

**Part I**  
**Model Theory**

Material on model theory is incomplete and experimental. It is currently simply an adaptation of Aldo Antonelli's notes on model theory, less those topics covered in the part on first-order logic (theories, completeness, compactness). It requires much more introduction, motivation, and explanation, as well as exercises, to be useful for a textbook. Andy Arana is at planning to work on this part specifically (issue #65).

# Chapter 1

## Basics of Model Theory

### 1.1 Reducts and Expansions

Often it is useful or necessary to compare languages which have symbols in common, as well as **structures** for these languages. The most common case is when all the symbols in a **language**  $\mathcal{L}$  are also part of a **language**  $\mathcal{L}'$ , i.e.,  $\mathcal{L} \subseteq \mathcal{L}'$ . An  $\mathcal{L}$ -**structure**  $\mathfrak{M}$  can then always be expanded to an  $\mathcal{L}'$ -**structure** by adding interpretations of the additional symbols while leaving the interpretations of the common symbols the same. On the other hand, from an  $\mathcal{L}'$ -structure  $\mathfrak{M}'$  we can obtain an  $\mathcal{L}$ -structure simply by “forgetting” the interpretations of the symbols that do not occur in  $\mathcal{L}$ .

mod:bas:red:  
defn:reduct **Definition 1.1.** Suppose  $\mathcal{L} \subseteq \mathcal{L}'$ ,  $\mathfrak{M}$  is an  $\mathcal{L}$ -**structure** and  $\mathfrak{M}'$  is an  $\mathcal{L}'$ -**structure**.  $\mathfrak{M}$  is the *reduct* of  $\mathfrak{M}'$  to  $\mathcal{L}$ , and  $\mathfrak{M}'$  is an *expansion* of  $\mathfrak{M}$  to  $\mathcal{L}'$  iff

1.  $|\mathfrak{M}| = |\mathfrak{M}'|$
2. For every **constant symbol**  $c \in \mathcal{L}$ ,  $c^{\mathfrak{M}} = c^{\mathfrak{M}'}$ .
3. For every **function symbol**  $f \in \mathcal{L}$ ,  $f^{\mathfrak{M}} = f^{\mathfrak{M}'}$ .
4. For every **predicate symbol**  $P \in \mathcal{L}$ ,  $P^{\mathfrak{M}} = P^{\mathfrak{M}'}$ .

mod:bas:red:  
prop:reduct **Proposition 1.2.** If an  $\mathcal{L}$ -**structure**  $\mathfrak{M}$  is a reduct of an  $\mathcal{L}'$ -**structure**  $\mathfrak{M}'$ , then for all  $\mathcal{L}$ -**sentences**  $\varphi$ ,

$$\mathfrak{M} \models \varphi \text{ iff } \mathfrak{M}' \models \varphi.$$

*Proof.* Exercise. □

**Problem 1.1.** Prove Proposition 1.2.

**Definition 1.3.** When we have an  $\mathcal{L}$ -structure  $\mathfrak{M}$ , and  $\mathcal{L}' = \mathcal{L} \cup \{P\}$  is the expansion of  $\mathcal{L}$  obtained by adding a single  $n$ -place **predicate symbol**  $P$ , and  $R \subseteq |\mathfrak{M}|^n$  is an  $n$ -place relation, then we write  $(\mathfrak{M}, R)$  for the expansion  $\mathfrak{M}'$  of  $\mathfrak{M}$  with  $P^{\mathfrak{M}'} = R$ .

## 1.2 Substructures

The **domain** of a **structure**  $\mathfrak{M}$  may be a subset of another  $\mathfrak{M}'$ . But we should obviously only consider  $\mathfrak{M}$  a “part” of  $\mathfrak{M}'$  if not only  $|\mathfrak{M}| \subseteq |\mathfrak{M}'|$ , but  $\mathfrak{M}$  and  $\mathfrak{M}'$  “agree” in how they interpret the symbols of the language at least on the shared part  $|\mathfrak{M}|$ . mod:bas:sub:  
sec

**Definition 1.4.** Given **structures**  $\mathfrak{M}$  and  $\mathfrak{M}'$  for the same language  $\mathcal{L}$ , we say that  $\mathfrak{M}$  is a **substructure** of  $\mathfrak{M}'$ , and  $\mathfrak{M}'$  an **extension** of  $\mathfrak{M}$ , written  $\mathfrak{M} \subseteq \mathfrak{M}'$ , iff mod:bas:sub:  
defn:substructure

1.  $|\mathfrak{M}| \subseteq |\mathfrak{M}'|$ ,
2. For each constant  $c \in \mathcal{L}$ ,  $c^{\mathfrak{M}} = c^{\mathfrak{M}'}$ ;
3. For each  $n$ -place **predicate symbol**  $f \in \mathcal{L}$   $f^{\mathfrak{M}}(a_1, \dots, a_n) = f^{\mathfrak{M}'}(a_1, \dots, a_n)$  for all  $a_1, \dots, a_n \in |\mathfrak{M}|$ .
4. For each  $n$ -place **predicate symbol**  $R \in \mathcal{L}$ ,  $\langle a_1, \dots, a_n \rangle \in R^{\mathfrak{M}}$  iff  $\langle a_1, \dots, a_n \rangle \in R^{\mathfrak{M}'}$  for all  $a_1, \dots, a_n \in |\mathfrak{M}|$ .

*Remark 1.* If the language contains no constant or **function symbols**, then any  $N \subseteq |\mathfrak{M}|$  determines a **substructure**  $\mathfrak{N}$  of  $\mathfrak{M}$  with **domain**  $|\mathfrak{N}| = N$  by putting  $R^{\mathfrak{N}} = R^{\mathfrak{M}} \cap N^n$ . mod:bas:sub:  
rem:substructure

## 1.3 Overspill

**Theorem 1.5.** *If a set  $\Gamma$  of sentences has arbitrarily large finite models, then it has an infinite model.* mod:bas:ove:  
sec

*Proof.* Expand the language of  $\Gamma$  by adding countably many new constants  $c_0, c_1, \dots$  and consider the set  $\Gamma \cup \{c_i \neq c_j : i \neq j\}$ . To say that  $\Gamma$  has arbitrarily large finite models means that for every  $m > 0$  there is  $n \geq m$  such that  $\Gamma$  has a model of cardinality  $n$ . This implies that  $\Gamma \cup \{c_i \neq c_j : i \neq j\}$  is finitely satisfiable. By compactness,  $\Gamma \cup \{c_i \neq c_j : i \neq j\}$  has a model  $\mathfrak{M}$  whose domain must be infinite, since it satisfies all inequalities  $c_i \neq c_j$ . □

**Proposition 1.6.** *There is no sentence  $\varphi$  of any first-order language that is true in a **structure**  $\mathfrak{M}$  if and only if the domain  $|\mathfrak{M}|$  of the **structure** is infinite.* mod:bas:ove:  
inf-not-fo

*Proof.* If there were such a  $\varphi$ , its negation  $\neg\varphi$  would be true in all and only the finite **structures**, and it would therefore have arbitrarily large finite models but it would lack an infinite model, contradicting [Theorem 1.5](#). □

## 1.4 Isomorphic Structures

mod:bas:iso:  
sec

First-order **structures** can be alike in one of two ways. One way in which they can be alike is that they make the same **sentences** true. We call such **structures** *elementarily equivalent*. But structures can be very different and still make the same **sentences** true—for instance, one can be **enumerable** and the other not. This is because there are lots of features of a **structure** that cannot be expressed in first-order languages, either because the language is not rich enough, or because of fundamental limitations of first-order logic such as the Löwenheim-Skolem theorem. So another, stricter, aspect in which **structures** can be alike is if they are fundamentally the same, in the sense that they only differ in the objects that make them up, but not in their structural features. A way of making this precise is by the notion of an *isomorphism*.

mod:bas:iso:  
defn:elem-equiv

**Definition 1.7.** Given two **structures**  $\mathfrak{M}$  and  $\mathfrak{M}'$  for the same **language**  $\mathcal{L}$ , we say that  $\mathfrak{M}$  is *elementarily equivalent* to  $\mathfrak{M}'$ , written  $\mathfrak{M} \equiv \mathfrak{M}'$ , if and only if for every **sentence**  $\varphi$  of  $\mathcal{L}$ ,  $\mathfrak{M} \models \varphi$  iff  $\mathfrak{M}' \models \varphi$ .

mod:bas:iso:  
defn:isomorphism

**Definition 1.8.** Given two **structures**  $\mathfrak{M}$  and  $\mathfrak{M}'$  for the same **language**  $\mathcal{L}$ , we say that  $\mathfrak{M}$  is *isomorphic* to  $\mathfrak{M}'$ , written  $\mathfrak{M} \simeq \mathfrak{M}'$ , if and only if there is a function  $h: |\mathfrak{M}| \rightarrow |\mathfrak{M}'|$  such that:

1.  $h$  is **injective**: if  $h(x) = h(y)$  then  $x = y$ ;
2.  $h$  is **surjective**: for every  $y \in |\mathfrak{M}'|$  there is  $x \in |\mathfrak{M}|$  such that  $h(x) = y$ ;
3. for every **constant symbol**  $c$ :  $h(c^{\mathfrak{M}}) = c^{\mathfrak{M}'}$ ;
4. for every  $n$ -place **predicate symbol**  $P$ :

mod:bas:iso:  
defn:iso-const

mod:bas:iso:  
defn:iso-pred

$$\langle a_1, \dots, a_n \rangle \in P^{\mathfrak{M}} \quad \text{iff} \quad \langle h(a_1), \dots, h(a_n) \rangle \in P^{\mathfrak{M}'};$$

mod:bas:iso:  
defn:iso-func

5. for every  $n$ -place **function symbol**  $f$ :

$$h(f^{\mathfrak{M}}(a_1, \dots, a_n)) = f^{\mathfrak{M}'}(h(a_1), \dots, h(a_n)).$$

mod:bas:iso:  
thm:isom

**Theorem 1.9.** *If  $\mathfrak{M} \simeq \mathfrak{M}'$  then  $\mathfrak{M} \equiv \mathfrak{M}'$ .*

*Proof.* Let  $h$  be an isomorphism of  $\mathfrak{M}$  onto  $\mathfrak{M}'$ . For any assignment  $s$ ,  $h \circ s$  is the composition of  $h$  and  $s$ , i.e., the assignment in  $\mathfrak{M}'$  such that  $(h \circ s)(x) = h(s(x))$ . By induction on  $t$  and  $\varphi$  one can prove the stronger claims:

- a.  $h(\text{Val}_s^{\mathfrak{M}}(t)) = \text{Val}_{h \circ s}^{\mathfrak{M}'}(t)$ .
- b.  $\mathfrak{M}, s \models \varphi$  iff  $\mathfrak{M}', h \circ s \models \varphi$ .

The first is proved by induction on the complexity of  $t$ .

1. If  $t \equiv c$ , then  $\text{Val}_s^{\mathfrak{M}}(c) = c^{\mathfrak{M}}$  and  $\text{Val}_{h \circ s}^{\mathfrak{M}'}(c) = c^{\mathfrak{M}'}$ . Thus,  $h(\text{Val}_s^{\mathfrak{M}}(t)) = h(c^{\mathfrak{M}}) = c^{\mathfrak{M}'}$  (by (3) of [Definition 1.8](#)) =  $\text{Val}_{h \circ s}^{\mathfrak{M}'}(t)$ .

2. If  $t \equiv x$ , then  $\text{Val}_s^{\mathfrak{M}}(x) = s(x)$  and  $\text{Val}_{h \circ s}^{\mathfrak{M}'}(x) = h(s(x))$ . Thus,  $h(\text{Val}_s^{\mathfrak{M}}(x)) = h(s(x)) = \text{Val}_{h \circ s}^{\mathfrak{M}'}(x)$ .
3. If  $t \equiv f(t_1, \dots, t_n)$ , then

$$\begin{aligned} \text{Val}_s^{\mathfrak{M}}(t) &= f^{\mathfrak{M}}(\text{Val}_s^{\mathfrak{M}}(t_1), \dots, \text{Val}_s^{\mathfrak{M}}(t_n)) \quad \text{and} \\ \text{Val}_{h \circ s}^{\mathfrak{M}'}(t) &= f^{\mathfrak{M}'}(\text{Val}_{h \circ s}^{\mathfrak{M}'}(t_1), \dots, \text{Val}_{h \circ s}^{\mathfrak{M}'}(t_n)). \end{aligned}$$

The induction hypothesis is that for each  $i$ ,  $h(\text{Val}_s^{\mathfrak{M}}(t_i)) = \text{Val}_{h \circ s}^{\mathfrak{M}'}(t_i)$ . So,

$$\begin{aligned} h(\text{Val}_s^{\mathfrak{M}}(t)) &= h(f^{\mathfrak{M}}(\text{Val}_s^{\mathfrak{M}}(t_1), \dots, \text{Val}_s^{\mathfrak{M}}(t_n))) \\ &= h(f^{\mathfrak{M}}(\text{Val}_{h \circ s}^{\mathfrak{M}'}(t_1), \dots, \text{Val}_{h \circ s}^{\mathfrak{M}'}(t_n))) && (1.1) \quad \text{mod:bas:iso:} \\ &= f^{\mathfrak{M}'}(\text{Val}_{h \circ s}^{\mathfrak{M}'}(t_1), \dots, \text{Val}_{h \circ s}^{\mathfrak{M}'}(t_n)) && (1.2) \quad \text{iso-1} \\ &= \text{Val}_{h \circ s}^{\mathfrak{M}'}(t) && \text{iso-2} \end{aligned}$$

Here, eq. (1.1) follows by induction hypothesis and eq. (1.2) by (5) of Definition 1.8.

Part (2) is left as an exercise.

If  $\varphi$  is a sentence, the assignments  $s$  and  $h \circ s$  are irrelevant, and we have  $\mathfrak{M} \models \varphi$  iff  $\mathfrak{M}' \models \varphi$ .  $\square$

**Problem 1.2.** Carry out the proof of (b) of Theorem 1.9 in detail. Make sure to note where each of the five properties characterizing isomorphisms of Definition 1.8 is used.

**Definition 1.10.** An *automorphism* of a structure  $\mathfrak{M}$  is an isomorphism of  $\mathfrak{M}$  onto itself.

**Problem 1.3.** Show that for any structure  $\mathfrak{M}$ , if  $X$  is a definable subset of  $\mathfrak{M}$ , and  $h$  is an automorphism of  $\mathfrak{M}$ , then  $X = \{h(x) : x \in X\}$  (i.e.,  $X$  is fixed under  $h$ ).

## 1.5 The Theory of a Structure

Every structure  $\mathfrak{M}$  makes some sentences true, and some false. The set of all the sentences it makes true is called its *theory*. That set is in fact a theory, since anything it entails must be true in all its models, including  $\mathfrak{M}$ .

**Definition 1.11.** Given a structure  $\mathfrak{M}$ , the *theory* of  $\mathfrak{M}$  is the set  $\text{Th}(\mathfrak{M})$  of sentences that are true in  $\mathfrak{M}$ , i.e.,  $\text{Th}(\mathfrak{M}) = \{\varphi : \mathfrak{M} \models \varphi\}$ .

We also use the term “theory” informally to refer to sets of sentences having an intended interpretation, whether deductively closed or not.

**Proposition 1.12.** For any  $\mathfrak{M}$ ,  $\text{Th}(\mathfrak{M})$  is complete.

*Proof.* For any **sentence**  $\varphi$  either  $\mathfrak{M} \models \varphi$  or  $\mathfrak{M} \models \neg\varphi$ , so either  $\varphi \in \text{Th}(\mathfrak{M})$  or  $\neg\varphi \in \text{Th}(\mathfrak{M})$ .  $\square$

mod:bas:thm:  
prop:equiv **Proposition 1.13.** *If  $\mathfrak{N} \models \varphi$  for every  $\varphi \in \text{Th}(\mathfrak{M})$ , then  $\mathfrak{M} \equiv \mathfrak{N}$ .*

*Proof.* Since  $\mathfrak{N} \models \varphi$  for all  $\varphi \in \text{Th}(\mathfrak{M})$ ,  $\text{Th}(\mathfrak{M}) \subseteq \text{Th}(\mathfrak{N})$ . If  $\mathfrak{N} \models \varphi$ , then  $\mathfrak{N} \not\models \neg\varphi$ , so  $\neg\varphi \notin \text{Th}(\mathfrak{N})$ . Since  $\text{Th}(\mathfrak{M})$  is complete,  $\varphi \in \text{Th}(\mathfrak{M})$ . So,  $\text{Th}(\mathfrak{N}) \subseteq \text{Th}(\mathfrak{M})$ , and we have  $\mathfrak{M} \equiv \mathfrak{N}$ .  $\square$

mod:bas:thm:  
remark:R *Remark 2.* Consider  $\mathfrak{R} = \langle \mathbb{R}, < \rangle$ , the **structure** whose domain is the set  $\mathbb{R}$  of the real numbers, in the **language** comprising only a 2-place **predicate symbol** interpreted as the  $<$  relation over the reals. Clearly  $\mathfrak{R}$  is **non-enumerable**; however, since  $\text{Th}(\mathfrak{R})$  is obviously consistent, by the Löwenheim-Skolem theorem it has an **enumerable** model, say  $\mathfrak{S}$ , and by **Proposition 1.13**,  $\mathfrak{R} \equiv \mathfrak{S}$ . Moreover, since  $\mathfrak{R}$  and  $\mathfrak{S}$  are not isomorphic, this shows that the converse of **Theorem 1.9** fails in general.

## 1.6 Partial Isomorphisms

**Definition 1.14.** Given two **structures**  $\mathfrak{M}$  and  $\mathfrak{N}$ , a *partial isomorphism* from  $\mathfrak{M}$  to  $\mathfrak{N}$  is a finite partial function  $p$  taking arguments in  $|\mathfrak{M}|$  and returning values in  $|\mathfrak{N}|$ , which satisfies the isomorphism conditions from **Definition 1.8** on its domain:

1.  $p$  is **injective**;
2. for every **constant symbol**  $c$ : if  $p(c^{\mathfrak{M}})$  is defined, then  $p(c^{\mathfrak{M}}) = c^{\mathfrak{N}}$ ;
3. for every  $n$ -place **predicate symbol**  $P$ : if  $a_1, \dots, a_n$  are in the domain of  $p$ , then  $\langle a_1, \dots, a_n \rangle \in P^{\mathfrak{M}}$  if and only if  $\langle p(a_1), \dots, p(a_n) \rangle \in P^{\mathfrak{N}}$ ;
4. for every  $n$ -place **function symbol**  $f$ : if  $a_1, \dots, a_n$  are in the domain of  $p$ , then  $p(f^{\mathfrak{M}}(a_1, \dots, a_n)) = f^{\mathfrak{N}}(p(a_1), \dots, p(a_n))$ .

That  $p$  is finite means that  $\text{dom}(p)$  is finite.

Notice that the empty function  $\emptyset$  is always a partial isomorphism between any two **structures**.

mod:bas:pis:  
defn:partialisom **Definition 1.15.** Two **structures**  $\mathfrak{M}$  and  $\mathfrak{N}$ , are *partially isomorphic*, written  $\mathfrak{M} \simeq_p \mathfrak{N}$ , if and only if there is a non-empty set  $I$  of partial isomorphisms between  $\mathfrak{M}$  and  $\mathfrak{N}$  satisfying the *back-and-forth* property:

1. (*Forth*) For every  $p \in I$  and  $a \in |\mathfrak{M}|$  there is  $q \in I$  such that  $p \subseteq q$  and  $a$  is in the domain of  $q$ ;
2. (*Back*) For every  $p \in I$  and  $b \in |\mathfrak{N}|$  there is  $q \in I$  such that  $p \subseteq q$  and  $b$  is in the range of  $q$ .

mod:bas:pis:  
thm:p-isom1 **Theorem 1.16.** *If  $\mathfrak{M} \simeq_p \mathfrak{N}$  and  $\mathfrak{M}$  and  $\mathfrak{N}$  are **enumerable**, then  $\mathfrak{M} \simeq \mathfrak{N}$ .*

*Proof.* Since  $\mathfrak{M}$  and  $\mathfrak{N}$  are **enumerable**, let  $|\mathfrak{M}| = \{a_0, a_1, \dots\}$  and  $|\mathfrak{N}| = \{b_0, b_1, \dots\}$ . Starting with an arbitrary  $p_0 \in I$ , we define an increasing sequence of partial isomorphisms  $p_0 \subseteq p_1 \subseteq p_2 \subseteq \dots$  as follows:

1. if  $n + 1$  is odd, say  $n = 2r$ , then using the Forth property find a  $p_{n+1} \in I$  such that  $p_n \subseteq p_{n+1}$  and  $a_r$  is in the domain of  $p_{n+1}$ ;
2. if  $n + 1$  is even, say  $n + 1 = 2r$ , then using the Back property find a  $p_{n+1} \in I$  such that  $p_n \subseteq p_{n+1}$  and  $b_r$  is in the range of  $p_{n+1}$ .

If we now put:

$$p = \bigcup_{n \geq 0} p_n,$$

we have that  $p$  is a an isomorphism between  $\mathfrak{M}$  and  $\mathfrak{N}$ . □

**Problem 1.4.** Show in detail that  $p$  as defined in [Theorem 1.16](#) is in fact an isomorphism.

**Theorem 1.17.** *Suppose  $\mathfrak{M}$  and  $\mathfrak{N}$  are **structures** for a purely relational language (a language containing only **predicate symbols**, and no **function symbols** or **constants**). Then if  $\mathfrak{M} \simeq_p \mathfrak{N}$ , also  $\mathfrak{M} \equiv \mathfrak{N}$ .* [mod:bas:pis:](#)  
[thm:p-isom2](#)

*Proof.* By induction on **formulas**, one shows that if  $a_1, \dots, a_n$  and  $b_1, \dots, b_n$  are such that there is a partial isomorphism  $p$  mapping each  $a_i$  to  $b_i$  and  $s_1(x_i) = a_i$  and  $s_2(x_i) = b_i$  (for  $i = 1, \dots, n$ ), then  $\mathfrak{M}, s_1 \models \varphi$  if and only if  $\mathfrak{N}, s_2 \models \varphi$ . The case for  $n = 0$  gives  $\mathfrak{M} \equiv \mathfrak{N}$ . □

*Remark 3.* If **function symbols** are present, the previous result is still true, but one needs to consider the isomorphism induced by  $p$  between the **substructure** of  $\mathfrak{M}$  generated by  $a_1, \dots, a_n$  and the **substructure** of  $\mathfrak{N}$  generated by  $b_1, \dots, b_n$ .

The previous result can be “broken down” into stages by establishing a connection between the number of nested quantifiers in a **formula** and how many times the relevant partial isomorphisms can be extended.

**Definition 1.18.** For any **formula**  $\varphi$ , the *quantifier rank* of  $\varphi$ , denoted by  $\text{qr}(\varphi) \in \mathbb{N}$ , is recursively defined as the highest number of nested quantifiers in  $\varphi$ . Two **structures**  $\mathfrak{M}$  and  $\mathfrak{N}$  are  *$n$ -equivalent*, written  $\mathfrak{M} \equiv_n \mathfrak{N}$ , if they agree on all sentences of quantifier rank less than or equal to  $n$ .

**Proposition 1.19.** *Let  $\mathcal{L}$  be a finite purely relational language, i.e., a language containing finitely many **predicate symbols** and **constant symbols**, and no **function symbols**. Then for each  $n \in \mathbb{N}$  there are only finitely many first-order sentences in the language  $\mathcal{L}$  that have quantifier rank no greater than  $n$ , up to logical equivalence.* [mod:bas:pis:](#)  
[prop:qr-finite](#)

*Proof.* By induction on  $n$ . □



**Definition 1.20.** Given a **structure**  $\mathfrak{M}$ , let  $|\mathfrak{M}|^{<\omega}$  be the set of all finite sequences over  $|\mathfrak{M}|$ . We use  $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$  to range over finite sequences of elements. If  $\mathbf{a} \in |\mathfrak{M}|^{<\omega}$  and  $a \in |\mathfrak{M}|$ , then  $\mathbf{a}a$  represents the *concatenation* of  $\mathbf{a}$  with  $a$ .

**Definition 1.21.** Given **structures**  $\mathfrak{M}$  and  $\mathfrak{N}$ , we define relations  $I_n \subseteq |\mathfrak{M}|^{<\omega} \times |\mathfrak{N}|^{<\omega}$  between sequences of equal length, by recursion on  $n$  as follows:

1.  $I_0(\mathbf{a}, \mathbf{b})$  if and only if  $\mathbf{a}$  and  $\mathbf{b}$  satisfy the same atomic **formulas** in  $\mathfrak{M}$  and  $\mathfrak{N}$ ; i.e., if  $s_1(x_i) = a_i$  and  $s_2(x_i) = b_i$  and  $\varphi$  is atomic with all **variables** among  $x_1, \dots, x_n$ , then  $\mathfrak{M}, s_1 \models \varphi$  if and only if  $\mathfrak{N}, s_2 \models \varphi$ .
2.  $I_{n+1}(\mathbf{a}, \mathbf{b})$  if and only if for every  $a \in A$  there is a  $b \in B$  such that  $I_n(\mathbf{a}a, \mathbf{b}b)$ , and vice-versa.

**Definition 1.22.** Write  $\mathfrak{M} \approx_n \mathfrak{N}$  if  $I_n(A, A)$  holds of  $\mathfrak{M}$  and  $\mathfrak{N}$  (where  $A$  is the empty sequence).

*mod:bas:pis: thm:b-n-f* **Theorem 1.23.** *Let  $\mathcal{L}$  be a purely relational language. Then  $I_n(\mathbf{a}, \mathbf{b})$  implies that for every  $\varphi$  such that  $\text{qr}(\varphi) \leq n$ , we have  $\mathfrak{M}, \mathbf{a} \models \varphi$  if and only if  $\mathfrak{N}, \mathbf{b} \models \varphi$  (where again  $\mathbf{a}$  satisfies  $\varphi$  if any  $s$  such that  $s(x_i) = a_i$  satisfies  $\varphi$ ). Moreover, if  $\mathcal{L}$  is finite, the converse also holds.*

*Proof.* The proof that  $I_n(\mathbf{a}, \mathbf{b})$  implies that  $\mathbf{a}$  and  $\mathbf{b}$  satisfy the same **formulas** of quantifier rank no greater than  $n$  is by an easy induction on  $\varphi$ . For the converse we proceed by induction on  $n$ , using [Proposition 1.19](#), which ensures that for each  $n$  there are at most finitely many non-equivalent **formulas** of that quantifier rank.

For  $n = 0$  the hypothesis that  $\mathbf{a}$  and  $\mathbf{b}$  satisfy the same quantifier-free **formulas** gives that they satisfy the same atomic ones, so that  $I_0(\mathbf{a}, \mathbf{b})$ .

For the  $n + 1$  case, suppose that  $\mathbf{a}$  and  $\mathbf{b}$  satisfy the same **formulas** of quantifier rank no greater than  $n + 1$ ; in order to show that  $I_{n+1}(\mathbf{a}, \mathbf{b})$  suffices to show that for each  $a \in |\mathfrak{M}|$  there is a  $b \in |\mathfrak{N}|$  such that  $I_n(\mathbf{a}a, \mathbf{b}b)$ , and by the inductive hypothesis again suffices to show that for each  $a \in |\mathfrak{M}|$  there is a  $b \in |\mathfrak{N}|$  such that  $\mathbf{a}a$  and  $\mathbf{b}b$  satisfy the same **formulas** of quantifier rank no greater than  $n$ .

Given  $a \in |\mathfrak{M}|$ , let  $\tau_n^a$  be set of **formulas**  $\psi(x, \mathbf{y})$  of rank no greater than  $n$  satisfied by  $\mathbf{a}a$  in  $\mathfrak{M}$ ;  $\tau_n^a$  is finite, so we can assume it is a single first-order **formula**. It follows that  $\mathbf{a}$  satisfies  $\exists x \tau_n^a(x, \mathbf{y})$ , which has quantifier rank no greater than  $n + 1$ . By hypothesis  $\mathbf{b}$  satisfies the same **formula** in  $\mathfrak{N}$ , so that there is a  $b \in |\mathfrak{N}|$  such that  $\mathbf{b}b$  satisfies  $\tau_n^a$ ; in particular,  $\mathbf{b}b$  satisfies the same **formulas** of quantifier rank no greater than  $n$  as  $\mathbf{a}a$ . Similarly one shows that for every  $b \in |\mathfrak{N}|$  there is a  $a \in |\mathfrak{M}|$  such that  $\mathbf{a}a$  and  $\mathbf{b}b$  satisfy the same **formulas** of quantifier rank no greater than  $n$ , which completes the proof.  $\square$

*mod:bas:pis: cor:b-n-f* **Corollary 1.24.** *If  $\mathfrak{M}$  and  $\mathfrak{N}$  are purely relational structures in a finite language, then  $\mathfrak{M} \approx_n \mathfrak{N}$  if and only if  $\mathfrak{M} \equiv_n \mathfrak{N}$ . In particular  $\mathfrak{M} \equiv \mathfrak{N}$  if and only if for each  $n$ ,  $\mathfrak{M} \approx_n \mathfrak{N}$ .*

## 1.7 Dense Linear Orders

**Definition 1.25.** A *dense linear ordering without endpoints* is a structure  $\mathfrak{M}$  for the language containing a single 2-place predicate symbol  $<$  satisfying the following sentences:

1.  $\forall x x < x$ ;
2.  $\forall x \forall y \forall z (x < y \rightarrow (y < z \rightarrow x < z))$ ;
3.  $\forall x \forall y (x < y \vee x = y \vee y < x)$ ;
4.  $\forall x \exists y x < y$ ;
5.  $\forall x \exists y y < x$ ;
6.  $\forall x \forall y (x < y \rightarrow \exists z (x < z \wedge z < y))$ .

**Theorem 1.26.** Any two *enumerable dense linear orderings without endpoints* are isomorphic. *mod:bas:dlo:  
thm:cantorQ*

*Proof.* Let  $\mathfrak{M}_1$  and  $\mathfrak{M}_2$  be *enumerable dense linear orderings without endpoints*, with  $<_1 = <^{\mathfrak{M}_1}$  and  $<_2 = <^{\mathfrak{M}_2}$ , and let  $\mathcal{I}$  be the set of all partial isomorphisms between them.  $\mathcal{I}$  is not empty since at least  $\emptyset \in \mathcal{I}$ . We show that  $\mathcal{I}$  satisfies the Back-and-Forth property. Then  $\mathfrak{M}_1 \simeq_p \mathfrak{M}_2$ , and the theorem follows by [Theorem 1.16](#).

To show  $\mathcal{I}$  satisfies the Forth property, let  $p \in \mathcal{I}$  and let  $p(a_i) = b_i$  for  $i = 1, \dots, n$ , and without loss of generality suppose  $a_1 <_1 a_2 <_1 \dots <_1 a_n$ . Given  $a \in |\mathfrak{M}_1|$ , find  $b \in |\mathfrak{M}_2|$  as follows:

1. if  $a <_2 a_1$  let  $b \in |\mathfrak{M}_2|$  be such that  $b <_2 b_1$ ;
2. if  $a_n <_1 a$  let  $b \in |\mathfrak{M}_2|$  be such that  $b_n <_2 b$ ;
3. if  $a_i <_1 a <_1 a_{i+1}$  for some  $i$ , then let  $b \in |\mathfrak{M}_2|$  be such that  $b_i <_2 b <_2 b_{i+1}$ .

It is always possible to find a  $b$  with the desired property since  $\mathfrak{M}_2$  is a dense linear ordering without endpoints. Define  $q = p \cup \{(a, b)\}$  so that  $q \in \mathcal{I}$  is the desired extension of  $p$ . This establishes the Forth property. The Back property is similar. So  $\mathfrak{M}_1 \simeq_p \mathfrak{M}_2$ ; by [Theorem 1.16](#),  $\mathfrak{M}_1 \simeq \mathfrak{M}_2$ . □

**Problem 1.5.** Complete the proof of [Theorem 1.26](#) by verifying that  $\mathcal{I}$  satisfies the Back property.

*Remark 4.* Let  $\mathfrak{S}$  be any *enumerable dense linear ordering without endpoints*. Then (by [Theorem 1.26](#))  $\mathfrak{S} \simeq \mathfrak{Q}$ , where  $\mathfrak{Q} = (\mathbb{Q}, <)$  is the *enumerable dense linear ordering* having the set  $\mathbb{Q}$  of the rational numbers as its domain. Now consider again the *structure*  $\mathfrak{R} = (\mathbb{R}, <)$  from [Remark 2](#). We saw that there is an *enumerable structure*  $\mathfrak{S}$  such that  $\mathfrak{R} \equiv \mathfrak{S}$ . But  $\mathfrak{S}$  is an *enumerable dense linear ordering without endpoints*, and so it is isomorphic (and hence

elementarily equivalent) to the **structure**  $\mathfrak{Q}$ . By transitivity of elementary equivalence,  $\mathfrak{R} \equiv \mathfrak{Q}$ . (We could have shown this directly by establishing  $\mathfrak{R} \simeq_p \mathfrak{Q}$  by the same back-and-forth argument.)

## Chapter 2

# Models of Arithmetic

### 2.1 Introduction

The *standard model* of arithmetic is the **structure**  $\mathfrak{N}$  with  $|\mathfrak{N}| = \mathbb{N}$  in which  $o$ ,  $l$ ,  $+$ ,  $\times$ , and  $<$  are interpreted as you would expect. That is,  $o$  is 0,  $l$  is the successor function,  $+$  is interpreted as addition and  $\times$  as multiplication of the numbers in  $\mathbb{N}$ . Specifically,

$$\begin{aligned}o^{\mathfrak{N}} &= 0 \\l^{\mathfrak{N}}(n) &= n + 1 \\+^{\mathfrak{N}}(n, m) &= n + m \\\times^{\mathfrak{N}}(n, m) &= nm\end{aligned}$$

Of course, there are structures for  $\mathcal{L}_A$  that have domains other than  $\mathbb{N}$ . For instance, we can take  $\mathfrak{M}$  with domain  $|\mathfrak{M}| = \{a\}^*$  (the finite sequences of the single symbol  $a$ , i.e.,  $\emptyset, a, aa, aaa, \dots$ ), and interpretations

$$\begin{aligned}o^{\mathfrak{M}} &= \emptyset \\l^{\mathfrak{M}}(s) &= s \frown a \\+^{\mathfrak{M}}(n, m) &= a^{n+m} \\\times^{\mathfrak{M}}(n, m) &= a^{nm}\end{aligned}$$

These two structures are “essentially the same” in the sense that the only difference is the **elements** of the **domains** but not how the **elements** of the **domains** are related among each other by the interpretation functions. We say that the two **structures** are *isomorphic*.

It is an easy consequence of the compactness theorem that any theory true in  $\mathfrak{N}$  also has models that are not isomorphic to  $\mathfrak{N}$ . Such structures are called *non-standard*. The interesting thing about them is that while the **elements** of a standard model (i.e.,  $\mathfrak{N}$ , but also all **structures** isomorphic to it) are exhausted by the values of the standard numerals  $\bar{n}$ , i.e.,

$$|\mathfrak{N}| = \{\text{Val}^{\mathfrak{N}}(\bar{n}) : n \in \mathbb{N}\}$$

that isn't the case in non-standard models: if  $\mathfrak{M}$  is non-standard, then there is at least one  $x \in |\mathfrak{M}|$  such that  $x \neq \text{Val}^{\mathfrak{M}}(\bar{n})$  for all  $n$ .

These non-standard elements are pretty neat: they are “infinite natural numbers.” But their existence also explains, in a sense, the incompleteness phenomena. Consider an example, e.g., the consistency statement for Peano arithmetic,  $\text{Con}_{\mathbf{PA}}$ , i.e.,  $\neg \exists x \text{Prf}_{\mathbf{PA}}(x, \ulcorner \perp \urcorner)$ . Since  $\mathbf{PA}$  neither proves  $\text{Con}_{\mathbf{PA}}$  nor  $\neg \text{Con}_{\mathbf{PA}}$ , either can be consistently added to  $\mathbf{PA}$ . Since  $\mathbf{PA}$  is consistent,  $\mathfrak{N} \models \text{Con}_{\mathbf{PA}}$ , and consequently  $\mathfrak{N} \not\models \neg \text{Con}_{\mathbf{PA}}$ . So  $\mathfrak{N}$  is *not* a model of  $\mathbf{PA} \cup \{\neg \text{Con}_{\mathbf{PA}}\}$ , and all its models must be nonstandard. Models of  $\mathbf{PA} \cup \{\neg \text{Con}_{\mathbf{PA}}\}$  must contain some **element** that serves as the witness that makes  $\exists x \text{Prf}_{\mathbf{PA}}(\ulcorner \perp \urcorner)$  true, i.e., a Gödel number of a **derivation** of a contradiction from  $\mathbf{PA}$ . Such an **element** can't be standard—since  $\mathbf{PA} \vdash \neg \text{Prf}_{\mathbf{PA}}(\bar{n}, \ulcorner \perp \urcorner)$  for every  $n$ .

## 2.2 Standard Models of Arithmetic

mod:mar:stm:  
sec

The language of arithmetic  $\mathcal{L}_A$  is obviously intended to be about numbers, specifically, about natural numbers. So, “the” standard model  $\mathfrak{N}$  is special: it is the model we want to talk about. But in logic, we are often just interested in structural properties, and any two **structures** that are isomorphic share those. So we can be a bit more liberal, and consider any **structure** that is isomorphic to  $\mathfrak{N}$  “standard.”

**Definition 2.1.** A **structure** for  $\mathcal{L}_A$  is *standard* if it is isomorphic to  $\mathfrak{N}$ .

mod:mar:stm:  
prop:standard-domain

**Proposition 2.2.** *If a structure  $\mathfrak{M}$  is standard, its domain is the set of values of the standard numerals, i.e.,*

$$|\mathfrak{M}| = \{\text{Val}^{\mathfrak{M}}(\bar{n}) : n \in \mathbb{N}\}$$

*Proof.* Clearly, every  $\text{Val}^{\mathfrak{M}}(\bar{n}) \in |\mathfrak{M}|$ . We just have to show that every  $x \in |\mathfrak{M}|$  is equal to  $\text{Val}^{\mathfrak{M}}(\bar{n})$  for some  $n$ . Since  $\mathfrak{M}$  is standard, it is isomorphic to  $\mathfrak{N}$ . Suppose  $g: \mathbb{N} \rightarrow |\mathfrak{M}|$  is an isomorphism. Then  $g(n) = g(\text{Val}^{\mathfrak{N}}(\bar{n})) = \text{Val}^{\mathfrak{M}}(\bar{n})$ . But for every  $x \in |\mathfrak{M}|$ , there is an  $n \in \mathbb{N}$  such that  $g(n) = x$ , since  $g$  is **surjective**.  $\square$

If a structure  $\mathfrak{M}$  for  $\mathcal{L}_A$  is standard, the elements of its **domain** can all be named by the standard numerals  $\bar{0}, \bar{1}, \bar{2}, \dots$ , i.e., the terms  $o, o', o'', \dots$ . Of course, this does not mean that the **elements** of  $|\mathfrak{M}|$  are the numbers, just that we can pick them out the same way we can pick out the numbers in  $|\mathfrak{N}|$ . explanation

**Problem 2.1.** Show that the converse of [Proposition 2.2](#) is false, i.e., give an example of a **structure**  $\mathfrak{M}$  with  $|\mathfrak{M}| = \{\text{Val}^{\mathfrak{M}}(\bar{n}) : n \in \mathbb{N}\}$  that is not isomorphic to  $\mathfrak{N}$ .

mod:mar:stm:  
prop:thq-standard

**Proposition 2.3.** *If  $\mathfrak{M} \models \mathbf{Q}$ , and  $|\mathfrak{M}| = \{\text{Val}^{\mathfrak{M}}(\bar{n}) : n \in \mathbb{N}\}$ , then  $\mathfrak{M}$  is standard.*

*Proof.* We have to show that  $\mathfrak{M}$  is isomorphic to  $\mathfrak{N}$ . Consider the function  $g: \mathbb{N} \rightarrow |\mathfrak{M}|$  defined by  $g(n) = \text{Val}^{\mathfrak{M}}(\bar{n})$ . By the hypothesis,  $g$  is **surjective**. It is also **injective**:  $\mathbf{Q} \vdash \bar{n} \neq \bar{m}$  whenever  $n \neq m$ . Thus, since  $\mathfrak{M} \models \mathbf{Q}$ ,  $\mathfrak{M} \models \bar{n} \neq \bar{m}$ , whenever  $n \neq m$ . Thus, if  $n \neq m$ , then  $\text{Val}^{\mathfrak{M}}(\bar{n}) \neq \text{Val}^{\mathfrak{M}}(\bar{m})$ , i.e.,  $g(n) \neq g(m)$ .

We also have to verify that  $g$  is an isomorphism.

1. We have  $g(o^{\mathfrak{N}}) = g(0)$  since,  $o^{\mathfrak{N}} = 0$ . By definition of  $g$ ,  $g(0) = \text{Val}^{\mathfrak{M}}(\bar{0})$ . But  $\bar{0}$  is just  $o$ , and the value of a term which happens to be a **constant symbol** is given by what the **structure** assigns to that **constant symbol**, i.e.,  $\text{Val}^{\mathfrak{M}}(o) = o^{\mathfrak{M}}$ . So we have  $g(o^{\mathfrak{N}}) = o^{\mathfrak{M}}$  as required.
2.  $g(r^{\mathfrak{N}}(n)) = g(n+1)$ , since  $r$  in  $\mathfrak{N}$  is the successor function on  $\mathbb{N}$ . Then,  $g(n+1) = \text{Val}^{\mathfrak{M}}(\overline{n+1})$  by definition of  $g$ . But  $\overline{n+1}$  is the same term as  $\bar{n}'$ , so  $\text{Val}^{\mathfrak{M}}(\overline{n+1}) = \text{Val}^{\mathfrak{M}}(\bar{n}')$ . By the definition of the value function, this is  $= r^{\mathfrak{M}}(\text{Val}^{\mathfrak{M}}(\bar{n}))$ . Since  $\text{Val}^{\mathfrak{M}}(\bar{n}) = g(n)$  we get  $g(r^{\mathfrak{N}}(n)) = r^{\mathfrak{M}}(g(n))$ .
3.  $g(+^{\mathfrak{N}}(n, m)) = g(n+m)$ , since  $+$  in  $\mathfrak{N}$  is the addition function on  $\mathbb{N}$ . Then,  $g(n+m) = \text{Val}^{\mathfrak{M}}(\overline{n+m})$  by definition of  $g$ . But  $\mathbf{Q} \vdash \overline{n+m} = (\bar{n} + \bar{m})$ , so  $\text{Val}^{\mathfrak{M}}(\overline{n+m}) = \text{Val}^{\mathfrak{M}}(\bar{n} + \bar{m})$ . By the definition of the value function, this is  $= +^{\mathfrak{M}}(\text{Val}^{\mathfrak{M}}(\bar{n}), \text{Val}^{\mathfrak{M}}(\bar{m}))$ . Since  $\text{Val}^{\mathfrak{M}}(\bar{n}) = g(n)$  and  $\text{Val}^{\mathfrak{M}}(\bar{m}) = g(m)$ , we get  $g(+^{\mathfrak{N}}(n, m)) = +^{\mathfrak{M}}(g(n), g(m))$ .
4.  $g(\times^{\mathfrak{N}}(n, m)) = \times^{\mathfrak{M}}(g(n), g(m))$ : Exercise.
5.  $\langle n, m \rangle \in <^{\mathfrak{N}}$  iff  $n < m$ . If  $n < m$ , then  $\mathbf{Q} \vdash \bar{n} < \bar{m}$ , and also  $\mathfrak{M} \models \bar{n} < \bar{m}$ . Thus  $\langle \text{Val}^{\mathfrak{M}}(\bar{n}), \text{Val}^{\mathfrak{M}}(\bar{m}) \rangle \in <^{\mathfrak{M}}$ , i.e.,  $\langle g(n), g(m) \rangle \in <^{\mathfrak{M}}$ . If  $n \not< m$ , then  $\mathbf{Q} \vdash \neg \bar{n} < \bar{m}$ , and consequently  $\mathfrak{M} \not\models \bar{n} < \bar{m}$ . Thus, as before,  $\langle g(n), g(m) \rangle \notin <^{\mathfrak{M}}$ . Together, we get:  $\langle n, m \rangle \in <^{\mathfrak{N}}$  iff  $\langle g(n), g(m) \rangle \in <^{\mathfrak{M}}$ .

□

**explanation**

The function  $g$  is the most obvious way of defining a mapping from  $\mathbb{N}$  to the domain of any other **structure**  $\mathfrak{M}$  for  $\mathcal{L}_A$ , since every such  $\mathfrak{M}$  contains **elements** named by  $\bar{0}$ ,  $\bar{1}$ ,  $\bar{2}$ , etc. So it isn't surprising that if  $\mathfrak{M}$  makes at least some basic statements about the  $\bar{n}$ 's true in the same way that  $\mathfrak{N}$  does, and  $g$  is also bijective, then  $g$  will turn into an isomorphism. In fact, if  $|\mathfrak{M}|$  contains no **elements** other than what the  $\bar{n}$ 's name, it's the only one.

**Proposition 2.4.** *If  $\mathfrak{M}$  is standard, then  $g$  from the proof of Proposition 2.3 is the only isomorphism from  $\mathfrak{N}$  to  $\mathfrak{M}$ .*

*mod:mar:stm:  
prop:thq-unique-iso*

*Proof.* Suppose  $h: \mathbb{N} \rightarrow |\mathfrak{M}|$  is an isomorphism between  $\mathfrak{N}$  and  $\mathfrak{M}$ . We show that  $g = h$  by induction on  $n$ . If  $n = 0$ , then  $g(0) = o^{\mathfrak{M}}$  by definition of  $g$ . But since  $h$  is an isomorphism,  $h(0) = h(o^{\mathfrak{N}}) = o^{\mathfrak{M}}$ , so  $g(0) = h(0)$ .

Now consider the case for  $n + 1$ . We have

$$\begin{aligned}
g(n + 1) &= \text{Val}^{\mathfrak{M}}(\overline{n + 1}) \text{ by definition of } g \\
&= \text{Val}^{\mathfrak{M}}(\overline{n'}) \\
&= \iota^{\mathfrak{M}}(\text{Val}^{\mathfrak{M}}(\overline{n})) \\
&= \iota^{\mathfrak{M}}(g(n)) \text{ by definition of } g \\
&= \iota^{\mathfrak{M}}(h(n)) \text{ by induction hypothesis} \\
&= h(\iota^{\mathfrak{N}}(n)) \text{ since } h \text{ is an isomorphism} \\
&= h(n + 1)
\end{aligned}$$

□

For any **denumerable** set  $X$ , there's a **bijection** between  $\mathbb{N}$  and  $X$ , so every explanation such set  $X$  is potentially the **domain** of a standard model. In fact, once you pick an object  $z \in X$  and a suitable function  $s: X \rightarrow X$  as  $\circ^{\mathfrak{X}}$  and  $\iota^{\mathfrak{X}}$ , the interpretation of  $+$ ,  $\times$ , and  $<$  is already fixed. Only functions  $s = \iota^{\mathfrak{X}}$  that are both **injective** and **surjective** are suitable in a standard model. It has to be **injective** since the successor function in  $\mathfrak{N}$  is, and that  $\iota$  is **injective** is expressed by a **sentence** true in  $\mathfrak{N}$  which  $\mathfrak{X}$  thus also has to make true. It has to be **surjective** because otherwise there would be some  $x \in X$  not in the domain of  $s$ , i.e., the **sentence**  $\forall x \exists y y' = x$  would be false—but it is true in  $\mathfrak{N}$ .

## 2.3 Non-Standard Models

We call a **structure** for  $\mathcal{L}_A$  standard if it is isomorphic to  $\mathfrak{N}$ . If a **structure** explanation isn't isomorphic to  $\mathfrak{N}$ , it is called non-standard.

**Definition 2.5.** A **structure**  $\mathfrak{M}$  for  $\mathcal{L}_A$  is *non-standard* if it is not isomorphic to  $\mathfrak{N}$ . The **elements**  $x \in |\mathfrak{M}|$  which are equal to  $\text{Val}^{\mathfrak{M}}(\overline{n})$  for some  $n \in \mathbb{N}$  are called *standard numbers* (of  $\mathfrak{M}$ ), and those not, *non-standard numbers*.

By [Proposition 2.2](#), any standard **structure** for  $\mathcal{L}_A$  contains only standard explanation **elements**. Consequently, a non-standard **structure** must contain at least one non-standard element. In fact, the existence of a non-standard **element** guarantees that the **structure** is non-standard.

**Proposition 2.6.** *If a structure  $\mathfrak{M}$  for  $\mathcal{L}_A$  contains a non-standard number,  $\mathfrak{M}$  is non-standard.*

*Proof.* Suppose not, i.e., suppose  $\mathfrak{M}$  standard but contains a non-standard number  $x$ . Let  $g: \mathbb{N} \rightarrow |\mathfrak{M}|$  be an isomorphism. It is easy to see (by induction on  $n$ ) that  $g(\text{Val}^{\mathfrak{N}}(\overline{n})) = \text{Val}^{\mathfrak{M}}(\overline{n})$ . In other words,  $g$  maps standard numbers of  $\mathfrak{N}$  to standard numbers of  $\mathfrak{M}$ . If  $\mathfrak{M}$  contains a non-standard number,  $g$  cannot be **surjective**, contrary to hypothesis. □

**Problem 2.2.** Recall that  $\mathbf{Q}$  contains the axioms

$$\begin{aligned} \forall x \forall y (x' = y' \rightarrow x = y) & \quad (Q_1) \\ \forall x 0 \neq x' & \quad (Q_2) \\ \forall x (x \neq 0 \rightarrow \exists y x = y') & \quad (Q_3) \end{aligned}$$

Give structures  $\mathfrak{M}_1, \mathfrak{M}_2, \mathfrak{M}_3$  such that

1.  $\mathfrak{M}_1 \models Q_1, \mathfrak{M}_1 \models Q_2, \mathfrak{M}_1 \not\models Q_3$ ;
2.  $\mathfrak{M}_2 \models Q_1, \mathfrak{M}_2 \not\models Q_2, \mathfrak{M}_2 \models Q_3$ ; and
3.  $\mathfrak{M}_3 \not\models Q_1, \mathfrak{M}_3 \models Q_2, \mathfrak{M}_3 \models Q_3$ ;

Obviously, you just have to specify  $0^{\mathfrak{M}_i}$  and  $1^{\mathfrak{M}_i}$  for each.

explanation

It is easy enough to specify non-standard structures for  $\mathcal{L}_A$ . For instance, take the structure with domain  $\mathbb{Z}$  and interpret all non-logical symbols as usual. Since negative numbers are not values of  $\bar{n}$  for any  $n$ , this structure is non-standard. Of course, it will not be a model of arithmetic in the sense that it makes the same sentences true as  $\mathfrak{N}$ . For instance,  $\forall x x' \neq 0$  is false. However, we can prove that non-standard models of arithmetic exist easily enough, using the compactness theorem.

**Proposition 2.7.** *Let  $\mathbf{TA} = \{\varphi : \mathfrak{N} \models \varphi\}$  be the theory of  $\mathbf{N}$ .  $\mathbf{TA}$  has an enumerable non-standard model.*

*Proof.* Expand  $\mathcal{L}_A$  by a new constant symbol  $c$  and consider the set of sentences

$$\Gamma = \mathbf{TA} \cup \{c \neq \bar{0}, c \neq \bar{1}, c \neq \bar{2}, \dots\}$$

Any model  $\mathfrak{M}^c$  of  $\Gamma$  would contain an element  $x = c^{\mathfrak{M}}$  which is non-standard, since  $x \neq \text{Val}^{\mathfrak{M}}(\bar{n})$  for all  $n \in \mathbb{N}$ . Also, obviously,  $\mathfrak{M}^c \models \mathbf{TA}$ , since  $\mathbf{TA} \subseteq \Gamma$ . If we turn  $\mathfrak{M}^c$  into a structure  $\mathfrak{M}$  for  $\mathcal{L}_A$  simply by forgetting about  $c$ , its domain still contains the non-standard  $x$ , and also  $\mathfrak{M} \models \mathbf{TA}$ . The latter is guaranteed since  $c$  does not occur in  $\mathbf{TA}$ . So, it suffices to show that  $\Gamma$  has a model.

We use the compactness theorem to show that  $\Gamma$  has a model. If every finite subset of  $\Gamma$  is satisfiable, so is  $\Gamma$ . Consider any finite subset  $\Gamma_0 \subseteq \Gamma$ .  $\Gamma_0$  includes some sentences of  $\mathbf{TA}$  and some of the form  $c \neq \bar{k}$ , but only finitely many. Suppose  $k$  is the largest number so that  $c \neq \bar{k} \in \Gamma_0$ . Define  $\mathfrak{N}_k$  by expanding  $\mathfrak{N}$  to include the interpretation  $c^{\mathfrak{N}_k} = k + 1$ .  $\mathfrak{N}_k \models \Gamma_0$ : if  $\varphi \in \mathbf{TA}$ ,  $\mathfrak{N}_k \models \varphi$  since  $\mathfrak{N}_k$  is just like  $\mathfrak{N}$  in all respects except  $c$ , and  $c$  does not occur in  $\varphi$ . And  $\mathfrak{N}_k \models c \neq \bar{n}$ , since  $n \leq k$ , and  $\text{Val}^{\mathfrak{N}_k}(c) = k + 1$ . Thus, every finite subset of  $\Gamma$  is satisfiable.  $\square$



## 2.4 Models of $\mathbf{Q}$

We know that there are non-standard **structures** that make the same **sentences** true as  $\mathfrak{N}$  does, i.e., is a model of **TA**. Since  $\mathfrak{N} \models \mathbf{Q}$ , any model of **TA** is also a model of **Q**. **Q** is much weaker than **TA**, e.g.,  $\mathbf{Q} \not\models \forall x \forall y (x+y) = (y+x)$ . Weaker theories are easier to satisfy: they have more models. E.g., **Q** has models which make  $\forall x \forall y (x+y) = (y+x)$  false, but those cannot also be models of **TA**, or **PA** for that matter. Models of **Q** are also relatively simple: we can specify them explicitly.

mod:mar:mdq:  
ex:model-K-of-Q

**Example 2.8.** Consider the **structure**  $\mathfrak{K}$  with domain  $|\mathfrak{K}| = \mathbb{N} \cup \{a\}$  and interpretations

$$\begin{aligned} 0^{\mathfrak{K}} &= 0 \\ \iota^{\mathfrak{K}}(x) &= \begin{cases} x+1 & \text{if } x \in \mathbb{N} \\ a & \text{if } x = a \end{cases} \\ +^{\mathfrak{K}}(x, y) &= \begin{cases} x+y & \text{if } x, y \in \mathbb{N} \\ a & \text{otherwise} \end{cases} \\ \times^{\mathfrak{K}}(x, y) &= \begin{cases} xy & \text{if } x, y \in \mathbb{N} \\ a & \text{otherwise} \end{cases} \\ <^{\mathfrak{K}} &= \{\langle x, y \rangle : x, y \in \mathbb{N} \text{ and } x < y\} \cup \{\langle x, a \rangle : x \in |\mathfrak{K}|\} \end{aligned}$$

To show that  $\mathfrak{K} \models \mathbf{Q}$  we have to verify that all axioms of **Q** are true in  $\mathfrak{K}$ . For convenience, let's write  $x^*$  for  $\iota^{\mathfrak{K}}(x)$  (the “successor” of  $x$  in  $\mathfrak{K}$ ),  $x \oplus y$  for  $+^{\mathfrak{K}}(x, y)$  (the “sum” of  $x$  and  $y$  in  $\mathfrak{K}$ ),  $x \otimes y$  for  $\times^{\mathfrak{K}}(x, y)$  (the “product” of  $x$  and  $y$  in  $\mathfrak{K}$ ), and  $x \odot y$  for  $\langle x, y \rangle \in <^{\mathfrak{K}}$ . With these abbreviations, we can give the operations in  $\mathfrak{K}$  more perspicuously as

$x$	$x^*$	$x \oplus y$	$m$	$a$	$x \otimes y$	$m$	$a$
$n$	$n+1$	$n$	$n+m$	$a$	$n$	$nm$	$a$
$a$	$a$	$a$	$a$	$a$	$a$	$a$	$a$

We have  $n \odot m$  iff  $n < m$  for  $n, m \in \mathbb{N}$  and  $x \odot a$  for all  $x \in |\mathfrak{K}|$ .

$\mathfrak{K} \models \forall x \forall y (x' = y' \rightarrow x = y)$  since  $*$  is **injective**.  $\mathfrak{K} \models \forall x 0 \neq x'$  since 0 is not a  $*$ -successor in  $\mathfrak{K}$ .  $\mathfrak{N} \models \forall x (x \neq 0 \rightarrow \exists y x = y')$  since for every  $n > 0$ ,  $n = (n-1)^*$ , and  $a = a^*$ .

$\mathfrak{K} \models \forall x (x + 0) = x$  since  $n \oplus 0 = n + 0 = n$ , and  $a \oplus 0 = a$  by definition of  $\oplus$ .  $\mathfrak{K} \models \forall x \forall y (x + y') = (x + y)'$  is a bit trickier. If  $n, m$  are both standard, we have:

$$(n \oplus m^*) = (n + (m + 1)) = (n + m) + 1 = (n \oplus m)^*$$

since  $\oplus$  and  $*$  agree with  $+$  and  $\iota$  on standard numbers. Now suppose  $x \in |\mathfrak{K}|$ . Then

$$(x \oplus a^*) = (x \oplus a) = a = a^* = (x \oplus a)^*$$

The remaining case is if  $y \in |\mathfrak{K}|$  but  $x = a$ . Here we also have to distinguish cases according to whether  $y = n$  is standard or  $y = b$ :

$$\begin{aligned}(a \oplus n^*) &= (a \oplus (n + 1)) = a = a^* = (x \oplus n)^* \\ (a \oplus a^*) &= (a \oplus a) = a = a^* = (x \oplus a)^*\end{aligned}$$

This is of course a bit more detailed than needed. For instance, since  $a \oplus z = a$  whatever  $z$  is, we can immediately conclude  $a \oplus a^* = a$ . The remaining axioms can be verified the same way.

$\mathfrak{K}$  is thus a model of  $\mathbf{Q}$ . Its “addition”  $\oplus$  is also commutative. But there are other sentences true in  $\mathfrak{N}$  but false in  $\mathfrak{K}$ , and vice versa. For instance,  $a \otimes a$ , so  $\mathfrak{K} \models \exists x x < x$  and  $\mathfrak{K} \not\models \forall x \neg x < x$ . This shows that  $\mathbf{Q} \not\models \forall x \neg x < x$ .

**Problem 2.3.** Prove that  $\mathfrak{K}$  from [Example 2.8](#) satisfies the remaining axioms of  $\mathbf{Q}$ ,

$$\forall x (x \times 0) = 0 \tag{Q6}$$

$$\forall x \forall y (x \times y') = ((x \times y) + x) \tag{Q7}$$

$$\forall x \forall y (x < y \leftrightarrow \exists z (x + z' = y)) \tag{Q8}$$

Find a [sentence](#) only involving  $\prime$  true in  $\mathfrak{N}$  but false in  $\mathfrak{K}$ .

**Example 2.9.** Consider the [structure](#)  $\mathfrak{L}$  with domain  $|\mathfrak{L}| = \mathbb{N} \cup \{a, b\}$  and interpretations  $\prime^{\mathfrak{L}} = *$ ,  $+^{\mathfrak{L}} = \oplus$  given by

[mod:mar:mdq:](#)  
[ex:model-L-of-Q](#)

$x$	$x^*$	$x \oplus y$	$m$	$a$	$b$
$n$	$n + 1$	$n$	$n + m$	$b$	$a$
$a$	$a$	$a$	$a$	$b$	$a$
$b$	$b$	$b$	$b$	$b$	$a$

Since  $*$  is [injective](#),  $0$  is not in its range, and every  $x \in |\mathfrak{L}|$  other than  $0$  is, axioms  $Q_1$ – $Q_3$  are true in  $\mathfrak{L}$ . For any  $x$ ,  $x \oplus 0 = x$ , so  $Q_4$  is true as well. For  $Q_5$ , consider  $x \oplus y^*$  and  $(x \oplus y)^*$ . They are equal if  $x$  and  $y$  are both standard, since then  $*$  and  $\oplus$  agree with  $\prime$  and  $+$ . If  $x$  is non-standard, and  $y$  is standard, we have  $x \oplus y^* = x = x^* = (x \oplus y)^*$ . If  $x$  and  $y$  are both non-standard, we have four cases:

$$\begin{aligned}a \oplus a^* &= b = b^* = (a \oplus a)^* \\ b \oplus b^* &= a = a^* = (b \oplus b)^* \\ b \oplus a^* &= b = b^* = (b \oplus y)^* \\ a \oplus b^* &= a = a^* = (a \oplus b)^*\end{aligned}$$

If  $x$  is standard, but  $y$  is non-standard, we have

$$\begin{aligned}n \oplus a^* &= n \oplus a = b = b^* = (n \oplus a)^* \\ n \oplus b^* &= n \oplus b = a = a^* = (n \oplus b)^*\end{aligned}$$

So,  $\mathfrak{L} \models Q_5$ . However,  $a \oplus 0 \neq 0 \oplus a$ , so  $\mathfrak{L} \not\models \forall x \forall y (x + y) = (y + x)$ .

**Problem 2.4.** Expand  $\mathfrak{L}$  of [Example 2.9](#) to include  $\otimes$  and  $\ominus$  that interpret  $\times$  and  $<$ . Show that your structure satisfies the remaining axioms of  $\mathbf{Q}$ ,

$$\forall x (x \times 0) = 0 \tag{Q6}$$

$$\forall x \forall y (x \times y') = ((x \times y) + x) \tag{Q7}$$

$$\forall x \forall y (x < y \leftrightarrow \exists z (x + z' = y)) \tag{Q8}$$

**Problem 2.5.** In  $\mathfrak{L}$  of [Example 2.9](#),  $a^* = a$  and  $b^* = b$ . Is there a model of  $\mathbf{Q}$  in which  $a^* = b$  and  $b^* = a$ ?

We've explicitly constructed models of  $\mathbf{Q}$  in which the non-standard [elements](#) live “beyond” the standard elements. [In fact](#), that much is required by the axioms. A non-standard [element](#)  $x$  cannot be  $\ominus 0$ . Otherwise, for some  $z$ ,  $x \oplus z^* = 0$  by  $Q8$ . But then  $0 = x \oplus z^* = (x \oplus z)^*$  by  $Q5$ , contradicting  $Q2$ . Also, for every  $n$ ,  $\mathbf{Q} \vdash \forall x (x < \bar{n}' \rightarrow (x = \bar{0} \vee x = \bar{1} \vee \dots \vee x = \bar{n}))$ , so we can't have  $a \ominus n$  for any  $n > 0$ .

## 2.5 Computable Models of Arithmetic

The standard model  $\mathfrak{N}$  has two nice features. Its domain is the natural numbers  $\mathbb{N}$ , i.e., its elements are just the kinds of things we want to talk about using the language of arithmetic, and the standard numeral  $\bar{n}$  actually picks out  $n$ . The other nice feature is that the interpretations of the non-logical symbols of  $\mathcal{L}_A$  are all *computable*. The successor, addition, and multiplication functions which serve as  $s^{\mathfrak{N}}$ ,  $+$  <sup>$\mathfrak{N}$</sup> , and  $\times^{\mathfrak{N}}$  are computable functions of numbers. (Computable by Turing machines, or definable by primitive recursion, say.) And the less-than relation on  $\mathfrak{N}$ , i.e.,  $<^{\mathfrak{N}}$ , is decidable.

Non-standard models of arithmetical theories such as  $\mathbf{Q}$  and  $\mathbf{PA}$  must contain non-standard elements. Thus their domains typically include [elements](#) in addition to  $\mathbb{N}$ . However, any countable [structure](#) can be built on any [denumerable](#) set, including  $\mathbb{N}$ . So there are also non-standard models with domain  $\mathbb{N}$ . In such models  $\mathfrak{M}$ , of course, at least some numbers cannot play the roles they usually play, since some  $k$  must be different from  $\text{Val}^{\mathfrak{M}}(\bar{n})$  for all  $n \in \mathbb{N}$ .

**Definition 2.10.** A structure  $\mathfrak{M}$  for  $\mathcal{L}_A$  is *computable* iff  $|\mathfrak{M}| = \mathbb{N}$  and  $s^{\mathfrak{M}}$ ,  $+$  <sup>$\mathfrak{M}$</sup> ,  $\times^{\mathfrak{M}}$  are computable functions and  $<^{\mathfrak{M}}$  is a decidable relation.

**Example 2.11.** Recall the structure  $\mathfrak{K}$  from [Example 2.8](#) Its domain was

$|\mathfrak{K}| = \mathbb{N} \cup \{a\}$  and interpretations

$$\begin{aligned} 0^{\mathfrak{K}} &= 0 \\ \iota^{\mathfrak{K}}(x) &= \begin{cases} x + 1 & \text{if } x \in \mathbb{N} \\ a & \text{if } x = a \end{cases} \\ +^{\mathfrak{K}}(x, y) &= \begin{cases} x + y & \text{if } x, y \in \mathbb{N} \\ a & \text{otherwise} \end{cases} \\ \times^{\mathfrak{K}}(x, y) &= \begin{cases} xy & \text{if } x, y \in \mathbb{N} \\ a & \text{otherwise} \end{cases} \\ <^{\mathfrak{K}} &= \{\langle x, y \rangle : x, y \in \mathbb{N} \text{ and } x < y\} \cup \{\langle x, a \rangle : x \in |\mathfrak{K}|\} \end{aligned}$$

But  $|\mathfrak{K}|$  is **denumerable** and so is equinumerous with  $\mathbb{N}$ . For instance,  $g: \mathbb{N} \rightarrow |\mathfrak{K}|$  with  $g(0) = a$  and  $g(n) = n + 1$  for  $n > 0$  is a **bijection**. We can turn it into an isomorphism between a new model  $\mathfrak{K}'$  of  $\mathbf{Q}$  and  $\mathfrak{K}$ . In  $\mathfrak{K}'$ , we have to assign different functions and relations to the symbols of  $\mathcal{L}_A$ , since different **elements** of  $\mathbb{N}$  play the roles of standard and non-standard numbers.

Specifically, 0 now plays the role of  $a$ , not of the smallest standard number. The smallest standard number is now 1. So we assign  $0^{\mathfrak{K}'} = 1$ . The successor function is also different now: given a standard number, i.e., an  $n > 0$ , it still returns  $n + 1$ . But 0 now plays the role of  $a$ , which is its own successor. So  $\iota^{\mathfrak{K}'}(0) = 0$ . For addition and multiplication we likewise have

$$\begin{aligned} +^{\mathfrak{K}'}(x, y) &= \begin{cases} x + y & \text{if } x, y > 0 \\ 0 & \text{otherwise} \end{cases} \\ \times^{\mathfrak{K}'}(x, y) &= \begin{cases} xy & \text{if } x, y > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

And we have  $\langle x, y \rangle \in <^{\mathfrak{K}'}$  iff  $x < y$  and  $x > 0$  and  $y > 0$ , or if  $y = 0$ .

All of these functions are computable functions of natural numbers and  $<^{\mathfrak{K}'}$  is a decidable relation on  $\mathbb{N}$ —but they are not the same functions as successor, addition, and multiplication on  $\mathbb{N}$ , and  $<^{\mathfrak{K}'}$  is not the same relation as  $<$  on  $\mathbb{N}$ .

**Problem 2.6.** Give a **structure**  $\mathfrak{L}'$  with  $|\mathfrak{L}'| = \mathbb{N}$  isomorphic to  $\mathfrak{L}$  of **Example 2.9**.

**explanation**

This example shows that  $\mathbf{Q}$  has computable non-standard models with domain  $\mathbb{N}$ . However, the following result shows that this is not true for models of  $\mathbf{PA}$  (and thus also for models of  $\mathbf{TA}$ ).

**Theorem 2.12** (Tennenbaum's Theorem).  *$\mathfrak{N}$  is the only computable model of  $\mathbf{PA}$ .*

## Chapter 3

# The Interpolation Theorem

### 3.1 Introduction

mod:int:int:  
sec The interpolation theorem is the following result: Suppose  $\models \varphi \rightarrow \psi$ . Then there is a **sentence**  $\chi$  such that  $\models \varphi \rightarrow \chi$  and  $\models \chi \rightarrow \psi$ . Moreover, every **constant symbol**, **function symbol**, and **predicate symbol** (other than  $=$ ) in  $\chi$  occurs both in  $\varphi$  and  $\psi$ . The **sentence**  $\chi$  is called an *interpolant* of  $\varphi$  and  $\psi$ .

The interpolation theorem is interesting in its own right, but its main importance lies in the fact that it can be used to prove results about definability in a theory, and the conditions under which combining two consistent theories results in a consistent theory. The first result is known as the Beth definability theorem; the second, Robinson's joint consistency theorem.

### 3.2 Separation of Sentences

mod:int:sep:  
sec A bit of groundwork is needed before we can proceed with the proof of the interpolation theorem. An interpolant for  $\varphi$  and  $\psi$  is a **sentence**  $\chi$  such that  $\varphi \models \chi$  and  $\chi \models \psi$ . By contraposition, the latter is true iff  $\neg\psi \models \neg\chi$ . A **sentence**  $\chi$  with this property is said to *separate*  $\varphi$  and  $\neg\psi$ . So finding an interpolant for  $\varphi$  and  $\psi$  amounts to finding a **sentence** that separates  $\varphi$  and  $\neg\psi$ . As so often, it will be useful to consider a generalization: a sentence that separates two *sets* of **sentences**.

**Definition 3.1.** A sentence  $\chi$  *separates* sets of sentences  $\Gamma$  and  $\Delta$  if and only if  $\Gamma \models \chi$  and  $\Delta \models \neg\chi$ . If no such **sentence** exists, then  $\Gamma$  and  $\Delta$  are *inseparable*.

The inclusion relations between the classes of models of  $\Gamma$ ,  $\Delta$  and  $\chi$  are represented below:

mod:int:sep:  
lem:sep1 **Lemma 3.2.** *Suppose  $\mathcal{L}_0$  is the language containing every **constant symbol**, **function symbol** and **predicate symbol** (other than  $=$ ) that occurs in both  $\Gamma$  and  $\Delta$ , and let  $\mathcal{L}'_0$  be obtained by the addition of infinitely many new **constant symbols**  $c_n$  for  $n \geq 0$ . Then if  $\Gamma$  and  $\Delta$  are inseparable in  $\mathcal{L}_0$ , they are also inseparable in  $\mathcal{L}'_0$ .*

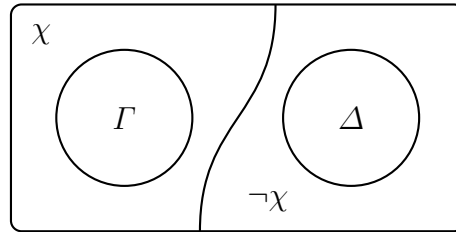


Figure 3.1:  $\chi$  separates  $\Gamma$  and  $\Delta$

*Proof.* We proceed indirectly: suppose by way of contradiction that  $\Gamma$  and  $\Delta$  are separated in  $\mathcal{L}'_0$ . Then  $\Gamma \models \chi[c/x]$  and  $\Delta \models \neg\chi[c/x]$  for some  $\chi \in \mathcal{L}_0$  (where  $c$  is a new **constant symbol**—the case where  $\chi$  contains more than one such new **constant symbol** is similar). By compactness, there are *finite* subsets  $\Gamma_0$  of  $\Gamma$  and  $\Delta_0$  of  $\Delta$  such that  $\Gamma_0 \models \chi[c/x]$  and  $\Delta_0 \models \neg\chi[c/x]$ . Let  $\gamma$  be the conjunction of all **formulas** in  $\Gamma_0$  and  $\delta$  the conjunction of all **formulas** in  $\Delta_0$ . Then

$$\gamma \models \chi[c/x], \quad \delta \models \neg\chi[c/x].$$

From the former, by Generalization, we have  $\gamma \models \forall x \chi$ , and from the latter by contraposition,  $\chi[c/x] \models \neg\delta$ , whence also  $\forall x \chi \models \neg\delta$ . Contraposition again gives  $\delta \models \neg\forall x \chi$ . By monotony,

$$\Gamma \models \forall x \chi, \quad \Delta \models \neg\forall x \chi,$$

so that  $\forall x \chi$  separates  $\Gamma$  and  $\Delta$  in  $\mathcal{L}_0$ . □

**Lemma 3.3.** *Suppose that  $\Gamma \cup \{\exists x \sigma\}$  and  $\Delta$  are inseparable, and  $c$  is a new **constant symbol** not in  $\Gamma$ ,  $\Delta$ , or  $\sigma$ . Then  $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$  and  $\Delta$  are also inseparable.* mod:int:sep:  
lem:sep2

*Proof.* Suppose for contradiction that  $\chi$  separates  $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$  and  $\Delta$ , while at the same time  $\Gamma \cup \{\exists x \sigma\}$  and  $\Delta$  are inseparable. We distinguish two cases:

1.  $c$  does not occur in  $\chi$ : in this case  $\Gamma \cup \{\exists x \sigma, \neg\chi\}$  is satisfiable (otherwise  $\chi$  separates  $\Gamma \cup \{\exists x \sigma\}$  and  $\Delta$ ). It remains so if  $\sigma[c/x]$  is added, so  $\chi$  does not separate  $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$  and  $\Delta$  after all.
2.  $c$  does occur in  $\chi$  so that  $\chi$  has the form  $\chi[c/x]$ . Then we have that

$$\Gamma \cup \{\exists x \sigma, \sigma[c/x]\} \models \chi[c/x],$$

whence  $\Gamma, \exists x \sigma \models \forall x (\sigma \rightarrow \chi)$  by the Deduction Theorem and Generalization, and finally  $\Gamma \cup \{\exists x \sigma\} \models \exists x \chi$ . On the other hand,  $\Delta \models \neg\chi[c/x]$  and hence by Generalization  $\Delta \models \neg\exists x \chi$ . So  $\Gamma \cup \{\exists x \sigma\}$  and  $\Delta$  are separable, a contradiction. □

### 3.3 Craig's Interpolation Theorem

mod:int:prf:  
sec

mod:int:prf:  
thm:interpol

**Theorem 3.4** (Craig's Interpolation Theorem). *If  $\models \varphi \rightarrow \psi$ , then there is a sentence  $\chi$  such that  $\models \varphi \rightarrow \chi$  and  $\models \chi \rightarrow \psi$ , and every constant symbol, function symbol, and predicate symbol (other than  $=$ ) in  $\chi$  occurs both in  $\varphi$  and  $\psi$ . The sentence  $\chi$  is called an interpolant of  $\varphi$  and  $\psi$ .*

*Proof.* Suppose  $\mathcal{L}_1$  is the language of  $\varphi$  and  $\mathcal{L}_2$  is the language of  $\psi$ . Let  $\mathcal{L}_0 = \mathcal{L}_1 \cap \mathcal{L}_2$ . For each  $i \in \{0, 1, 2\}$ , let  $\mathcal{L}'_i$  be obtained from  $\mathcal{L}_i$  by adding the infinitely many new constant symbols  $c_0, c_1, c_2, \dots$ .

If  $\varphi$  is unsatisfiable,  $\exists x x \neq x$  is an interpolant. If  $\neg\psi$  is unsatisfiable (and hence  $\psi$  is valid),  $\exists x x = x$  is an interpolant. So we may assume also that both  $\varphi$  and  $\neg\psi$  are satisfiable.

In order to prove the contrapositive of the Interpolation Theorem, assume that there is no interpolant for  $\varphi$  and  $\psi$ . In other words, assume that  $\{\varphi\}$  and  $\{\neg\psi\}$  are inseparable in  $\mathcal{L}_0$ .

Our goal is to extend the pair  $(\{\varphi\}, \{\neg\psi\})$  to a maximally inseparable pair  $(\Gamma^*, \Delta^*)$ . Let  $\varphi_0, \varphi_1, \varphi_2, \dots$  enumerate the sentences of  $\mathcal{L}_1$ , and  $\psi_0, \psi_1, \psi_2, \dots$  enumerate the sentences of  $\mathcal{L}_2$ . We define two increasing sequences of sets of sentences  $(\Gamma_n, \Delta_n)$ , for  $n \geq 0$ , as follows. Put  $\Gamma_0 = \{\varphi\}$  and  $\Delta_0 = \{\neg\psi\}$ . Assuming  $(\Gamma_n, \Delta_n)$  are already defined, define  $\Gamma_{n+1}$  and  $\Delta_{n+1}$  by:

1. If  $\Gamma_n \cup \{\varphi_n\}$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ , put  $\varphi_n$  in  $\Gamma_{n+1}$ . Moreover, if  $\varphi_n$  is an existential formula  $\exists x \sigma$  then pick a new constant symbol  $c$  not occurring in  $\Gamma_n, \Delta_n, \varphi_n$  or  $\psi_n$ , and put  $\sigma[c/x]$  in  $\Gamma_{n+1}$ .
2. If  $\Gamma_{n+1}$  and  $\Delta_n \cup \{\psi_n\}$  are inseparable in  $\mathcal{L}'_0$ , put  $\psi_n$  in  $\Delta_{n+1}$ . Moreover, if  $\psi_n$  is an existential formula  $\exists x \sigma$ , then pick a new constant symbol  $c$  not occurring in  $\Gamma_{n+1}, \Delta_n, \varphi_n$  or  $\psi_n$ , and put  $\sigma[c/x]$  in  $\Delta_{n+1}$ .

Finally, define:

$$\Gamma^* = \bigcup_{n \geq 0} \Gamma_n, \quad \Delta^* = \bigcup_{n \geq 0} \Delta_n.$$

By simultaneous induction on  $n$  we can now prove:

mod:int:prf:  
part-a

1.  $\Gamma_n$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ ;

mod:int:prf:  
part-b

2.  $\Gamma_{n+1}$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ .

The basis for (1) is given by Lemma 3.2. For part (2), we need to distinguish three cases:

1. If  $\Gamma_0 \cup \{\varphi_0\}$  and  $\Delta_0$  are separable, then  $\Gamma_1 = \Gamma_0$  and (2) is just (1);
2. If  $\Gamma_1 = \Gamma_0 \cup \{\varphi_0\}$ , then  $\Gamma_1$  and  $\Delta_0$  are inseparable by construction.

3. It remains to consider the case where  $\varphi_0$  is existential, so that  $\Gamma_1 = \Gamma_0 \cup \{\exists x \sigma, \sigma[c/x]\}$ . By construction,  $\Gamma_0 \cup \{\exists x \sigma\}$  and  $\Delta_0$  are inseparable, so that by [Lemma 3.3](#) also  $\Gamma_0 \cup \{\exists x \sigma, \sigma[c/x]\}$  and  $\Delta_0$  are inseparable.

This completes the basis of the induction for (1) and (2) above. Now for the inductive step. For (1), if  $\Delta_{n+1} = \Delta_n \cup \{\psi_n\}$  then  $\Gamma_{n+1}$  and  $\Delta_{n+1}$  are inseparable by construction (even when  $\psi_n$  is existential, by [Lemma 3.3](#)); if  $\Delta_{n+1} = \Delta_n$  (because  $\Gamma_{n+1}$  and  $\Delta_n \cup \{\psi_n\}$  are separable), then we use the induction hypothesis on (2). For the inductive step for (2), if  $\Gamma_{n+2} = \Gamma_{n+1} \cup \{\varphi_{n+1}\}$  then  $\Gamma_{n+2}$  and  $\Delta_{n+1}$  are inseparable by construction (even when  $\varphi_{n+1}$  is existential, by [Lemma 3.3](#)); and if  $\Gamma_{n+2} = \Gamma_{n+1}$  then we use the inductive case for (1) just proved. This concludes the induction on (1) and (2).

It follows that  $\Gamma^*$  and  $\Delta^*$  are inseparable; if not, by compactness, there is  $n \geq 0$  that separates  $\Gamma_n$  and  $\Delta_n$ , against (1). In particular,  $\Gamma^*$  and  $\Delta^*$  are consistent: for if the former or the latter is inconsistent, then they are separated by  $\exists x x \neq x$  or  $\forall x x = x$ , respectively.

We now show that  $\Gamma^*$  is maximally consistent in  $\mathcal{L}'_1$  and likewise  $\Delta^*$  in  $\mathcal{L}'_2$ . For the former, suppose that  $\varphi_n \notin \Gamma^*$  and  $\neg\varphi_n \notin \Gamma^*$ , for some  $n \geq 0$ . If  $\varphi_n \notin \Gamma^*$  then  $\Gamma_n \cup \{\varphi_n\}$  is separable from  $\Delta_n$ , and so there is  $\chi \in \mathcal{L}'_0$  such that both:

$$\Gamma^* \models \varphi_n \rightarrow \chi, \quad \Delta^* \models \neg\chi.$$

Likewise, if  $\neg\varphi_n \notin \Gamma^*$ , there is  $\chi' \in \mathcal{L}'_0$  such that both:

$$\Gamma^* \models \neg\varphi_n \rightarrow \chi', \quad \Delta^* \models \neg\chi'.$$

By propositional logic,  $\Gamma^* \models \chi \vee \chi'$  and  $\Delta^* \models \neg(\chi \vee \chi')$ , so  $\chi \vee \chi'$  separates  $\Gamma^*$  and  $\Delta^*$ . A similar argument establishes that  $\Delta^*$  is maximal.

Finally, we show that  $\Gamma^* \cap \Delta^*$  is maximally consistent in  $\mathcal{L}'_0$ . It is obviously consistent, since it is the intersection of consistent sets. To show maximality, let  $\sigma \in \mathcal{L}'_0$ . Now,  $\Gamma^*$  is maximal in  $\mathcal{L}'_1 \supseteq \mathcal{L}'_0$ , and similarly  $\Delta^*$  is maximal in  $\mathcal{L}'_2 \supseteq \mathcal{L}'_0$ . It follows that either  $\sigma \in \Gamma^*$  or  $\neg\sigma \in \Gamma^*$ , and either  $\sigma \in \Delta^*$  or  $\neg\sigma \in \Delta^*$ . If  $\sigma \in \Gamma^*$  and  $\neg\sigma \in \Delta^*$  then  $\sigma$  would separate  $\Gamma^*$  and  $\Delta^*$ ; and if  $\neg\sigma \in \Gamma^*$  and  $\sigma \in \Delta^*$  then  $\Gamma^*$  and  $\Delta^*$  would be separated by  $\neg\sigma$ . Hence, either  $\sigma \in \Gamma^* \cap \Delta^*$  or  $\neg\sigma \in \Gamma^* \cap \Delta^*$ , and  $\Gamma^* \cap \Delta^*$  is maximal.

Since  $\Gamma^*$  is maximally consistent, it has a model  $\mathfrak{M}'_1$  whose **domain**  $|\mathfrak{M}'_1|$  comprises all and only the elements  $c^{\mathfrak{M}'_1}$  interpreting the **constant symbols**—just like in the proof of the completeness theorem (??). Similarly,  $\Delta^*$  has a model  $\mathfrak{M}'_2$  whose **domain**  $|\mathfrak{M}'_2|$  is given by the interpretations  $c^{\mathfrak{M}'_2}$  of the **constant symbols**.

Let  $\mathfrak{M}_1$  be obtained from  $\mathfrak{M}'_1$  by dropping interpretations for **constant symbols**, **function symbols**, and **predicate symbols** in  $\mathcal{L}'_1 \setminus \mathcal{L}'_0$ , and similarly for  $\mathfrak{M}_2$ . Then the map  $h: M_1 \rightarrow M_2$  defined by  $h(c^{\mathfrak{M}'_1}) = c^{\mathfrak{M}'_2}$  is an isomorphism in  $\mathcal{L}'_0$ , because  $\Gamma^* \cap \Delta^*$  is maximally consistent in  $\mathcal{L}'_0$ , as shown. This follows because any  $\mathcal{L}'_0$ -sentence either belongs to both  $\Gamma^*$  and  $\Delta^*$ , or to neither: so  $c^{\mathfrak{M}'_1} \in P^{\mathfrak{M}'_1}$  if and only if  $P(c) \in \Gamma^*$  if and only if  $P(c) \in \Delta^*$  if and only if



$c^{\mathfrak{M}'_2} \in P^{\mathfrak{M}'_2}$ . The other conditions satisfied by isomorphisms can be established similarly.

Let us now define a model  $\mathfrak{M}$  for the language  $\mathcal{L}_1 \cup \mathcal{L}_2$  as follows:

1. The domain  $|\mathfrak{M}|$  is just  $|\mathfrak{M}'_2|$ , i.e., the set of all elements  $c^{\mathfrak{M}'_2}$ ;
2. If a predicate symbol  $P$  is in  $\mathcal{L}_2 \setminus \mathcal{L}_1$  then  $P^{\mathfrak{M}} = P^{\mathfrak{M}'_2}$ ;
3. If a predicate  $P$  is in  $\mathcal{L}_1 \setminus \mathcal{L}_2$  then  $P^{\mathfrak{M}} = h(P^{\mathfrak{M}'_2})$ , i.e.,  $\langle c_1^{\mathfrak{M}'_2}, \dots, c_n^{\mathfrak{M}'_2} \rangle \in P^{\mathfrak{M}}$  if and only if  $\langle c_1^{\mathfrak{M}'_1}, \dots, c_n^{\mathfrak{M}'_1} \rangle \in P^{\mathfrak{M}'_1}$ .
4. If a predicate symbol  $P$  is in  $\mathcal{L}_0$  then  $P^{\mathfrak{M}} = P^{\mathfrak{M}'_2} = h(P^{\mathfrak{M}'_1})$ .
5. Function symbols of  $\mathcal{L}_1 \cup \mathcal{L}_2$ , including constant symbols, are handled similarly.

Finally, one shows by induction on formulas that  $\mathfrak{M}$  agrees with  $\mathfrak{M}'_1$  on all formulas of  $\mathcal{L}'_1$  and with  $\mathfrak{M}'_2$  on all formulas of  $\mathcal{L}'_2$ . In particular,  $\mathfrak{M} \models I^* \cup \Delta^*$ , whence  $\mathfrak{M} \models \varphi$  and  $\mathfrak{M} \models \neg\psi$ , and  $\not\models \varphi \rightarrow \psi$ . This concludes the proof of Craig's Interpolation Theorem.  $\square$

### 3.4 The Definability Theorem

mod:int:def:  
sec

One important application of the interpolation theorem is Beth's definability theorem. To define an  $n$ -place relation  $R$  we can give a formula  $\chi$  with  $n$  free variables which does not involve  $R$ . This would be an *explicit* definition of  $R$  in terms of  $\chi$ . We can then say also that a theory  $\Sigma(P)$  in a language containing the  $n$ -place predicate symbol  $P$  explicitly defines  $P$  if it contains (or at least entails) a formalized explicit definition, i.e.,

$$\Sigma(P) \models \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow \chi(x_1, \dots, x_n)).$$

But an explicit definition is only one way of defining—in the sense of determining completely—a relation. A theory may also be such that the interpretation of  $P$  is fixed by the interpretation of the rest of the language in any model. The definability theorem states that whenever a theory fixes the interpretation of  $P$  in this way—whenever it *implicitly defines*  $P$ —then it also explicitly defines it.

**Definition 3.5.** Suppose  $\mathcal{L}$  is a language not containing the predicate symbol  $P$ . A set  $\Sigma(P)$  of sentences of  $\mathcal{L} \cup \{P\}$  *explicitly defines*  $P$  if and only if there is a formula  $\chi(x_1, \dots, x_n)$  of  $\mathcal{L}$  such that

$$\Sigma(P) \models \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow \chi(x_1, \dots, x_n)).$$

**Definition 3.6.** Suppose  $\mathcal{L}$  is a language not containing the predicate symbols  $P$  and  $P'$ . A set  $\Sigma(P)$  of sentences of  $\mathcal{L} \cup \{P\}$  *implicitly defines*  $P$  if and only if

$$\Sigma(P) \cup \Sigma(P') \models \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow P'(x_1, \dots, x_n)),$$

where  $\Sigma(P')$  is the result of uniformly replacing  $P$  with  $P'$  in  $\Sigma(P)$ .

In other words, for any model  $\mathfrak{M}$  and  $R, R' \subseteq |\mathfrak{M}|^n$ , if both  $(\mathfrak{M}, R) \models \Sigma(P)$  and  $(\mathfrak{M}, R') \models \Sigma(P')$ , then  $R = R'$ ; where  $(\mathfrak{M}, R)$  is the **structure**  $\mathfrak{M}'$  for the expansion of  $\mathcal{L}$  to  $\mathcal{L} \cup \{P\}$  such that  $P^{\mathfrak{M}'} = R$ , and similarly for  $(\mathfrak{M}, R')$ .

**Theorem 3.7** (Beth Definability Theorem). *A set  $\Sigma(P)$  of  $\mathcal{L} \cup \{P\}$ -formulas implicitly defines  $P$  if and only if  $\Sigma(P)$  explicitly defines  $P$ .*

*Proof.* If  $\Sigma(P)$  explicitly defines  $P$  then both

$$\begin{aligned}\Sigma(P) \models & \quad \forall x_1 \dots \forall x_n [(P(x_1, \dots, x_n) \leftrightarrow \chi(x_1, \dots, x_n))] \\ \Sigma(P') \models & \quad \forall x_1 \dots \forall x_n [(P'(x_1, \dots, x_n) \leftrightarrow \chi(x_1, \dots, x_n))]\end{aligned}$$

and the conclusion follows. For the converse: assume that  $\Sigma(P)$  implicitly defines  $P$ . First, we add **constant symbols**  $c_1, \dots, c_n$  to  $\mathcal{L}$ . Then

$$\Sigma(P) \cup \Sigma(P') \models P(c_1, \dots, c_n) \rightarrow P'(c_1, \dots, c_n).$$

By compactness, there are finite sets  $\Delta_0 \subseteq \Sigma(P)$  and  $\Delta_1 \subseteq \Sigma(P')$  such that

$$\Delta_0 \cup \Delta_1 \models P(c_1, \dots, c_n) \rightarrow P'(c_1, \dots, c_n).$$

Let  $\theta(P)$  be the conjunction of all **sentences**  $\varphi(P)$  such that either  $\varphi(P) \in \Delta_0$  or  $\varphi(P') \in \Delta_1$  and let  $\theta(P')$  be the conjunction of all **sentences**  $\varphi(P')$  such that either  $\varphi(P) \in \Delta_0$  or  $\varphi(P') \in \Delta_1$ . Then  $\theta(P) \wedge \theta(P') \models P(c_1, \dots, c_n) \rightarrow P'c_1 \dots c_n$ . We can re-arrange this so that each **predicate symbol** occurs on one side of  $\models$ :

$$\theta(P) \wedge P(c_1, \dots, c_n) \models \theta(P') \rightarrow P'(c_1, \dots, c_n).$$

By Craig's Interpolation Theorem there is a **sentence**  $\chi(c_1, \dots, c_n)$  not containing  $P$  or  $P'$  such that:

$$\theta(P) \wedge P(c_1, \dots, c_n) \models \chi(c_1, \dots, c_n); \quad \chi(c_1, \dots, c_n) \models \theta(P') \rightarrow P'(c_1, \dots, c_n).$$

From the former of these two entailments we have:  $\theta(P) \models P(c_1, \dots, c_n) \rightarrow \chi(c_1, \dots, c_n)$ . And from the latter, since an  $\mathcal{L} \cup \{P\}$ -model  $(\mathfrak{M}, R) \models \varphi(P)$  if and only if the corresponding  $\mathcal{L} \cup \{P'\}$ -model  $(\mathfrak{M}, R) \models \varphi(P')$ , we have  $\chi(c_1, \dots, c_n) \models \theta(P) \rightarrow P(c_1, \dots, c_n)$ , from which:

$$\theta(P) \models \chi(c_1, \dots, c_n) \rightarrow P(c_1, \dots, c_n).$$

Putting the two together,  $\theta(P) \models P(c_1, \dots, c_n) \leftrightarrow \chi(c_1, \dots, c_n)$ , and by monotony and generalization also

$$\Sigma(P) \models \forall x_1 \dots \forall x_n (P(x_1, \dots, x_n) \leftrightarrow \chi(x_1, \dots, x_n)). \quad \square$$

## Chapter 4

# Lindström's Theorem

### 4.1 Introduction

In this chapter we aim to prove Lindström's characterization of first-order logic as the maximal logic for which (given certain further constraints) the Compactness and the Downward Löwenheim-Skolem theorems hold (?? and ??). First, we need a more general characterization of the general class of logics to which the theorem applies. We will restrict ourselves to *relational* languages, i.e., languages which only contain **predicate symbols** and individual constants, but no **function symbols**.

### 4.2 Abstract Logics

mod:lin:alg:  
sec

**Definition 4.1.** An *abstract logic* is a pair  $\langle L, \models_L \rangle$ , where  $L$  is a function that assigns to each **language**  $\mathcal{L}$  a set  $L(\mathcal{L})$  of **sentences**, and  $\models_L$  is a relation between **structures** for the **language**  $\mathcal{L}$  and **elements** of  $L(\mathcal{L})$ . In particular,  $\langle F, \models \rangle$  is ordinary first-order logic, i.e.,  $F$  is the function assigning to the **language**  $\mathcal{L}$  the set of first-order **sentences** built from the constants in  $\mathcal{L}$ , and  $\models$  is the satisfaction relation of first-order logic.

Notice that we are still employing the same notion of **structure** for a given **language** as for first-order logic, but we do not presuppose that **sentences** are built up from the basic symbols in  $\mathcal{L}$  in the usual way, nor that the relation  $\models_L$  is recursively defined in the same way as for first-order logic. So for instance the definition, being completely general, is intended to capture the case where **sentences** in  $\langle L, \models_L \rangle$  contain infinitely long conjunctions or disjunction, or quantifiers other than  $\exists$  and  $\forall$  (e.g., “there are infinitely many  $x$  such that ...”), or perhaps infinitely long quantifier prefixes. To emphasize that “**sentences**” in  $L(\mathcal{L})$  need not be ordinary **sentences** of first-order logic, in this chapter we use **variables**  $\alpha, \beta, \dots$  to range over them, and reserve  $\varphi, \psi, \dots$  for ordinary first-order **formulas**.

**Definition 4.2.** Let  $\text{Mod}_L(\alpha)$  denote the class  $\{\mathfrak{M} : \mathfrak{M} \models_L \alpha\}$ . If the language needs to be made explicit, we write  $\text{Mod}_L^{\mathcal{L}}(\alpha)$ . Two structures  $\mathfrak{M}$  and  $\mathfrak{N}$  for  $\mathcal{L}$  are *elementarily equivalent in*  $\langle L, \models_L \rangle$ , written  $\mathfrak{M} \equiv_L \mathfrak{N}$ , if the same sentences from  $L(\mathcal{L})$  are true in each.

**Definition 4.3.** An abstract logic  $\langle L, \models_L \rangle$  for the language  $\mathcal{L}$  is *normal* if it satisfies the following properties:

1. (*L-Monotony*) For languages  $\mathcal{L}$  and  $\mathcal{L}'$ , if  $\mathcal{L} \subseteq \mathcal{L}'$ , then  $L(\mathcal{L}) \subseteq L(\mathcal{L}')$ .
2. (*Expansion Property*) For each  $\alpha \in L(\mathcal{L})$  there is a finite subset  $\mathcal{L}'$  of  $\mathcal{L}$  such that the relation  $\mathfrak{M} \models_L \alpha$  depends only on the reduct of  $\mathfrak{M}$  to  $\mathcal{L}'$ ; i.e., if  $\mathfrak{M}$  and  $\mathfrak{N}$  have the same reduct to  $\mathcal{L}'$  then  $\mathfrak{M} \models_L \alpha$  if and only if  $\mathfrak{N} \models_L \alpha$ .
3. (*Isomorphism Property*) If  $\mathfrak{M} \models_L \alpha$  and  $\mathfrak{M} \simeq \mathfrak{N}$  then also  $\mathfrak{N} \models_L \alpha$ .
4. (*Renaming Property*) The relation  $\models_L$  is preserved under renaming: if the language  $\mathcal{L}'$  is obtained from  $\mathcal{L}$  by replacing each symbol  $P$  by a symbol  $P'$  of the same arity and each constant  $c$  by a distinct constant  $c'$ , then for each structure  $\mathfrak{M}$  and sentence  $\alpha$ ,  $\mathfrak{M} \models_L \alpha$  if and only if  $\mathfrak{M}' \models_L \alpha'$ , where  $\mathfrak{M}'$  is the  $\mathcal{L}'$ -structure corresponding to  $\mathcal{L}$  and  $\alpha' \in L(\mathcal{L}')$ .
5. (*Boolean Property*) The abstract logic  $\langle L, \models_L \rangle$  is closed under the Boolean connectives in the sense that for each  $\alpha \in L(\mathcal{L})$  there is a  $\beta \in L(\mathcal{L})$  such that  $\mathfrak{M} \models_L \beta$  if and only if  $\mathfrak{M} \not\models_L \alpha$ , and for each  $\alpha$  and  $\beta$  there is a  $\gamma$  such that  $\text{Mod}_L(\gamma) = \text{Mod}_L(\alpha) \cap \text{Mod}_L(\beta)$ . Similarly for atomic formulas and the other connectives.
6. (*Quantifier Property*) For each constant  $c$  in  $\mathcal{L}$  and  $\alpha \in L(\mathcal{L})$  there is a  $\beta \in L(\mathcal{L})$  such that

$$\text{Mod}_L^{\mathcal{L}'}(\beta) = \{\mathfrak{M} : (\mathfrak{M}, a) \in \text{Mod}_L^{\mathcal{L}}(\alpha) \text{ for some } a \in |\mathfrak{M}|\},$$

where  $\mathcal{L}' = \mathcal{L} \setminus \{c\}$  and  $(\mathfrak{M}, a)$  is the expansion of  $\mathfrak{M}$  to  $\mathcal{L}$  assigning  $a$  to  $c$ .

7. (*Relativization Property*) Given a sentence  $\alpha \in L(\mathcal{L})$  and symbols  $R, c_1, \dots, c_n$  not in  $\mathcal{L}$ , there is a sentence  $\beta \in L(\mathcal{L} \cup \{R, c_1, \dots, c_n\})$  called the *relativization* of  $\alpha$  to  $R(x, c_1, \dots, c_n)$ , such that for each structure  $\mathfrak{M}$ :

$$(\mathfrak{M}, X, b_1, \dots, b_n) \models_L \beta \text{ if and only if } \mathfrak{N} \models_L \alpha,$$

where  $\mathfrak{N}$  is the substructure of  $\mathfrak{M}$  with domain  $|\mathfrak{N}| = \{a \in |\mathfrak{M}| : R^{\mathfrak{M}}(a, b_1, \dots, b_n)\}$  (see Remark 1), and  $(\mathfrak{M}, X, b_1, \dots, b_n)$  is the expansion of  $\mathfrak{M}$  interpreting  $R, c_1, \dots, c_n$  by  $X, b_1, \dots, b_n$ , respectively (with  $X \subseteq M^{n+1}$ ).

**Definition 4.4.** Given two abstract logics  $\langle L_1, \models_{L_1} \rangle$  and  $\langle L_2, \models_{L_2} \rangle$  we say that the latter is *at least as expressive* as the former, written  $\langle L_1, \models_{L_1} \rangle \leq \langle L_2, \models_{L_2} \rangle$ , if for each **language**  $\mathcal{L}$  and **sentence**  $\alpha \in L_1(\mathcal{L})$  there is a **sentence**  $\beta \in L_2(\mathcal{L})$  such that  $\text{Mod}_{L_1}^{\mathcal{L}}(\alpha) = \text{Mod}_{L_2}^{\mathcal{L}}(\beta)$ . The logics  $\langle L_1, \models_{L_1} \rangle$  and  $\langle L_2, \models_{L_2} \rangle$  are *equivalent* if  $\langle L_1, \models_{L_1} \rangle \leq \langle L_2, \models_{L_2} \rangle$  and  $\langle L_2, \models_{L_2} \rangle \leq \langle L_1, \models_{L_1} \rangle$ .

*Remark 5.* First-order logic, i.e., the abstract logic  $\langle F, \models \rangle$ , is normal. In fact, the above properties are mostly straightforward for first-order logic. We just remark that the expansion property comes down to extensionality, and that the relativization of a **sentence**  $\alpha$  to  $R(x, c_1, \dots, c_n)$  is obtained by replacing each **subformula**  $\forall x \beta$  by  $\forall x (R(x, c_1, \dots, c_n) \rightarrow \beta)$ . Moreover, if  $\langle L, \models_L \rangle$  is normal, then  $\langle F, \models \rangle \leq \langle L, \models_L \rangle$ , as can be shown by induction on first-order **formulas**. Accordingly, with no loss in generality, we can assume that every first-order **sentence** belongs to every normal logic.

### 4.3 Compactness and Löwenheim-Skolem Properties

mod.lin:lsp:  
sec

We now give the obvious extensions of compactness and Löwenheim-Skolem to the case of abstract logics.

**Definition 4.5.** An abstract logic  $\langle L, \models_L \rangle$  has the *Compactness Property* if each set  $\Gamma$  of  $L(\mathcal{L})$ -**sentences** is satisfiable whenever each finite  $\Gamma_0 \subseteq \Gamma$  is satisfiable.

**Definition 4.6.**  $\langle L, \models_L \rangle$  has the *Downward Löwenheim-Skolem property* if any satisfiable  $\Gamma$  has an **enumerable** model.

The notion of partial isomorphism from [Definition 1.15](#) is purely “algebraic” (i.e., given without reference to the **sentences** of the language but only to the constants provided by the **language**  $\mathcal{L}$  of the **structures**), and hence it applies to the case of abstract logics. In case of first-order logic, we know from [Theorem 1.17](#) that if two **structures** are partially isomorphic then they are elementarily equivalent. That proof does not carry over to abstract logics, for induction on **formulas** need not be available for arbitrary  $\alpha \in L(\mathcal{L})$ , but the theorem is true nonetheless, provided the Löwenheim-Skolem property holds.

mod.lin:lsp:  
thm:abstract-p-isom

**Theorem 4.7.** *Suppose  $\langle L, \models_L \rangle$  is a normal logic with the Löwenheim-Skolem property. Then any two **structures** that are partially isomorphic are elementarily equivalent in  $\langle L, \models_L \rangle$ .*

*Proof.* Suppose  $\mathfrak{M} \simeq_p \mathfrak{N}$ , but for some  $\alpha$  also  $\mathfrak{M} \models_L \alpha$  while  $\mathfrak{N} \not\models_L \alpha$ . By the Isomorphism Property we can assume that  $|\mathfrak{M}|$  and  $|\mathfrak{N}|$  are disjoint, and by the Expansion Property we can assume that  $\alpha \in L(\mathcal{L})$  for a finite **language**  $\mathcal{L}$ . Let  $\mathcal{I}$  be a set of partial isomorphisms between  $\mathfrak{M}$  and  $\mathfrak{N}$ , and with no loss of generality also assume that if  $p \in \mathcal{I}$  and  $q \subseteq p$  then also  $q \in \mathcal{I}$ .

$|\mathfrak{M}|^{<\omega}$  is the set of finite sequences of **elements** of  $|\mathfrak{M}|$ . Let  $S$  be the ternary relation over  $|\mathfrak{M}|^{<\omega}$  representing concatenation, i.e., if  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in |\mathfrak{M}|^{<\omega}$  then

$S(\mathbf{a}, \mathbf{b}, \mathbf{c})$  holds if and only if  $\mathbf{c}$  is the concatenation of  $\mathbf{a}$  and  $\mathbf{b}$ ; and let  $T$  be the ternary relation such that  $T(\mathbf{a}, b, \mathbf{c})$  holds for  $b \in M$  and  $\mathbf{a}, \mathbf{c} \in |\mathfrak{M}|^{<\omega}$  if and only if  $\mathbf{a} = a_1, \dots, a_n$  and  $\mathbf{c} = a_1, \dots, a_n, b$ . Pick new 3-place predicate symbols  $P$  and  $Q$  and form the structure  $\mathfrak{M}^*$  having the universe  $|\mathfrak{M}| \cup |\mathfrak{M}|^{<\omega}$ , having  $\mathfrak{M}$  as a substructure, and interpreting  $P$  and  $Q$  by the concatenation relations  $S$  and  $T$  (so  $\mathfrak{M}^*$  is in the language  $\mathcal{L} \cup \{P, Q\}$ ).

Define  $|\mathfrak{N}|^{<\omega}$ ,  $S'$ ,  $T'$ ,  $P'$ ,  $Q'$  and  $\mathfrak{N}^*$  analogously. Since by hypothesis  $\mathfrak{M} \simeq_p \mathfrak{N}$ , there is a relation  $I$  between  $|\mathfrak{M}|^{<\omega}$  and  $|\mathfrak{N}|^{<\omega}$  such that  $I(\mathbf{a}, \mathbf{b})$  holds if and only if  $\mathbf{a}$  and  $\mathbf{b}$  are isomorphic and satisfy the back-and-forth condition of Definition 1.15. Now, let  $\mathfrak{M}$  be the structure whose domain is the union of the domains of  $\mathfrak{M}^*$  and  $\mathfrak{N}^*$ , having  $\mathfrak{M}^*$  and  $\mathfrak{N}^*$  as substructures, in the language with one extra binary predicate symbol  $R$  interpreted by the relation  $I$  and predicate symbols denoting the domains  $|\mathfrak{M}^*|$  and  $|\mathfrak{N}^*|$ .

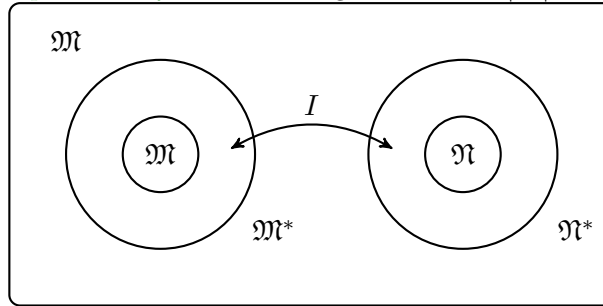


Figure 4.1: The structure  $\mathfrak{M}$  with the internal partial isomorphism.

The crucial observation is that in the language of the structure  $\mathfrak{M}$  there is a first-order sentence  $\theta_1$  true in  $\mathfrak{M}$  saying that  $\mathfrak{M} \models_L \alpha$  and  $\mathfrak{N} \not\models_L \alpha$  (this requires the Relativization Property), as well as a first-order sentence  $\theta_2$  true in  $\mathfrak{M}$  saying that  $\mathfrak{M} \simeq_p \mathfrak{N}$  via the partial isomorphism  $I$ . By the Löwenheim-Skolem Property,  $\theta_1$  and  $\theta_2$  are jointly true in an enumerable model  $\mathfrak{M}_0$  containing partially isomorphic substructures  $\mathfrak{M}_0$  and  $\mathfrak{N}_0$  such that  $\mathfrak{M}_0 \models_L \alpha$  and  $\mathfrak{N}_0 \not\models_L \alpha$ . But enumerable partially isomorphic structures are in fact isomorphic by Theorem 1.16, contradicting the Isomorphism Property of normal abstract logics.  $\square$

#### 4.4 Lindström's Theorem

**Lemma 4.8.** *Suppose  $\alpha \in L(\mathcal{L})$ , with  $\mathcal{L}$  finite, and assume also that there is an  $n \in \mathbb{N}$  such that for any two structures  $\mathfrak{M}$  and  $\mathfrak{N}$ , if  $\mathfrak{M} \equiv_n \mathfrak{N}$  and  $\mathfrak{M} \models_L \alpha$  then also  $\mathfrak{N} \models_L \alpha$ . Then  $\alpha$  is equivalent to a first-order sentence, i.e., there is a first-order  $\theta$  such that  $\text{Mod}_L(\alpha) = \text{Mod}_L(\theta)$ .*

*Proof.* Let  $n$  be such that any two  $n$ -equivalent structures  $\mathfrak{M}$  and  $\mathfrak{N}$  agree on the value assigned to  $\alpha$ . Recall Proposition 1.19: there are only finitely many

mod:lin:prf:  
sec

mod:lin:prf:  
lem:lindstrom

first-order **sentences** in a finite **language** that have quantifier rank no greater than  $n$ , up to logical equivalence. Now, for each fixed **structure**  $\mathfrak{M}$  let  $\theta_{\mathfrak{M}}$  be the conjunction of all first-order **sentences**  $\alpha$  true in  $\mathfrak{M}$  with  $\text{qr}(\alpha) \leq n$  (this conjunction is finite), so that  $\mathfrak{N} \models \theta_{\mathfrak{M}}$  if and only if  $\mathfrak{N} \equiv_n \mathfrak{M}$ . Then put  $\theta = \bigvee \{\theta_{\mathfrak{M}} : \mathfrak{M} \models_L \alpha\}$ ; this disjunction is also finite (up to logical equivalence).

The conclusion  $\text{Mod}_L(\alpha) = \text{Mod}_L(\theta)$  follows. In fact, if  $\mathfrak{N} \models_L \theta$  then for some  $\mathfrak{M} \models_L \alpha$  we have  $\mathfrak{N} \models \theta_{\mathfrak{M}}$ , whence also  $\mathfrak{N} \models_L \alpha$  (by the hypothesis of the lemma). Conversely, if  $\mathfrak{N} \models_L \alpha$  then  $\theta_{\mathfrak{N}}$  is a disjunct in  $\theta$ , and since  $\mathfrak{N} \models \theta_{\mathfrak{N}}$ , also  $\mathfrak{N} \models_L \theta$ .  $\square$

*mod:lin:prf:  
thm:lindstrom*

**Theorem 4.9** (Lindström's Theorem). *Suppose  $\langle L, \models_L \rangle$  has the Compactness and the Löwenheim-Skolem Properties. Then  $\langle L, \models_L \rangle \leq \langle F, \models \rangle$  (so  $\langle L, \models_L \rangle$  is equivalent to first-order logic).*

*Proof.* By [Lemma 4.8](#), it suffices to show that for any  $\alpha \in L(\mathcal{L})$ , with  $\mathcal{L}$  finite, there is  $n \in \mathbb{N}$  such that for any two **structures**  $\mathfrak{M}$  and  $\mathfrak{N}$ : if  $\mathfrak{M} \equiv_n \mathfrak{N}$  then  $\mathfrak{M}$  and  $\mathfrak{N}$  agree on  $\alpha$ . For then  $\alpha$  is equivalent to a first-order **sentence**, from which  $\langle L, \models_L \rangle \leq \langle F, \models \rangle$  follows. Since we are working in a finite, purely relational **language**, by [Theorem 1.23](#) we can replace the statement that  $\mathfrak{M} \equiv_n \mathfrak{N}$  by the corresponding algebraic statement that  $I_n(\emptyset, \emptyset)$ .

Given  $\alpha$ , suppose towards a contradiction that for each  $n$  there are **structures**  $\mathfrak{M}_n$  and  $\mathfrak{N}_n$  such that  $I_n(\emptyset, \emptyset)$ , but (say)  $\mathfrak{M}_n \models_L \alpha$  whereas  $\mathfrak{N}_n \not\models_L \alpha$ . By the Isomorphism Property we can assume that all the  $\mathfrak{M}_n$ 's interpret the constants of the language by the same objects; furthermore, since there are only finitely many atomic **sentences** in the language, we may also assume that they satisfy the same atomic **sentences** (we can take a subsequence of the  $\mathfrak{M}_n$ 's otherwise). Let  $\mathfrak{M}$  be the union of all the  $\mathfrak{M}_n$ 's, i.e., the unique minimal **structure** having each  $\mathfrak{M}_n$  as a substructure. As in the proof of [Theorem 4.7](#), let  $\mathfrak{M}^*$  be the extension of  $\mathfrak{M}$  with **domain**  $|\mathfrak{M}| \cup |\mathfrak{M}|^{<\omega}$ , in the expanded **language** comprising the concatenation predicates  $P$  and  $Q$ .

Similarly, define  $\mathfrak{N}_n$ ,  $\mathfrak{N}$  and  $\mathfrak{N}^*$ . Now let  $\mathfrak{M}$  be the **structure** whose **domain** comprises the **domains** of  $\mathfrak{M}^*$  and  $\mathfrak{N}^*$  as well as the natural numbers  $\mathbb{N}$  along with their natural ordering  $\leq$ , in the **language** with extra predicates representing the **domains**  $|\mathfrak{M}|$ ,  $|\mathfrak{N}|$ ,  $|\mathfrak{M}|^{<\omega}$  and  $|\mathfrak{N}|^{<\omega}$  as well as predicates coding the domains of  $\mathfrak{M}_n$  and  $\mathfrak{N}_n$  in the sense that:

$$\begin{aligned} |\mathfrak{M}_n| &= \{a \in |\mathfrak{M}| : R(a, n)\}; & |\mathfrak{N}_n| &= \{a \in |\mathfrak{N}| : S(a, n)\}; \\ |\mathfrak{M}_n|^{<\omega} &= \{a \in |\mathfrak{M}|^{<\omega} : R(a, n)\}; & |\mathfrak{N}_n|^{<\omega} &= \{a \in |\mathfrak{N}|^{<\omega} : S(a, n)\}. \end{aligned}$$

The **structure**  $\mathfrak{M}$  also has a ternary relation  $J$  such that  $J(n, \mathbf{a}, \mathbf{b})$  holds if and only if  $I_n(\mathbf{a}, \mathbf{b})$ .

Now there is a **sentence**  $\theta$  in the **language**  $\mathcal{L}$  augmented by  $R, S, J$ , etc., saying that  $\leq$  is a discrete linear ordering with first but no last element and such that  $\mathfrak{M}_n \models \alpha$ ,  $\mathfrak{N}_n \not\models \alpha$ , and for each  $n$  in the ordering,  $J(n, \mathbf{a}, \mathbf{b})$  holds if and only if  $I_n(\mathbf{a}, \mathbf{b})$ .

Using the Compactness Property, we can find a model  $\mathfrak{M}^*$  of  $\theta$  in which the ordering contains a non-standard element  $n^*$ . In particular then  $\mathfrak{M}^*$  will

contain substructures  $\mathfrak{M}_{n^*}$  and  $\mathfrak{N}_{n^*}$  such that  $\mathfrak{M}_{n^*} \models_L \alpha$  and  $\mathfrak{N}_{n^*} \not\models_L \alpha$ . But now we can define a set  $\mathcal{I}$  of pairs of  $k$ -tuples from  $|\mathfrak{M}_{n^*}|$  and  $|\mathfrak{N}_{n^*}|$  by putting  $\langle \mathbf{a}, \mathbf{b} \rangle \in \mathcal{I}$  if and only if  $J(n^* - k, \mathbf{a}, \mathbf{b})$ , where  $k$  is the length of  $\mathbf{a}$  and  $\mathbf{b}$ . Since  $n^*$  is non-standard, for each standard  $k$  we have that  $n^* - k > 0$ , and the set  $\mathcal{I}$  witnesses the fact that  $\mathfrak{M}_{n^*} \simeq_p \mathfrak{N}_{n^*}$ . But by [Theorem 4.7](#),  $\mathfrak{M}_{n^*}$  is  $L$ -equivalent to  $\mathfrak{N}_{n^*}$ , a contradiction.  $\square$

## Photo Credits



# Bibliography