcal{A} \vDash \exists x \bigwedge_{i=1}^{k} p_i(x) \neq 0$ and so $\mathcal{A} \vDash \alpha$ as required.

ACF_p is model complete as an extension of ACF, and it has a prime model — the algebraic closure of the prime field of characteristic p. Thus by theorem 5.5.2, ACF_p is complete. ■

5.5.8 Definition

The **model completion** of an $\mathcal{L}$-theory T_0 is an extension $T \subseteq \mathcal{L}^0$, such that $T + D\mathcal{A}$ is complete for every $\mathcal{A} \vDash T_0$.

Obviously the model completion of a theory, if one exists, is model complete (since a model of T is a model of T_0). T is also model compatible with T_0, since if $\mathcal{A} \vDash T_0$ then since $T + D\mathcal{A}$ is consistent, there exists a $\mathcal{B} \vDash T + D\mathcal{A}$ so $\mathcal{B}$ is a T-model in which one can embed $\mathcal{A}$ (and trivially every T model can be embedded in itself, a T_0 model). If a model completion exists, it is unique: suppose T and T' are model completions of T_0. Since both theories a model compatible with T_0, they are with each other. And further since they are model complete and therefore inductive, we showed in an exercise that $T + T'$ is also model compatible with T. So if $\mathcal{A} \vDash T$ then it can be embedded into some $\mathcal{B} \vDash T + T'$. Since T is model complete this means $\mathcal{A} \preceq \mathcal{B}$ and so $\mathcal{A} \equiv \mathcal{B}$, meaning $\mathcal{A} \vDash T'$. By symmetry we have that every T'-model is a T-model and thus $\mathrm{Md}\, T = \mathrm{Md}\, T'$, so $T = T'$.

5.5.9 Example

ACF is the model completion of the theory T_J of integral domains, and thus also the theory T_F of fields (since if $\mathcal{A} \vDash T_F$ then $\mathcal{A} \vDash T_J$ so $\mathsf{ACF} + D\mathcal{A}$ is complete). To show this, let $\mathcal{A} \vDash T_J$, and since ACF is model complete so too is $T := \mathsf{ACF} + D\mathcal{A} \subseteq \mathcal{L}A$ T also has a prime model, the closure $\overline{\mathcal{A}}$, and so is complete as required.

5.5.10 Definition

$\mathcal{A} \vDash T$ is **existentially closed** in T, for short $\exists$-closed, if $\mathcal{A} \subseteq_{ec} \mathcal{B}$ for every $\mathcal{B} \vDash T$ where $\mathcal{A} \subseteq \mathcal{B}$.

For example every algebraically closed field is $\exists$-closed in the theory of fields. Let $\mathcal{A} \vDash \mathsf{ACF}$ and $\mathcal{A} \subseteq \mathcal{B} \vDash T_F$, and then let $\mathcal{C}$ be an algebraic extension of $\mathcal{B}$ then by the model completeness of ACF we have $\mathcal{A} \preceq \mathcal{C}$, and thus by lemma 5.4.11, we have that $\mathcal{A} \subseteq_{ec} \mathcal{B}$. The following lemma, in a sense, generalizes the fact that every field is embeddable in an algebraically closed field.

5.5.11 Lemma

Let T be a $\forall\exists$-theory over some countable language $\mathcal{L}$. Then every infinite T-model $\mathcal{A}$ can be extended to a T-model $\mathcal{A}^*$ which is $\exists$-closed in T and $|\mathcal{A}| = |\mathcal{A}^*|$.

We assume for simplicity that $\mathcal{A}$ is countably infinite, if $\mathcal{A}$ is uncountable then the following proof proceeds similarly but utilizing an ordinal enumeration instead of a normal one when necessary. Since $\mathcal{A}$ is countable, so too is $\mathcal{L}A$, so let $\alpha_0, \alpha_1, \dots$ be an enumeration of the $\exists$-sentences of $\mathcal{L}A$. Let $\mathcal{A}_0 = \mathcal{A}_A$, and inductively define $\mathcal{A}_{n+1}$ to be an extension of $\mathcal{A}_n$ in $\mathcal{L}A$ such that $\mathcal{A}_{n+1} \vDash T + \alpha_n$ if such an extension exists, otherwise have $\mathcal{A}_{n+1} := \mathcal{A}_n$. Since T is inductive, $\mathcal{B}_0 := \bigcup_{n=0}^{\infty} \mathcal{A}_n \vDash T$. If $\alpha = \alpha_n$ is valid in some extension $\mathcal{B} \vDash T$ of $\mathcal{B}_0$, then $\mathcal{A}_{n+1} \vDash \alpha$ and so $\mathcal{B}_0 \vDash \alpha$ as it is an $\exists$-sentence. Now we repeat this construction with an enumeration of all the $\exists$-sentences in $\mathcal{L}B_0$ to obtain a $\mathcal{L}B_0$-structure $\mathcal{B}_1 \vDash T$. In such a way we get a sequence $\mathcal{B}_0 \subseteq \mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \cdots$

of $\mathcal{L}B_n$-structures $\mathcal{B}_{n+1} \vDash T$. Let $\mathcal{A}^*$ be the $\mathcal{L}$-reduct of $\bigcup_{n=0}^{\infty} \mathcal{B}_n \vDash T$. Now suppose $\mathcal{A}^* \subseteq \mathcal{B} \vDash T$, and assume $\mathcal{B} \vDash \exists \vec{x} \beta(\vec{a}, \vec{x})$ for $\vec{a} \in (A^*)^n$. This means that $\vec{a} \in B_m^n$ and $\mathcal{B}_m \vDash \beta(\vec{a}, \vec{b})$ for suitable m, and so $\bigcup_{n=0}^{\infty} \mathcal{B}_n \vDash \beta(\vec{a}, \vec{b})$ so $\mathcal{A}^* \vDash \exists \vec{x} \beta(\vec{a}, \vec{x})$. Notice that a countable amount of countable unions were performed to construct $\mathcal{A}^*$ and thus it too is countable. ■

5.5.12 Theorem (Lindström's Criterion)

A countable κ-categorical $\forall\exists$-theory T without finite models is model complete.

Since all T-models are infinite, T has a model of cardinality κ by the Löwenheim-Skolem theorems. And by the above lemma, T has a model which is $\exists$-closed in the theory. This means that all T-models of cardinality κ are $\exists$-closed as they are all isomorphic, and so $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} \subseteq_{ec} \mathcal{B}$ for all T-models $\mathcal{A}, \mathcal{B}$ of cardinality κ. By the remark after theorem 5.5.3, this means that T is model complete. ■

By Vaught's test, any theory satisfying Lindström's criterion is also complete.

5.5.13 Example

The following theories are therefore model complete:

(1) The $\aleph_0$-categorical theory of atomless Boolean algebras. An atomless Boolean algebra is a Boolean algebra $\mathcal{B}$ where for every $a \neq 0$ (where $0 = x \cap \neg x$) there is some $b \neq 0$ such that $b < a$ ($<$ is the partial lattice order $a \leq b \iff a \cup b = b \iff a \cap b = a$).

(2) The $\aleph_1$-categorical theory of $\mathbb{Q}$-vector spaces. This is as a rational vector space of cardinality $\aleph_1$ has a basis of cardinality $\aleph_1$ and so all rational vector spaces of cardinality $\aleph_1$ are isomorphic.

(3) The $\aleph_1$-categorical theory of ACF_p for any prime p or $p = 0$. Notice that we get the model completeness of ACF_p in a method independent of ACF, and this implies the model completeness of ACF: if $\mathcal{A}, \mathcal{B} \vDash \mathsf{ACF}$ and $\mathcal{A} \subseteq \mathcal{B}$ then both fields have the same characteristic $p \geq 0$, and so $\mathcal{A}, \mathcal{B} \vDash \mathsf{ACF}_p$. Since ACF_p is model complete, we have that $\mathcal{A} \preceq \mathcal{B}$, so ACF is also model complete.

Exercise

Prove that of the four theories DO_{ij}, only DO_{00} is model complete. Moreover show that DO_{00} is the model completion of DO.

Notice that $[1, \infty) \subseteq [0, \infty)$ are both DO_{10}-models and $[1, \infty) \vDash \forall x\, 1 \leq x$, but $[0, \infty) \nvDash \forall x\, 1 \leq x$ so we do not have that $[1, \infty) \preceq [0, \infty)$, meaning DO_{10} is not model complete. Similar counterexamples can be found for DO_{01} and DO_{11}. Now, DO_{00} is a $\forall\exists$-theory, since $\neg\mathsf{L}$ and $\neg\mathsf{R}$ are $\forall\exists$-sentences and DO is a $\forall\exists$-theory. Furthermore we have shown that it is $\aleph_0$-categorical and has no finite models, so by Lindström's criterion, it is model complete. Now let $\mathcal{A} \vDash \mathsf{DO}$, we will suppose $\mathcal{A} \vDash \mathsf{DO}_{10}$, then we can extend it to a DO_{00}-model $\mathbb{Q} + \mathcal{A}$ where $p < a$ for all $p \in \mathbb{Q}$ and $a \in A$. This is a prime model of $\mathsf{DO}_{00} + D\mathcal{A}$, and therefore $\mathsf{DO}_{00} + D\mathcal{A}$ is complete (since it is model complete as an extension of DO_{00} and has a prime model), meaning DO_{00} is the model completion of DO.

Exercise

Let T be the theory of torsion-free divisible abelian groups. Show that T is the model completion of the theory T_0 of torsion-free abelian groups.

Let $\mathcal{A} \vDash T_0$ then we can extend it to a torsion-free divisible abelian group as follows: define the equivalence relation $\sim$ on $A \times \mathbb{N}$ by $\frac{a}{n} \sim \frac{b}{m}$ if and only if $am = bn$. Then define addition on the quotient by $\frac{a}{n} + \frac{b}{m} = \frac{am+bn}{nm}$. This is a torsion-free divisible abelian group, it is torsion free since $n\frac{a}{n} = \frac{a}{1}$ (which we view as a), which has the same order as a, ie. infinite. And it is divisible since $n\frac{a}{nm} = \frac{a}{m}$. This is a prime model for $T + D\mathcal{A}$ since every divisible extension of $\mathcal{A}$ must define its equivalence of $\frac{a}{n}$ (ie. there exists some b such that $nb = a$, so map $\frac{a}{n}$ to b. This defines an embedding). And since T is model complete (we have already shown that it is $\aleph_1$-categorical as a T-model is essentially a $\mathbb{Q}$-vector space, so its model completeness follows from Lindström's criterion), so too is $T + D\mathcal{A}$, therefore it is complete as required.

Exercise

T^* is called the **model companion** provided that T and T^* are model compatible and T^* is model complete. Show that if T^* exists, it is unique and further that every $\mathcal{A} \vDash T^*$ is $\exists$-closed in T.

Suppose T_1 and T_2 are two model companions of T, then T, T_1, T_2 are all model compatible, and therefore so too is $T_1 + T_2$. So let $\mathcal{A} \vDash T_1$, then $T_1 + T_2 + D\mathcal{A}$ is consistent so let $\mathcal{B} \vDash T_1 + T_2 + D\mathcal{A}$, meaning without loss of generality $\mathcal{A} \subseteq \mathcal{B}$. Since $\mathcal{B} \vDash T_1, T_2$ and T_1 is model complete, $\mathcal{A} \preceq \mathcal{B}$ and in particular $\mathcal{A} \equiv \mathcal{B}$, so $\mathcal{A} \vDash T_2$. We get the converse by symmetry, so we have $\operatorname{Md} T_1 = \operatorname{Md} T_2$ as required.

Let $\mathcal{A} \vDash T^*$ and $\mathcal{B} \vDash T$ such that $\mathcal{A} \subseteq \mathcal{B} \vDash T$. Since T and T^* are model compatible, there exists a $\mathcal{C} \vDash T^*$ such that $\mathcal{B} \subseteq \mathcal{C}$, and since $\mathcal{A} \preceq \mathcal{C}$ by model completeness of T^*, we have that $\mathcal{A} \subseteq_{ec} \mathcal{C}$ as required.

5.6 Quantifier Elimination

We say that a theory T *allows quantifier elimination* if for every $\varphi \in \mathcal{L}$ there exists a quantifier-free φ' such that $\varphi \equiv_T \varphi'$. By theorem 5.5.3, a theory which allows quantifier elimination is model complete since quantifier-free formulas are also $\forall$-formulas. Therefore a theory which allows quantifier elimination must be a $\forall\exists$-theory.

Since $\{\exists, \neg, \wedge\}$ is a complete bundle in the sense that every formula generated by this bundle is equivalent to one generated by $\{\forall, \neg, \wedge\}$ and vice versa, to show that a theory allows quantifier elimination we need only show that $\exists x\alpha$ is equivalent to a quantifier-free formula for α quantifier-free. This is as we can induct over the construction over formulas and convert them to quantifier-free equivalents.

We can put α into DNF, and since $\exists$ distributes over disjunctions, we can assume that α is a conjunction of literals. If any of these literals do not contain x, then its prefix $\exists x$ can be disregarded as $\exists x(\alpha \wedge \beta) \equiv \exists x\alpha \wedge \beta$ for $x \notin var\beta$, so we can further assume that x occurs in each literal. Furthermore we can assume that no literal is of the form $x = t$ for $x \notin vart$ as $\exists x(x = t \wedge \alpha) \equiv \alpha\frac{t}{x}$, and thus the quantifier has been eliminated. Using bound renaming we can assume that $x \neq \boldsymbol{v}_0$ and so neither $x = x$ nor $x \neq x$ are literals occurring in any conjunction, as they can be replaced with $\top$ and $\bot$ respectively (the reason for $x \neq \boldsymbol{v}_0$ is as $\top := \boldsymbol{v}_0 = \boldsymbol{v}_0$ and $\bot := \boldsymbol{v}_0 \neq \boldsymbol{v}_0$. Alternatively we could simply define $\top$ and $\bot$ as new prime formulas).

Call an $\exists$-formula *simple* if it is of the form $\exists x \bigwedge_i \alpha_i$ where α_i is a conjunction of literals, and for every i, $x \in var\alpha_i$. Then we have proven the following

5.6.1 Theorem

T allows quantifier elimination if and only if every simple $\exists$-formula $\exists x \bigwedge_i \alpha_i$ is equivalent in T to some quantifier-free formula. We can further assume that none of the literals α_i are $x = x$, $x \neq x$, or $x = t$ with $x \in vart$.

5.6.2 Example

$T = \mathsf{DO}_{00}$ allows quantifier elimination. Notice that

$$y \not< z \equiv_T z < y \vee z = y, \quad z \neq y \equiv_T z < y \vee y < z$$

and since $(\alpha \vee \beta) \wedge \gamma \equiv (\alpha \wedge \gamma) \vee (\beta \wedge \gamma)$, we may assume that the conjunctions α_i contain no negations. Thus we deal only with formulas of the form

$$\exists x(y_1 < x \wedge \cdots \wedge y_n < x \wedge x < z_1 \wedge \cdots \wedge x < z_m)$$

This is equivalent to $\bot$ if $x = y_i$ or $x = z_i$ for any i. If $n = 0$ or $m = 0$ this is equivalent to $\top$. Otherwise it is equivalent to $\bigwedge_{i,j} y_i < z_j$ which is quantifier-free.

Notice that DO itself does not allow quantifier elimination as $\alpha(y) := \exists x\, x < y$ has no quantifier-free equivalent. If it were then for any $\mathcal{A}, \mathcal{B} \vDash \mathsf{DO}$ where $\mathcal{A} \subseteq \mathcal{B}$ and $a \in A$, $\mathcal{B} \vDash \alpha(a) \implies \mathcal{A} \vDash \alpha(a)$. But let $\mathcal{B} \vDash \mathsf{DO}_{00}$ and $\mathcal{A} \vDash \mathsf{DO}_{10}$ and let a be $\mathcal{A}$'s right edge element (eg. $A = [1, \infty) \cap \mathbb{Q}$ and $B = \mathbb{Q}$, $a = 1$).

And SO does not allow quantifier elimination as it is not a $\forall\exists$-theory. And the same holds for SO_{ij}.

5.6.3 Example

Let ZGE be the theory in $\mathcal{L}\{0,1,+,-,<,2|,3|,\dots\}$, whose axioms include the axioms for ordered abelian groups as well as

$$\forall x(0<x\leftrightarrow 1\le x),\quad \forall x(m|x\leftrightarrow \exists y\, my=x),\quad \vartheta_m:=\forall x\bigvee_{k=0}^{m-1} m|x+k \qquad (m=2,3,\dots)$$

The reducts of ZGE-models to $\mathcal{L}:=\mathcal{L}\{0,1,+,-,<\}$ are called *$\mathbb{Z}$-groups*. ϑ_m states that for any $\mathbb{Z}$-group G, G/mG is cyclic of order m. Let ZG be the $\mathcal{L}$-reduct theory of ZGE (which is well-defined as $m|$ is explicitly defined in ZGE), its models are precisely all $\mathbb{Z}$-groups since ZGE is a definitorial and therefore conservative extension of ZG. Notice that $\vdash_{\mathsf{ZG}} \forall x\eta_n$ for every n where $\eta_n:=0\le x<n\to\bigvee_{k=0}^{n-1} x=k$.

We will now show that ZGE allows quantifier-elimination. Notice that

$$t\ne s\equiv_{\mathsf{ZGE}} s<t\vee t<s,\quad m\nmid t\equiv_{\mathsf{ZGE}}\bigvee_{i=1}^{m-1} m|t+i,\quad m|t\equiv_{\mathsf{ZGE}} m|-t$$

And so we can assume that the kernel of a simple $\exists x$-formula is a conjunction of formulas of the form $n_ix=t_i^0, n_i'x<t_i^1, t_i^2<n_i''x, m_i|n_i'''x+t_i^3$ where $x\notin \mathrm{var}t_i^j$. Since we have that $t<s\equiv_{\mathsf{ZGE}} nt<ns$ and $m|t\equiv_{\mathsf{ZGE}} nm|nt$ for $n\ne 0$ we can assume that $n_i=n_i'=n_i''=n_i'''=n>1$ for some n by multiplying both sides of each literal by a suitable value. This means that we can asssume that a simple $\exists$-formula is of the form

$$\exists x\left(\bigwedge_{i=1}^{k_0} nx=t_i^0\wedge\bigwedge_{i=1}^{k_1} t_i^1<nx\wedge\bigwedge_{i=1}^{k_2} nx<t_i^2\wedge\bigwedge_{i=1}^{k_3} m_i|nx+t_i^3\right)$$

Let $y=nx$ and $m_0=n$, we get that this is equivalent to

$$\exists y\left(\bigwedge_{i=1}^{k_0} y=t_i^0\wedge\bigwedge_{i=1}^{k_1} t_i^1<y\wedge\bigwedge_{i=1}^{k_2} y<t_i^2\wedge\bigwedge_{i=1}^{k_3} m_i|y+t_i^3\wedge m_0|y\right)$$

By the above theorem, we can assume $k_0=0$ and so by substituting x for y and setting $t_0^3=0$, this is equivalent to

$$\exists x\left(\bigwedge_{i=1}^{k_1} t_i^1<x\wedge\bigwedge_{i=1}^{k_2} x<t_i^2\wedge\bigwedge_{i=0}^{k_3} m_i|x+t_i^3\right)$$

Let m be the least common multiple of $m_0,\dots,m_{k_3}$. We split into cases:

Case 1: $k_1,k_2=0$. Then the formula is $\exists x\bigwedge_{i=0}^{k_3} m_i|x+t_i^3$, which is equivalent to $\bigvee_{j=1}^{m}\bigwedge_{i=0}^{k_3} m_i|x+t_i^3$. This is since if there exists an x such that $m_i|x+t_i^3$ for all $0\le i\le k_3$, then there must exist some $1\le j\le m$ which satifies this: by ϑ_m there exists such a j where $m|x+(m-j)$ and so $m|x-j$ and so $m_i|x-j$ for all relevant i, and so $m_i|x+t_i^3-(x-j)=j+t_i^3$ as required.

Case 2: $k_1\ne 0$. Let j be as above, then this is equivalent to

$$\bigvee_{\mu=1}^{k_1}\left(\bigwedge_{i=1}^{k_1} t_i^1\le t_\mu^1\wedge\bigvee_{j=1}^{m}\left(\bigwedge_{i=1}^{k_2} t_\mu^1+j<t_i^2\wedge\bigwedge_{i=0}^{k_3} m_i|t_\mu^1+j+t_i^3\right)\right)$$

This implies the formula, as we can take $x=t_\mu^1+j$ for the μ which is satisfied. Now suppose x witnesses the formula, then if $\bigwedge_{i=1}^{k_1} t_i^1\le t_\mu^1$ holds (which it must for some μ of course), then the rest of the disjuncts hold as well. We can take the same j as before, where $m_i|j+t_i^3$. To show $t_\mu^1+j<t_i^2$ it is sufficient to show $t_\mu^1+j\le x$. Suppose the converse then $0<x-t_\mu^1<j$, which means $x-t_\mu^1=k$ for some $k<j$ by η_j, so $x=k+t_\mu^1$. So $m_i|t_\mu^1+j-x=j-k$ for all i, but then $m|j-k<m$.

Case 3: $k_1=0$ and $k_2\ne 0$. This is similar to the previous case but using the smallest term among t_i^2 instead of the largest of t_i^1.

5.6.4 Corollary

ZGE is model complete. Furthermore, ZGE and ZG are both complete and decidable.

Since $\mathbb{Z}$ is a prime model of ZGE, it is complete due to it being model complete (by quantifier elimination) and having a prime model. Since ZGE is an explicit extension of ZG, it being complete implies so too is ZG. These theories are complete and axiomatizable and so they are decidable. ■

5.6.5 Definition

Let $X \subseteq \mathcal{L}$ be a set of $\mathcal{L}$-formulas, then it is a **boolean basis** for $\mathcal{L}$ if every formula $\varphi \in \mathcal{L}$ is in $\langle X \rangle_T$, ie. every formula is equivalent in T to a boolean combination of formulas from X.

Let $\mathcal{M}, \mathcal{M}'$ be $\mathcal{L}$-models, write $\mathcal{M} \equiv_X \mathcal{M}'$ to mean that $\mathcal{M} \vDash \varphi \iff \mathcal{M}' \vDash \varphi$ for all $\varphi \in X$.

5.6.6 Theorem (Basis Theorem for Formulas)

Let T be an $\mathcal{L}$-theory and $X \subseteq \mathcal{L}$. Then suppose that $\mathcal{M} \vDash_X \mathcal{M}' \implies \mathcal{M} \equiv \mathcal{M}'$ for all $\mathcal{M}, \mathcal{M}' \vDash T$. Then X is a boolean basis for $\mathcal{L}$ in T.

Let $\alpha \in \mathcal{L}$ and define $Y_\alpha := \{\gamma \in \langle X \rangle \mid \alpha_T \gamma\}$. Then as in the proof of the Basis Theorem for Sentences, show that $Y_\alpha \vdash \alpha$ arguing analogously but considering $\mathcal{L}$-models instead of $\mathcal{L}$-structures. ■

5.6.7 Definition

Call a theory T **substructure complete** if for all models $\mathcal{A} \subseteq \mathcal{B}$ where $\mathcal{B} \vDash T$, $T + D\mathcal{A}$ is complete. This is equivalent to saying that T is the model completion of $T^\forall$.

5.6.8 Theorem

For every $\mathcal{L}$-theory T, the following are equivalent:

(1) T allows quantifier elimination (2) T is substructure complete

(1) $\implies$ (2): let $\mathcal{A}$ be a substructure of a T-model, $\varphi(\vec{x}) \in \mathcal{L}$ and $a \in A^n$ such that $\mathcal{A} \vDash \varphi(\vec{a})$. Let $\mathcal{B} \vDash T, D\mathcal{A}$ so we can assume $\mathcal{A} \subseteq \mathcal{B}$. Since T allows quantifier elimination, we can assume that φ is quantifier-free and so $\mathcal{B} \vDash \varphi(\vec{a})$. And since $\mathcal{B} \vDash T$ is arbitrary we have $D\mathcal{A} \vdash_T \varphi(\vec{a})$, so $D\mathcal{A} + T = Th\mathcal{A}_A$, meaning it is complete.
(2) $\implies$ (1): let X be the set of literals in $\mathcal{L}$, then saying that T allows quantifier elimination is the same as saying that X is a boolean basis for $\mathcal{L}$ in T. So we will prove that for all $\mathcal{M}, \mathcal{M}' \vDash T$, $\mathcal{M} \equiv_X \mathcal{M}' \implies \mathcal{M} \equiv \mathcal{M}'$. Suppose $\mathcal{M} = (\mathcal{A}, w) \vDash T$ and suppose $\mathcal{M} \vDash \varphi(\vec{x})$ where $x \neq \varnothing$, let $a_i = x_i^w$. Denote $\mathcal{A}^E$ the substructure of $\mathcal{A}$ generated by $E := \{a_1, \ldots, a_n\}$. By (2), $T + D\mathcal{A}^E$ is complete and consistent with $\varphi(\vec{a})$ since $\mathcal{A}_A$ satisfies $T + D\mathcal{A}^E + \varphi(\vec{a})$, this means $D\mathcal{A}^E \vdash_T \varphi(\vec{a})$ as it is complete. Now we showed that this means $D\mathcal{A}^E \cap \mathcal{L}E \vdash_T \varphi(\vec{a})$ (proven in an exercise, boils down to the coincidence theorem). By the compactness theorem there exist literals $\lambda_1(\vec{a}), \ldots, \lambda_n(\vec{a}) \in D\mathcal{A}^E$ such that $\bigwedge_{i=1}^k \lambda_i(\vec{a}) \vdash_T \varphi(\vec{a})$. Since $\vec{a}$ does not occur in T, this means $\bigwedge_{i=1}^k \lambda_i(\vec{x}) \vdash_T \varphi(\vec{x})$. And since $\mathcal{M} \vDash \bigwedge_{i=1}^k \lambda_i(\vec{x})$ (since $\vec{x}^{\mathcal{M}} = \vec{a}$ and these are all literals satisfied by $\mathcal{A}^E$), and $\mathcal{M} \equiv_X \mathcal{M}'$, we have that $\mathcal{M}' \vDash \bigwedge_{i=1}^k \lambda_i(\vec{x})$ and so $\mathcal{M}' \vDash \varphi(\vec{x})$. And so by symmetry we have $\mathcal{M} \vDash \varphi(\vec{x}) \iff \mathcal{M}' \vDash \varphi(\vec{x})$, ie. $\mathcal{M} \equiv \mathcal{M}'$. ■

5.6.9 Corollary

An $\forall$-theory T allows quantifier elimination if and only if it is model complete.

This is since $\mathcal{A} \subseteq \mathcal{B} \vDash T \implies \mathcal{A} \vDash T$, so (2) in the above theorem is satisfied if and only if $T + D\mathcal{A}$ is complete for all $\mathcal{A} \vDash T$, ie. that T is model complete. ■

5.6.10 Theorem

ACF allows quantifier elimination.

We must show that ACF is substructure complete, meaning it is the model completion of $\mathsf{ACF}^{\forall}$. But we already showed that $\mathsf{ACF}^{\forall}$ is simply the theory of integral domains and indeed ACF is its model completion. ∎

5.7 Reduced Products and Ultraproducts

5.7.1 Definition

A **filter** on a set I is a non-empty collection $\mathcal{F} \subseteq \mathcal{P}(I)$ which is closed under intersections and also upward closed (meaning if $A \in \mathcal{F}$ and $A \subseteq B$, then $B \in \mathcal{F}$). A filter is *proper* if $\mathcal{F} \neq \mathcal{P}(I)$, which is equivalent to $\varnothing \notin \mathcal{F}$.

By upward closure, obviously $I \in \mathcal{F}$. And $\mathcal{F}$ being a filter is equivalent to

$$A \cap B \in \mathcal{F} \iff A \in \mathcal{F} \text{ and } B \in \mathcal{F}$$

every filter satisfies this, and a collection which satisfies this is a filter: such a collection is obviously closed under intersections, and if $A \in \mathcal{F}$ and $A \subseteq \mathcal{F}$ then $A \cap B = A \in \mathcal{F}$, so $B \in \mathcal{F}$.

For a set $K \subseteq I$, $\{J \subseteq I \mid K \subseteq K\}$ is a filter, the *principal filter generated by* K, and it is proper if and only if $K \neq \varnothing$. If I is infinite, the set of all cofinite subsets of I ($\{J \subseteq I \mid I \setminus J \text{ is finite}\}$) is also a proper filter, called the *Frechét filter* of I. This is a filter since $A \cap B$ is cofinite if and only if A and B are, since $I \setminus (A \cap B) = (I \cap A) \cup (I \cap B)$.

5.7.2 Definition

A filter $\mathcal{F} \subseteq \mathcal{P}(I)$ is an **ultrafilter** if it also satisfies $I \setminus A \in \mathcal{F} \iff A \notin \mathcal{F}$. An ultrafilter is obviously proper, and if it contains the Frechét filter (ie. contains all cofinite sets), it is *nontrivial*.

Suppose $\mathcal{F}$ is a proper filter. Then it is an ultrafilter if and only if $A \cup B \in \mathcal{F}$ if and only if $A \in \mathcal{F}$ or $B \in \mathcal{F}$. An ultrafilter satisfies this property: if $A \cup B \in \mathcal{F}$ and suppose $A, B \notin \mathcal{F}$ then $A^c, B^c \in \mathcal{F}$ so $(A \cup B) \cap A^c \cap B^c = \varnothing \in \mathcal{F}$ in contradiction. And suppose $A \in \mathcal{F}$ or $B \in \mathcal{F}$ then $A, B \subseteq A \cup B$ so by upward closure $A \cup B \in \mathcal{F}$. And if $\mathcal{F}$ satisfies this property, then suppose $A^c \in \mathcal{F}$ then since the filter is proper $A \notin \mathcal{F}$. And if $A \notin \mathcal{F}$, since $A \cup A^c \in \mathcal{F}$ either A or A^c is in the filter, so $A^c \in \mathcal{F}$.

This property can be extended inductively to $\mathcal{F}$ is an ultrafilter if and only if $A_1 \cup \cdots \cup A_n \in \mathcal{F}$ if and only if $A_i \in \mathcal{F}$ for some $1 \leq i \leq n$.

Ultrafilters are maximal proper filters (the converse also holds). If $\mathcal{F}$ is an ultrafilter and $\mathcal{F} \subseteq \mathcal{F}'$, then let $J \in \mathcal{F}'$, if $J \notin \mathcal{F}$ then $J^c \in \mathcal{F} \subseteq \mathcal{F}'$ which contradicts $\mathcal{F}'$ being proper.

5.7.3 Proposition

Every trivial ultrafilter over an infinite set I is of the form $\{J \subseteq I \mid i \in J\}$ for some $i \in I$. Such a filter is called the **principal ultrafilter** generated by i.

Firstly this is indeed a trivial ultrafilter: $i \in A \cap B$ if and only if $i \in A$ and $i \in B$ so it is a filter, $i \in I \setminus A$ if and only if $i \notin A$ so it is an ultrafilter, and $I \setminus \{i\}$ is cofinite yet does not exist in the filter so it is trivial.

Now suppose $\mathcal{F}$ is a trivial ultrafilter over I, then let J be cofinite not in $\mathcal{F}$. Suppose that $J = I \setminus \{i_1, \dots, i_n\}$ and so $\{i_1, \dots, i_n\} \in U$, and by above this means there exists some $i \in \{i_1, \dots, i_n\}$ such that $\{i\} \in \mathcal{F}$. So by upward closure we have $\{J \subseteq I \mid i \in J\} \subseteq \mathcal{F}$ and since the left is an ultrafilter and thus maximal, we have equality. ∎

This proof showed that if $\{i\} \in \mathcal{F}$ for an ultrafilter, then $\mathcal{F}$ is the principal ultrafilter generated by i. In particular if I is finite, all of its ultrafilters are trivial (equivalently principal) since $I = \{i_1, \dots, i_n\} \in \mathcal{F}$ and so some $\{i_j\} \in \mathcal{F}$.

We utilize the propositional compactness theorem to prove the following

5.7.4 Theorem (The Ultrafilter Theorem)

Every subset $F \subseteq \mathcal{P}(I)$ can be extended to an ultrafilter $\mathcal{U}$ if it has the **finite intersection property**: for every $M_1, \dots, M_n \in F$, their intersection is nonempty.

For every $J \subseteq I$ define a propositional variable p_J. Then we define the axiom system

$$X: \qquad p_{M\cap N} \leftrightarrow p_M \wedge p_N, \quad p_{M^c} \leftrightarrow \neg p_M, \quad p_J \qquad (M, N \subseteq I, J \in F)$$

if $w \vDash X$ then define $\mathcal{U} := \{J \subseteq I \mid w \vDash p_J\}$. This is obviously an ultrafilter containing F. So all we must do is prove that every finite subset of X is satisfiable, for which it is sufficient to prove the ultrafilter theorem for finite I. But this is easy, let $F = \{M_1, \dots, M_n\}$, $\varnothing \neq D = M_1 \cap \cdots \cap M_n$ and $i \in D$, then the principal ultrafiler generated by i contains F. ■

We know define reduced and ultraproducts, this is a lengthier definition so stay vigilant. Let $(\mathcal{A}_i)_{i\in I}$ be a family of $\mathcal{L}$-structures, and F a proper filter on the indexing set I. We define an equivalence relation $\approx_F$ on the domain of $\mathcal{B} = \prod_{i\in I} \mathcal{A}_i$ by

$$(a_i)_{i\in I} \approx_F (b_i)_{i\in I} \iff \{i \in I \mid a_i = b_i\} \in F$$

In order to make notation more readable, we will use single letters like a or b to denote elements of B, like $(a_i)_{i\in I}$. This is indeed an equivalence relation: define $I_{a=b} := \{i \in I \mid a_i = b_i\}$, obviously $I_{a=a} = I \in F$ so $\approx_F$ is reflexive, and $I_{a=b} = I_{b=a}$ so $\approx_F$ is symmetric, and $I_{a=b} \cap I_{b=c} \subseteq I_{a=c}$ so $\approx_F$ is transitive (since $I_{a=b}, I_{b=c} \in F$ implies $I_{a=b} \cap I_{b=c} \in F$ which implies in turn that $I_{a=c} \in F$).

Furthermore $\approx_F$ is a congruence on the algebraic reduct of $\mathcal{B}$: let f be an n-ary function symbol in $\mathcal{L}$ and suppose $\vec{a} \approx_F \vec{b}$ (where $\vec{a} = (a^1, \dots, a^n)$ and $\vec{b} = (b^1, \dots, b^n)$). Define $I_{\vec{a}=\vec{b}} = \bigcap_{i=1}^n I_{a^i=b^i}$, and so $I_{\vec{a}=\vec{b}} \in F$ since $I_{a^i=b^i} \in F$ for every i. And surely $I_{\vec{a}=\vec{b}} \subseteq I_{f^{\mathcal{B}}\vec{a}=f^{\mathcal{B}}\vec{b}}$, so $f^{\mathcal{B}}\vec{a} \approx_F f^{\mathcal{B}}\vec{b}$.

So let $C = \{a/F \mid a \in B\}$ be the partition of B with respect to $\approx_F$ where a/F denotes the equivalence class of a with respect to $\approx_F$, meaning $a/F = b/F$ if and only if $a \approx_F b$ or equivalently $I_{a=b} \in F$. This is the domain of an $\mathcal{L}$-structure $\mathcal{C}$ where for function symbols f we define $f^{\mathcal{C}}(\vec{a}/F) := (f^{\mathcal{B}}\vec{a})/F$. This definition is well-defined since f is a congruence. And of course for constant symbols c, $c^{\mathcal{C}} := c^{\mathcal{B}}/F$.

For a relation symbol r, define $I_{r\vec{a}} := \left\{i \in I \mid r^{\mathcal{A}_i}\vec{a}_i\right\}$ where $\vec{a}_i := (a_i^1, \dots, a_i^n)$. Then define $r^{\mathcal{C}}\vec{a}/F \iff I_{r\vec{a}} \in F$. This is well defined: if $\vec{a} \approx_F \vec{b}$ then $I_{\vec{a}=\vec{b}} \in F$, and $I_{r\vec{a}} \cap I_{\vec{a}=\vec{b}} \subseteq I_{r\vec{b}}$, so since F is a filter this implies $I_{r\vec{b}} \in F$ as well.

This structure $\mathcal{C}$ is called the *reduced product* of $\{\mathcal{A}_i\}_{i\in I}$ by F, and is denoted $\prod_{i\in I}^F \mathcal{A}_i$ or $\prod_{i\in I} \mathcal{A}_i/F$.

Let $\mathcal{C} = \prod_{i\in I}^F \mathcal{A}_i$, and let $w\colon Var \longrightarrow B = \prod_{i\in I} A_i$ be a valuation, define the valuation $w_i\colon Var \longrightarrow A_i$ defined by $x \mapsto (x^w)_i$, so $x^w = (x^{w_i})_{i\in I}$. Term induction gives $t^w = (t^{w_i})_{i\in I}$ for all terms t. Then define the valuation $w/F\colon Var \longrightarrow C$ by $x^{w/F} = x^w/F$, and by term induction we have $t^{w/F} = t^w/F$. Indeed: $(f\vec{t})^{w/F} = f^{\mathcal{C}}(\vec{t}^{w/F}) = f^{\mathcal{C}}(\vec{t}^w/F) = (f^{\mathcal{B}}\vec{t}^w)/F = (f\vec{t})^w/F$. Let $w'\colon Var \longrightarrow C$ be a valuation, then define $w\colon Var \longrightarrow B$ by choosing $x^w \in x^{w'}$, and so $x^{w'} = x^w/F$, meaning $w' = w/F$. So every valuation on C is of the form w/F for some valuation w on B.

Let $w\colon Var \longrightarrow B$ and $\alpha \in \mathcal{L}$ then define $I_\alpha^w := \{i \in I \mid \mathcal{A}_i \vDash \alpha[w_i]\}$. Then we have that $I_{\exists x\beta}^w \subseteq I_\beta^{w'}$ where $w' = w_x^a$ for some $a \in B$. Indeed let $i \in I_{\exists x\beta}^w$ so $\mathcal{A}_i \vDash \exists x\beta[w_i]$, then there exists some $a_i \in A_i$ such that $\mathcal{A}_i \vDash \beta[w_i{}_x^{a_i}]$. For $i \notin I_{\exists x\beta}^w$ choose any a_i. Then define $a = (a_i)_{i\in I}$ and so for every $i \in I_{\exists x\beta}^w$, $\mathcal{A}_i \vDash \beta[w_i{}_x^{a_i}] = \beta[w_i']$, so $i \in I_\beta^{w'}$.

The case of particular interest is when F is an ultrafilter, in such a case $\prod_{i\in I}^F \mathcal{A}_i$ is called an *ultraproduct* of $\{\mathcal{A}_i\}_{i\in I}$. If $\mathcal{A}_i = \mathcal{A}$ for all $i \in I$ then we write $\mathcal{A}^I/F$ in place of $\prod_{i\in I}^F \mathcal{A}$ and this is called an *ultrapower* of $\mathcal{A}$. An important theorem regarding ultrapowers but not proven here is that $\mathcal{A} \equiv \mathcal{B}$ if and only if $\mathcal{A}$ and $\mathcal{B}$ have isomorphic unltrapowers.

5.7.5 Theorem (Łoś's Ultraproduct Theorem)

Let $\mathcal{C} = \prod_{i\in I}^F \mathcal{A}_i$ be an ultraproduct of the $\mathcal{L}$-structures $\mathcal{A}_i$. Then for all formulas $\alpha \in \mathcal{L}$ and all valuations $w\colon Var \longrightarrow \prod_{i\in I} A_i$,

$$\mathcal{C} \vDash \alpha[w/F] \iff I_\alpha^w \in F \quad (:= \{i \in I \mid \mathcal{A}_i \vDash \alpha[w_i]\} \in F)$$

We prove this by induction on α. For equations $t_1 = t_2$,

$$\mathcal{C} \vDash t_1 = t_2[w/F] \iff t_1^{w/F} = t_2^{w/F} \iff t_1^w/F = t_2^w/F \qquad \text{(since } t^{w/F} = t^w/F)$$

$$\begin{aligned} &\iff \{i \in I \mid t_1^{w_i} = t_2^{w_i}\} \in F && \text{(since } t^w = (t^{w_i})_{i\in I}) \\ &\iff \{i \in I \mid \mathcal{A}_i \vDash t_1 = t_2[w_i]\} \iff I^w_{t_1=t_2} \in F \end{aligned}$$

For prime formulas $r\vec{t}$:

$$\begin{aligned} \mathcal{C} \vDash r\vec{t}[w/F] &\iff r^{\mathcal{C}}\vec{t}^{w/F} \iff r^{\mathcal{C}}\vec{t}^{w}/F && \text{(since } t^{w/F} = t^w/F) \\ &\iff I_{r\vec{t}^w} \in F \iff \left\{i \in I \mid r^{\mathcal{A}_i}\vec{t}_i^w\right\} = \left\{i \in I \mid r^{\mathcal{A}_i}\vec{t}^{w_i}\right\} && \text{(since } t_i^w = t^{w_i}) \\ &\iff \left\{i \in I \mid \mathcal{A}_i \vDash r\vec{t}[w_i]\right\} \in F \iff I^w_{r\vec{t}} \in F \end{aligned}$$

For conjunctions:

$$\begin{aligned} \mathcal{C} \vDash \alpha\wedge\beta[w/F] &\iff \mathcal{C} \vDash \alpha, \beta[w/F] \\ &\iff I^w_\alpha, I^w_\beta \in F && \text{(induction hypothesis)} \\ &\iff I^w_\alpha \cap I^w_\beta \in F && \text{(filter property)} \\ &\iff I^w_{\alpha\wedge\beta} \in F && \text{(since } I^w_\alpha \wedge I^w_\beta = I^w_{\alpha\wedge\beta}) \end{aligned}$$

For negations: $\mathcal{C} \vDash \neg\alpha[w/F] \iff \mathcal{C} \nvDash \alpha[w/F] \iff I^w_\alpha \notin F$, and since F is an ultrafilter this is equivalent to $\iff I \setminus I^w_\alpha \in F$. And $I\setminus I^w_\alpha = \{i \in I \mid \mathcal{A}_i \nvDash \alpha[w_i]\} = \{i \in I \mid \mathcal{A}_i \vDash \neg\alpha[w_i]\} = I^w_{\neg\alpha}$, so this is equivalent to $I^w_{\neg\alpha} \in F$ as required.

For $\forall x\alpha$ we first show that $I^w_{\forall x\alpha} \in F$ implies $\mathcal{C} \vDash \forall x\alpha$. Let $a \in \prod_{i\in I} A_i$ and $w' := w^a_x$. Since $I^w_{\forall x\alpha} \subseteq I^{w'}_\alpha$, we have that $I^{w'}_\alpha \in F$ so $\mathcal{C} \vDash \alpha[w^a_x/F]$ by the induction hypothesis. But since a is arbitrary, $\mathcal{C} \vDash \forall x\alpha[w/F]$.

To prove the converse, this is equivalent to $I^w_{\forall x\alpha} \notin F \implies \mathcal{C} \nvDash \forall x\alpha$. Since F is an ultrafilter this is equivalent to $I^w_{\exists x\beta} \in F \implies \mathcal{C} \vDash \exists x\beta[w/F]$ with $\beta := \neg\alpha$. If $I^w_{\exists x\beta} \in F$ then since $I^w_{\exists x\beta} \subseteq I^{w'}_\beta$ for $w' = w^a_x$ for some $a \in B$, $I^{w'}_\beta \in F$ and so $\mathcal{C} \vDash \beta[w'/F]$, and so $\mathcal{C} \vDash \exists x\beta[w/F]$ as required. ■

For sentences α since the valuation does not affect its satisfiability in a structure,

$$\prod_{i\in I}^F \mathcal{A}_i \vDash \alpha \iff \{i \in I \mid \mathcal{A}_i \vDash \alpha\} \in F$$

for an ultrafilter F.

Notice that we can define the embedding $\iota\colon a \mapsto (a)_{i\in I}/F$ from $\mathcal{A}$ to $\mathcal{A}^I/F$. This is an elementary embedding: firstly it is injective since $(a)_{i\in I} \approx_F (b)_{i\in I}$ if and only if $\{i \in I \mid a = b\} \in F$, and this set is either I or $\varnothing$. Since F is proper, this is if and only if $a = b$. And it is an embedding since $f\iota\vec{a} = f(a_1)/F\cdots(a_n)/F = (f\vec{a})/F = \iota f\vec{a}$, and $r\iota\vec{a} = r(a_1)/F\cdots(a_n)/F \iff \{i \in I \mid r\vec{a}\} \in F$ which is either I or $\varnothing$ so this is if and only if $r\vec{a}$. Notice that if w/F maps $x_i \mapsto \iota a_i$ then all w_j are equal and map $x_i \mapsto a_i$, since then $(x_i^{w_j})_{j\in I}/F = (a_i)_{j\in I}/F = \iota a_i = x_i^{w/F}$. Thus

$$\mathcal{A}^I/F \vDash \alpha(\iota a_1, \ldots, \iota a_n) \iff \{i \in I \mid \mathcal{A} \vDash \alpha(a_1, \ldots, a_n)\} \in F$$

and this set is either I or $\varnothing$ so $\mathcal{A} \vDash \alpha(a_1, \ldots, a_n)$. Therefore $\mathcal{A} \preceq \mathcal{A}^I/F$.

This also gives us a purely model-theoretic proof of the compactness theorem:

5.7.6 Theorem (The Compactness Theorem)

Let $X \subseteq \mathcal{L}$ and let I be the set of all finite subsets of X. If every $i \in I$ has a model $(\mathcal{A}_i, w_i)$ then there exists an ultrafilter F on I such that $\prod_{i\in I}^F \mathcal{A}_i \vDash X[w/F]$ where $x^w = (x^{w_i})_{i\in I}$. Meaning that if every finite subset of X has a model, then so too does X.

For every $\alpha \in X$, define $J_\alpha := \{i \in I \mid \alpha \in i\}$, and then define $E := \{J_\alpha \mid \alpha \in X\}$. E has the finite intersection property, since $\{\alpha_1, \ldots, \alpha_n\} \in J_{\alpha_1} \cap \cdots \cap J_{\alpha_n}$, and so by The Ultrafilter Theorem there exists an ultrafilter F on I such that $E \subseteq F$. If $\alpha \in X$ and $i \in J_\alpha$ (meaning $\alpha \in i$), then $\mathcal{A}_i \vDash \alpha[w_i]$ and so $J_\alpha \subseteq I^w_\alpha$, so $I^w_\alpha \in F$. So by Łoś's Ultraproduct Theorem this means $\prod_{i\in I}^F \mathcal{A}_i \vDash \alpha[w/F]$ as required. ■

Let us define $\boldsymbol{K}_{\mathcal{L}}$ to be the class of all $\mathcal{L}$-structures. Recall that a class of structures is Δ-*elementary* if it is the class of models of some first order theory T, $\operatorname{Md} T$. And it is *elementary* if it is the class of some finitely axiomatizable first order theory.

5.7.7 Theorem

Let $\boldsymbol{K} \subseteq \boldsymbol{K}_{\mathcal{L}}$, then

(1) $\boldsymbol{K}$ is Δ-elementary if and only if $\boldsymbol{K}$ is closed under elementary equivalence and ultraproducts,

(2) $\boldsymbol{K}$ is elementary if and only if $\boldsymbol{K}$ is closed under elementary equivalence and ultraproducts and $\boldsymbol{K}^c$ is closed under ultraproducts.

(1) Obviously if $\boldsymbol{K}$ is Δ-elementary then it is closed under elementary equivalence, since if $\mathcal{A} \equiv \mathcal{B}$ and $\mathcal{A} \vDash T$ then $\mathcal{B} \vDash T$. And if $\mathcal{A}_i \vDash T$ for all $i \in I$ then $\prod_{i\in I}^F \mathcal{A}_i \vDash T$ if and only if for every $\alpha \in T$, $\{i \in I \mid \mathcal{A}_i \vDash \alpha\} \in F$. But this is just I, so we have the required.

Let us define $T := Th\boldsymbol{K}$, and so we claim that $\boldsymbol{K} = \mathrm{Md}\, T$. Obviously $\boldsymbol{K} \subseteq \mathrm{Md}\, T$, so let $\mathcal{A} \vDash T$, and let I be the set of all finite subsets of $Th\mathcal{A}$. For every $i = \{\alpha_1, \dots, \alpha_n\} \in I$ there exists some $\mathcal{A}_i \in \boldsymbol{K}$ such that $\mathcal{A}_i \vDash i$. Otherwise $\bigvee_{i=1}^n \neg\alpha_i \in T$, which contradicts $i \subseteq Th\mathcal{A}$ (since $\mathcal{A} \vDash \alpha_i, \neg\alpha_i$ for some i). So every finite subset of $Th\mathcal{A}$ is satisfied by some $\mathcal{A}_i$ and thus by the above theorem there exists an ultrafilter F such that $\prod_{i\in I}^F \mathcal{A}_i \vDash Th\mathcal{A}$, and so $\prod_{i\in I}^F \mathcal{A}_i \equiv \mathcal{A}$. But $\boldsymbol{K}$ is closed under ultraproducts so $\prod_{i\in I}^F \mathcal{A}_i \in \boldsymbol{K}$, and it is closed under equivalences so $\mathcal{A} \in \boldsymbol{K}$. So $\mathcal{A} \vDash T \iff \mathcal{A} \in \boldsymbol{K}$, meaning $\boldsymbol{K}$ is Δ-elementary.

(2) If $\boldsymbol{K}$ is elementary then both $\boldsymbol{K}$ and $\boldsymbol{K}^c$ are Δ-elementary (as $\boldsymbol{K} = \mathrm{Md}\,\alpha \implies \boldsymbol{K}^c = \mathrm{Md}\,\neg\alpha$) and so both are closed under ultraproducts and equivalence. By above, $\boldsymbol{K} = \mathrm{Md}\, S$ for some $S \subseteq \mathcal{L}^0$. Let I be the set of all finite nonempty subsets of S, then there exists some $i = \{\alpha_1, \dots, \alpha_n\} \in I$ such that $\mathrm{Md}\, i \subseteq \boldsymbol{K}$. Otherwise let $\mathcal{A}_i \vDash i$ but $\mathcal{A}_i \in \boldsymbol{K}^c$ for every $i \in I$. Again by the above compactness theorem there exists an ultraproduct $\mathcal{C}$ such that $\mathcal{C} \vDash i$ for every $i \in I$ and since $\boldsymbol{K}^c$ is closed under ultraproducts, $\mathcal{C} \in \boldsymbol{K}^c$. And so $\mathcal{C} \vDash S$, meaning $\mathcal{C} \in \boldsymbol{K}$ in contradiction. But since $\boldsymbol{K} = \mathrm{Md}\, S \subseteq \mathrm{Md}\, i$, we have that $\boldsymbol{K} = \mathrm{Md}\, i = \mathrm{Md} \bigwedge_{i=1}^n \alpha_i$, meaning it is elementary. ∎

5.7.8 Example

Let $\boldsymbol{K}$ be the Δ-elementary class of fields of characteristic 0. We claim that $\boldsymbol{K}$ is not elementary, and we will prove it by showing that $\boldsymbol{K}^c$ is not closed under ultraproducts. Let $\mathcal{P}_i$ be the prime field of characteristic p_i, and let F be a nontrivial ultrafilter on $\mathbb{N}$. Then $\prod_{i\in I}^F \mathcal{P}_i$ has characteristic 0, since $\{i \in I \mid \mathcal{P}_i \vDash \neg\mathtt{char}_p\}$ for any prime p is cofinite and thus belongs to F. So $\prod_{i\in I}^F \mathcal{P}_i \vDash \neg\mathtt{char}_p$, meaning it is a field of characteristic 0 (it is a field since the class of fields is Δ-elementary and therefore closed under ultraproducts).

Notice that for the minimal filter $F = \{I\}$, $a \approx_F b \iff \{i \in I \mid a_i = b_i\} \in F \iff a_i = b_i$ for all $i \in I$, thus $\prod_{i\in I}^{\{I\}} \mathcal{A}_i \cong \prod_{i\in I} \mathcal{A}_i$. So whatever we can prove on reduced products holds for direct products too. From here on, all filters are assumed to be proper.

5.7.9 Theorem

Let $\mathcal{C} = \prod_{i\in I}^F \mathcal{A}_i$ be a reduced product, $w\colon Var \longrightarrow \prod_{i\in I} \mathcal{A}_i$, and α a Horn formula. Then $I_\alpha^w \in F \implies \mathcal{C} \vDash \alpha[w/F]$. In particular if α is a Horn sentence, $\{i \in I \mid \mathcal{A}_i \vDash \alpha\} \implies \mathcal{C} \vDash \alpha$.

We prove this by induction on Horn formulas. For prime formulas both directions of the conditional hold, since in our proof of Łoś's Ultraproduct Theorem for the step on prime formulas, no ultraproduct property was used. Since F is proper $I_{\neg\alpha}^w \in F \implies I_\alpha^w \notin F \implies \mathcal{C} \nvDash \alpha[w/F] \implies \mathcal{C} \vDash \neg\alpha[w/F]$ for α prime, and so the conditional holds for all literals. Now suppose the condition holds for prime α and basic Horn formula β and suppose $I_{\alpha\to\beta}^w$, then if $\mathcal{C} \vDash \alpha[w/F]$ then $I_\alpha^w \in F$ since α is prime. Since $I_\alpha^w \cap I_{\alpha\to\beta}^w \subseteq I_\beta^w$, we get that $I_\beta^w \in F$ so $\mathcal{C} \vDash \beta[w/F]$ by induction, meaning $\mathcal{C} \vDash \alpha \to \beta[w/F]$. Induction on $\wedge$ and $\forall$ proceed similar to Łoś, and $\exists$ follows from $I_{\exists x\beta}^w \subseteq I_\beta^{w_x^a}$ for some $a \in \prod_{i\in I} A_i$. ∎

This means that the model classes of Horn theories are closed under reduced products, and in particular direct products. The converse holds as well: every class of models closed under reduced products is a Horn theory, but this is significantly more challenging to prove.

6 Index