Model Theory, Universal Algebra and Order

J. Adler, J. Schmid, M. Sprenger

January 18, 2006

Contents

Preface

This is the preface. It is an unnumbered chapter. The `markboth` TeX field at the beginning of this paragraph sets the correct page heading for the Preface portion of the document. The preface does not appear in the table of contents.

Introduction

As the title *Model Theory, Universal Algebra and Order* insinuates, the present module is mainly compiled from three parts. So why not make three different modules or papers? The answer may be put in one single word: *Structure*. What is structure in the context of mathematics? There are several answers to this, but the one we would like to lie emphasis on is to understand structure as *giving a form to something otherwise formless, thereby exhibiting an implicit meaning.*

This is all very nice and eloquent, but how does it translate to the fields of Mathematics lending their names for the title of this module? For Model Theory, the answer is a very direct one: Model Theory *is* the Theory of structures, where structures are sets with additional operations, relations and especially designated elements called constants. First Order Logic provides languages to speak about these structures on a formal level, and the connection between structures and language is so close that we are able to draw conclusions from one to the other and vice versa.

There is another meaning of structure which comes close to Order: Structure as providing a way to arrange something otherwise formless. In this sense, ordered sets are sets equipped with a binary relation (the *order*) structuring the underlying set. On the other hand, ordered sets are none other than examples of structures in the sense of the previous paragraph.

Remains the question where structure comes in with Universal Algebra: Algeras in the sense of Universal Algebra are again structures in the sense of Model Theory, the language specified by the *type* of the algebra. Moreover, as in the case of ordered sets, the fundamental operations lay a structure upon the universe of the algebras, even some sort of *division into parts* if we consider congruences.

Thus, the three fields Model Theory, Universal Algebra and Order Theory share a close relation with the notion of structure in one or more senses of the word. But this is clearly not the only relation between them. Universal Algebra may be regarded as the theory of structures for functional languages, i.e. first–

order languages without relation symbols. Ordered sets, on the other hand, are relational structures, i.e. a first–order language being able to describe their basic properties must be equipped with at least one relation symbol.

Therefore we could regard Universal Algebra and Order Theory as specializations or sub–branches of Model Theory. But one look at the table of contents reveals that ordered sets are mentioned quite early in the development. The reason for this lies in the fact that, being very fundamental and rather simple in their definition, ordered sets qualify as nice examples for structures in the sense of Model Theory. Moreover the possibility to show algebraic aspects shows that we even need not limit ourselves to examples of relational structures. Ordered sets serve at least in two ways[1] : As examples of structures for simple languages, and as examples of algebras in the sense of Universal Algebra.

The last introductory remark concerns the role of Set Theory: Being a (classical) branch of Mathematics, Model Theory is based on Set Theory the same way as are Group Theory or Calculus, i.e. proofs of Model Theoretical assumptions are done in the framework of Zermelo–Fränkel Set Theory. However, this very Set Theory is formulated in a very simple formal language which can be regarded from the point of view of first–order logic, so in principle, Set Theory is an example of a theory in the sense of this module and allows thus for structures and models. This slightly disturbing (seeming) circularity will be partially treated in the appendix of this module.

How To Use This Module

In principle there is nothing to say against using the three parts which make up this module separately, i.e. the reader may restrict her or his attention exclusively to the chapters and sections dealing with Model Theory or Ordered Sets. However, the three subjects of this module are interdependent in several ways, so the best procedure to ensure a deeper understanding surely is to read them all.

[1] There are more, of course: E.g., whenever any kind of structure is involved, *substructures* will have to be considered as well, and these again, when looked at from the point of view of set–inclusion, form an ordered set displaying nice properties.

Chapter 1

The General Background

The aim of this chapter is to provide some general background knowledge of concepts which will be used in the different contexts.

1.1 Set Theory, Featuring the Axiom of Choice

We will now introduce what little of set theory is needed to grasp the set theoretical notions used in this paper.

The notation we use for Set Theoretical notions is common standard, e.g. $\cup$ and $\bigcup$ will stand for (binary and arbitrary, respectively) union of sets, likewise $\cap$ and $\bigcap$ for intersection, $\subseteq$ for the subset–relation (equality not excluded) and $\mathcal{P}(S)$ will denote the power–set (the set of all subsets) of the set S.

A set is called **finite** if for some natural number n, there is a bijective function (to be defined below) from this set into set $\{0, \dots, n-1\}$, and **infinite** otherwise. Infinite sets are further distinguished by their "degree of infiniteness", a concept formalized by the **cardinality**, a notion which unfortunately is too complicated to be introduced in detail for the moment. The same holds for the notion of (proper) **classes**. Every set is a class, but there are classes which no longer are sets; those are called proper classes. For the moment it is enough to think of classes as collections which are too big (or too general) to be sets; the classic "infamous" example is the "set" of all sets, which rather turns out to be the *class* of all sets. Other examples include the class of all cardinal numbers, the class of all singleton sets, the class of all groups, rings, fields etc, the class of $\mathcal{L}$–structures for any formal language $\mathcal{L}$, and any $\mathrm{Mod}\,\Sigma$ for any consistent set of sentences Σ (as will be shown later).

Clearly, the classical axioms of set theory are not to be questioned nor scrutinized. Thus we consider the implications of the existence of infinite sets such as ω and we will never worry about what axiomatic complications may be involved when dealing with constructions such as set–unions, powersets or even instantiations of the schema of replacements. For the moment we may as well simply have faith in the fact that set theory *can* be axiomatized by giving a (rather simple, as far as syntax is concerned) set of sentences (axioms) in a formal language which in its simplicity is almost boring, since its sole non–logical constituent is the binary relation–symbol $\in$.

Whenever we mention a model for a set of sentences, we are building this model in our universe, which, in turn, is actually relying on another theory formalized in first–order logic. From this point of view, Model Theory deals with translations from arbitrary theories into Set Theory, and consistency–arguments should always be relativized to the (unverified) consistency of Set Theory. But clearly nobody who is serious about dealing with Model Theory is keeping this "detail" in mind; you would not expect somebody working with real–valued calculus to handle numbers as infinitely nested intervals either, would you?

Still, we feel obliged to make a few remarks on behalf of the axiom which in itself was and still is subject of arguments about constructivism: The Axiom of Choice. As should be well remembered, the Axiom of Choice (AC) postulates something along the line of

> "The cartesian product of a non–empty family of non–empty sets is non–empty."

or

> "Given a set of sets all of which are non–empty, there exists a function which, for each of these sets, picks one element."

Although intuition cries for undisputed acceptance of these statements when the sets involved are from everyday's experience (products of the sets of natural, rational, real numbers or finite sets), things tend to be less clear for exotic cases, i.e. when the sets involved are way beyond the horizon of countability, let alone of finiteness. It should be well known to any mathematician that the AC is neither provable nor refutable from the rest of the axioms of set theory which constitute the so called *axiomatization of Zermelo and Fraenkel*, provided this axiomatization is in fact consistent, which in itself is still open to discussion.

We will not argue about the acceptability of the AC. In fact, the AC is central to some of the constructions we will use, e.g. the ultraproducts in chapter 7. We will make use of the AC not in its incarnations mentioned above, but in form of the so called Zorn's Lemma (ZL). Unfortunately, although ZL is in general easier to use than the AC[1] and closer to the construct aimed at, formulating ZL is a

[1] This phenomenon mirrors the fact that the AC involves less structure and thus less in-

bit more complicated since it involves notions from Order Theory (see Chapter 3 for the definitions of "ordered set", "chain", "upper bound" and "maximal element"):

> "If every sub–chain on some ordered set P has an upper bound, then P has a maximal element."

An instructing exercise in set theory is to show that the AC and ZL are in fact equivalent.

Another notion of set theory that we will use is the cardinality of a set and cardinal numbers. For the purposes of this module, think of two sets **having the same cardinality** if and only if there is an exact correspondence (a *bijection*) between their elements; i.e. we could write down two lists of their respective elements of the same "length". The length of this list would then correspond to the **cardinality** of the sets.

We write **card** X for the cardinality of a set X. In this sense $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{Z}$ have the same cardinality (they all are **countable**), while $\mathbb{R}$ is "bigger" than all of them (**uncountable**) and $\mathcal{P}(\mathbb{N})$ again has the same cardinality as $\mathbb{R}$. We will use the symbol $\aleph_0$ for the countable cardinality.

Since not any two sets have the same cardinality, we need to compare the cardinalities (or sizes) of sets, and thus we write $\operatorname{card} X \leq \operatorname{card} Y$ if and only if there is an injective function (to be defined below) from X to Y.

1.2 Functions, Mappings and Operations

We assume that the reader is familiar with the concept of a function (or mapping or map). For the sake of completeness, we note that a function is specified by two sets X and Y and a set of ordered pairs

$$f = \{\ldots, \langle x, y\rangle, \ldots\} \subseteq X \times Y$$

such that for any $x \in X$ there is a unique $y \in Y$ such that $\langle x, y\rangle \in f$, i.e.

$$\text{for all } x \in X \text{ there is a } y \in Y \text{ such that } \langle x, y\rangle \in f$$

and

$$\langle x_1, y_1\rangle, \langle x_2, y_2\rangle \in f \text{ and } x_1 = x_2 \text{ implies } y_1 = y_2.$$

formation on the side of the structures, while when applying ZL, a major part of work is already done by "preparing" the structure under consideration, i.e. by finding the requested constructs that fulfill the premises of ZL. With the AC, the premises are rather weak, they do not presuppose a structure. Thus, information mostly comes later when trying to use choice functions or the like to find the eventual result originally aimed at.

In this module, for f as above, we will write

- **domain** or **source** of f, $\operatorname{dom} f$, for X, and **co–domain** or **target** of f, $\operatorname{cod} f$, for Y;
- the **image** of f, $\operatorname{im} f$, for the set $\{f(x)\,;\, x \in \operatorname{dom} f\}$, and if $Z \subseteq \operatorname{dom} f$, the **image of** Z **under** f, $f[Z]$, for $\{f(x)\,;\, x \in Z\}$;
- the **value of** x **under** f, $f(x)$, for the unique y such that $\langle x, y\rangle \in f$, provided $x \in \operatorname{dom} f$; in this case f is said to **map** x **to** y, which is denoted by $f(x) = y$ or $f : x \mapsto y$;
- $f : X \longrightarrow Y$ to denote that f has domain X and co–domain Y;
- id_X for the **identity map** on X, i.e. the *unique map* $\mathrm{id}_X : X \longrightarrow X$ with $\mathrm{id}_X(x) = x$ for all $x \in X$.

The following distinct cases are assumed to be well–known:

- A function f is said to be **injective** (or **one–to–one**) if and only if

$$\text{for any } x_1, x_2 \in \operatorname{dom} f, f(x_1) = f(x_2) \text{ implies } x_1 = x_2;$$

- for $Z \subseteq Y$, f is said to be **onto** Z if and only if

$$\text{for any } y \in Z, f(x) = y \text{ for some } x \in \operatorname{dom} f,$$

  i.e. if and only if $\operatorname{im} F = Z$;
- if f is onto $\operatorname{cod} f$, i.e. if $\operatorname{im} f = \operatorname{cod} f$, then f is simply called **onto** or **surjective**;
- f is said to be **bijective** or a **bijection** if and only if f is both injective and surjective.

If f is injective, then there is a canonical **inverse function**

$$f^{-1} : \operatorname{im} f \longrightarrow \operatorname{dom} f$$

given by

$$f^{-1}(y) = x \text{ iff } f(x) = y.$$

If f is injective, f^{-1} is also injective, and of course $\operatorname{dom} f^{-1} = \operatorname{cod} f$ if and only if f is onto and hence a bijection.

Please remember that, for a finite set X and $f : X \longrightarrow X$, f is injective if and only if f is surjective if and only if f is bijective.

The notion of **operations** will be used to denote the special case where $\operatorname{dom} f = (\operatorname{cod} f)^n$ for some $n \in \mathbb{N}$ (for the notion of direct product, see 1.3 below). Therefore, operations are the only functions which can be iterated, by which we mean the following: if $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$, then $g \circ f$ stands for the **product of** f **and** g, i.e. $g \circ f : X \longrightarrow Z$ with $g \circ f(x) = g(f(x))$. $\circ$ is **associative** in the sense that $f \circ (g \circ h) = (f \circ g) \circ h$, provided the notation is meaningful regarding domains and co–domains. If $f : X \longrightarrow X$, then the **iterations** $f^n : X \longrightarrow X$ $(n \in \mathbb{N})$ are defined inductively over n by $f^0 = \mathrm{id}_{\operatorname{dom} f}$ and $f^{n+1} = f \circ f^n$.

1.3 Cartesian Products and Projections

Cartesian or direct products are a widespread technique of constructing new structures from given ones. If a collection of structures shares some specific property, then products constructed from these structures may or may not share this property, which is expressed by saying that the property is / is not preserved under the construction of direct products. Both cases will be documented in the different chapters of this module. For the time being, all we want to do is recall the definitions and fix some notations in the context of direct products.

For $K \neq \emptyset$, let $\{A_k \,;\, k \in K\}$ be any collection of sets. Their **direct product** $\prod_{k\in K} A_k$ is defined to be the set of all maps $a : K \longrightarrow \bigcup_{k\in K} A_k$ satisfying $a(k) \in A_k$ for all $k \in K$. $\prod_{k\in K} A_k$ is obviously empty whenever at least one of the sets A_k is empty and the axiom of choice implies the converse to hold as well.

The A_k's are called the **factors** of the direct product $\prod_{k\in K} A_k$. A direct product where all the factors are identical (or, as will be the case in the context of universal algebras, isomorphic) is called a **direct power** and is written as A^I instead of $\prod_{i\in I} A_i$.

The following notational conventions will be useful: Maps in $\prod_{k\in K} A_k$ will be written by listing their values, i.e. as $\langle \ldots, a(k), \ldots \rangle$ or $\langle a_k \,;\, k \in K \rangle$, and referred to as K**-tuples**. If K is finite, $\operatorname{card} K = m$, we use the usual notation for m-tuples $\langle a_1, \ldots, a_m \rangle$, and write also $A_1 \times \ldots \times A_m$ instead of $\prod_{k\in\{1,\ldots,m\}} A_k$. If the length m of an m-tuple is clear from the context, we may simplify $\langle a_1, \ldots, a_m \rangle$ to $\vec{a}$. This is especially handy if the inputs of some fundamental operation f are concerned where the arity is clear; in this case, we will sometimes write $f(\vec{a})$ instead of the more precise $f(a_1, \ldots, a_r)$.

It should be remembered that from a cartesian product we may extract the single factors via the so called **(canonical) projections**, denoted by π_j, i.e.

$$\pi_j \ : \ \prod_{i \in I} A_i \longrightarrow A_j, \quad \pi_j(\langle a_i \ ; \ i \in I\rangle) := a_j$$

If for all $i \in I$ $f_i \ : \ B \longrightarrow A_i$, we will use the "family–notation" $\langle f_i \ ; \ i \in I\rangle$ to denote the function from Y to $\prod_{i \in I} X_i$ defined by

$$\langle f_i \ ; \ i \in I\rangle : \ Y \longrightarrow \prod_{i \in I} X_i \qquad \langle f_i \ ; \ i \in I\rangle \ (y) := \langle f_i(y) \ ; \ i \in I\rangle \ .$$

Thus, in the light of the aforementioned, $\langle \pi_i \ ; \ i \in I\rangle$ is nothing other than $\mathrm{id}_{\prod_{i \in I} X_i}$.

In the case of direct powers, there is something like a counterpart of the canonical projections, the **(canonical) embedding**, mostly denoted by ι. Thus $\iota \ : \ A \longrightarrow A^I$ is defined by $\iota(a) := \langle a \ ; \ i \in I\rangle = \langle \ldots, a, a, a, \ldots\rangle$. The image $\iota[A]$ of A under the embedding ι is a specially denominated subset of the direct product, the **diagonal** Δ_{A^I} **of** A^I:

$$\Delta_{A^I} := \iota(A) = \{a \in A^I \ ; \ \pi_i(a) = \pi_j(a) \text{ for all } i, j \in I\} \ .$$

Do not be confused if in the further developments you will find the symbolization Δ_A: This is just a short form for Δ_{A^2} which is common for the identity equivalence relation on A.

1.4 Equivalence Relations

Equivalence relations (or equivalences for short) occur in many verses of the mathematician's lore. Dividing a collection in parts, thereby respecting certain rules, is found e.g. wherever functions from one set to another are under scrutiny. Belonging to the same part of the partitioning is the same as "being equivalent" under some equivalence.

Since equivalence relations are binary relation on some set, we must first fix the notions concerning binary relations. If X is any set, then a **binary relation** ϑ on X is a set of ordered pairs of elements[2] of X, $\vartheta \subseteq X \times X$. Because binary

[2] In Set Theory, relations are regarded as given by their extensions, so a relation is the same as the set of all pairs that are defined to stand in relation. This characteristic of relations is sometimes called the **extensional view**, as opposed to the intensional view, where relations are identified with their meaning, so that relations actually may be extensionally equal but still intensionally different. In order to distinguish relations intensionally, we must look at their *meaning*, and we are bound to leave the field of Mathematics and get involved with philosophical questions. As an example, consider the set of non–negative reals, $\mathbb{R}_0^+$, and the relations R_1 and R_2 given by xR_1y if and only if $y = x^2$ and xR_2y if and only if $x = \sqrt{y}$. Extensionally, the relations are both equal to the set $\{\langle z, z^2\rangle \ ; \ z \in \mathbb{R}_0^+\}$, so they may not be distinguished from the extensional point of view. But intensionally they differ,

relations are sets, the operations set–union, –intersection etc. can be applied to them.

Infix notation is common practice with binary relations as you may know from earlier experiences with $\leq$ or $\in$. Therefore, we usually write $x\vartheta y$ rather than $\langle x, y\rangle \in \vartheta$.

Definition 1.4.1 A binary relation ϑ on some set X is called an **equivalence relation** iff, for all $x, x_1, x_2, x_3 \in X$, the following three conditions are met:

(i) $x\vartheta x$ (reflexivity)

(ii) $x_1\vartheta x_2 \Longrightarrow x_2\vartheta x_1$ (symmetry)

(iii) $[x_1\vartheta x_2$ and $x_2\vartheta x_3] \Longrightarrow x_1\vartheta x_3$ (transitivity)

Eq X is used to denote the set of all equivalence relations on X.

Please remember that for $\vartheta \subseteq X \times X$, X is called the **carrier** of ϑ.

The notation $x \equiv y \mod \vartheta$ for $x\vartheta y$ is also common for equivalence relations. The set $[x]_\vartheta :=\{y \in X\ ;\ y\vartheta x\}$ is called the ϑ–**equivalence class** or the ϑ–**block** of x. If the actual choice of ϑ is clear from the context, the subscript $_\vartheta$ is sometimes dropped, thus $[x]$ will stand for $[x]_\vartheta$.

The definition of equivalence directly implies that equivalence–classes $[x]$ and $[y]$ (for $x \neq y$) are either disjoint or identical, thus dividing X into parts that have no elements in common. This is what is meant by "$\{[x]\ ;\ x \in X\}$ is a **partitioning** of X", and consequently the equivalence classes are **partitions** of the set X. The set $\{[x]_\vartheta\ ;\ x \in X\}$ is called the **quotient (set) of** X **relative to** ϑ and is denoted by X/ϑ.

The process of assigning to an element $x \in X$ its equivalence class $[x]$ defines a function $\pi_\vartheta : X \longrightarrow X/\vartheta$ which is usually called the **(canonical) projection** or **(canonical) map** (associated to ϑ). It is easy to see that π_ϑ is surjective.

Every equivalence uniquely determines a partitioning of the underlying set X. Interestingly, the other direction works equally well: Any partitioning of X gives rise to a unique equivalence relation by defining to elements equivalent if and only if they belong to the same partition. Partitioning and equivalences are *dual notions.*

Example 1.4.2 To mention but two elementary examples, on any set X, there are a smallest (w.r.t. $\subseteq$) equivalence relation $\{\langle x, x\rangle\ ;\ x \in X\}$ and a largest equivalence relation $X \times X$. It is common practice to denote these by Δ_X and ∇_X respectively.

---

since calculating the square–root is in general more complex than calculating the square, so we might say that R_2 is of higher complexity than R_1. Of course the same remarks apply to functions as well, since functions are nothing but special kinds of binary relations.

A very natural place to look for equivalences is in the context of functions (mappings, morphisms). By declaring to be equivalent all those arguments $x \in X$ which are mapped by a function $f : X \longrightarrow Y$ to the same function–value $y_0 = f(x_0)$, we are defining an equivalence relation, usually called the **kernel** of f, denoted by $\ker(f)$; i.e.

$$\ker(f) = \{\langle x_1, x_2\rangle \in X \times X \;;\; f(x_1) = f(x_2)\}\ .$$

Vice versa, starting from an equivalence relation ϑ, its natural projection is a function whose kernel is exactly ϑ. To sum up:

Remark 1.4.3 Every equivalence relation ϑ gives rise to a function π_ϑ (the **canonical projection** associated to ϑ) defined by $\pi_\vartheta(x) := [x]_\vartheta$, which has exactly ϑ as its kernel.

We will come across a multitude of equivalence relations in the following chapters, but mostly will impose some sort of "structure" in the form of operations or relations upon the carriers, and the equivalences of interest will respect this additional structure, i.e. they will be compatible (to be defined later) with the operations and relations; in this case we will call the equivalence relations *congruences* (cf. Homomorphism Theorem 10.1.5).

1.5 Reduced Products

The constructions of the last two sections combine nicely to a variant of products, the quotients of direct products under equivalence relations.

Definition 1.5.1 If $\langle X_i \;;\; i \in I\rangle$ is a family of sets taking indices in some (non–empty) set I, and if $\sim$ is an equivalence relation on the direct product $\prod_{i\in I} X_i$, then the set $\prod_{i\in I} X_i \;/ \sim = \{[\langle x_i \;;\; i \in I\rangle]_\sim \;;\; \langle x_i \;;\; i \in I\rangle \in \prod_{i\in I} X_i\}$ of equivalence–classes under $\sim$ is called the **reduced product of the family** $\langle X_i \;;\; i \in I\rangle$ **under** $\sim$.

If $X_i = X_j =: X$ for all $i, j \in I$, the reduced product under $\sim$ is called a **reduced power** under $\sim$ and is denoted by $X^I / \sim$.

Exercise 1.5.2 (To test your ability to see the obvious:) Show that, for a family $\langle X_i \;;\; i \in I\rangle$ $(I \neq \emptyset)$ and $\pi_i : \prod_{i\in I} X_i \longrightarrow X_i$ the canonical projection onto the ith component, there is a very natural bijective correspondence between $\prod_{i\in I} X_i \,/ \ker \pi_i$ and X_i.

Of course, both in model theory and in universal algebra, we will be more specific in our choice of equivalence relations. There is a procedure applicable to both fields which involves the notion of a filter over the index set.

Definition 1.5.3 Let S be any non–empty set. A system $\mathcal{F} \subseteq \mathcal{P}(S)$ of subsets of S is called a **filter over** S iff the following conditions are met:

(i) $U_1, U_2 \in \mathcal{F} \Longrightarrow U_1 \cap U_2 \in \mathcal{F}$,

(ii) $U \in \mathcal{F}, U \subseteq V \subseteq S \Longrightarrow V \in \mathcal{F}$, and

(iii) $\emptyset \notin \mathcal{F} \neq \emptyset$.

To put it in mathematical prose, filters are sets of subsets of some set, and they are

(i) closed under super–sets ("upper–closed"),

(ii) closed under intersection,

(iii) a proper, non–empty subset of the power–set.

Examples of filters are easy to find: If $x \in X$, then $\{Y \subseteq X \,;\, x \in Y\}$ is a filter over X; also $\{X\}$ is a filter over X.

Filters give rise to equivalence–relations in a canonical way.

Exercise 1.5.4 Show that for any filter $\mathcal{F}$ over the set I, the relation $\sim_{\mathcal{F}}$ defined on the direct product $\prod_{i \in I} X_i$ by

$$\langle x_i \,;\, i \in I\rangle \sim_{\mathcal{F}} \langle y_i \,;\, i \in I\rangle \ \text{ iff } \ \{i \in I \,;\, x_i = y_i\} \in \mathcal{F}$$

is an equivalence.

For a less trivial example of a filter, let S be an infinite set and $\mathcal{P}_{\text{cof}}S$ be the set of all subsets $Y \subseteq S$ such that $S \setminus Y$ is finite.

Exercise 1.5.5 Show that $\mathcal{P}_{\text{cof}}S$ is a filter on S.

Expressed in terms of $\sim_{\mathcal{P}_{\text{cof}}S}$, we see that $\langle x_s \,;\, s \in S\rangle \sim_{\mathcal{P}_{\text{cof}}S} \langle y_s \,;\, s \in S\rangle$ if and only if $\{s \in S \,;\, x_s = y_s\} \in \mathcal{P}_{\text{cof}}S$, i.e. if and only if $\{s \in S \,;\, x_s \neq y_s\}$ is finite, i.e. if and only if x and y agree on *almost all* components.

Whenever we construct a reduced product over some equivalence $\sim_{\mathcal{F}}$ which stems from a filter in the sense of 1.5.4, we will denote the resulting reduced product by $\prod_{s \in S} X_s \,/\mathcal{F}$ (instead of $\prod_{s \in S} X_s \,/ \sim_{\mathcal{F}}$) and call it the **reduced product under** $\mathcal{F}$. Accordingly, the canonical projection $\pi_{\sim_{\mathcal{F}}}$ will be written as $\pi_{\mathcal{F}}$. For the reduced power, similar notations apply.

Finally, as a counterpart to the canonical projection $\pi_{\mathcal{F}}$ in the case of the reduced power $X^S/\mathcal{F}$, we would like you to check as an exercise that the **canonical embedding** $\iota : X \longrightarrow X^S/\mathcal{F}$ defined by $\iota(x) := [\langle x \, ; \, s \in S \rangle]$ is injective for any filter $\mathcal{F}$.

Chapter 2

Logic Through the Looking Glass

The aim of this chapter is to review the concepts of first–order logic that are indispensable for the understanding of the present paper. For the sake of brevity we will omit all proves and leave out some details of definition, especially where these details do not present one but of many possibilities to describe the desired notion.

2.1 What You Should Remember From First–Order Logic

Dealing with model theory the way we plan to relies on a certain amount of knowledge of the (more or less) basic concepts from first–order logic. Therefore, we suggest that the reader becomes acquainted with the following notions, all of which have been dealt with in detail in Module N4.1:

1. Formal languages for first–order predicate calculus, henceforth simply called formal langages.

2. Structures for formal languages.

3. Terms, formulae, sentences and proofs in formal languages.

4. Satisfaction of a formula in a structure under a valuation.

5. A model of a formula or a set of formulae.

Figure 2.1: Kurt Gödel

6. The term–structure of a formal language, as it is used in the proof of completeness.

The notational conventions, abbreviations and definitions we will follow (unless otherwise stated) are:

- Formal languages (denoted by the calligraphic letter $\mathcal{L}$ with or without subscripts) are characterized by their *non–logical symbols* which consist of constant–, relation– and function–symbols. All other symbols used to build well–formed expressions constituting the syntax are common to all languages. In fact, the actual form of the non–logical symbols does not matter, since a language is fully specified by the *arity–functions* μ for relation– and λ for function–symbols and the index set K for constant–symbols.

- Formal languages are *containing equality as a logical symbol* $\doteq$; thus we should call our formal languages "formal languages for first–order logic with equality". Accordingly, an appropriate set of axioms for the notion of formal deductions is assumed to include axioms for equality (cf. 2.1 below).

- Structures are denoted by calligraphic letters such as $\mathcal{A}, \mathcal{B}, \mathcal{C} \ldots$, formulae by lower case Greek letters $\alpha, \beta, \ldots, \varphi, \vartheta, \ldots$ and sets of formulae by capital Greek letters $\Sigma, \Phi \ldots$.

- Formalisation of first–order logic and its proofs is done by using a Hilbert–style axiom system, which consists of axioms and inference rules. In this system proofs are understood to be finite tupels of formulae obeying certain constraints; the axioms are

- For a formal language $\mathcal{L}$, $\mathrm{Tm}\,\mathcal{L}$ denotes the set of $\mathcal{L}$–terms, $\mathrm{Fml}\,\mathcal{L}$ denotes the set of $\mathcal{L}$–formulae, $\mathrm{Sen}\,\mathcal{L}$ denotes the set of $\mathcal{L}$–sentences, i.e. formulae without free variables.

- Also, $\mathrm{CT}(\mathcal{L})$ denotes the term–structure of $\mathcal{L}$, i.e. the set of $\mathcal{L}$–terms that do not contain any variables. Moreover, if $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$, then $\mathrm{CT}(\mathcal{L})\,/\Sigma$ is used to denote the set of equivalence–classes of $\mathrm{CT}(\mathcal{L})$ under the equivalence relation of *equality provable from* Σ, where two terms $t_1, t_2 \in \mathrm{CT}(\mathcal{L})$ are equivalent iff $\Sigma \vdash t_1 \doteq t_2$.

- Sentences suffice, i.e. results are mostly of interest when formulated for sentences instead of the more general formulae; nevertheless formulae containing free variables will take a central role in proofs and lemmata.

- As is common practice, $\vdash$ denotes (syntactical) provability or deducibility, $\models$ denotes the relation of satisfaction of a formula. ($\models$ will also be used to denote semantical implication, which, by Goedel's Completeness Theorem, is in fact equivalent to the syntactical provability.)

- Consistency is a property of sets of sentences, namely that not everything is provable from a certain set. A set of sentences that is not consistent is called inconsistent.

Equally vital for the understanding of the concepts introduced in this book are the following properties of first–order logic, proved in the very same source:

1. **Compactness:**
   A set Σ of sentences has a model iff every finite subset of Σ has a model.

2. **Completeness:**
   If $\Sigma \vdash \varphi$ and $\mathcal{A} \models \Sigma$, then $\mathcal{A} \models \varphi$

3. **Correctness:**
   If $\Sigma \not\vdash \varphi$, then there is a model of Σ which is not a model of φ

Being a rather fundamental branch of mathematics, Model Theory relies on little or no prerequisites. Nevertheless, there are some basic concepts from Set Theory and algebra, the latter because of the large supply of examples drawn from classic theories of groups, rings or fields. Knowledge of the differences between a proper class and a set, of the notion of cardinalities and concerning the possibilities of axiomatizations of Set Theory are thus of some help, but not really necessary for the understanding of our treatment of Model Theory.

Now for the good news: The reader is NOT assumed to be familiar with any of the concepts of Universal Algebra or Order Theory. Model theoretical constructions will be introduced in just enough detail to leave room for lots of exercises. Since we are not introducing something entirely new to the world of mathematics, but are merely trying to show ways of getting used to some ideas, the emphasis will lie on providing the students with opportunities to get their hands dirty and write their own detailed proofs to deepen the understanding of the concepts formerly unknown to them.

2.2 Syntax

Among the main subjects of discourse in this module are the constructs based on languages: On an intuitive level, a (formal) language (from the syntactical point of view) is a collection of strings formed from a fixed supply of (pairwise distinct!) elements called symbols by a fixed set of rules (*grammar*).

Definition 2.2.1 A **formal language** is defined by a triple $\mathcal{L} = \langle \lambda, \mu, K \rangle$ where, for some sets I and J, $\lambda : I \longrightarrow \mathbb{N}$ and $\mu : J \longrightarrow \mathbb{N}$, and K is some set. I, J and K are the sets of indices of **relation–, function– and constant–symbols** respectively, and λ and μ are the arity functions of the relation– and function–symbols, respectively; i.e. $\lambda(j)$ is the arity (number of arguments) of the relation symbol R_i, while $\mu(j)$ is the arity of the function symbol f_j. Moreover we assume we are given

- countably infinitely many **variables** $v_0, v_1, \ldots$
- an **equality–symbol** $\doteq$
- **logical connectives** $\wedge, \neg, \forall$
- **auxiliary symbols** (brackets)

Definition 2.2.2 The $\mathcal{L}$**–terms** of the formal language $\mathcal{L} = \langle \lambda, \mu, K \rangle$ are defined inductively as follows:

- every variable v_i and, for every $k \in K$, the constant symbol c_k are $\mathcal{L}$–terms
- if $j \in J$ and $t_1, \ldots, t_{\mu(j)}$ are terms, then $f_j(t_1, \ldots, t_{\mu(j)})$ is a $\mathcal{L}$–term

A $\mathcal{L}$–term is called **variable–free** if it does not contain any variables, i.e. iff it is built up using constant– and function–symbols exclusively.

Brackets will be placed to provide unique reading of the terms. Also, freeing ourselves from the chains of the somewhat bulky prefix notation, we will tend to use infix notation where binary function symbols (i.e. of arity 2) are concerned.

Definition 2.2.3 The **$\mathcal{L}$–formulae** of the formal language $\mathcal{L} = \langle \lambda, \mu, K \rangle$ are defined as follows:

- if $t, t', t_1, \ldots, t_{\lambda(i)}$ are terms, then $t \doteq t'$ and $R_i(t_1, \ldots, t_{\lambda(i)})$ are $\mathcal{L}$–formulae (so called **atomic–formulae**)
- if φ and ψ are $\mathcal{L}$–formulae and v_n is a variable, then $\neg\varphi$, $\varphi \wedge \psi$ and $\forall v_n \varphi$ are $\mathcal{L}$–formulae.

(The same notational conventions will be followed that were mentioned for terms.)

A **negatomic $\mathcal{L}$–formula** is a $\mathcal{L}$–formula which is either atomic or the negation $\neg\vartheta$ of an atomic formula ϑ. A variable–free negatomic $\mathcal{L}$–formula is of course a negatomic formula not containing any variables, i.e. negatomic formulae built up from variable–free terms exclusively.

The **scope** of the **(universal) quantifier** $\forall v_n$ in $\forall v_n \varphi$ is φ. An occurrence[1] of a variable v_n is called **free** if it does not lie in the scope of a quantifier $\forall v_n$. An occurrence that is not free is called **bound**. We talk of variables being free or bound in a formula to indicate that there are free or bound occurrences of this variable, respectively. A formula containing no free occurrences of variables is called a **sentence**.

The symbol $\equiv$ is used to denote syntactical equality ("equal as strings") of terms or formulae.

The following abbreviations will be used:

- $\varphi \vee \psi :\equiv \neg(\neg\varphi \wedge \psi)$
- $\varphi \rightarrow \psi :\equiv (\neg\varphi) \vee \psi$
- $\varphi \leftrightarrow \psi : (\varphi \wedge \psi) \vee ((\neg\varphi) \wedge (\neg\psi))$
- $\exists v_n \varphi :\equiv \neg\forall v_n \neg\varphi$

For a formal language $\mathcal{L}$, we let $\mathrm{Tm}\,\mathcal{L}$, $\mathrm{Fml}\,\mathcal{L}$ and $\mathrm{Sen}\,\mathcal{L}$ denote the set of all $\mathcal{L}$–terms, $\mathcal{L}$–formulae and $\mathcal{L}$–sentences, respectively.

Logic deals with proofs and deductions, so next we want to start building well–formed sequences of formulae that we will call (formal) proofs. A formal proof is an object which has to obey certain rules of formation as well, but since the actual way we restrict ourselves by such rules is (1) by no means unique and (2) strictly a matter of syntactic preferences and proof theoretic intentions, an explicit description of a formal system of proof cannot be our goal here. Let us,

[1] A **variable occurrence** is best thought of as the information describing what variable appears at which position inside a string.

for the moment, just fix that we are given a set of formulae called **axioms** and a set of **rules**, where a rule is an "instruction" how to prolong a given proof to a longer one. Then, for any set Σ of formulae (sometimes called **premises**), we define a **(formal) proof** from Σ to be a finite sequence of formulae such that each component of this sequence is either an axiom, an element of Σ or is deducible from components with smaller indices by application of one of the rules. We moreover call a sequence a **proof of φ from Σ** iff it is a proof from Σ and φ is its last component, and we denote this fact by $\Sigma \vdash \varphi$. If there is a proof of φ from the empty set of premises, then we write $\vdash \varphi$ and call φ a **theorem**. (This notion of formal definition of provability is called a **Hilbert–style axiomatization of first–order logic**, thus hinting at the fact that there are other ways of doing this formalizing.)

The actual implementation of this notion of "formal proof" is actually rather arbitrary, although we are to pay attention that completeness and correctness (see below) are complied with, since otherwise there is no sense in even start trying to do model theory. But showing faith in our system of proof being both complete and correct, we may impose some further properties such as recursiveness or recursive enumerability of the set of rules and axioms.

2.3 Semantics

To enable us to do some model theory, we most certainly need to know what a model is. Our first aim is to establish an intimate connection between our notion of proof (formalized in whatever which way) and the notion of satisfaction and truth in some model, to be defined in an instant. We therefore need the definition of the semantical constructs we will deal with, together with the means of interpreting the syntactical elements:

Definition 2.3.1 For a formal language $\mathcal{L} = \langle \lambda, \mu, K \rangle$, a **$\mathcal{L}$–structure** is a family $\mathcal{A} = \langle A, \langle R_i^{\mathcal{A}} \rangle_{i \in I}, \langle f_j^{\mathcal{A}} \rangle_{j \in J}, \langle c_k^{\mathcal{A}} \rangle_{k \in K} \rangle$ where

- $A =: |\mathcal{A}|$ is a non-empty set, the **universe** of $\mathcal{A}$,
- for every $i \in I$, $R_i^{\mathcal{A}} \subseteq A^{\lambda(i)}$
- for every $j \in J$, $f_j^{\mathcal{A}} : A^{\mu(j)} \longrightarrow A$
- for every $k \in K$, $c_k^{\mathcal{A}} \in A$.

$R_i^{\mathcal{A}}$, $f_j^{\mathcal{A}}$ and $c_k^{\mathcal{A}}$ are called the **interpretation** of R_i, f_j and c_k, respectively, in $\mathcal{A}$.

Definition 2.3.2 A **variable assignment function**, or **assignment** or **valuation**, into the $\mathcal{L}$–structure $\mathcal{A}$ is a function h from the set of variables into $|\mathcal{A}|$. If x is a variable and $a \in |\mathcal{A}|$, then the **modified assignment** $h[\binom{v}{a}]$ is defined to be the same assignment as h, except that $h(x) = a$. Clearly, a modified assignment is an assignment.

Definition 2.3.3 For a $\mathcal{L}$–term t, a $\mathcal{L}$–structure $\mathcal{A}$ and a valuation h, the **interpretation** (or **meaning**) $t^{\mathcal{A}}[h]$ of t in $\mathcal{A}$ under h is defined by[2]

$$(c_k)^{\mathcal{A}}[h] := c_k^{\mathcal{A}}, \qquad (v_n)^{\mathcal{A}}[h] := h(v_n)$$

$$(f_j(t_1, \ldots, t_{\mu(j)}))^{\mathcal{A}}[h] := f_j^{\mathcal{A}}(t_1^{\mathcal{A}}[h], \ldots, t_{\mu(j)}^{\mathcal{A}}[h]).$$

Definition 2.3.4 For a $\mathcal{L}$–formula φ, a $\mathcal{L}$–structure $\mathcal{A}$ and a valuation h, we define the relation $\models$ of **satisfaction**, notation $\mathcal{A} \models \varphi[h]$ ("$\mathcal{A}$ satisfies φ under h), by Noetherian induction as follows:

- $\mathcal{A} \models t_1 \doteq t_2[h]$ iff $t_1^{\mathcal{A}}[h] = t_2^{\mathcal{A}}[h]$,
- $\mathcal{A} \models R_i(t_1, \ldots, t_n)[h]$ iff $\langle t_1^{\mathcal{A}}[h], \ldots, t_n^{\mathcal{A}}[h]\rangle \in R_i^{\mathcal{A}}$,
- $\mathcal{A} \models \neg\varphi[h]$ iff not $\mathcal{A} \models \varphi[h]$,
- $\mathcal{A} \models \varphi \wedge \psi[h]$ iff $[\mathcal{A} \models \varphi[h]$ and $\mathcal{A} \models \psi[h]]$,
- $\mathcal{A} \models \forall v_n \varphi$ iff for all $a \in |\mathcal{A}|$, $\mathcal{A} \models \varphi[h\binom{v_n}{a}]$.

If not $\mathcal{A} \models \varphi[h]$, then we write $\mathcal{A} \not\models \varphi[h]$.

The notions of satisfaction, validity in a structure, validity and model generalize to *sets of* formulae in the obvious way, e.g. $\mathcal{A} \models \Sigma$ iff, for all $\varphi \in \Sigma$, $\mathcal{A} \models \varphi$.

Definition 2.3.5 Let φ be a $\mathcal{L}$–formula, $\mathcal{A}$ a $\mathcal{L}$–structure and h a valuation into $\mathcal{A}$. If $\mathcal{A} \models \varphi[h]$ for all valuations h, we say that φ is **valid in** $\mathcal{A}$, or $\mathcal{A}$ **is a model of** φ, notation $\mathcal{A} \models \varphi$. If a $\mathcal{L}$–sentence α is valid in all $\mathcal{L}$–structures, then we call α **valid**, notation $\models \varphi$.

Definition 2.3.6 For a $\mathcal{L}$–sentence α and a set of $\mathcal{L}$–sentences Σ, we say that φ is a **semantic consequence of** or **follows semantically** from Σ or Σ **implies** φ **semantically**, notation[3] $\Sigma \models \alpha$, iff, for all $\mathcal{L}$–structures, $\mathcal{A} \models \varphi$ provided $\mathcal{A} \models \Sigma$.

---

[2] Do not let yourself be confused by some symbols appearing on both sides of the defining equation. This is merely a justification for baptizing the structure's relations, functions and constants the way we did.

[3] Some textbooks use $\Vdash$ instead of $\models$ to distinguish notationally the two notions of satisfaction and semantical implication which, in the light of completeness, turn out to be not that different.

2.4 Completeness

We have thus defined the definitional foundation on which to build the main body of mathematical logic. To show that a notion of proof is suitable for our purposes, we must show that it is accurate enough in the sense that the provability is an exact formalization of the semantical implication. This is the crucial point in the main theorems of any introductory lecture on first–order logic, and it is faced mostly in the form of the so called Completeness Theorem and the Correctness Theorem:

Theorem 2.4.1 (Completeness) Let $\mathcal{L}$ be a formal language, $\Sigma \subset \mathrm{Sen}\,\mathcal{L}$ and $\alpha \in \mathrm{Sen}\,\mathcal{L}$. Then

$$\Sigma \models \alpha \text{ implies } \Sigma \vdash \alpha.$$

Theorem 2.4.2 (Correctness) Let $\mathcal{L}$ be a formal language, $\Sigma \subset \mathrm{Sen}\,\mathcal{L}$ and $\alpha \in \mathrm{Sen}\,\mathcal{L}$. Then

$$\Sigma \vdash \alpha \text{ implies } \Sigma \models \alpha.$$

Since our definition of sentence or even theorem does not mention meaningfulness in the context of contradiction, we might want point out the following distinction:

Definition 2.4.3 $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ is called **inconsistent** or **contradictory** iff $\Sigma \vdash \alpha$ for all $\alpha \in \mathrm{Sen}\,\mathcal{L}$. If Σ is not inconsistent, we call it **consistent**.

There are several equivalent formulations of this. Since the above is the most common, we chose it as a definition, but actually we would prefer the following characterization which could serve as a definition of inconsistency as well:

> A set Σ of sentences is inconsistent iff $\Sigma \vdash \neg v_0 \doteq v_0$ (where v_0 is simply the first of our infinitely many variables).

The point in preferring the latter formulation to most others is that it is less depending on the language than the former. So what we should be doing first thing we learned that there are inconsistent sets is actually to prove that inconsistency does not rely on the language. But, having been taught that first–order logic satisfies completeness, we might argue semantically, and clearly having a model is *not* a matter of language.

Please note the slight shift in formulation after we defined satisfaction: We no longer talk about formulae but restrict our attention to sentences. This is justified by a result that can be found in most lectures and textbooks dealing with mathematical logic and may be summarized with by "sentences suffice", i.e. results about provability and satisfaction remain valid when formulae are replaced by their universal closure, i.e. the formulae prefixed by just enough universal quantifiers to bind all free variable occurrences. Since this is subject to mere syntactical investigation, we restrain from going into more details at this point.

Completeness and correctness attain an even more impressive formulation using the notion of consistency:

Theorem 2.4.4 (Strong Completeness) Let $\mathcal{L}$ be a formal language, $\Sigma \subset \mathrm{Sen}\,\mathcal{L}$. Then

$$\Sigma \text{ has a model iff } \Sigma \text{ is consistent.}$$

One direct consequence of Completeness must be mentioned at this point, since it will be used and, moreover, reformulated semantically in the course of further action: The Compactness Theorem. Compactness should ring a whole Christmas trees full of bells to the reader who has ever dealt with topology, order theory, Set Theory or a somewhat more universal breed of algebra than mere group theory. In math, compactness is usually denoting the possibility of reducing an infinite "something" to a finite part of itself without losing information or characteristics of the original entity. The most famous (or notorious) instantiation of this phenomenon is without reasonable doubt the compactness of topological spaces. Roughly speaking is a topological space called compact if, whenever it can be written as the union of subsets (which are bound to some further constraints, but which we need not concern with here), then it is already equal to the union of a *finite* subset of this collection of sets. Compactness in the context of first–order logic is similar in the overall statement, but contrary to topological compactness, it is a property true for all systems of first–order logic fitting the definition we gave above, thus we do not speak of the compactness of some systems, but we merely formulate the following result:

Theorem 2.4.5 (Compactness) Let $\mathcal{L}$ be a formal language, $\Sigma \subset \mathrm{Sen}\,\mathcal{L}$. Then

$$\Sigma \text{ has a model iff every finite subset of } \Sigma \text{ has a model.}$$

The proof of this theorem, which provides a lot of simplification on the part of consistency arguments for arbitrary sets of sentences, is almost disappointingly simple once we are equipped with completeness: Inconsistency of Σ would mean deduction of something contradictory, and deductions are defined to be *finite* sequences!

2.5 Term Structures

The proves of the theorems of correctness and completeness rely on the actual formalization of "formal prove". Still, certain techniques are bound to be used despite this freedom of choice, and at least one of these techniques even will pop up now and then in model theory as a means of constructing certain minimal models. The reader familiar with correctness proves will most surely feel remembered by parts of the next

Definition 2.5.1 Let $\mathcal{L}$ be a formal language and $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$.

1. If V is a set of variables, $\operatorname{Tm}_V \mathcal{L}$ denotes the set of $\mathcal{L}$–terms t such that all variables in t are contained in V. On the set $\operatorname{Tm}_V \mathcal{L}$, define the equivalence relation $\cong_\Sigma$ by $t_1 \cong_\Sigma t_2$ iff $\Sigma \vdash t_1 \doteq t_2$. Moreover, for $t \in \operatorname{Tm}_V \mathcal{L}$, let $[t]$ denote the $\cong_\Sigma$–equivalence class of t.
2. Define the **term–structure over** V **modulo** Σ to be the structure $\operatorname{Tm}_V \mathcal{L}/\Sigma$ defined by
   - $|\operatorname{Tm}_V \mathcal{L}/\Sigma| := \{[t]\ ;\ t \in \operatorname{Tm}\mathcal{L}\}$
   - $\langle [t_1], \ldots, [t_{\lambda(i)}]\rangle \in R_i^{\operatorname{Tm}_V \mathcal{L}/\Sigma}$ iff $\Sigma \vdash R_i(t_1, \ldots, t_{\lambda(i)})$
   - $f_j^{\operatorname{Tm}_V \mathcal{L}/\Sigma}([t_1], \ldots, [t_{\mu(j)}]) := [f_j(t_1, \ldots, t_{\mu(j)})]$
   - $c_k^{\operatorname{Tm}_V \mathcal{L}/\Sigma} := [c_k]$
3. $\operatorname{CT}(\mathcal{L}) := \operatorname{CT}(\mathcal{L})\,/\emptyset$

If $V = \emptyset$, then $\operatorname{Tm}_V \mathcal{L}$ is the set of all closed $\mathcal{L}$–terms and often denoted by $\operatorname{CT}(\mathcal{L})$. The *closed term structure* $\operatorname{CT}(\mathcal{L})\,/\Sigma$ is in fact of great importance since it constitutes the model constructed in most proofs of the correctness theorem (over a language which is enriched by enough constant symbols to *witness* all existential sentences).

If V is the set of all variables, then of course $\operatorname{Tm}_V \mathcal{L}$ is of course $\operatorname{Tm}\mathcal{L}$.

Clearly, calling $\operatorname{Tm}_V \mathcal{L}/\Sigma$ the term–*structure* is a misuse of terminology and might be slightly misleading, since $\operatorname{Tm}_V \mathcal{L}/\Sigma$ need not be a structure as defined above. But with some additional precautions these obstacles may be circumnavigated. We simply have to take care that there are actually any terms of the desired kind, i.e. we must provide at least one constant or one variable for the term–structure to be a real structure. This proviso lacks importance in the context of the proofs of completeness, since there we are dealing with sets of sentences that, by some preceding constructions, are sentences of a language we expanded by adding constants that actually witness all possible existentially quantified formulae provable from our original set of sentences Σ. We do not expect for the reader to understand every single detail of this rather talkative exposé (or even believe all of it), but we like to point out that the constructions we will encounter in our developments circling the Löwenheim–Skolem–Theorems are close relatives of the term–structures defined above.

Chapter 3

Ordered Sets

In this chapter we present a first approach to the subject of order and partially ordered sets. Being one of the most fundamental notions in mathematics, order is elementary for Model Theory and omnipresent throughout model theoretic considerations. Constructs stemming from ordered sets are used in model theoretic treatments. Ordered sets themselves are used as examples for structures and algebras.

Moreover, ordered sets offer two different aspects, a relational aspect (as sets equipped with a binary relation) and an algebraic aspect (as sets with binary operations, i.e. as *algebras*, cf Section 9.1, Definition 9.1.2). These two aspects interact nicely and provide the opportunity to reformulate results proved for one aspect in the context of the other.

3.1 Order and (Semi-)Lattices

To establish nomenclature and notation, we include the relevant definitions.

Definition 3.1.1 Given any set X, an **order(–relation)** on X is a binary relation ρ on X which is

- reflexive ($x\rho x$ for all $x \in X$),
- antisymmetric ($x\rho y$ and $y\rho x$ together imply $x = y$ for all $x, y \in X$) and
- transitive ($x\rho y$ and $y\rho z$ together imply $x\rho z$ for all $x, y, z \in X$).

The pair $\mathbf{X} := \langle X, \rho \rangle$ is then called an **ordered set** or **poset**, the latter expression reflecting the older nomenclature *partially ordered set*. Orders are customarily denoted by symbols like $\leq$, $\preceq$, $\sqsubseteq$ or similar. X is called the **carrier** or **universe** of $\mathbf{X}$. If the order is clear from the context, we will sometimes not distinguish between X and $\mathbf{X}$.

Definition 3.1.2 Let $\leq$ be an order on the set X.

1. $\leq$ is **total** iff $x \leq y$ or $y \leq x$ for all $x, y \in X$, in which case $\langle X, \leq \rangle$ is called a **totally ordered set** or **chain**.

2. The **dual order** of $\leq$ is the order $\geq$ on X defined by $x \geq y$ if and only if $y \leq x$ (for all $x, y \in X$).

3. The **strict** order associated to $\leq$, usually denoted by the symbol $<$, is defined by $x < y$ if and only if both $x \leq y$ and $x \neq y$.

Let us establish a brief connection to Logic and Model Theory. We notice that in a formal language $\mathcal{L}$ with at least one binary relation–symbol, we can express the conditions for reflexivity, antisymmetry and transitivity by $\mathcal{L}$–sentences α_r, α_a and α_t respectively. Thus, a poset is an $\mathcal{L}$–structure $\mathrm{Str}\, X$ satisfying $\{\alpha_r, \alpha_a, \alpha_t\}$. Clearly, totality is expressible as well, so the class of posets and the class of totally ordered sets are both describable in terms of first–order logic. In later chapters we will denote this fact by calling these classes *elementary* (cf. Chapter 4, Definition 4.1.6).

Example 3.1.3

1. Define the binary relation ρ on $\mathbb{R}$ by $x\rho y$ iff there exists $z \in \mathbb{R}$ such that $x + z^2 = y$. Then ρ is an order on $\mathbb{R}$ – in fact, it is the natural one, $x\rho y$ iff $x \leq y$.

2. Define δ on $\mathbb{N}$ by $x\delta y$ iff there exists $z \in \mathbb{N}$ such that $xz = y$. δ is the **divisibility order** on $\mathbb{N}$, usually denoted by $x|y$.

3. In our context, a very common type of ordered set has its carrier X consisting of certain designated subsets of some set U. Its order is then defined by the subset relation $\subseteq$ in U.

4. (For those with some exposure to Set Theory.) A set α is an ordinal iff every element of α is also a subset of α (i.e. α is a *transitive* set) and the binary relation $\in$ is a total order on α.[1]

[1] Readers familiar with some of the more exotic variants of set theory will notice that we implicitly presupposed the axiom of foundation to hold for our set theoretical universe. This is common practice, which is why we mention it only in this footnote.

5. For any formal language $\mathcal{L}$ we have a natural order on the set of $\mathcal{L}$–expressions, the order of *subexpression*. Specializing to terms and formulae, we might consider sets of terms or formulae to be ordered by the sub–term– or sub–formula–relation, respectively. Even if this example seems to be a bit farfetched, later definitions and results will justify its mention here.

Definition 3.1.4 Let S be a subset of an ordered set $\langle X, \leq \rangle$.

1. $s \in S$ is called **greatest element** of S if $x \leq s$ for all $x \in S$;
   $s \in S$ is called **least element** of S if $s \leq x$ for all $x \in S$.

2. $s \in S$ is called a **maximum** of S if, for all $x \in S$, $s \leq x$ implies $s = x$;
   $s \in S$ is called a **minimum** of S if, for all $x \in S$, $x \leq s$ implies $s = x$.

Clearly, by antisymmetry, greatest and least elements of S are unique, provided they exist. Any greatest element of S is a maximum of S, and any least element of S is a minimum of S. Maxima and minima need not exist for a given subset, and even if they exist, they need not be unique. (Exercise: Prove these statements!)

Example 3.1.5

1. $\mathbb{N}$ with the natural order $\leq$ has a least element as does any subset of $\mathbb{N}$. But greatest elements and maxima only exist for finite, non–empty subsets of $\mathbb{N}$.

2. Consider some non–empty set X with ρ being the *diagonal* of X, i.e. $\rho = \{\langle x, x \rangle \,;\, x \in X\}$. Then, ρ is indeed an order on X. Any element $x \in X$ is both a minimum and a maximum for any subset $S \subseteq X$ with $x \in S$, but only singleton subsets $\{x\}$ have a greatest and a least element.

Later experiences will teach us that not all notions of interest are expressible in first–order logic, not even in such a (seemingly!) simple theoretical context as the one of ordered sets. For instance, the statement "every subset has a supremum" exceeds the limits of first–order logic. Another such notion is given in the following definition.

Definition 3.1.6 An order $\leq$ on some set X is called **noetherian** or **well–founded** if, for every non–empty subset $S \subseteq X$, there exists a minimal element of S. We then say that $\langle X, \leq \rangle$ is a **noetherian order** or a **noetherian poset**.

Figure 3.1: Emmy Noether (1882-1935)

In structural induction, we made use of a fundamental property of noetherian orders before we even knew about them. The next proposition is a justification in retrospect.

Proposition 3.1.7 (Noetherian induction) Let $\langle X, \leq \rangle$ be a noetherian poset. Assume $S \subseteq X$ is non–empty and for all $x \in X$,

$$[\text{for all } y,\ \text{if } y \leq x \text{ and } y \neq x, \text{ then } y \in S] \text{ implies } x \in S.$$

Then $S = X$.

Proof. By way of contradiction, assume $\langle X, \leq \rangle$ is noetherian and $S \subseteq X$ satisfies

$$[\text{for all } y,\ \text{if } y \leq x \text{ and } y \neq x, \text{ then } y \in S] \text{ implies } x \in S, \qquad (*)$$

but $S \neq X$. Then $X \smallsetminus S \neq \emptyset$, i.e. we find a minimal element $x_0 \in X \smallsetminus S$. Then, for any y with $y \leq x_0$ and $y \neq x_0$ we have $y \in S$ by minimality, but this implies $x_0 \in S$ by $(*)$, contradicting the choice of x_0. ∎

Natural induction is a special form of noetherian induction; structural induction for terms or formulae of formal languages is another example. If you know your way through Set Theory and ordinals, you will be familiar with the fact that every ordinal is noetherian (ordered by $\in$), and if we assume the Axiom of Foundation to hold, every set is noetherian.

Using structural or natural induction as a tool to prove statements is only valid if the respective order (sub–formula, sub–term, $\leq$ etc.) is noetherian. For the set $\mathbb{N}$ of natural numbers and $\leq$ this is well-established, and for constructs in the context of formal languages, we must rely on the *finiteness* of the expressions.

Let us have a look at some notions which are vital for the algebraization of orders (cf. Chapter 11).

Definition 3.1.8 In a poset $\langle X, \leq \rangle$, an **upper bound** of $S \subseteq X$ is an element $u \in X$ such that $s \leq u$ for all $s \in S$. A **least upper bound** of S is an upper bound u of S satisfying $u \leq v$ for every upper bound v of S. It is clear by the antisymmetry of $\leq$ that least upper bounds are unique whenever they exist (the proof is left as an exercise). A least upper bound of S is also called the **supremum** of S and written Sup S (or $\mathrm{Sup}_{\leq} S$ to emphasize the order relation). If Sup X exists, it is called the **greatest element** or **top** of $\mathbf{X}$ and is sometimes written $\top_{\mathbf{X}}$.
Analogically, a **lower bound** of S (in X) is an element $l \in X$ such that $l \leq s$ for all $s \in S$. The notions of a **greatest lower bound**, of S alias **infimum** of S, and a **least element** of X, alias **bottom** of X, are defined analogically with notations Inf S (or $\mathrm{Inf}_{\leq} S$) and $\perp_{\mathbf{X}}$.

Note that the special cases Sup $\emptyset$ and Inf $\emptyset$ coincide with $\perp_{\mathbf{X}}$ and $\top_{\mathbf{X}}$ respectively, provided they exist. (Exercise: why?)

Definition 3.1.9 For $\mathbf{X} = \langle X, \leq \rangle$ and $u, v \in X$ we define the **interval** $[u, v]$ to be the set $\{x \in X\ ;\ u \leq x \leq v\}$ (which is empty unless $u \leq v$); similarly, the **lower end** determined by u is $(u] := \{x \in X\ ;\ x \leq u\}$ and correspondingly the **upper end** is $[v) := \{x \in X\ ;\ v \leq x\}$. We say that u is a **lower cover** of v (and v an **upper cover** of u) iff $[u, v] = \{u, v\}$; this situation is frequently denoted by $u \prec v$ or $v \succ u$. An upper cover of $\perp_{\mathbf{X}}$ is called an **atom** of $\mathbf{X}$, similarly, a **coatom** is a lower cover of $\top_{\mathbf{X}}$. Elements $u, v \in X$ are said to be **comparable** iff $u \leq v$ or $v \leq u$, **incomparable** otherwise, the latter situation being denoted by $\|$. A subset $S \subseteq X$ such that $u \| v$ for all $u, v \in S$ is called an **antichain**.

Example 3.1.10

1. In Example 3.1.3 1 the set $\{r \in \mathbb{R};\ 2\rho r^2\}$ has no infimum while Inf $V = 0$ for $V = \{1/n; n \in \mathbb{N}\}$. There are no covers in $(\mathbb{R}, \rho)$, and no incomparables since $(\mathbb{R}, \rho)$ is a chain.

2. In Example 3.1.3 2 we have $\perp_{\mathbb{N}} = 1$ and $\top_{\mathbb{N}} = 0$. The atoms of $(\mathbb{N}, \delta)$ are precisely the prime numbers (which form an antichain), while there are no coatoms.

Exercise 3.1.11
Prove the statements in Example 3.1.10; especially, show that there are no covers in $(\mathbb{R}, \rho)$, Inf $\{1/n; n \in \mathbb{N}\} = 0$ and that in Example 2, the atoms are exactly the prime numbers while there are no coatoms.

Ordered sets are *relational structures*, i.e. they are sets endowed with some fundamental relations, as opposed to algebras which are sets endowed with some fundamental operations. Under certain circumstances however, an order relation may be described by a (binary) operation on the same carrier set, and vice versa. The key fact is that the supremum of any two elements of an ordered set is uniquely determined whenever it exists, and so is the infimum. We capture this situation in the following definition.

Definition 3.1.12 An ordered set $\mathbf{S} = \langle S, \leq \rangle$ is called a Sup –**semilattice** iff $\text{Sup}_{\leq}\{x, y\}$ exists for all $x, y \in S$; it is called an Inf –**semilattice** iff $\text{Inf}_{\leq}\{x, y\}$ exists for all $x, y \in S$. An ordered set that is both a Sup –semilattice and an Inf –semilattice is called a **lattice**.

Note that in a Sup –semilattice $\mathbf{S}$ the supremum of any *finite, non–empty* subset $U \subseteq S$ exists (why?), and so does the infimum of any finite, non–empty subset in an Inf –semilattice.

Example 3.1.13 Let $U \neq \emptyset$ be any set and pick a nonempty proper subset U_0 of U. Define $S := \{Z \subseteq U\ ;\ U_0 \not\subseteq Z\}$. The ordered set $\langle S, \subseteq \rangle$ is an Inf –semilattice but not a Sup –semilattice, and we get $\text{Inf}\,\{Z_1, Z_2\} = Z_1 \cap Z_2$. Under the dual order ($Z_1 \leq Z_2$ iff $Z_1 \supseteq Z_2$) S will be a Sup –semilattice but not an Inf –semilattice. (Exercise: What is the corresponding semilattice operation in this case?)

Clearly, suprema and infima are closely connected, and thus, in an ordered set where they both exist, we observe the following properties.

Lemma 3.1.14 For all x, y,

$$\begin{aligned} \text{Sup}\,\{x, \text{Inf}\,\{x, y\}\} &= x; \text{ and} \\ \text{Inf}\,\{x, \text{Sup}\,\{x, y\}\} &= x. \end{aligned}$$

Proof. Exercise. ∎

Example 3.1.15

1. For the divisibility order δ on $\mathbb{N}$ (Example 3.1.3 1), Sup and Inf of any finite subset of $\mathbb{N}$ exist, since $\text{Inf}\,\{m, n\}$ =g.c.d.[2] of m and n resp. $\text{Sup}\,\{m, n\}$ = l.c.m.[3] of m and n. (Exercise: Verify this!)

2. (cf. Example 3.1.3 2) A special type of lattice is given by a collection $\mathbf{L}$ of subsets of some set X such that $U, V \in \mathcal{L}$ implies that $U \cap V$ and $U \cup V$ both are in $\mathcal{L}$. The suprema and infima in the lattice resulting

[2] g.c.d. of m and n: greatest common divisor of m and n, i.e. the greatest number d such that both m/d and n/d are integers.

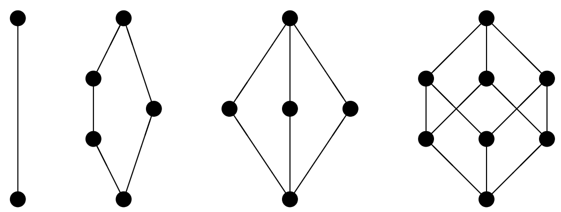
[3] l.c.m. of m and n: lowest common multiple of m and n, i.e. the lowest number $d > 0$ such that both d/m and d/n are integers.

from ordering **L** by set–inclusion coincide, in this case, with set union and intersection.

3. (See also 9.3.4.) To prevent you from forming the impression that set–union and –intersection are the only examples of lattice–suprema and –infima, consider the following examples: If X is a group (ring, vector space, topological space, boolean algebra), then the collection of sub–groups (sub–rings, linear sub–spaces, closed sub–spaces, sub–algebras) ordered by $\subseteq$ is a lattice. The infimum corresponds with set–intersection, again, whereas the supremum is a little bit more complicated than mere set–union, since e.g. the union of two subgroups need not again be a sub–group; rather can Sup of two sub–groups be shown to be exactly the smallest sub–group containing the two sub–groups. If we look at this from another angle this is exactly the sub–group generated by (sub–ring generated by, linear hull of, topological closure of, sub–algebra generated by) the set–union of the sub–groups (sub–rings, linear sub–spaces, closed sub–spaces, sub–algebras).

A important tool for working with orders are **(Hasse) diagrams**, especially for finite orders or finite parts of arbitrary orders. Given such an order P, its diagram consists of of small circles in the plane, representing the elements of P, and straight line segments, representing the covering relation in P. More precisely, the circles representing $u, v \in P$ are joined by a line segment exactly if $u \prec v$, and in this case the circle representing v must be strictly above the circle representing u (in the natural orientation of the plane). Moreover, no other circles are incident with this line segment.

Example 3.1.16 Here are some examples:



2 $\qquad$ $\mathbf{N_5}$ $\qquad$ $\mathbf{M_3}$ $\qquad$ $\mathbf{B}_3$

2 is the smallest non–trivial ordered set. It has 2 elements and clearly is a lattice (as are the other three examples depicted).$\mathbf{N_5}$ and $\mathbf{M_3}$ are lattices which will be of special interest in the context of modularity and distributivity (cf. Section 11.2). Finally, $\mathbf{B}_3$ is also known as the *boolean algebra with* 3 *atoms*, which hints at the fact that it is another kind of *algebraic structure*; simple calculations show that it is also a lattice.

3.2 Complete Lattices and Closure Operators

Completeness is another important property lattices sometimes have. It is not an algebraic property in the sense that, although it is formalized in the language of the lattice's order–relation, it is not expressible using the algebraic aspect (Sup and Inf) of the lattices. Moreover, completeness is not expressible by means of first–order logic.

Definition 3.2.1 A lattice $\mathbf{L} = \langle L, \leq \rangle$ is **complete** iff $\mathrm{Sup}_{\leq} A$ and $\mathrm{Inf}_{\leq} A$ exist in L for *any* subset $A \subseteq L$.

In particular, every complete lattice L has a least element $\bot_{\mathbf{L}}$ and a greatest element $\top_{\mathbf{L}}$. (Exercise: Why?)

Example 3.2.2

1. Every finite lattice is a complete lattice.

2. While the set $\mathbb{R}$ of real numbers (with the usual order) is a lattice, it is not a complete lattice. However, the *extended reals* $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ with $-\infty < x < +\infty$ for all $x \in \mathbb{R}$ and the natural order within $\mathbb{R}$ constitute a complete lattice.

3. The power set $\mathcal{P}(X)$ of any set X is easily seen to be a complete lattice under the order of set–inclusion. (Exercise: Prove this by finding the appropriate Sup and Inf .)

4. (See Example 3.1.31) $\mathbb{N}$ with the order of divisibility is a complete lattice. Again, as an exercise, find Sup and Inf appropriate for *any* subset.

In spite of being defined in terms of order exclusively, completeness is preserved by lattice isomorphisms, thus it is a *property of algebras* in the sense that preserving the *algebraic* structure of a lattice ensures the preservance of completeness as well.

The task of recognizing complete lattices is cut in half by the following result.

Proposition 3.2.3 For an order $\mathbf{L} = \langle L, \leq \rangle$, thwe following are equivalent:

(i) L is a complete lattice;

(ii) for any subset $S \subseteq L$, $\mathrm{Inf}_{\leq} S$ exists in L;

(iii) for any subset $S \subseteq L$, $\mathrm{Sup}_{\leq} S$ exists in L.

Proof. Clearly, any complete lattice satisfies (ii) and (iii). Conversely, an ordered set satisfying (ii) and (iii) is clearly a complete lattice.

Now assume (ii) and take any $S \subseteq L$. Let $U := \{u \in L\, ;\, u \geq s$ for all $s \in S\}$. Then $\operatorname{Inf} U = \operatorname{Sup} S$ in L, so the arbitrary subset S has a supremum in L and therefore **L** satisfies (iii). Hence (ii) implies (iii), hence by the above observation (ii) implies (i).

The dual argument works to show that (iii) implies (i). ■

Complete lattices arise in many important situations as special collections of subsets of some base set X. This fact motivates the following definition.

Definition 3.2.4 A collection $\mathcal{C} \subseteq \mathcal{P}(X)$ of subsets of a set X is a **closure system** (on X) iff $\bigcap \mathcal{S} \in \mathcal{C}$ for any subcollection $\mathcal{S} \subseteq \mathcal{C}$.

Note that $X \in \mathcal{C}$ for any closure system $\mathcal{C}$ on X since X coincides with the set intersection of the *empty* subcollection of $\mathcal{C}$. As an immediate consequence of Proposition 3.2.3, every closure system $\mathcal{C}$ is a complete lattice under set inclusion as order relation. The infimum in $\langle \mathcal{C}, \subseteq \rangle$ coincides with set–intersection, but the supremum is generally different from set–union, e.g. $\operatorname{Sup}\{C_1, C_2\}$ in $\mathcal{C}$ is given by $\bigcap \{C \in \mathcal{C}\, ;\, C_1 \cup C_2 \subseteq C\}$ which may *properly* contain $C_1 \cup C_2$ as a subset.

Examples of closure systems can be found throughout the whole field of mathematics, e.g. wherever we consider sub–constructs (sub–groups, sub–rings, closed sub–spaces, linear sub–spaces, sub–algebras etc.) under the aspect of finding the smallest sub–construct which contains a given subset of the whole.

The next definition provides an alternative way to describe closure systems.

Definition 3.2.5 Let X be any set. A map $C : \mathcal{P}(X) \longrightarrow \mathcal{P}(X)$ is a **closure operator** on X iff for any $U, V \subseteq X$ we have

(i) $U \subseteq C(U)$ (C is *extensive*),

(ii) $C(C(U)) = C(U)$ (C is *idempotent*) and

(iii) $U \subseteq V$ implies $C(U) \subseteq C(V)$ (C is *monotonic*).

$\operatorname{span} X$, the linear sub–space spanned (generated) by some subset X of some vector space V could serve as a natural example.

While closure *systems* emphasize the static aspect (closure systems considered as the collection of the *final states* of some process), closure *operators* are the dynamical counterpart (the process itself). As we shall see, they present different sides of the same coin:
Given a closure system $\mathcal{C}$ on X, define the function $C_{\mathcal{C}} : \mathcal{P}(X) \longrightarrow \mathcal{P}(X)$ by

$$C_{\mathcal{C}}(U) := \bigcap \{A \in \mathcal{C}\, ;\, U \subseteq A\}$$

for any $U \subseteq X$. Conversely, given a closure operator C, define

$$\mathcal{C}_C := \{A \subseteq X \; ; \; A = C(A)\} \; .$$

It is not difficult to show that $C_\mathcal{C}$ is a closure operator and that $\mathcal{C}_C$ is a closure system; moreover

$$C_{\mathcal{C}_C} = C \text{ and } \mathcal{C}_{C_\mathcal{C}} = \mathcal{C}.$$

(Exercise: Work out detailed proofs for these statements!).

Although *complete* lattices from the point of view of their definition do not differ much from lattices, the omission of the finiteness–condition for subsets for which suprema and infima are to exist is a fundamental difference from the point of view of first–order logic. There is no set Σ of $\mathcal{L}$–sentences — for an appropriate language $\mathcal{L}$ — such that an $\mathcal{L}$–structure is a complete lattice if and only if it is a model of Σ. However, for the moment we are not in the position to prove this, since we would use semantical arguments not yet introduced. The complete proof will be given in Proposition 11.4.3.

Chapter 4

First Steps in Model Theory

In this chapter, we provide a overview of the basic concepts you are bound to come across when studying Model Theory.

4.1 Introducing Mod and Th

This section is intended to provide some introductory remarks and definitions that will be central in any aspect of Model Theory as it will be presented in the following chapters.

The starting point of model theory, if there is any, is best located in the vicinity of the two operators Mod and Th. Since model theory basically concerns algebraic properties common to classes of models of formal theories, the transition from syntax to semantics has to be a fixed and smooth one. The notion of satisfaction or validity in a model is too detailed and clumsy for such a purpose. There must be ways of describing satisfaction in the context of whole sets of sentences and classes of models. This leads to the definition of the two class–operators Mod and Th. Let us begin by fixing some nomenclature intended to make our lives a lot easier.

Definition 4.1.1 For a formal language $\mathcal{L}$, define the class $\mathrm{Str}\mathcal{L}$ by

$$\mathrm{Str}\mathcal{L} := \{\mathcal{A}; \mathcal{A} \text{ is an } \mathcal{L}\text{-structure}\}\,.$$

Note that $\mathrm{Str}\mathcal{L}$ is a *proper class*[1] and may be regarded as the semantical

[1] Thus generally, $\mathrm{Str}\,\mathcal{L}$ is not a set. This exemplifies a phenomenon occurring throughout mathematics, when dealing with the structure of sets is the focus of concern and not the

counterpart to Sen $\mathcal{L}$.

Working with sets of sentences and classes of models, questions such as the following arise:

- Does a given set of sentences describe a class of structures that shows nice semantical / categorical properties?
- Does a given class of models have a nice description in a language of first–order logic?

Thus, what we are looking for are semantical / syntactical counterparts of notions defined strictly in the syntactical / semantical realm respectively. We are therefore in need of some means to interconnect the two realms of syntax and semantics, a task that is now fulfilled using the two operators Mod and Th.

Note that the following definition consists of mutually dual parts:

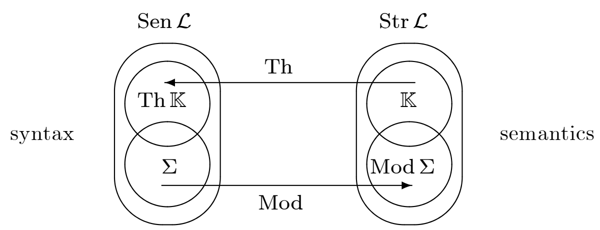
Definition 4.1.2 Let $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$ (a sub*class*!) and $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$.

(i) $\mathrm{Th}(\mathbb{K}) := \{\sigma \in \mathrm{Sen}\,\mathcal{L}\,;\, \mathcal{A} \models \sigma$ for every $\mathcal{A} \in \mathbb{K}\}$ ("theory of $\mathbb{K}$")

(ii) $\mathrm{Mod}(\Sigma) := \{\mathcal{A} \in \mathrm{Str}\,\mathcal{L}\,;\, \mathcal{A} \models \sigma$ for every $\sigma \in \Sigma\}$ ("model class of Σ")

Thus, Th assigns to a class $\mathbb{K}$ the set of sentences valid in all structures in $\mathbb{K}$, while Mod assigns to a set of sentences the class of structures in which all the sentences in Σ are valid. We note that:[2]

$$\begin{array}{rccc} \mathrm{Th}: & \mathcal{P}(\mathrm{Str}\,\mathcal{L}) & \to & \mathcal{P}(\mathrm{Sen}\,\mathcal{L}) \\ \mathrm{Mod}: & \mathcal{P}(\mathrm{Sen}\,\mathcal{L}) & \to & \mathcal{P}(\mathrm{Str}\,\mathcal{L}) \end{array}$$



To keep notation legible and bracketing reasonable we simply write $\mathrm{Th}\,\mathbb{K}$ and $\mathrm{Mod}\,\Sigma$ instead of the more accurate $\mathrm{Th}(\mathbb{K})$ and $\mathrm{Mod}(\Sigma)$.

Exercise 4.1.3 Prove the following equivalences:

$$\begin{array}{rcl} \mathcal{A} \in \mathrm{Mod}\,\Sigma & \text{iff} & \mathcal{A} \models \Sigma \\ \sigma \in \mathrm{Th}\{\mathcal{A}\} & \text{iff} & \mathcal{A} \models \sigma \end{array}$$

elements of the sets. The lack of information on the side of the elements allows in principle to construct the class of all sets as a derivative of the class of all $\mathcal{L}$–structure, thus the latter being a set would imply an antinomy along the line of *Russell's Paradox*; see also Section **??**.

[2] NB: (Th, Mod) is what algebraists call a *Galois–connection*.

There are several simplifications you are bound to come across in the context of Mod and Th: First, brackets and the symbol $\circ$ for the operational product are often dropped. Thus, for $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$, $\alpha \in \mathrm{Sen}\,\mathcal{L}$, $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$ and $\mathcal{A} \in \mathrm{Str}\,\mathcal{L}$, we regard expressions like

$$\mathrm{Th}\,\mathrm{Mod}\,\mathbb{K}, \quad \mathrm{Th}\,\mathrm{Mod}\,\alpha, \quad \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}, \quad \mathrm{Mod}\,\mathrm{Th}\,\mathcal{A}$$

as equally meaningful as their clumsy, full–blown equivalents

$$\mathrm{Th} \circ \mathrm{Mod}(\mathbb{K}), \quad \mathrm{Th} \circ \mathrm{Mod}(\{\alpha\}), \quad \mathrm{Mod} \circ \mathrm{Th}(\mathbb{K}), \quad \mathrm{Mod} \circ \mathrm{Th}(\{\mathcal{A}\}),$$

although the latter notations are the more exact. Since unique reading is ensured with the shortened and more readable forms, we will stick to these in further developments.

Lemma 4.1.4 For $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ and $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$,

(i) $\Sigma \subseteq \mathrm{Th}\,\mathrm{Mod}\,\Sigma$ and $\mathbb{K} \subseteq \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$.

(ii) Mod and Th are **antimonotonic**, i.e.
If $\Sigma_1 \subseteq \Sigma_2$, then $\mathrm{Mod}\,\Sigma_1 \supseteq \mathrm{Mod}\,\Sigma_2$.
If $\mathbb{K}_1 \subseteq \mathbb{K}_2$, then $\mathrm{Th}\,\mathbb{K}_1 \supseteq \mathrm{Th}\,\mathbb{K}_2$.

(iii) $\mathrm{Th}\,\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K} = \mathrm{Th}\,\mathbb{K}$ and $\mathrm{Mod}\,\mathrm{Th}\,\mathrm{Mod}\,\Sigma = \mathrm{Mod}\,\Sigma$.

Proof.

(i) $\sigma \in \Sigma$ implies $\mathcal{A} \models \sigma$ for every $\mathcal{A} \in \mathrm{Mod}\,\Sigma$, thus $\sigma \in \mathrm{Th}\,\mathrm{Mod}\,\Sigma$. Similarly for $\mathbb{K}$.

(ii) $\mathcal{A} \in \mathrm{Mod}\,\Sigma_2$ implies $\mathcal{A} \models \Sigma_2$, thus, since $\Sigma_1 \subseteq \Sigma_2$, we have $\mathcal{A} \models \Sigma_1$, i.e. $\mathcal{A} \in \mathrm{Mod}\,\Sigma_1$. Similarly for $\mathbb{K}_1, \mathbb{K}_2$.

(iii) $\mathbb{K} \subseteq \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$ by (i). Thus $\mathrm{Th}\,\mathbb{K} \supseteq \mathrm{Th}\,\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$, using (ii). On the other hand, using (i) we have $\mathrm{Th}\,\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K} \supseteq \mathrm{Th}\,\mathbb{K}$. Putting it all together, we have $\mathrm{Th}\,\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K} = \mathrm{Th}\,\mathbb{K}$. Similarly for the second claim.

■

Exercise 4.1.5 Bridge the gaps in the proof to Lemma 4.1.4.

Lemma 4.1.4 implies that both Mod Th and Th Mod are closure operators (cf. Definition 3.2.5). (To find a proof for this makes another exercise.) As always in the context of closure operators, the "closed entities" deserve some special attention:

Definition 4.1.6 Let $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ and $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$. We call

(i) Σ a **theory** iff $\Sigma = \mathrm{Th}\,\mathrm{Mod}\,\Sigma$.

(ii) $\mathbb{K}$ a **elementary class** iff $\mathbb{K} = \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$.

The use of "elementary" instead of "elementary class" is common practice.

Considering the products Th Mod and Mod Th as operators on $\mathrm{Sen}\,\mathcal{L}$ and $\mathrm{Str}\,\mathcal{L}$ respectively, we see that theories and elementary classes are the respective fixed points of these operators.

Exercise 4.1.7 What about $\emptyset$ and $\mathrm{Str}\,\mathcal{L}$? Are those elementary classes? And what about $\emptyset$ and $\mathrm{Sen}\,\mathcal{L}$ as sets of $\mathcal{L}$–sentences? Are they theories?

To subsequently motivate these definitions, we observe that being a theory means being exactly the set of sentences that hold in all structures of some appropriate class of $\mathcal{L}$–structures, while being elementary is equivalent to being the class of all models of some set of sentences:

Corollary 4.1.8 For $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ and $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$,

(i) Σ is a theory iff $\Sigma = \mathrm{Th}\,\mathbb{K}$ for some $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$.

(ii) $\mathbb{K}$ is elementary iff $\mathbb{K} = \mathrm{Mod}\,\Sigma$ for some $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$.

Exercise 4.1.9 Prove corollary 4.1.8.

Some authors use the word **axiomatizable** instead of elementary to emphasize the fact that, for an elementary class $\mathbb{K}$, there is a set Σ of sentences axiomatizing exactly this class $\mathbb{K}$. The elements of Σ are called **axioms**, not to be confused with the axioms of first–order logic upon which our notion of formal deduction is based.

Clearly, elementary classes are a widespread phenomenon and the reader will have come across a few examples, maybe even without being aware of it, e.g. the class of all sets, the class of all groups, rings, fields etc. By providing the respective axioms we implicitely prove (or are being told) that these classes are indeed elementary. (Exercise: What would be axioms appropriate for the class of all set?)

The task of providing a set of axioms for some given class is, from the point of view of Model Theory, secondary to the task of finding means of characterizing classes as being or failing to be elementary *without falling back to syntactic notions*. Thus we will, in later sections, successfully hunt for algebraic notions to tell elementary classes from the ones which are not. Since we are mainly dealing with model–theoretic aspects, we therefore prefer the terminology "elementary" to "axiomatizable". Nevertheless, calling non–elementary classes non–axiomatizable puts the focus on the fact that such classes evade description by the means of first–order logic.

For the time being, given $\mathcal{L}$–structures $\mathcal{A}$ and $\mathcal{B}$, the only way to find distinguishing properties of these two structures is via the formal language $\mathcal{L}$, i.e. some formulae that are satisfied in only one of the two structures. If, by means of the language, the two structures cannot be told from each other, we will say that they are elementary equivalent:

Definition 4.1.10 $\mathcal{L}$–structures $\mathcal{A}$ and $\mathcal{B}$ are called **elementary equivalent** iff $\mathrm{Th}\,\mathcal{A} = \mathrm{Th}\,\mathcal{B}$; notation $\mathcal{A} \equiv \mathcal{B}$.

The situation compares to other fields of math, e.g. algebra where categorizing groups is only of interest up to the level of isomorphisms, i.e. isomorphic groups are regarded as identical. In the context of first–order logic, the resolution is as grainy as the notion of elementary equivalence, so elementary equivalent structures are one and the same where satisfaction of formulae is concerned, since they satisfy the same $\mathcal{L}$–sentences.

The following observation is almost too obvious to mention:

Lemma 4.1.11 If $\mathcal{A}, \mathcal{B} \in \mathrm{Str}\,\mathcal{L}$, then

$$\mathcal{A} \equiv \mathcal{B} \text{ iff } [\mathcal{A} \models \varphi \text{ iff } \mathcal{B} \models \varphi \text{ for all } \varphi \in \mathrm{Sen}\,\mathcal{L}].$$

The proof is left as an exercise.

Elementary equivalence is "language depending", as is exemplified by the fact that $\mathbb{Q}$ and $\mathbb{R}$ are elementary equivalent as ordered sets, but not as fields. (A proof for this is yet beyond our facilities; the necessary tools will be provided below.) Still there is no risk of confusion since, whenever necessary, we will always make clear which language is considered

The operators Mod and Th are mutually dual "interfaces" between syntax and semantics. Notions defined for one realm are "translated" via these operators to the other realm. E.g. elementary equivalence is clearly rooted in syntax, and we may wonder in what form it is expressed using semantic notions. We will come back to this in Section 8.3.

For a class of $\mathcal{L}$–structures being an elementary class means being **axiomatizable**, i.e. being definable by a set of $\mathcal{L}$–formulae. Since looking for an appropriate set of formulae may take forever, we start looking for other criteria to decide whether we need not even start the search in the first place. Thereby we try to stay in the semantical realm rather than using syntax. This will be the subject of Chapters 7 and 8.

The following is obvious:

Lemma 4.1.12 Every elementary class $\mathbb{K}$ is **closed under elementary equivalence**, i.e.

$$\text{If } \mathcal{A} \in \mathbb{K} \text{ and } \mathcal{B} \equiv \mathcal{A}, \text{ then } \mathcal{B} \in \mathbb{K}.$$

We leave the proof as an easy exercise to get used to dealing with the notions involved.

Note that the converse of Lemma 4.1.12 is not true, as we shall later see examples of classes of $\mathcal{L}$–structures that are not elementary, but are still closed under $\equiv$(cf. Examples 8.1.7 and 8.1.9) .

4.2 Expanding and Restricting Languages

Expanding languages is a rather simple concept that poses no real problems of understanding, yet it is of importance to model theoretic constructions. In principle, we could expand a language $\mathcal{L}$ to another language $\overline{\mathcal{L}}$ by adding new relation–, function– or constant–symbols (or some of all, for that). So (we will not elaborate a detailed definition here since we are lacking a precise definition of a formal language to begin with) we will say that $\overline{\mathcal{L}}$ **expands** $\mathcal{L}$ if $\overline{\mathcal{L}}$ contains all non–logical symbols of $\mathcal{L}$ (plus some more, eventually). Consequently, $\mathcal{L}$ is called a **sub–language** of $\overline{\mathcal{L}}$.

In practice, the only kind of expansion of a language we will regularly use is expansion by constants. Thus an expansion $\overline{\mathcal{L}}$ of $\mathcal{L}$ and $\mathcal{L}$ itself will comprise the same relation– and function–symbols, but $\overline{\mathcal{L}}$ contains constant symbols that are not present in $\mathcal{L}$. Since our main usage of languages is to formalize structures, you could say that an expansion by constant–symbols provides the larger vocabulary by having names for elements that were "nameless" in the sub–language.

Expansion of a language carries over to most of the syntactical concepts in the following sense: Suppose $\overline{\mathcal{L}}$ expands $\mathcal{L}$. Some tedious but instructing experiences in Noetherian Induction show that

- $\mathrm{Tm}\,\mathcal{L} \subseteq \mathrm{Tm}\,\overline{\mathcal{L}}$
- $\mathrm{Fml}\,\mathcal{L} \subseteq \mathrm{Fml}\,\overline{\mathcal{L}}$
- $\mathrm{Sen}\,\mathcal{L} \subseteq \mathrm{Sen}\,\overline{\mathcal{L}}$

Clearly, any $\overline{\mathcal{L}}$–structure $\overline{\mathcal{A}}$ can be made to a $\mathcal{L}$–structure, simply by "forgetting" about the interpretations of the non–logical symbols added when forming the expansion $\overline{\mathcal{L}}$. Almost with the same directness we may "expand" any $\mathcal{L}$–structure to a $\overline{\mathcal{L}}$–structure, this time by simply adding the missing interpretations by arbitrary definitions.

Example 4.2.1 If $\mathcal{L}_0$ is the language of groups consisting of a binary function–symbol $+$ and a constant–symbol 0, and $\mathcal{L}_1$ is language of rings (and fields) consisting of two binary function–symbol $+$ and $\cdot$ and two constant–symbols 0 and 1, then $\mathcal{L}_1$ is an expansion of $\mathcal{L}_0$. Every $\mathcal{L}_1$–structure $\mathcal{A}$ is also a $\mathcal{L}_0$–structure, as is exemplified by any ring $\langle R, +, \cdot, 0, 1\rangle$ carrying also the structure of a group $\langle R, +, 0\rangle$, the "group–aspect" resulting from simply forgetting about the interpretations $\cdot$ and 1 of the second function– and constant–symbol, respectively.

Conversely, any group (i.e. any $\mathcal{L}_0$–structure) can be expanded to a structure for the expanded language $\mathcal{L}_1$ by "making up" arbitrary interpretations for the new non–logical symbols $\cdot$ and 1. Nobody expects such arbitrary interpretations to define a ring, but they *do expand* the group to a $\mathcal{L}_1$–structure.

Arbitrary adaptations of course are not the real source of inspiration for the notion of expansion, quite on the contrary, the constructions you are most likely to encounter while dwelling the realms of model theory stem from the syntactical approach of adding constant–symbols for all elements of some structure:

Definition 4.2.2 Let $\mathcal{L}$ be a formal language.

1. If C is a set of constant–symbols not already in $\mathcal{L}$, then $\mathcal{L}_C$ will denote the expansion of $\mathcal{L}$ resulting from adding the constant–symbols $c \in C$ to $\mathcal{L}$.

2. If X is any set, then $\mathcal{L}_X$ will denote the expansion of $\mathcal{L}$ resulting from adding, for every element $x \in X$, a new constant–symbol c_x to $\mathcal{L}$.

3. If $\mathcal{A}$ is a $\mathcal{L}$–structure, then $\mathcal{L}_\mathcal{A} := \mathcal{L}_{|\mathcal{A}|}$.

For 3, the $\mathcal{L}$–structure $\mathcal{A}$ implicitly carries the natural interpretation which makes it into a $\mathcal{L}_\mathcal{A}$–structure $\overline{\mathcal{A}}$ by defining simply $c_a^{\overline{\mathcal{A}}} := a$ for any $a \in |\mathcal{A}|$. To symbolize this, we will use the notation $\overline{\mathcal{A}} := \langle \mathcal{A}, c_a\rangle_{a\in|\mathcal{A}|}$.

We add to our list of consequences of expansions

- $\operatorname{Th}\mathcal{A} \subseteq \operatorname{Th}\overline{\mathcal{A}}$

Taking one step back, one might recall a remark we made concerning term–structures in the absence of constant–symbols, and one might even start to wonder what $\mathrm{CT}(\mathcal{L}_\mathcal{A})/\operatorname{Th}\mathcal{A}$ looks like ...

The title of this section not only mentions expansion but also restriction of languages, the latter of which we still owe you any word about.

Let's look at a set Σ of $\mathcal{L}$–sentences for some language $\mathcal{L}$. If $\mathcal{L}$ is infinite in the sense that $\mathcal{L}$ has infinitely many non–logical symbols, while Σ is finite, we may with all the right in the world say that most symbols in the language are a waste of time, they will never be used while we deal with Σ. So we might just as well, instead of carrying the full weight of $\mathcal{L}$, restrict our attention to the relevant part of $\mathcal{L}$, i.e. the symbols actually occurring in Σ. We thus define

Definition 4.2.3 If $\mathcal{L}$ is a formal language, $\varphi \in \mathrm{Fml}\,\mathcal{L}$ and $\Sigma \subseteq \mathrm{Fml}\,\mathcal{L}$, we denote by $\mathcal{L}(\varphi)$ the sub–language of $\mathcal{L}$ which comprises exactly the constant–, relation– and function–symbols occurring in φ, and consequently by $\mathcal{L}(\Sigma)$ the sub–language of $\mathcal{L}$ consisting exactly of the non–logical–symbols occurring in some formula in Σ.

Par abus de language we could write $\mathcal{L}(\Sigma) = \bigcup_{\varphi \in \Sigma} \mathcal{L}(\varphi)$, implying the union is "intelligent" in that it joins not the mere languages but their sets of constant–, relation– and function–symbols, respectively.

It is noteworthy that, by syntactical considerations, we can prove that if $\Sigma \vdash \varphi$ with $\varphi \notin \mathcal{L}(\Sigma)$, then $\vdash \varphi$. The prove of this uses induction on the length of the deduction of φ from Σ and is a nice exercise to refresh the technical skills for syntax matters.

Also, note that we did not discard $\mathcal{L}$ in the definition of $\mathcal{L}(\varphi)$ entirely for the sole reason of ontological soundness, i.e. we need our formula to be a formula built from symbols that come from some collection accessible to us, and this collection we call $\mathcal{L}$. But the more natural way is, of course, to just have a look at a formula and collect the non–logical symbols, counting on our instinct for math to be able to tell function– from relation– from constant–symbols and these again from logical connectives. Mostly you will not come upon "language–declarations" before doing some math.

4.3 Size Does Matter

In this section, we are going to take a closer look at a construction used mostly to prove the correctness theorem of first–order logic. Put in "everyday's math language", proves for Gödels correctness theorem mostly encapsulate the construction of some model built from syntactical notions, the so–called *term–structure*, by taking the set of variable–free terms of the given formal language $\mathcal{L}$ and dividing this set into equivalence classes under the congruence relation of "being provably equal". The main problem encountered when doing so is to ensure that there really are any variable–free terms, and that there are enough of them to provide examples for any existential sentence formally deducible.

Maybe you recall, from predicate–calculus, that every consistent set of $\mathcal{L}$–formulae may be extended to a set of $\mathcal{L}'$–formulae which is maximal–consistent and has witnesses. Just in case you do not and for the sake of completeness, we sketch the process of providing such an expansion of $\mathcal{L}$ and the set of formulae, and moreover outline the construction of the resulting term–structure in the expanded language.

But since we do not want to bother to give a lecture on notions introduced somewhere else in much more accuracy and with better the motivational framework, we rather try to slightly generalize the result and bend the main focus towards the goals of model theory.

Since the Löwenheim–Skolem theorems introduced in the next sections will deal with the existence of models with universes of desired cardinalities, we first have a look at how the cardinality of a language translates to the cardinalities of certain sets of syntactical constructs:

Definition 4.3.1 For a formal language $\mathcal{L}$, we let **card** $\mathcal{L}$ denote the cardinality of the set of non–logical symbols of $\mathcal{L}$, i.e. the cardinality of the (disjoint) union of the sets of function–, relation– and constant –symbols of $\mathcal{L}$. Especially, we say that $\mathcal{L}$ is **finite** (**infinite**) if card $\mathcal{L}$ is finite (infinite), and similarly for $\mathcal{L}$ being **(un)countable**.[3]

For any set X, let X^* denote the set of all finite families of elements of (or *strings over* X).

Lemma 4.3.2 For any infinite cardinal κ, $\operatorname{card} X \leq \kappa$ iff $\operatorname{card} X^* \leq \kappa$. Moreover, if X is infinite, then $\operatorname{card} X^* = \operatorname{card} X$.

Proof. (Using Set Theory:) If κ is an infinite cardinal, then the following holds:

- $\kappa^n = \kappa$ for all $n \in \mathbb{N}$;
- $\gamma * \kappa = \kappa$ for all $\gamma \leq \kappa$.

Now, if $\operatorname{card} X \leq \kappa$, then

$$\begin{aligned} \operatorname{card}(X^*) &= \sum_{n\in\mathbb{N}} \operatorname{card}(X^n) \\ &= \sum_{n\in\mathbb{N}} \operatorname{card} X = \aleph_0 * \operatorname{card} X \\ &\leq \aleph_0 * \kappa = \kappa. \end{aligned}$$

Also, since $\operatorname{card} X^* \geq \operatorname{card} X$, the other direction follows easily. ■

Lemma 4.3.3 Let $\mathcal{L}$ be a formal language. Then

[3] The notion of countability will — in the course of this paper — be used to stand for *countable or finite*. So, unlike some other authors in the area of logic or Set Theory, for a set to be countable there has to exist an injective mapping from the set to $\mathbb{N}$ which does not necessarily have to be surjective!

1. $\operatorname{card}\operatorname{Tm}\mathcal{L} \leq \operatorname{card}\operatorname{Fml}\mathcal{L} = \operatorname{card}\operatorname{Sen}\mathcal{L}$;
2. If $\mathcal{L}$ is countable, then so are $\operatorname{Tm}\mathcal{L}$, $\operatorname{Fml}\mathcal{L}$ and $\operatorname{Sen}\mathcal{L}$;
3. if $\operatorname{card}\mathcal{L}$ is infinite, then $\operatorname{card}\operatorname{Tm}\mathcal{L} \leq \operatorname{card}\mathcal{L} = \operatorname{card}\operatorname{Fml}\mathcal{L}$;
4. for any set X, if $\operatorname{card}\mathcal{L} \leq \operatorname{card}X$, then $\operatorname{card}\operatorname{Tm}\mathcal{L}_X = \operatorname{card}\operatorname{Fml}\mathcal{L} = \operatorname{card}X$.

Proof.

1. This is an easy exercise: find 1–1–functions from $\operatorname{Tm}\mathcal{L}$ into $\operatorname{Fml}\mathcal{L}$, from $\operatorname{Fml}\mathcal{L}$ into $\operatorname{Sen}\mathcal{L}$ and from $\operatorname{Sen}\mathcal{L}$ into $\operatorname{Fml}\mathcal{L}$.
2. Since every variable is a term, $\operatorname{Tm}\mathcal{L}$ is infinite.

■

Next we like to introduce the construction by Skolem mentioned in this section's title. The main idea behind these Skolem–functions is that pure existential sentences must be verified by an example already in the set of sentences under consideration, thus ensuring that we are equipped with enough constant –symbols to construct the desired syntactical models.

First we need some auxiliary syntactical notions:

Definition 4.3.4 A $\mathcal{L}$–formula φ is called a **property**(–formula) if φ has at most one free variable. The set of all $\mathcal{L}$–properties will be denoted by $\operatorname{Prop}\mathcal{L}$.

If φ is a property formula with free variable x, then we define, for any $\mathcal{L}$–structure $\mathcal{A}$, the **extension** of φ in $\mathcal{A}$, denoted by $|\varphi|_{\mathcal{A}}$, by

$$|\varphi|_{\mathcal{A}} := \{a \in |\mathcal{A}| \; ; \; \mathcal{A} \models \varphi[h\tbinom{x}{a}] \text{ for any valuation } h \text{ into } \mathcal{A}\} \; .$$

So the extension of a property is simply the set of all elements which “have this property”. Note that for properties not having any free variables, the extension in a structure is either empty or the whole universe.

Definition 4.3.5 A set of sentences $\Sigma \subseteq \operatorname{Fml}\mathcal{L}$ is said to **have witnesses** if for every property $\varphi \in \operatorname{Prop}\mathcal{L}$, there is a constant–symbol c in $\mathcal{L}$ such that $\exists x\varphi \to \varphi\,(x/c) \in \Sigma$.

Lemma 4.3.6 Let $\Sigma_0 \subseteq \Sigma_1 \subseteq \Sigma_2 \subseteq \ldots$ be a infinite ascending chain of deductively closed, consistent sets of $\mathcal{L}$–sentences. Then $\bigcup_{n\in\mathbb{N}} \Sigma_n$ is deductively closed and consistent.

Proof. If $\Sigma' := \bigcup_{n\in\mathbb{N}} \Sigma_n$ were inconsistent, then (by compactness) we would find an inconsistent finite subset $\{\varphi_1,\ldots,\varphi_n\}\subseteq \Sigma'$ and thus $\Sigma_{i_1},\ldots,\Sigma_{i_n}$ with $\{\varphi_1,\ldots,\varphi_n\}\subseteq \Sigma_{i_1}\cup\ldots\cup\Sigma_{i_n}$. But then $\Sigma_{i_1}\cup\ldots\cup\Sigma_{i_n} = \Sigma_m$ would be inconsistent for some m, contradicting the assumptions.

Σ' is deductively closed since, again by compactness, if $\Sigma' \vdash \varphi$, then there is again a finite subset $\{\varphi_1,\ldots,\varphi_n\}$ with $\{\varphi_1,\ldots,\varphi_n\}\vdash \varphi$, so again we find $\Sigma_m \vdash \varphi$ for some m, but then, since Σ_m is deductively closed, $\varphi \in \Sigma_m \subseteq \Sigma'$. ■

Lemma 4.3.7 If $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$, then

$$\Sigma' := \Sigma \cup \{\exists x\varphi \to \varphi\,(x/c_\varphi)\ ;\ \varphi \in \mathrm{Prop}\,\mathcal{L}\}$$

is consistent.

Proof. This is proved in most Textbooks dealing with the proof of completeness of first–order logic via Henkin–style structures. To get the idea, please consult the literature. ■

Most of the proof of the following theorem has been done in ..., but for the sake of self–completeness and the cardinality–related observation in part (iv), we are going to give a proof, omitting certain tedious details.

Theorem 4.3.8 If $\mathcal{L}$ is a formal language and $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ is consistent, then there is an extension $\mathcal{L}'$ of $\mathcal{L}$ and a set Σ' of $\mathcal{L}'$–sentences such that

(i) $\Sigma \subseteq \Sigma'$;

(ii) Σ' has witnesses;

(iii) Σ' is consistent and deductively closed;

(iv) $\mathrm{card}\,\mathcal{L}' = \max\{\aleph_0, \mathrm{card}\,\mathcal{L}\}$.

Proof. Define, for $n \in \mathbb{N}$, formal languages $\mathcal{L}_n$ expanding $\mathcal{L}$ and $\Sigma_n \subseteq \mathrm{Sen}\,\mathcal{L}_n$ as follows: Set $\mathcal{L}_0 := \mathcal{L}$ and $\Sigma_0 := \Sigma$. Now suppose we are given $\mathcal{L}_i$ and Σ_i. Then we set

$$\mathcal{L}_{i+1} := \langle \mathcal{L}, c_\varphi \rangle_{\varphi \in \mathrm{Prop}\,\mathcal{L}_i}$$

and

$$\Sigma_{i+1} := \mathrm{Ded}(\Sigma_i \cup \{\exists x\varphi \to \varphi\,(x/c_\varphi)\ ;\ \varphi \in \mathrm{Prop}\,\mathcal{L}_i\}),$$

where it is understood that the new constants c_φ are pairwise distinct. Finally, we set $\mathcal{L}' := \bigcup_{n\in\mathbb{N}} \mathcal{L}_n$ and $\Sigma' := \bigcup_{n\in\mathbb{N}} \Sigma_n$.

Then clearly (i) holds. As for (ii), if $\varphi \in \mathrm{Prop}\,\mathcal{L}'$, then $\varphi \in \mathrm{Prop}\,\mathcal{L}_i$ for some $i \in \mathbb{N}$, and thus $\exists x\varphi \to \varphi\,(x/c_\varphi) \in \Sigma_{i+1} \subseteq \Sigma'$. (iii) follows directly from Lemma 4.3.6.

To show (iv), we first note that if $\mathcal{L}$ is countable, then $\mathrm{Fml}\,\mathcal{L}$ is countable and so all the expansions $\mathcal{L}_i$ of $\mathcal{L}$ are countable. But then $\mathcal{L}'$ is the countable union of countable sets and thus, by Set Theory, $\mathcal{L}'$ is countable. On the other hand, if $\mathrm{card}\,\mathcal{L} > \aleph_0$, then $\mathrm{card}\,\mathcal{L}_i = \mathrm{card}\,\mathcal{L}$ for all expansions $\mathcal{L}_i$ of $\mathcal{L}$, and again set theories tells us that $\mathrm{card}\,\mathcal{L}' = \mathrm{card}\,\mathcal{L}$. ■

Theorem 4.3.9 If $\mathcal{L}$ is any formal language and $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ is consistent, then there is a model $\mathcal{A}$ of Σ with $\mathrm{card}\,|\mathcal{A}| \leq \max\{\aleph_0, \mathrm{card}\,\mathcal{L}\}$.

Proof. Let $\mathcal{L}$ be a formal language and $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ consistent. By 4.3.8, we find a language $\mathcal{L}' \supseteq \mathcal{L}$ and a consistent theory $\Sigma' \subseteq \mathrm{Sen}\,\mathcal{L}'$ such that $\mathcal{L}'$ has witnesses. Moreover, since Σ' is consistent, we have $\mathrm{Mod}\,\Sigma' \neq \emptyset$ and thus, for some $\mathcal{A} \in \mathrm{Mod}\,\Sigma'$, $\Sigma' \subseteq \mathrm{Th}\,\mathcal{A}$. So w.l.o.g we assume that Σ' is complete (and hence maximal consistent).

Now, by correctness, $\mathrm{CT}(\mathcal{L})'\,/\Sigma'$ is a model of Σ', and

$$\mathrm{card}\,|\,\overset{\prime}{\mathrm{CT}}(\mathcal{L})\,/\Sigma'| \leq \mathrm{card}\,\mathrm{Tm}\,\mathcal{L}' = \mathrm{card}\,\mathcal{L}'.$$

■

So we see that, unless we're dealing with a very rich language with lots of names and symbols for functions and relations, we may find models for consistent sets of sentences that are rather small:

Corollary 4.3.10 If $\mathcal{L}$ is finite or countable, then every consistent set of $\mathcal{L}$–sentences has a countable (finite or infinite) model.

So we see that there are countable models of the theory of the complex or real numbers considered as fields. What's more of a surprise, the above corollary implies the existence of a *countable model of ZFC Set Theory*, a theory dealing with cardinals as mind-bogglingly uncountable as $\aleph_{\aleph_0}^{\aleph_{\aleph_0}}$, thus such a countable model must have an element having exactly the same first–order properties provable for $\aleph_{\aleph_0}^{\aleph_{\aleph_0}}$ in ZFC!

Nevertheless, you might be left with the feeling of a somewhat empty stomach, since what we did is to build a model made up from purely syntactical constructs. So our knowledge about this model is not deeper than our knowledge of the syntax of first–order logic, and there we have to live with the boundaries set by incompleteness/undecidability. What would (even if only mentally) more reassuring is the construction of a smaller model, given a structure which is, in the light of the above results, fitted with a universe full of redundant entities. By reducing this model to its necessary small part, we could probably have a better grasp of its properties by using (non–first–order) results known about the bigger structure. This is exactly what the Downward Löwenheim–Skolem Theorem is all about, and we will deal with it right now.

4.4 Meet the Morphisms!

When considering classes of mathematical constructs sharing some general structural common ground, it is common mathematical practice to have a closer look at mappings between them preserving these properties. These mappings are often[4] called *(homo–)morphisms*. We are now looking at such mappings in the context of model–theory, which means the properties to be preserved are given by first–order logic, i.e. the interpretations of the symbols:

Definition 4.4.1 Let $\mathcal{A}$, $\mathcal{B}$ be $\mathcal{L}$–structures. A map $\eta : |\mathcal{A}| \longrightarrow |\mathcal{B}|$ is a $\mathcal{L}$–**homomorphism** from $\mathcal{A}$ into $\mathcal{B}$ iff

- for all relation–symbols R_i and all $a_1, \ldots, a_{\lambda(i)} \in |\mathcal{A}|$,

$$\langle a_1, \ldots, a_{\lambda(i)} \rangle \in R_i^{\mathcal{A}} \text{ implies } \langle \eta(a_1), \ldots, \eta(a_{\lambda(i)}) \rangle \in R_i^{\mathcal{B}};$$

- for all function–symbols f_j and all $a_1, \ldots, a_{\mu(j)} \in |\mathcal{A}|$,

$$\eta(f_j^{\mathcal{A}}(a_1, \ldots, a_{\mu(j)}) = f_j^{\mathcal{B}}(\eta(a_1), \ldots, \eta(a_{\mu(j)}));$$

- for all constant–symbols c_k,

$$\eta(c_k^{\mathcal{A}}) = c_k^{\mathcal{B}}.$$

$\mathcal{A}$ is called the **source** (or **domain**) and $\mathcal{B}$ the **target** (or **co–domain**) of η. If η is surjective, $\mathcal{B}$ is called a **homomorphic image** of $\mathcal{A}$. We write $\hom \mathcal{A}\mathcal{B}$ for the set of all homomorphisms from $\mathcal{A}$ into $\mathcal{B}$.

To simplify the notation, we introduce the following abbreviation: If X, Y and S are sets and $\eta : X \longrightarrow Y$, then the function $\eta^S : X^S \longrightarrow Y^S$ is defined by

$$\eta^{(S)}(\langle x_s \, ; \, s \in S \rangle) := \langle \eta(x_s) \, ; \, s \in S \rangle \, .$$

Especially, if $n \in \mathbb{N}$,

$$\eta^{(n)}(x_1, \ldots, x_n) := \langle \eta(x_1), \ldots, \eta(x_n) \rangle,$$

[4] At least this is true if algebraic structures are involved. Of course continuous functions fall in this category also, as do order–preserving functions.

where the notation

$$\underbrace{\eta \times \eta \times \ldots \times \eta}_{n \text{ factors}} \text{ for } \eta^{(n)}$$

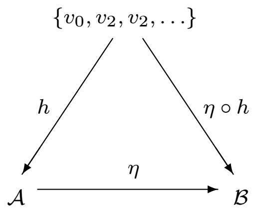
is equally popular.

As simple examples we see, for any $\mathcal{L}$–structure $\mathcal{A}$, that the *identity–map* $\mathrm{id}_{\mathcal{A}}$ defined by $\mathrm{id}_{\mathcal{A}}(a) = a$ for all $a \in |\mathcal{A}|$, is a homomorphism. Moreover, it is easy to see that the composition $\eta \circ \rho$ of two homomorphism is again a homomorphism. Another example involves the notion of the direct product of structures, since then we see that the projections are homomorphisms as well. (This will become clear when products will actually be needed.)

Homomorphic behavior of a map, as defined on functions and constants, translates to terms in general:

Lemma 4.4.2 If $\mathcal{A}, \mathcal{B}$ are $\mathcal{L}$–structures and $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ is a $\mathcal{L}$–homomorphism, then for any $\mathcal{L}$–term t and any valuation h into $\mathcal{A}$,

$$\eta(t^{\mathcal{A}}[h]) = t^{\mathcal{B}}[\eta \circ h].$$



Proof. By induction on t.

- If $t \equiv x$ is a variable, then

$$\begin{aligned} \eta(t^{\mathcal{A}}[h]) &= \eta(h(x)) \\ &= \eta \circ h(x) = t^{\mathcal{B}}[\eta \circ h]. \end{aligned}$$

- If $t \equiv c_k$ is a constant–symbol, then

$$\begin{aligned} \eta(t^{\mathcal{A}}[h]) &= \eta(c_k^{\mathcal{A}}) \\ &= c_k^{\mathcal{B}} = t^{\mathcal{B}}[\eta \circ h] \end{aligned}$$

since valuations are negligible for constant–symbols.

- If $t \equiv f_j(t_1, \ldots, t_{\mu(j)})$ for terms $t_1, \ldots, t_{\mu(j)}$, then

$$\begin{aligned}
\eta(t^{\mathcal{A}}[h]) &= \eta(f_j^{\mathcal{A}}(t_1^{\mathcal{A}}[h], \ldots, t_{\mu(j)}^{\mathcal{A}}[h])) \\
&= f_j^{\mathcal{B}}(\eta(t_1^{\mathcal{A}}[h]), \ldots, \eta(t_{\mu(j)}^{\mathcal{A}}[h])) \\
&= f_j^{\mathcal{B}}(t_1^{\mathcal{B}}[\eta \circ h], \ldots, t_{\mu(j)}^{\mathcal{B}}[\eta \circ h]) \text{ (by ind. hyp.)} \\
&= t^{\mathcal{B}}[\eta \circ h].
\end{aligned}$$

■

Characterizing the homomorphisms according to their *universal* behavior is the next step we take. For example, looking at groups, we call the injective homomorphisms *monomorphisms* and surjective homomorphisms *epimorphisms*. Now being injective or surjective is not exactly a *universal* property, since it can be verified in a rather local setting involving but the domain and the co–domain of the homomorphism. But there are indeed more general ways of expressing the central ideas:

Definition 4.4.3 Let $\mathcal{A}$ and $\mathcal{B}$ be $\mathcal{L}$–structures and $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ a $\mathcal{L}$–homomorphism.

1. η is called a $\mathcal{L}$–**monomorphism** (or **mono**) if, for any $\mathcal{L}$–structure $\mathcal{C}$ and any $\rho_1, \rho_2 : \mathcal{C} \longrightarrow \mathcal{A}$,

$$\eta \circ \rho_1 = \eta \circ \rho_2 \text{ implies } \rho_1 = \rho_2.$$

$$\mathcal{C} \underset{\rho_2}{\overset{\rho_1}{\rightrightarrows}} \mathcal{A} \xrightarrow{\eta} \mathcal{B}$$

2. η is called a $\mathcal{L}$–**epimorphism** (or **epi**) if, for any $\mathcal{L}$–structure $\mathcal{C}$ and any $\rho_1, \rho_2 : \mathcal{B} \longrightarrow \mathcal{C}$,

$$\rho_1 \circ \eta = \rho_2 \circ \eta \text{ implies } \rho_1 = \rho_2.$$

$$\mathcal{A} \xrightarrow{\eta} \mathcal{B} \underset{\rho_2}{\overset{\rho_1}{\rightrightarrows}} \mathcal{C}$$

3. η is called a $\mathcal{L}$–**isomorphism** (or **iso**) if there is a $\rho : \mathcal{B} \longrightarrow \mathcal{A}$ such that

$$\rho \circ \eta = \mathrm{id}_{\mathcal{A}} \text{ and } \eta \circ \rho = \mathrm{id}_{\mathcal{B}}.$$

$$\mathcal{A} \underset{\eta}{\overset{\eta^{-1}}{\leftrightarrows}} \mathcal{B}$$

4. η is called a **(isomorphic) $\mathcal{L}$–embedding** if η is an isomorphism onto a substructure of $\mathcal{B}$.

$\mathcal{A}$ and $\mathcal{B}$ are called **isomorphic**, notation $\mathcal{A} \cong \mathcal{B}$, if there is a $\mathcal{L}$–isomorphism $\eta : \mathcal{A} \longrightarrow \mathcal{B}$. $\mathcal{A}$ is called **(isomorphically) embeddable into $\mathcal{B}$** if there is an embedding $\eta : \mathcal{A} \longrightarrow \mathcal{B}$.

There might be some confusion arising since in classical algebra, e.g. group–monomorphisms are *defined* to be injective homomorphisms. Since we are inclined to take a somewhat more universal attitude, we used the universal (or categorical) properties as defining statements. So for the present lecture, we distinguish between mono and 1–1, between epi and surjective. Of course there are connections:

Remark 4.4.4 1. injective homomorphisms are monos,

2. surjective homomorphisms are epis,

3. isos are injective and surjective.

But be aware of the fact that for a map to be an $\mathcal{L}$–isomorphism, it needs more than just being injective and surjective. As an exercise, study the (simple) setting for $\mathcal{L}$ being the language having $\leq$ as only non–logical symbol, and consider the two $\mathcal{L}$–structures $\mathcal{A} := \langle\{(0,1),(1,0)\},\leq\rangle$ and $\mathcal{B} := \langle\{(0,0),(1,1)\},\leq\rangle$, where in both cases $\leq$ is the point–wise ordering. It will present no real difficulty to find a $\mathcal{L}$–homomorphism that is both injective and surjective, but still $\mathcal{A}$ and $\mathcal{B}$ are not isomorphic.

The following provides simple paraphrasing of the conditions for a map to be an isomorphism:

Lemma 4.4.5 For a homomorphism $\eta : \mathcal{A} \longrightarrow \mathcal{B}$, the following are equivalent:

(i) η is a $\mathcal{L}$–isomorphism;

(ii) η is injective and surjective and η^{-1} is a $\mathcal{L}$–homomorphism;

(iii) η is injective and surjective and

$$R_i^{\mathcal{A}}(a_1,\ldots,a_{\lambda(i)}) \text{ iff } R_i^{\mathcal{B}}(\eta(a_1),\ldots,\eta(a_{\lambda(i)}))$$

for all relation–symbols R_i of $\mathcal{L}$ and all $a_1,\ldots,a_{\lambda(i)} \in |\mathcal{A}|$.

Proof. Exercise. ■

In later chapters we will be concerned with algebras, which are roughly speaking structures without relations. It is not hard to see that in this setting, isos correspond exactly to bijective homomorphisms.

Isomorphic structures can be converted one into the other by a "renaming" of their elements, provided by the isomorphism considered. For any isomorphism $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ the inverse map η^{-1} is an isomorphism from $\mathcal{B}$ to $\mathcal{A}$ (Exercise: Prove this).

Throughout mathematical fields, finding an isomorphism between two structures is a way of showing that within the theory at hand, we cannot distinguish these two structures, they look exactly the same, which serves as a motivation to drop the distinction between them and actually *regarding them as one and the same.* This is what is often expressed by the idiom "unique up to isomorphism", as in "up to isomorphism, there is exactly one group of 7 elements". The general setting we are concerned with is no different, only are we not content with having found a 1–1 and onto homomorphism as e.g. group theorists have the advantage of, we must be a little more careful (while staying even more *universal*).

Now that we know about isomorphisms, it's time to look at some consequences accompanying this notion. If isomorphic should have any prospect of standing for "indistinguishable", then surely first–order logic is not allowed to tell isomorphic structures apart. In other words, isomorphisms should not only preserve algebraic properties but first–order–properties as well. So clearly the following makes a lot of sense:

Theorem 4.4.6 For any two $\mathcal{L}$–structures $\mathcal{A}$ and $\mathcal{B}$,

$$\text{If } \mathcal{A} \cong \mathcal{B}, \text{ then } \mathcal{A} \equiv \mathcal{B}.$$

Proof. First, we note that for any valuation $h_{\mathcal{A}}$ into $\mathcal{A}$, $\eta \circ h_{\mathcal{A}}$ is a valuation into $\mathcal{B}$ and, conversely, for any valuation $h_{\mathcal{B}}$ into $\mathcal{B}$, there is a valuation $h_{\mathcal{A}}$ into $\mathcal{A}$ with $h_{\mathcal{B}} = \eta \circ h_{\mathcal{A}}$.

Next we realize that, by 4.4.2, for any $\mathcal{L}$–term t and for any two valuations $h_{\mathcal{A}}$ into $\mathcal{A}$ and $h_{\mathcal{B}}$ into $\mathcal{B}$,

$$t^{\mathcal{B}}[\eta \circ h_{\mathcal{A}}] = \eta(t^{\mathcal{A}}[h_{\mathcal{A}}])$$

and

$$t^{\mathcal{A}}[\eta^{-1} \circ h_{\mathcal{B}}] = \eta^{-1}(t^{\mathcal{B}}[h_{\mathcal{B}}])$$

giving us a bijective correspondence on terms.

Stepping over to $\mathcal{L}$–formulae, we have to show that, for all $\varphi \in \operatorname{Fml}\mathcal{L}$ and all valuations $h_\mathcal{A}, h_\mathcal{B}$ into $\mathcal{A}$, $\mathcal{B}$ respectively,

$$\mathcal{B} \models \varphi[\eta \circ h_\mathcal{A}] \text{ iff } \mathcal{A} \models \varphi[h_\mathcal{A}].$$

You will not be too surprised to hear that this can be done via structural induction over φ. So we have to consider the following cases:

- If $\varphi \equiv t_1 \doteq t_2$ for $\mathcal{L}$–terms t_1, t_2, then for any valuation h into $\mathcal{A}$ we have

$$\begin{aligned} \mathcal{A} \models \varphi[h] \quad & \text{iff} \quad t_1^\mathcal{A}[h] = t_2^\mathcal{A}[h] \\ & \text{iff} \quad \eta(t_1^\mathcal{A}[h]) = \eta(t_2^\mathcal{A}[h]) \text{ (η is 1–1 and onto)} \\ & \text{iff} \quad t_1^\mathcal{B}[\eta \circ h] = t_2^\mathcal{B}[\eta \circ h] \text{ (by the above)} \\ & \text{iff} \quad \mathcal{B} \models \varphi[\eta \circ h]. \end{aligned}$$

- The case where $\varphi \equiv R_i(t_1, \ldots, t_n)$ is left as an exercise.

- If $\varphi \equiv \neg\vartheta$ for a $\mathcal{L}$–formula ϑ, then for any valuation h into $\mathcal{A}$ we have

$$\begin{aligned} \mathcal{A} \models \varphi[h] \quad & \text{iff} \quad \mathcal{A} \not\models \vartheta[h] \\ & \text{iff} \quad \mathcal{B} \not\models \vartheta[\eta \circ h] \text{ (by ind. hyp.)} \\ & \text{iff} \quad \mathcal{B} \models \varphi[\eta \circ h]. \end{aligned}$$

- Again the case $\varphi \equiv \vartheta \wedge \psi$ is left as an exercise to the reader.

- If $\varphi \equiv \forall x \vartheta$ for a variable x and a $\mathcal{L}$–formula ϑ, then for any valuation h into $\mathcal{A}$ we have

$$\begin{aligned} \mathcal{A} \models \varphi[h] \quad & \text{iff} \quad \mathcal{A} \models \vartheta[h\tbinom{x}{a}] \text{ for all } a \in |\mathcal{A}| \\ & \text{iff} \quad \mathcal{A} \models \vartheta[h\tbinom{x}{\eta^{-1}(b)}] \text{ for all } b \in |\mathcal{B}| \\ & \text{iff} \quad \mathcal{B} \models \vartheta[\eta \circ (h\tbinom{x}{\eta^{-1}(b)})] \text{ for all } b \in |\mathcal{B}| \\ & \text{iff} \quad \mathcal{B} \models \vartheta[(\eta \circ h)\tbinom{x}{b}] \text{ for all } b \in |\mathcal{B}| \\ & \text{iff} \quad \mathcal{B} \models \varphi[\eta \circ h]. \end{aligned}$$

■

To hope for the converse to hold as well is being over–optimistic and not justified, as later examples will show. So in general, elementary equivalent structures need not be isomorphic; our first–order languages lack a mechanism to express certain properties which are preserved by isomorphisms. If we carefully study the definition of isomorphisms, we might even spot the crucial point where formal languages for first–order logic fail to provide the necessary constraints: Since relation–, function– and constant–symbols are syntactical elements, homomorphic behavior is implemented in first–order languages. So we only might get in trouble where the existence of a unique inverse morphism is demanded, or to regard it in a more local setting, with bijectivity. Still, if the structures under consideration are finite, there will be no problem, but with infinite structures, elementary equivalence will prove to be properly weaker than isomorphism.

For the rest of this section, we need an auxiliary notion which will help us to capture $\mathcal{L}$–isomorphisms for the finite case using the first–order language $\mathcal{L}$. For any $n \in \mathbb{N}$ we are going to define a set of $\mathcal{L}$–formulae as follows: ${\Gamma^*}_n$ contains exactly the following formulae

- $v_{m_1} \doteq v_{m_2}$ and $\neg v_{m_1} \doteq v_{m_2}$ for any $m_1, m_2 \in \{1, \ldots, n\}$;
- $c_k \doteq v_m$ and $\neg c_k \doteq v_m$ for any $m \in \{1, \ldots, n\}$;
- $f_j(v_{m_1}, \ldots, v_{m_{\mu(j)}}) \doteq v_i$ and $\neg f_j(v_{m_1}, \ldots, v_{m_{\mu(j)}}) \doteq v_i$ for any $m_1, \ldots, m_{\mu(j)} \in \{1, \ldots, n\}$;
- $R_i(v_{m_1}, \ldots, v_{m_{\lambda(i)}})$ and $\neg R_i(v_{m_1}, \ldots, v_{m_{\lambda(i)}})$ for any $m_1, \ldots, i_{\lambda(i)} \in \{1, \ldots, n\}$.

We note that (1) only the n variables $v_1, \ldots, v_n$ are occurring in formulae in ${\Gamma^*}_n$ and (2) ${\Gamma^*}_n$ is finite whenever $\mathcal{L}$ is finite.

If $\mathcal{A}$ is a $\mathcal{L}$–structure and h a valuation in $\mathcal{A}$, then $\Gamma^*(\mathcal{A}, h)$ is given by

$$\Gamma^*(\mathcal{A}, h) := \{\varphi \in {\Gamma^*}_{\mathrm{card}\,|\mathcal{A}|} \,;\, \mathcal{A} \models \varphi[h]\}.$$

Lemma 4.4.7 Let $\mathcal{A}, \mathcal{B} \in \mathrm{Str}\,\mathcal{L}$ be finite. A map $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ is an isomorphism iff for any valuation h into $\mathcal{A}$, $\Gamma^*(\mathcal{A}, h) = \Gamma^*(\mathcal{B}, \eta \circ h)$.

Proof. Assume $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ is an isomorphism and h is a valuation into $\mathcal{A}$. So $\mathrm{card}\,|\mathcal{A}| = \mathrm{card}\,|\mathcal{B}|$. So we only have to consider formulae in ${\Gamma^*}_{\mathrm{card}\,|\mathcal{A}|}$.

Also, for any structure $\mathcal{C}$, any valuation h into $\mathcal{C}$ and any $\varphi \in {\Gamma^*}_{\mathrm{card}\,|\mathcal{C}|}$ *which is not a negation*, we have $\neg\varphi \in \Gamma^*(\mathcal{C}, h)$ iff $\varphi \notin \Gamma^*(\mathcal{C}, h)$, so we can actually restrict our attention to the non–negated formulae.

Let $\varphi \in \Gamma^*(\mathcal{A}, h)$.

- if $\varphi \equiv v_{m_1} \doteq v_{m_2}$, then we see that

$$\begin{aligned} \varphi \in \Gamma^*(\mathcal{A}, h) \quad &\text{iff} \quad h(v_{m_1}) = h(v_{m_2}) \\ &\text{iff} \quad \eta \circ h(v_{m_1}) = \eta \circ h(v_{m_2}) \text{ since } \eta \text{ is a bijection} \\ &\text{iff} \quad \varphi \in \Gamma^*(\mathcal{B}, \eta \circ h) \end{aligned}$$

- similar for the other cases.

For the other direction, we first note that $\Gamma^*(\mathcal{A}, h) = \Gamma^*(\mathcal{B}, \eta \circ h)$ for any h implies $\text{card}\,|\mathcal{A}| = \text{card}\,|\mathcal{B}|$. Now let h be a valuation into $\mathcal{A}$ which is 1–1 on $\{v_1, \ldots, v_{\text{card}\,|\mathcal{A}|}\}$. Then for η with $\Gamma^*(\mathcal{A}, h) = \Gamma^*(\mathcal{B}, \eta \circ h)$, η is a bijection since for $a, b \in |\mathcal{A}|$, $a \neq b$, we find v_i, v_j, $i, j \in \{1, \ldots, \text{card}\,|\mathcal{A}|\}$ with $h(v_i) = a$, $h(v_j) = b$, so $\neg v_i \doteq v_j \in \Gamma^*(\mathcal{A}, h)$ and thus $\neg v_i \doteq v_j \in \Gamma^*(\mathcal{B}, \eta \circ h)$, i.e. $\eta(a) = \eta(h(v_i)) \neq \eta(h(v_j)) = \eta(b)$. Since a, b were arbitrary, η is 1–1 and thus by finiteness a bijection. Moreover, for any constant–symbol c_k, $c_k^{\mathcal{A}} = h(v_i)$ for some $i \in \{1, \ldots, \text{card}\,|\mathcal{A}|\}$, so $c_k \doteq v_i \in \Gamma^*(\mathcal{A}, h) = \Gamma^*(\mathcal{B}, \eta \circ h)$, so $c_k^{\mathcal{B}} = \eta \circ h(v_i) = \eta(c_k^{\mathcal{A}})$. Similar argumentations for function– and relation–symbols show that η is an isomorphism. ■

Exercise 4.4.8 Write out the details of the above proof.

Lemma 4.4.9 If $\mathcal{L}$ is a *finite* formal language[5] and $\mathcal{A}$ is a *finite* $\mathcal{L}$–structure, then there is a $\mathcal{L}$–sentence $\gamma_{\mathcal{A}}$ such that for any $\mathcal{B} \in \text{Str}\,\mathcal{L}$,

$$\mathcal{B} \models \gamma_{\mathcal{A}} \text{ iff } \mathcal{A} \cong \mathcal{B}.$$

Proof. Let h be a valuation into $\mathcal{A}$ which is 1–1 on $\{v_1, \ldots, v_{\text{card}\,|\mathcal{A}|}\}$. Since the language is finite, $\Gamma^*(\mathcal{A}, h)$ is finite, say $\Gamma^*(\mathcal{A}, h) = \{\psi_1, \ldots, \psi_k\}$. Let the $\mathcal{L}$–formula $\varphi_{\mathcal{A}}$ be given by $\varphi_{\mathcal{A}} :\equiv \psi_1 \wedge \ldots \wedge \psi_k$. Moreover, for $n \in \mathbb{N}$, let $\varphi_{\leq n}$ be the sentence expressing that there are *at most* n elements in the universe. Now let $\text{card}\,|\mathcal{A}| = n$. We claim that the desired $\mathcal{L}$–sentence $\gamma_{\mathcal{A}}$ is

$$\gamma_{\mathcal{A}} \equiv \varphi_{\leq n} \wedge \exists v_1 \ldots \exists v_n \varphi_{\mathcal{A}}.$$

It's easy to see that $\mathcal{A} \models \gamma_{\mathcal{A}}$. Now assume $\mathcal{A} \cong \mathcal{B}$. Then by Theorem 4.4.6, $\mathcal{A} \equiv \mathcal{B}$, and $\mathcal{A} \models \gamma_{\mathcal{A}}$, so $\mathcal{B} \models \gamma_{\mathcal{A}}$.

Conversely, if $\mathcal{B} \models \gamma_{\mathcal{A}}$, then $\mathcal{B}$ has at most n elements and is thus finite. Also $\mathcal{B} \models \exists v_1 \ldots \exists v_n \varphi_{\mathcal{A}}$, so $\mathcal{B} \models \varphi_{\mathcal{A}}[h']$ for some valuation h' into $\mathcal{B}$. Let $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ satisfy $\eta(h(v_i)) := h'(v_i)$ for all $i \in \{1, \ldots, \text{card}\,|\mathcal{A}|\}$. (This uniquely defines η since h is supposed to be 1–1 on $\{v_1, \ldots, v_n\}$. Now let $\psi \in {\Gamma^*}_n$. If $\psi \in \Gamma^*(\mathcal{A}, h)$, then $\mathcal{B} \models \psi[\eta \circ h]$ by the choice of h' and since $h' = \eta \circ h$. On the other hand, if $\psi \notin \Gamma^*(\mathcal{A}, h)$, then (1) if ψ is not negated, $\neg\psi \in \Gamma^*\,\mathcal{A}h$ and thus $\mathcal{B} \models \neg\psi[\eta \circ h]$, i.e. $\mathcal{B} \not\models \psi[\eta \circ h]$, or (2) if ψ is negated, say $\psi \equiv \vartheta$, the same argument holds usingϑ instead of $\neg\psi$. So together we get, for any $\psi \in {\Gamma^*}_n$,

$$\mathcal{B} \models \psi[\eta \circ h] \text{ iff } \psi \in \Gamma^*(\mathcal{A}, h).$$

[5] So $\mathcal{L}$ has only finitely many non–logical symbols, cf. Definition 4.3.1.

Since $\mathcal{B}$ is finite and has at most card $|\mathcal{A}|$ elements, $\Gamma^*(\mathcal{B}, \eta \circ h)$ is a subset of ${\Gamma^*}_{\mathrm{card}\,|\mathcal{A}}$, so we get $\Gamma^*(\mathcal{A}, h) = \Gamma^*(\mathcal{B}, \eta \circ h)$. But by Lemma 4.4.7, η is an isomorphism. ■

Proposition 4.4.10 If $\mathcal{A}$ and $\mathcal{B}$ are *finite* $\mathcal{L}$–structures, then

$$\mathcal{A} \cong \mathcal{B} \text{ iff } \mathcal{A} \equiv \mathcal{B}.$$

Proof. Clearly we only have to show the right–to–left half (the other direction is exactly 4.4.6). So assume $\mathcal{A} \ncong \mathcal{B}$. Define

$$\mathcal{E} := \{\eta \,:\, \mathcal{A} \longrightarrow \mathcal{B}\,;\, \eta \text{ is bijective}\}\ .$$

Now none of the $\eta \in \mathcal{E}$ is a isomorphism, so we find, for every $\eta \in \mathcal{E}$,

a constant–symbol c_k such that $\eta(c_k^{\mathcal{A}}) \neq c_k^{\mathcal{B}}$

or a function–symbol f_j and $a_1, \ldots, a_{\mu(j)} \in |\mathcal{A}|$ with

$$\eta(f_j^{\mathcal{A}}(a_1, \ldots, a_{\mu(j)})) \neq f_j^{\mathcal{B}}(\eta(a_1), \ldots, \eta(a_{\mu(j)}))$$

or a relation–symbol R_i and $a_1, \ldots, a_{\lambda(i)} \in |\mathcal{A}|$ such that

$$\text{not } [R_i^{\mathcal{A}}(a_1, \ldots, a_{\lambda(i)}) \text{ iff } R_i^{\mathcal{B}}(\eta(a_1), \ldots, \eta(a_{\lambda(i)}))].$$

Let $\mathcal{L}_1$ be the sub–language of $\mathcal{L}$ containing only one such constant–, function– or relation–symbol for every $\eta \in \mathcal{E}$. Then $\mathcal{L}_1$ is finite. Clearly every $\mathcal{L}$–isomorphism would be a $\mathcal{L}_1$–isomorphism, so also in $\mathcal{L}_1$, $\mathcal{A}$ and $\mathcal{B}$ are not isomorphic. Applying Lemma ■

Chapter 5

The Löwenheim–Skolem Theorems

The main goal of this chapter is to demonstrate that there are limits to the extent in which first–order theories influence the cardinality of their models. We will be show that, roughly speaking, the number of non–logical symbols of a language marks the lower boundary for the number of elements in the models, and that for theories with infinite models, there is no upper limit to the cardinality of models.

5.1 Cardinality

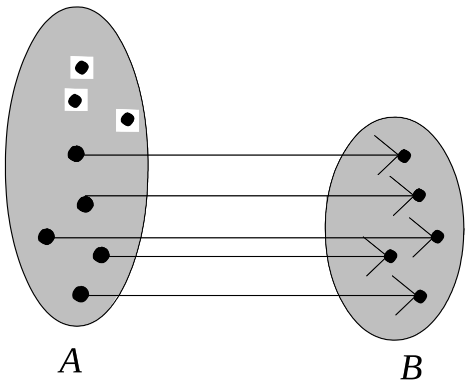
The aim of this section is to explain the basic notion of the *cardinality* of a set. It is not meant to be an introduction to the theory of cardinals, since we are interested in an intuitive understanding of the concept of cardinality and in its precise definition in the context of Set Theory (which is where cardinals originated).

The Löwenheim–Skolem theorems deal with the size a model of a theory may

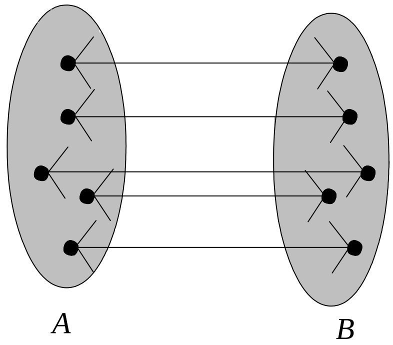
Figure 5.1: Leopold Löwenheim (1878–1957)

Figure 5.2: Georg Cantor (1845–1918)

have. Therefore, we need a way to *measure* a model's size and to compare it with other models. A rather straightforward way to measure a collection's size is clearly to count the entities that belong to that collection. Then, we attach the number we found by this counting as the collection's measure. Even if we do not want to deal with the (abstract) concept of numbers, we are still in a position to compare collections with respect to the number of their elements by making a one–to–one assignment of their elements: The larger collection will always have some elements which cannot be assigned.



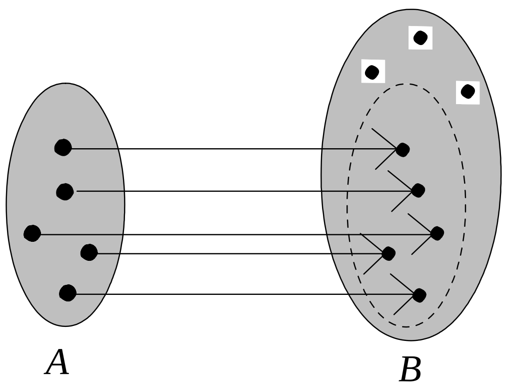
Set Theory (which, as mentioned above, is *not* the focus of this section) provides the notion of *bijective mappings* (cf. Section 1.2) which fulfill the task of such a one–to–one assignment. If there is a bijective mapping from some set A onto a set B, then these sets are said to be *equipotent*, i.e. they have the same number of elements, the same size.



A *cardinal (number)*, then, is a representative for the equivalence class of all sets having the same size, chosen in a canonical fashion. *So cardinals are special sets representing all the sets to which they are equipotent.* If there is a bijection between a set X and a cardinal κ, then we say that X is of cardinality κ, card $=$ κ, alluding to the fact that two distinct cardinals may never be equipotent. In fact, Set Theory proves that there is *at most* one κ for a given set X such that card $= \kappa$. In Set Theoretic contexts where the Axiom of Choice is assumed,"at most one κ" can be replaced by "exactly one κ".

Exercise 5.1.1 Show that equipotency, as introduced above, is an equivalence relation.

We now know how to find out if two sets have the same size; but there needs to be also a way to compare sets with different cardinalities. Our intuitive understanding of a bijective assignment as an abstract representation of counting suggests the idea of injective mappings representing the assignment of a whole set A onto a *part* of a set B (for which we not necessarily have to use the whole of B). So there may be some elements left in B, therefore B must be at least as big as A.



We will agree on the fact that A is at most of the size of B, or A is at most as big as B, if we find an injective mapping from A to B.

Exercise 5.1.2 Turning the above reasoning upside down, what kind of mappings from B to A would in an equally plausible way constitute the fact that B is at least of the size of A?

Exercise 5.1.3 Show that if A and B are equipotent, then both A is at least of the size of B and B is at least of the size of A.

There are a few questions which usually turn up when students are first confronted with this way of counting elements. First, you may wonder in what this method differs from everyday counting methods. Actually, it differs not at all, certainly not as long as we stay within certain limits regarding the nature of the collections we are considering. The step from comparing collections to assigning an abstract representative for the number of its elements can be found both in the process of counting apples in a basket and in the measuring of the cardinality of sets.

Cardinals are not arbitrary as sets, all to the contrary, they are required to satisfy certain requirements we will not further elaborate here. But as a consequence of the rather special form of cardinals, we have to make sure that there are cardinal numbers any kind of sets, even for rather large sets such as the set of natural or real numbers. As a matter of fact, in most mathematical contexts, there is a firm belief in the Axiom of Choice, and it is exactly this axiom which — within the Set Theoretical foundations of Mathematics — provides us with cardinals for *any kind of set*. Therefore, we will never have to worry to run out of cardinals.

Another question might be whether we could ever run out of sets. In other words: Is there an upper bound to the size of sets? The answer is clearly no, and the reason lies in the (axiomatically given) fact that for any set A, the power set $\mathcal{P}(A)$ of A (i.e. the collection of all of A's subsets) is again a set. Moreover, $\mathcal{P}(A)$ can be shown to be strictly larger than A, i.e. there is an injection from A into $\mathcal{P}(A)$, but not vice versa, and clearly there is no bijection between those two sets. The argument to show this involves a technique which can be applied equally well to several other contexts and therefore has a name: The *Diagonal Argument*. Its application to Set Theory goes back to Cantor and can be summed up as follows: If we assume that there is a bijection η between A and $\mathcal{P}(A)$, then we look at the subset D of A given by $D := \{a \in A\,;\ a \notin \eta(a)\}$. Since D is a subset of A, there must be some $d \in A$ with $\eta(d) = D$. But this leads to a contradiction since then $d \in D$ if and only if $d \notin D$.

Exercise 5.1.4 How do we come to this contradiction?

Yet another question: Since we accept the existence of *infinite* sets, how can they be distinguished from *finite* sets? The answer can be given in two ways, which in Set Theory must be distinguished because of their different axiomatic foundations, but which for our purposes can be regarded as equivalent: A set is infinite if it is at least as large as the set of natural numbers. On the other hand, a set is finite if it is not infinite. Alternatively, a set is (Dedekind–) infinite if there is a bijection onto one of its proper subsets. Consequently a set is (Dedekind–) finite if it is not (Dedekind–) infinite.

The finite cardinal numbers bear the same names as their respective natural numbers. The infinite cardinals, on the other hand, are designated by the hebrew symbol $\aleph$ (aleph) with the appropriate index: $\aleph_0$ is the first infinite cardinal, $\aleph_1$ the second, etc. Sets of cardinality $\aleph_0$ are called *countable* or *countably infinite*, a nomenclature which originates from the idea that those sets are of the same cardinality as the set of natural numbers $\mathbb{N}$ and could therefore be numbered or counted by the natural numbers. Infinite sets which are not countable are, of course, uncountable. Just as there are different sizes of finite sets, there are different levels of infinity.

Exercise 5.1.5 Which of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{N}^2$, $\mathcal{P}(\mathbb{N})$ are countable?

We are all quite used to do calculations with finite sets, such as to build the sum / difference / product of their numbers of elements. To be more precise, we are able to calculate the number of elements in constructs such as the set union or direct product using elementary arithmetic on the numbers of elements in the argument sets of the constructions. E.g. for finite sets A and B, the number of elements in the direct product $A \times B$ is the product of the number of elements in A and the number of elements in B,

$$\mathrm{card}(A \times B) = (\mathrm{card}\, A \cdot \mathrm{card}\, B),$$

or the number of elements in a union of sets $A_1, ..., A_n$ is at most the sum of the numbers of elements in the A_i's,

$$\mathrm{card}(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} (\mathrm{card}\, A_i)$$

As for finite sets, there is a way to do calculations with infinite cardinals, but some results are quite different from the ones for finite sets. (The proofs of the following statements can be found in any textbook on Set Theory.)

The sum of two cardinals is defined as the cardinality of the disjoint union of two sets of the cardinality corresponding to the summand cardinals: The sum of 2 and 3 is 5 because the disjoint union of a two-element set and a three-element set is a five-element set. Applying this (not too precise) definition to infinite cardinals, we may find that the disjoint union of two infinite sets can always be mapped bijectively onto the larger of the two sets. Therefore, the sum of two infinite cardinals is always equal to the larger of the two cardinals.This may be generalized to more than two (or even infinitely many) argument sets: If X_n ($n \in \mathbb{N}$) are countably many infinite sets of cardinality $\leq \kappa$, then $\bigcup_{n\in\mathbb{N}} X_n$ and $\prod_{n\in\mathbb{N}} X_n$ again have cardinality $\leq \kappa$. These facts will be used in the next section.

5.2 Cardinality and Languages

The Löwenheim-Skolem Theorems are statements about how the possible size of a model of a theory in some language $\mathcal{L}$ depends on the number of non–logical symbols in this language. To clarify this point, we introduce the notion of the cardinality of a language and that of a model.

Definition 5.2.1 For a formal language $\mathcal{L}$, the **cardinality of** $\mathcal{L}$, notation $\parallel \mathcal{L} \parallel$, is given by

$$\parallel \mathcal{L} \parallel := \mathrm{card}(\mathrm{Fml}\,\mathcal{L}).$$

The motivation for this seemingly arbitrary definition will become clear once the Löwenheim-Skolem Theorems are formulated and proven.

The definition of the cardinality of a model is slightly more straightforward.

Definition 5.2.2 For a structure $\mathcal{A} \in \mathrm{Str}\,\mathcal{L}$, the **cardinality of** $\mathcal{A}$, notation $\mathrm{card}\,\mathcal{A}$, is defined to be the cardinality of the universe of $\mathcal{A}$:

$$\mathrm{card}\,\mathcal{A} := \mathrm{card}\,|\mathcal{A}|.$$

Lemma 5.2.3 If $\mathcal{L}$ is a formal language given by index-sets I, J and K, then

$$\parallel \mathcal{L} \parallel = \max\{\mathrm{card}\,I, \mathrm{card}\,J, \mathrm{card}\,K, \aleph_0\}.$$

Proof. Because we lack the formal apparatus to give a precise proof, we will not go into details here, but give an outline of the reasoning:
An $\mathcal{L}$-formula φ is a finite sequence of symbols, therefore there are at most as many formulae as there are sequences of symbols. We defined a formal language to have countably many variables $v_0, v_1, ...$ (that is why we have to explicitely include $\aleph_0$ in the equation), so the cardinality of the set of terms is at least $\aleph_0$. Since every constant symbol is a term, and, for every function symbol f_j,

$f_j(v_0, ..., v_{\mu(j)-1})$ is also a term, the cardinality of the set of terms is at least

$$\kappa := \operatorname{card}(J \cup K \cup \{v_i; i \in \mathbb{N}\}) = \max\{\operatorname{card} J, \operatorname{card} K, \aleph_0\}.$$[1]

On the other hand, every term is a finite sequence of variables, function symbols and constant symbols (and some auxiliary symbols like brackets), hence there are at most as many terms as there are sequences of this kind. But the cardinality of the set of sequences of variables, constant or function symbols is $\operatorname{card} \bigcup_{n \in \mathbb{N}}((J \cup K \cup \{v_i; i \in \mathbb{N}\})^n)$. From the remarks at the end of the previous section, we know that this is again κ. This proves that the set of terms has cardinality $\max\{\operatorname{card} J, \operatorname{card} K, \aleph_0\}$. Similar arguments applied to the set of $\mathcal{L}$–formulae prove the claim of the lemma. ∎

For the next result, the reader is kindly invited to recall the Completeness Theorem for First Order Logic. One way to formulate this theorem is by the following statement:

Theorem 5.2.4 Every consistent set of sentences has a model.

The Downward Löwenheim–Skolem Theorem may be seen as the following sharpening of the statement of Theorem 5.2.4:

Theorem 5.2.5 (Extended Completeness Theorem) If $\Sigma \subseteq \operatorname{Sen} \mathcal{L}$ is consistent, then there is a model $\mathcal{A}$ of Σ such that $\operatorname{card} \mathcal{A} \leq \operatorname{card} \mathcal{L}$.

Instead of proving the Extended Completeness Theorem 5.2.5 directly, we are going to analyze the steps necessary to verify the Completeness Theorem 5.2.4 under the aspect of cardinality. Such a proof in most cases consists of the following steps:

1. [**Max**] Show that every consistent set Σ of $\mathcal{L}$-sentences can be expanded to a maximally consistent set of sentences.
2. [**Wit**] Show that every consistent set Σ of $\mathcal{L}$-sentences can be expanded to a consistent set Σ' of $\mathcal{L}'$-sentences, where $\mathcal{L}'$ extends $\mathcal{L}$ by adding new constants such that every existential sentence has a witness.
3. [CT $\mathcal{L}'/\Sigma$] The term-structure of $\mathcal{L}'$ modulo Σ is a model for the original consistent set Σ of $\mathcal{L}$-sentences.

Let us now analyze these steps under the aspect of cardinality. Moreover, let us assume that we start with a language $\mathcal{L}$ that has cardinality κ.

[1] Without loss of generality, we assume the sets I, J, K and $\{v_i; i \in \mathbb{N}\}$ to be pairwise disjoint.

1. Clearly, every (consistent or not) set Σ of $\mathcal{L}$-sentences has a cardinality at most κ since a set of sentences is a subset of the set of formulae. The same holds for a maximally consistent set of $\mathcal{L}$-sentences. Therefore, in step (Max), we do not exceed the cardinality κ given by the language.

2. This step consists of an infinite number of expansions $\mathcal{L}_i$ $(i \in \mathbb{N})$ of the language $\mathcal{L}$. If $\mathcal{L}_i$ is given, $\mathcal{L}_{i+1}$ is constructed by adding a new constant symbol $c_{\exists x\varphi}$ for every existential sentence $\exists x\varphi \in \operatorname{Sen}\mathcal{L}$, so

$$\| \mathcal{L}_{i+1} \| = \max\{\operatorname{card} I, \operatorname{card} J, \operatorname{card}(K_i \cup \{c_{\exists x\varphi}; \exists x\varphi \in \operatorname{Sen}\mathcal{L}_i\}), \aleph_0\},$$

   which by simple cardinal arithmetic is easily seen to be κ. (The exact nature of the sets Σ_{i+1} of $\mathcal{L}_{i+1}$-sentences is of no importance to our considerations since these sets are clearly sets of formulae whose cardinality is bounded by κ as well. However, please note that the languages $\mathcal{L}_i$ differ only in their sets of constant symbols, and that $\operatorname{Fml}\mathcal{L}_{i+1} \supseteq \operatorname{Fml}\mathcal{L}_i$.)

   Finally, $\mathcal{L}'$ is defined to be $\mathcal{L}$ expanded by the sets of new constant symbols introduced in the constructions of the $\mathcal{L}_i$'s. From this, $\| \mathcal{L}' \|$ can easily be seen not to exceed κ since it is defined to be the cardinality of a union of countably many sets whose cardinality is κ,

$$\| \mathcal{L}' \| = \operatorname{card}(\bigcup_{i\in\mathbb{N}} \| \mathcal{L}_i \|) \leq \operatorname{card}(\aleph_0 \times \kappa) = \kappa.$$

   From this we see that also the constructions in this step do not cross the cardinality boundary set by $\mathcal{L}$.

3. By definition, the term–structure $\operatorname{CT}\mathcal{L}'/\Sigma'$ has as universe the set $\operatorname{CT}\mathcal{L}'$ of closed terms of $\mathcal{L}'$ modulo the equivalence relation $\approx_{\Sigma'}$ given by

$$t_1 \approx_{\Sigma'} t_2 \text{ iff } \Sigma' \vdash t_1 \doteq t_2,$$

   i.e. an element of $\operatorname{CT}\mathcal{L}'/\Sigma'$ is an $\approx_{\Sigma'}$–equivalence class.
   Since $t \doteq t \in \operatorname{Fml}\mathcal{L}'$ for every $t \in \operatorname{CT}\mathcal{L}'$, we find that

$$\operatorname{card}(\operatorname{CT}\mathcal{L}') \leq \operatorname{card}(\operatorname{Fml}\mathcal{L}')$$

   and because the assignment $t \mapsto [t]_{\approx_{\Sigma'}}$ defines a surjection from $\operatorname{CT}\mathcal{L}'$ onto $\operatorname{CT}\mathcal{L}'/\Sigma'$,

$$\operatorname{card}(\operatorname{CT}\mathcal{L}'/\Sigma') \leq \operatorname{card}(\operatorname{CT}\mathcal{L}').$$

Figure 5.3: Thoralf Albert Skolem (1887–1963)

We conclude

$$\mathrm{card}(\mathrm{CT}\,\mathcal{L}'/\Sigma') \leq \mathrm{card}(\mathrm{CT}\,\mathcal{L}') \leq \mathrm{card}(\mathrm{Fml}\,\mathcal{L}') = \parallel \mathcal{L}' \parallel = \kappa.$$

In other words, the constructed model's cardinality is at most κ.

5.3 The Downward Löwenheim–Skolem Theorem

Let us now combine and apply the results from the previous section.

Theorem 5.3.1 (Downward Löwenheim–Skolem Theorem)
If $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ has a model, then Σ has a model whose cardinality is at most $\parallel \mathcal{L} \parallel$.

Proof. This is nothing else than Theorem 5.2.5 which we just proved. ■

The following special case of Theorem 5.3.1 is merely of historical interest, since it is thus that the Downward Löwenheim–Skolem Theorem was originally formulated.

Corollary 5.3.2 If $\mathcal{L}$ is countable (finite or infinite), then $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ has a model iff Σ has a countable model.

The proofs of the following direct consequences of the Downward Löwenheim–Skolem Theorem 5.3.1 are left as exercises.

Corollary 5.3.3 If $\mathcal{L}$ is countable (finite or infinite), then Σ has a countable model.

Corollary 5.3.4 If a sentence $\alpha \in \mathrm{Sen}\,\mathcal{L}$ has any model at all, then α also has a countable model.

Corollary 5.3.5 if $\mathcal{A} \in \operatorname{Str} \mathcal{L}$, then there is a $\mathcal{B} \in \operatorname{Str} \mathcal{L}$ with $\mathcal{B} \equiv \mathcal{A}$ and $\operatorname{card} \mathcal{B} \leq \| \mathcal{L} \|$.

Corollary 5.3.6 A (consistent) theory which has only uncountable models cannot be axiomatized by countably many axioms or in a countable language.

The following corollaries present more example–related results (the proofs of which are omitted since they follow simply from having a closer look at the respective languages): Remember that ZFC is the Set Theory of Zermelo–Fränkel with the Axiom of Choice formulated in the (strikingly simple) language consisting of nothing but the binary relation–symbol $\in$.

Corollary 5.3.7 If ZFC is consistent, then ZFC has a countable model.

Unfortunately, the premise cannot be omitted, since it is not known whether ZFC is consistent or not. But still, since a whole generation of mathematicians successfully develop their branch in the world based on Set Theoretic notions and thus on ZFC, we may, at least for the moment, assume that ZFC is consistent and later, if ever we should stumble over an inconsistency, try to fix the leak. So, if we plant our beliefs in this ground, the above corollary states that there is a countable model for ZFC; that is, there is a set (!) $\mathcal{M}$ equipped with a binary relation $\tilde{\in}$ which mirrors the whole universe of sets, with the "element of"–relation $\in$ modelled by the binary relation $\tilde{\in}$. As elements of this set, we find *all ordinals, cardinals, sets we can prove to exist according to* ZFC and the like. Among other, we may therefore even find $\mathcal{M}$ itself inside this model $\mathcal{M}$. But this would imply that $\mathcal{M}$ cannot possibly be a set in the sense of ZFC. This seeming paradoxon is resolved with the realization that, e.g., *being uncountable* in the model $\mathcal{M}$ is not the same as being uncountable in the universe of sets; alternatively, if you prefer a less platonistic point of view, the meaning of *being uncountable* depends on the context — or model — this expression is interpreted in. For more on this subject, we refer the reader to any of the books on (axiomatic) Set Theory mentioned in the bibliography.

Corollary 5.3.8 There exists a countable algebraically closed field (as a sub–ring of $\mathbb{C}$).

Corollary 5.3.9 There exists a countable infinite Boolean algebra.

The first of these results confronts us with the fact that, algebraically speaking, the field of complex numbers is far too big for the purpose of solving any given equation. It is also noteworthy that the witnesses added in the iterative process of extending a given subset are solutions of equations, and the witness–extentions are algebraic extensions.

Figure 5.4: Alfred Tarski (1902–1983)

For the second example, the astonishing fact is that we easily find examples of *finite* or *uncountable* boolean algebras by looking at the powerset of some set. For a *countable infinite* boolean algebra, we cannot proceed the same way, since for finite sets, the powerset is finite, and for infinite sets, the powerset is uncountable. But we know by the main theorem of this section that a countable boolean algebra must exists.[2]

5.4 The Upward Löwenheim–Skolem Theorem

In this section, we are going lo look in the opposite direction of what we have just arrived at: While the Downward Löwenheim–Skolem Theorem shows the existence of *small* models (where of course the exact meaning of *small* depends on the formal language that is used for the formalization), the Upward Löwenheim–Skolem Theorem will prove the existence of *arbitrarily large models*, under certain provisos. Once we have proved this upward variant, there will be two main conclusions to be drawn from this:

- For every first–order theory there are *non–standard models*, provided this theory has at least one infinite model; e.g. there are uncountable models of Paeano–arithmetic.
- First–order formal languages are *incapable of expressing cardinalities outsizeing their own cardinality.* This is exemplified by the fact that when considering $\mathbb{Q}$ as a structure for the language having as sole non–logical symbol the binary relation–symbol $\leq$, $\mathrm{Th}\,\mathbb{Q}$ has uncountable models; the (rather poor, in terms of expressive power) language of order is unable to distinguish countable from uncountable.

[2] A direct approach to this is to show that $\{s \in \mathbb{N}; s \text{ finite or } \mathbb{N} \setminus s \text{ finite}\}$ is a countable infinite boolean algebra. But there are more refined variants of this result which no longer can be verified by such a direct construction, i.e. the existence of an atom–free countable boolean algebra.

We have to face the question about a method to *force* structures to exceed certain cardinalities. As in the previous section, the tool to ensure great sizes of models is the language itself (cf. 4.2.2 for the definition of $\mathcal{L}_C$):

Lemma 5.4.1 Let $\mathcal{L}$ be a formal language, and let $C := \{c_\beta \; ; \; \beta < \kappa\}$ be a set of cardinality κ of (pairwise distinct) constant–symbols (which may or may not belong to $\mathcal{L}$). Let $\mathcal{L}_C$ be the formal language $\mathcal{L}$ with the elements of C added as constants (if necessary), and let $\Gamma \subseteq \operatorname{Sen} \mathcal{L}_C$ be given by $\Gamma := \{\neg c \doteq c' \; ; \; c, c' \in C, \; c \neq c'\}$. Then for any $\mathcal{L}_C$–structure $\mathcal{A}$, if $\mathcal{A} \models \Gamma$, then $\operatorname{card} \mathcal{A} \geq \kappa$.

Proof. Exercise. ∎

Proposition 5.4.2 Let $\mathcal{L}$ be a formal language and $\Sigma \subseteq \operatorname{Sen} \mathcal{L}$. Moreover, let C be an infinite set of constant–symbols not in $\mathcal{L}$. Then

$$\Sigma \cup \{\neg c \doteq c' \; ; \; c, c' \in C, \; c \neq c'\} \text{ has a model iff } \Sigma \text{ has a model.}$$

Proof. Let $\Gamma := \{\neg c \doteq c' \; ; \; c, c' \in C, \; c \neq c'\}$.
If $\Sigma \cup \Gamma$ has a model, say $\mathcal{A}$, then by lemma 5.4.1, $\operatorname{card} \mathcal{A} \geq \operatorname{card} C$, so $\mathcal{A}$ is an infinite model, and it is clearly also a model of Σ.
Conversely, if Σ has an infinite model $\mathcal{A}$, then for any finite $\Gamma_0 \subseteq \Gamma$ and any finite $\Sigma_0 \subseteq \Sigma$, $\mathcal{A} \models \Sigma_0 \cup \Gamma_0$, where the new constant–symbols in Γ_0 are interpreted as arbitrary, pairwise distinct elements of $\mathcal{A}$. It follows that any finite subset of $\Sigma \cup \Gamma$ has a model, so by compactness, $\Sigma \cup \Gamma$ itself has a model. ∎

As an exercise, analyze the proof of Proposition 5.4.2 to find out why the model of Σ has to be infinite.

Theorem 5.4.3 (Upward Löwenheim–Skolem Theorem)
Let $\mathcal{L}$ be a formal language and $\mathcal{A} \in \operatorname{Str} \mathcal{L}$ be infinite. Then, for any $\kappa \geq \| \mathcal{L} \|$ there is an $\mathcal{L}$–structure $\mathcal{B}$ with $\mathcal{B} \equiv \mathcal{A}$ and $\operatorname{card} \mathcal{B} = \kappa$.

Proof. Let $\Sigma := \operatorname{Th} \mathcal{A}$, let C be a set of κ distinct new constant–symbols not in $\mathcal{L}$, and $\Gamma := \{\neg c \doteq c' \; ; \; c, c' \in C, \; c \neq c'\}$.
By Proposition 5.4.2, $\Sigma \cup \Gamma$ has a model. Thus, by the Downward Löwenheim–Skolem Theorem 5.3.1, $\Sigma \cup \Gamma$ has a model $\mathcal{B}$ with $\operatorname{card} \mathcal{B} \leq \operatorname{card}(C \cup \| \mathcal{L} \|) = \kappa$, and from Lemma 5.4.1, we know that $\operatorname{card} \mathcal{B} \geq \kappa$. Therefore $\operatorname{card} \mathcal{B} = \kappa$, and clearly $\mathcal{B} \equiv \mathcal{A}$ because $\mathcal{B} \models \operatorname{Th} \mathcal{A}$. ∎

Corollary 5.4.4 No theory has exclusively countable infinite models.

In the context of axiomatizations of $\mathbb{N}$, this implies that if we find the classical countable infinite model as a model of any axiomatization, we also find models of any uncountable infinite cardinality! Clearly these models no longer qualify as *classical* models since they simply have too many elements.

Corollary 5.4.5 There is no theory such that any two infinite models are isomorphic.

Proof. ... because "isomorphic" clearly implies "of the same cardinality". ■

Exercise 5.4.6 Show that there is a set Σ of $\mathcal{L}$–sentences (for an adequate language $\mathcal{L}$) such that $\mathcal{A}$ is a model for Σ iff $\mathcal{A}$ has either exactly one or infinitely many elements.
(Hint: Consider $\Sigma := \{(\forall x \forall y\ x \dot{=} y) \vee \neg c \dot{=} c'; c, c' \in C, c \neq c'\}$ for adequate C.)

We conclude this section by hinting at the possibility of an alternative, more semantically oriented proof of the Upward Löwenheim–Skolem Theorem 5.4.3. But for the moment, however, we do not have the necessary tools to perform this aproach and therefore postpone it to Section 7.4.

Chapter 6

Theories

Since both the definition of “elementary” as well as that of “theory” rely on the operators Th and Mod, we are confronted with a syntactical notion defined in semantical terms and a semantical notion defined in syntactical terms. Naturally, one might ask if these circumnavigations are necessary. Hence, we will look for more direct ways of expressing that some set of $\mathcal{L}$–sentences is a theory or some class of $\mathcal{L}$–structures is elementary, allthewhile we will remain in the realm the respective notion belongs to. In Chapter 8 we will see that, contrary to the seeming symmetry of the definitions, the complexities of the two tasks differ considerably; the one concerning theories is rather simple and direct, while the characterization of elementary classes involves non–trivial constructions and results.

6.1 Theories and Complete Theories

A generally accepted way to represent the fact that a sentence α is deducible from the empty set of premises is

$$\vdash \alpha,$$

thereby insinuating that, in order to prove α, we do not have to make any additional assumptions whatsoever. We thus see that very small sets of premises may have considerable logical implications. The general rule is: “More assumptions imply more conclusions”. However, there is always the possibility that a set Σ of sentences already contains every sentence which might be deduced from it. Such a set of sentences is said to be deductively closed.

Definition 6.1.1 Let $\mathcal{L}$ be a formal language.

(i) For $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$, define $\operatorname{Ded}\Sigma$, the **deductive closure** of Σ, to be the set of all sentences deducible from Σ,

$$\operatorname{Ded}\Sigma := \{\alpha \in \operatorname{Sen}\mathcal{L}\,;\, \Sigma \vdash \alpha\}\ .$$

(ii) $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$ is called **deductively closed** iff $\Sigma = \operatorname{Ded}\Sigma$.

From Definition 6.1.1 it follows immediately that $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$ is deductively closed if and only if for any $\alpha \in \operatorname{Aus}(L)$,

$$\Sigma \vdash \alpha \text{ iff } \alpha \in \Sigma.$$

Let $\mathcal{L}$ be a formal language, $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$, and consider an $\mathcal{L}$–sentence α with

$$\alpha \in \operatorname{Th}\operatorname{Mod}\Sigma.$$

The definition of Th (cf. Definition 4.1.2) implies that $\mathcal{A} \models \alpha$ for any $\mathcal{A} \in \operatorname{Mod}\Sigma$, and with the definition of Mod, we find that $\mathcal{A} \models \Sigma \Longrightarrow \mathcal{A} \models \alpha$ for any $\mathcal{A} \in \operatorname{Str}\mathcal{L}$, i.e. $\Sigma \models \alpha$.

Gödel's Completeness Theorem 2.4.1 states that this last statement is equivalent to $\Sigma \vdash \alpha$, which is the same as $\alpha \in \operatorname{Ded}\Sigma$.

This hints at the way to prove the following Lemma, which reformulates Gödel's Completeness Theorem by using the operators Mod, Th and Ded.

Lemma 6.1.2 For $\Sigma \subseteq \operatorname{Aus}(L)$,

$$\operatorname{Th}\operatorname{Mod}\Sigma = \operatorname{Ded}\Sigma.$$

Proof. Exercise. ■

Theorem 6.1.3 (Theories)
$\Sigma \subseteq \operatorname{Sen}\mathcal{L}$ is a theory iff Σ is deductively closed.

Proof. Follows immediately from Lemma 6.1.2. ■

Thus, the Theorem 6.1.3 characterizes theories as the fixed points of the operator Ded, and thereby provides a definition of *theories* via the purely syntactical notion of "deductively closed".

Example 6.1.4

(i) The "smallest" $\mathcal{L}$-theory for any formal language $\mathcal{L}$ is obtained from $\Sigma = \emptyset$. It follows that

$$\operatorname{Th}\operatorname{Mod}\emptyset = \operatorname{Ded}\emptyset = \text{“the set of all theorems of } \mathcal{L}\text{”} \neq \emptyset$$

(ii) On the other hand, we can see that the "largest" $\mathcal{L}$-theory is $\operatorname{Sen}\mathcal{L}$, where

$$\operatorname{Th}\operatorname{Mod}\operatorname{Sen}\mathcal{L} = \operatorname{Th}\emptyset = \operatorname{Sen}\mathcal{L}.$$

Please keep in mind that the smallest theory is *not* the empty set. We can, however, say that the smaller the theory, the larger the associated elementary class of models, and vice versa. The largest theory $\operatorname{Sen}\mathcal{L}$ is special in the sense that it is inconsistent, and, moreover, that it is the *only* inconsistent $\mathcal{L}$–theory (Proof: Exercise). Thus, it is only natural to ask if there are any maximal theories among the consistent ones. Due to the contra–variant behaviour[1] of the operators Th and Mod, we must have a closer look at the theories of small classes of models which are yet not too small (i.e. not empty).

Definition 6.1.5 $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$ is called **complete** iff $\Sigma = \operatorname{Th}\{\mathcal{A}\}$ for a single $\mathcal{L}$-structure $\mathcal{A}$.

The proof of the following Lemma is a simple exercise.

Lemma 6.1.6 Complete sets of sentences are theories.

The semantic characterization of complete sets of $\mathcal{L}$–sentences is as simple (and accurate, as will be shown below!) as it can be, i.e. complete sets of $\mathcal{L}$–sentences are characterized by a single $\mathcal{L}$-structure. The syntactical characterization is, among other results, provided by the next theorem. For further applications, statement (v) is of special importance.

Theorem 6.1.7 (Complete theories)
For a consistent theory Σ, the following statements are equivalent:

(i) Σ is complete.

(ii) Σ is maximally consistent.

(iii) For any $\alpha \in \operatorname{Sen}\mathcal{L}$: either $\Sigma \vdash \alpha$ or $\Sigma \vdash \neg\alpha$ (but not both).

(iv) For any $\mathcal{A}, \mathcal{B} \in \operatorname{Mod}\Sigma$: $\mathcal{A} \equiv \mathcal{B}$.

(v) $\Sigma = \operatorname{Th}\{\mathcal{A}\}$ for any $\mathcal{A} \in \operatorname{Mod}\Sigma$.

Proof (Excerpt). *(i) ⇒ (ii)*: Consider $\Sigma = \operatorname{Th}\{\mathcal{A}\}$ and $\alpha \notin \Sigma$. Then, $\mathcal{A} \not\models \alpha$, so $\mathcal{A} \models \neg\alpha$, but then $\Sigma \cup \{\alpha\}$ is inconsistent. Thus, we have shown that any set

[1] I.e. $\Sigma_1 \subseteq \Sigma_2 \Longrightarrow \operatorname{Mod}\Sigma_1 \supseteq \operatorname{Mod}\Sigma_2$; similarly for Th.

of $\mathcal{L}$–sentences expanding Σ is inconsistent, thus, Σ is maximally consistent.

(ii) $\Rightarrow$ *(iii)*: If $\Sigma \nvdash \alpha$, then, thanks to the maximal consistency of Σ we have $\neg\alpha \in \Sigma$ and so clearly $\Sigma \vdash \neg\alpha$. ■

Exercise 6.1.8 i Find proofs for the remaining implications in Theorem 6.1.7 .

ii Where is the consistency of Σ being used? Which implications still hold if consistency is dropped as a premise?

The following examples are presented without proofs:

Example 6.1.9 (Algebraically closed fields of characteristic 0)
Let $\mathcal{L} = \{+, -, \cdot, 0, 1\}$ be the formal language for rings and fields.

Let F be an algebraically closed field[2] of characteristic[3] 0. It is possible (though not for us, at the moment) to show that $\mathrm{Th}(F) = \mathrm{Th}(\mathbb{C})$. Thus, the theory of algebraically closed fields of characteristic 0 is complete, and any algebraically closed field of characteristic 0 is elementary equivalent to the field of complex numbers $\mathbb{C}$. In other words, to verify that a sentence α in the language $\mathcal{L}$ of rings holds for all algebraically closed fields of characteristic 0, we just have to show that $\mathbb{C}$ is a model of α (a task which, depending on the complexity of α, may involve quite a lot of functional calculus).

Example 6.1.10 (Order theories)
(For the less algebraically inclined.) Let $L_0 = \{<\}$, i.e. L_0 contains $<$ as the only non–logical symbol. Consider the set Σ :

α_1:	$\forall x$	$\neg(x < x)$	Irreflexivity
α_2:	$\forall x, y, z$	$(x < y \land y < z \rightarrow x < z)$	Transitivity
α_3:	$\forall x, y$	$(x < y \lor x \doteq y \lor y < x)$	Totality or linearity
α_4:	$\forall x, y$	$(x < y \rightarrow \exists z\, (x < z \land z < y))$	Density
α_5:	$\forall x \exists y, z$	$(y < x \land x < z)$	No endpoints

An $\mathcal{L}$–structure $\mathcal{A} \in \mathrm{Mod}\,\Sigma$ is called a **d**ense **o**rder **w**ithout **e**ndpoints, and we thus designate the theory of $\mathrm{Mod}\,\Sigma$ by Σ_{DOWE}.

It is quite plausible that both $\mathbb{Q}$ and $\mathbb{R}$ are dense orders without endpoints. Clearly, though, they are quite different; more precisely, they are not isomorphic, since $\mathbb{Q}$ is countable while $\mathbb{R}$ is not. However, with the expressive power of $\mathcal{L}$,

[2] I.e. in F, every polynom $f(x)$ can be written as a product of linear and constant polynoms: $f(x) = c(x - \xi_1) \cdots (x - \xi_n)$. From earlier experiences in elementary math, you probably remember that this is one of the most prominent features of the complex numbers.

[3] The characteristic of a field is the order of 1, the neutral element of multiplication, in the (additive!) group $\langle F, +, 0\rangle$. Thereby "characteristic 0" stands for "infinite characteristic". Note that a finite field always has characteristic different from 0, whereas the converse is false.

we are indeed unable to tell them from one another, since it can be shown (cf. Section 6.2) that $\Sigma_{\mathrm{DOWE}} = \mathrm{Th}(\mathbb{Q})$. Especially, Σ_{DOWE} is complete and any two models of Σ_{DOWE} are thus elementary equivalent.

Exercise 6.1.11

i Verify that a set Σ of sentences having both finite and infinite models is never complete.

ii Verify that a set Σ of sentences having finite models of different cardinalities is never complete.

iii Show that this does not hold for Σ having infinite models of different cardinalities.

Let us now have a look at a nice property of complete theories: If one complete theory is included in the other, then they are the same!

Corollary 6.1.12 i For complete theories $\Sigma, \Theta \subseteq \mathrm{Sen}\,\mathcal{L}$:

$$\Sigma \subseteq \Theta \text{ iff } \Sigma = \Theta.$$

ii For $\mathcal{L}$–structures $\mathcal{A}, \mathcal{B}$:

$$\mathrm{Th}\,\mathcal{A} \subseteq \mathrm{Th}\,\mathcal{B} \text{ iff } \mathcal{A} \equiv \mathcal{B}.$$

Proof.

i Let $\Sigma, \Theta \subseteq \mathrm{Sen}\,\mathcal{L}$ be complete and $\Sigma \subseteq \Theta$. Let $\alpha \in \mathrm{Sen}\,\mathcal{L} \setminus \Sigma$. Then, $\neg\alpha \in \Sigma$, thus $\neg\alpha \in \Theta$ and thus $\alpha \notin \Theta$. Therefore $\Sigma = \Theta$.

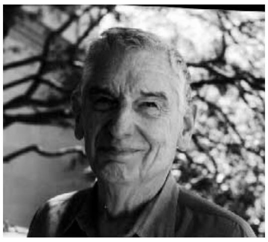
ii Almost too easy to be an exercise.

■

6.2 Proving Completeness of Theories: An Example

In this section we intend to give a proof of the completeness of the *theory of dense orders without endpoints*, Σ_{DOWE}. This theory is axiomatized by a finite set $\{\alpha_1, \ldots, \alpha_5\}$ of axioms in the language $\mathcal{L}$ having the binary relation symbol $\leq$ as sole non–logical component. The proof of the completeness of Σ_{DOWE} roughly consists states and uses a corollary of the upward Löwenheim–Skolem

Theorem and then shows that, up to isomorphism, there is only one countable dense order without endpoints. Proving the latter statement, we will introduce a technique widely used in model theory, the *back–and–forth–construction*.



Robert Lawson Vaught (1926–2002)

Corollary 6.2.1 (Łoś-Vaught Test) Assume $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$ has only infinite models and, for some cardinal $\kappa \geq \max\{\aleph_0, \mathrm{card}\,\mathcal{L}\}$, any two models of Σ having cardinality κ are isomorphic. Then Σ is complete.

Proof. Let $\mathcal{A}$ and $\mathcal{B}$ be any two models of Σ. Then by the Löwenheim–Skolem Theorems (either upward or downward, depending on the cardinality of $\mathcal{A}$ and $\mathcal{B}$), we find models $\mathcal{A}'$ and $\mathcal{B}'$ having cardinality κ such that $\mathcal{A} \equiv \mathcal{A}'$ and $\mathcal{B} \equiv \mathcal{B}'$. But then $\mathcal{A}' \cong \mathcal{B}'$ by assumption, whence $\mathcal{A}' \equiv \mathcal{B}'$, and we conclude

$$\mathcal{A} \equiv \mathcal{A}' \equiv \mathcal{B}' \equiv \mathcal{B}.$$

So we showed that any two models of Σ are elementary equivalent, which is equivalent to Σ being complete. ■

As we mentioned earlier, elementary equivalence is — as a criterion of identification — coarser than isomorphism. In the following we will show that, up to isomorphism and thus also up to elementary equivalence, there is only one countable dense order without endpoints. Using Vaught's Test we may thus conclude that Σ_{DOWE} is complete and, $\langle \mathbb{Q}, \leq \rangle$ being a model of Σ_{DOWE}, any dense order without endpoints is elementary equivalent to $\langle \mathbb{Q}, \leq \rangle$. But we do not claim that $\mathbb{R}$ and $\mathbb{Q}$ are *isomorphic* as dense orders without endpoints (they are not of the same cardinality), nor do we claim that any two *uncountable* dense orders without endpoints, *even if they have the same cardinality*, are isomorphic.

From 6.1.10, recall the axioms for the theory Σ_{DOWE} of dense orders without endpoints.

Proposition 6.2.2 Any two countable models of Σ_{DOWE} are isomorphic.

Proof. The proof is done using the **back–and–forth–construction** which delivers an isomorphism via (in the general case *transfinite*) induction.

Suppose $\mathcal{A}$ and $\mathcal{B}$ are countable models of Σ_{DOWE}. So we find enumerations[4] $\langle a_n \,;\, n \in \mathbb{N}\rangle$ of $|\mathcal{A}|$ and $\langle b_n \,;\, n \in \mathbb{N}\rangle$ of $|\mathcal{B}|$. We will now define new enumerations $\langle a'_n \,;\, n \in \mathbb{N}\rangle$ of $|\mathcal{A}|$ and $\langle b'_n \,;\, n \in \mathbb{N}\rangle$ of $|\mathcal{B}|$ such that the assignment $a'_n \mapsto b'_n$ is order–preserving and thus defines an isomorphism from $\mathcal{A}$ to $\mathcal{B}$. This again is done by a countable variant of *transfinite induction*:

Assume that for any $i, j < n$, $a'_i < a'_j$ iff $b'_i < b'_j$. We distinguish the following two cases:

- if n is even, say $n = 2m$, we define $a'_n := a_k$ such that k is minimal with $a_k \notin \{a'_0, \ldots, a'_{n-1}\}$. Then we have to consider the following three subcases:
  - $a'_n < a'_i$ for all $i < n$, in which case we choose b'_n such that $b'_n < b'_i$ for all $i < n$, which can be achieved since $\mathcal{B}$ has no endpoints; or
  - $a'_n > a'_i$ for all $i < n$, in which case we choose b'_n such that $b'_n > b'_i$ for all $i < n$, which can be achieved for the same reason; or
  - $a'_i < a'_n < a'_{i+1}$ for some $i < n$, in which case we choose b'_n such that $b'_i < b'_n < b'_{i+1}$ for all $i < n$, which can be achieved since $\leq$ is dense on $\mathcal{B}$.
- if n is odd, say $n = 2m+1$, we define $b'_n := b_k$ such that k is minimal with $b_k \notin \{b'_0, \ldots, b'_{n-1}\}$. Again we are to face three subcases which are dealt with as above, switching the a's and b's and $\mathcal{A}$ and $\mathcal{B}$ respectively.

To show that the assignment $a'_n \mapsto b'_n$ does indeed define an isomorphism between $\mathcal{A}$ and $\mathcal{B}$ (i.e. an order–preserving bijection with order–preserving inverse) is left as an exercise. ■

As an illustration, let's have a look at how this back–and–forth mechanism could work in detail. Suppose we are given an enumeration $\langle q_n \,;\, n \in \mathbb{N}\rangle$ of $\mathbb{Q}$ and an enumeration $\langle b_n \,;\, n \in \mathbb{N}\rangle$ of $\mathcal{B} = \langle (0,1) \cap \mathbb{Q}, \leq\rangle$. Note that $\mathcal{B}$ thus defined is a countable models of Σ_{DOWE}. The construction starts at $n = 0$, where we are to take q_0 and assign it to any b_i (no constraints so far). So let's say $q'_0 = q_0 = 0$, and for b'_0 we chose b_{220364}, which happens to be, say, $5/13$. The next step ($n = 1$) goes from $\mathcal{B}$ to $\mathbb{Q}$. So we are to take for b'_1 the b_i with smallest i not yet dealt with, which is clearly b_0, say $19/69$. Since $b'_1 < b'_0$, we must take for q'_1 some rational number smaller that q'_0, so let's say $q'_1 = -1$, which happens to be some q_i, say q_1, by some fancy coincidence. The next step is $n = 2$, so we start with $\mathbb{Q}$ again, taking for q'_2 the q_i with smallest i not yet treated, which happens to be q_2, say $q_2 = 1$. Then $q'_2 > q'_0, q'_1$, so we have to find a b'_2 which is

[4] An enumeration of a set X is a bijection from $\mathbb{N}$ onto X, i.e. a map from $\mathbb{N}$ to X, which justifies to write the enumeration as $\langle x_n \,;\, n \in \mathbb{N}\rangle$.

strictly greater than both b'_0 and b'_1, which could be 37/95 or any other rational number strictly between 5/13 and 1. And so on ... .

Exercise 6.2.3 Run the first few steps with your favorite enumerations of $\mathbb{Q}$ and $(0,1) \cap \mathbb{Q}$. (If you do not have a favorite enumeration of $(0,1) \cap \mathbb{Q}$, take the enumeration of $\mathbb{Q}$ and restrict it to the interval $(0,1)$.)

The above proof may be paraphrased as follows: A countable dense order without endpoints has an overall similarity in the sense that any non–empty open sub–interval looks exactly like the set itself, so the same holds for any two open sub–intervals. By assigning the first element of $\mathcal{A}$ to some b'_0, we divide both $\mathcal{A}$ and $\mathcal{B}$ into two sub–intervals which have this mutual similarity. But then we may continue, using the same argument for these sub–intervals, and so on. In a way you could say that dense orders without endpoints have a flexibility which allows them to be stretched without altering the structural information.

Now, for the case of uncountable dense orders without endpoints, we consider the following example:

Example 6.2.4 Let A be the subset of $\mathbb{R}$ given by $A := \{\zeta \in \mathbb{R}\,;\, 0 < \zeta < 1\} \cup \{q \in \mathbb{Q}\,;\, 1 \leq q < 2\}$. Then

- A is uncountable since it is the union of a countable and an uncountable set;
- $\mathcal{A} = \langle A, \leq \rangle$ is a dense order without endpoints since the axioms $\alpha_1, \ldots, \alpha_5$ are satisfied in $\mathcal{A}$ (Exercise: verify this!);
- but $\mathcal{A}$ is not isomorphic to $\mathbb{R}$! To see this, first note that an isomorphism, being a homomorphism after all, is bound to map intervals onto intervals since it preserves the order $\leq$; now, assuming $\eta : \mathbb{R} \longrightarrow \mathcal{A}$ is onto and order–preserving, we find a $\zeta_0 \in \mathbb{R}$ such that $\eta(\zeta_0) = 3/2$. But then η must map the (open) interval $\{\zeta \in \mathbb{R}\,;\, \zeta_0 < \zeta\}$ to the set $\{q \in \mathbb{Q}\,;\, 3/2 < q < 2\}$. But then ζ cannot be 1–1 since it maps an uncountable set onto a countable one.

Exercise 6.2.5 An example somewhat different would be to let $A := \mathbb{R} \setminus \{0\}$ and to show that $\langle \mathbb{R}, \leq \rangle$ and $\mathcal{A} = \langle A, \leq \rangle$ are not isomorphic. Does that work? How about $\langle \mathbb{Q}, \leq \rangle$ and $\langle A \cap \mathbb{Q}, \leq \rangle$?

Chapter 7

Ultraproducts

In this chapter we will take our first steps towards the definition of a technique for building new models. This technique will unite the two tools of *products* and *quotients* of structures, notions which readers are familiar with from earlier experiences in algebra. Combining these notions, we will introduce *ultraproducts*, which provide models for theories that are no longer elementary equivalent to the models the construction is based upon.

The results we aim at are not basic and require some definitional input, which we give in the following section.

7.1 Ultrafilters

For the following section, readers should recall the definition of cartesian products and their associated projections from Section 1.3).

Definition 7.1.1 Let S be a set and, for $s \in S$, let $\mathcal{A}_s$ be an $\mathcal{L}$–structure. Then, the **direct product** $\prod_{s\in S}\mathcal{A}_s$ is defined as the following $\mathcal{L}$–structure:

- The universe is the cartesian product of the universes, i.e.

$$|\prod_{s\in S}\mathcal{A}_s| := \prod_{s\in S}|\mathcal{A}_s|$$

- The relation–, function– and constant–symbols are interpreted by component, i.e., writing $\mathcal{B}$ for $\prod_{s\in S}\mathcal{A}_s$,

  for all relation–symbols R_i and all $a_1,\ldots,a_{\lambda(i)} \in |\prod_{s\in S}\mathcal{A}_s|$,

$$R_i^{\mathcal{B}}(a_1,\ldots,a_{\lambda(i)}) \text{ iff, for all } s \in S,\ R_i^{\mathcal{A}_s}(\pi_s(a_1),\ldots,\pi_s(a_{\lambda(i)}));$$

for all function–symbols f_j and all $a_1, \ldots, a_{\mu(j)} \in |\prod_{s\in S} \mathcal{A}_s|$,

$$f_j^{\mathcal{B}}(a_1, \ldots, a_{\mu(j)}) = \langle f^{\mathcal{A}_s}(\pi_s(a_1), \ldots, \pi_s(a_{\mu(j)})) \; ; \; s \in S\rangle; \text{ and,}$$

for all constant–symbols c_k,

$$c_k^{\mathcal{B}} = \langle c_k^{\mathcal{A}_s} \; ; \; s \in S\rangle \; .$$

Of course the conventions in notation used for cartesian products (cf. Section 1.3) apply to direct products of structures as well.

We may be tempted to try building new models for theories from old ones by simply using direct products of structures. However, as the following simple counter examples show, direct products are too generous, in the sense that, in general, they no longer belong to the given class of structures.

Example 7.1.2 The direct product of a family of fields is, in general, no longer a field (as an exercise, verify this by examining $\mathbb{Z}_2 \times \mathbb{Z}_2$ for divisors of zero). The direct product of a family of total orders is, in general, no longer a total order (e.g. $\mathbb{N} \times \mathbb{N}$).

The solution is to not consider the product *per se*, but merely an appropriate quotient, i.e. we are going to work with reduced products. To define the equivalence suitable for our purposes, we need the following notion of ultrafilters.

Definition 7.1.3 (Ultrafilter) Let S be any non–empty set. Then, a filter $\mathcal{U} \subseteq \mathcal{P}(S)$ over S (cf. 1.5.3) is called an **ultrafilter over** S iff

(iv) $V \subseteq S$ implies [either $V \in \mathcal{U}$ or $S \setminus V \in \mathcal{U}$].

We have attached the number (iv) to the clause to remind you that in Definition 1.5.3, we already fixed three clauses which define a system $\mathcal{F}$ as a filter:

(i) $U_1, U_2 \in \mathcal{F} \Longrightarrow U_1 \cap U_2 \in \mathcal{F}$,

(ii) $U \in \mathcal{F}, U \subseteq V \subseteq S \Longrightarrow V \in \mathcal{F}$, and

(iii) $\emptyset \notin \mathcal{F} \neq \emptyset$.

Remark 7.1.4 An ultrafilter $\mathcal{U}$ over S is best thought of as a system of *sufficiently large subsets*, and with this in mind the above clauses can be read in the following way.

1. (i) The intersection of two sufficiently large sets is still sufficiently large.

   (ii) A set containing a sufficiently large subset is itself sufficiently large.

(iii) The empty set is not sufficiently large.

(iv) Any set is either sufficiently large or else its complement is sufficiently large.

Remark 7.1.5 Excluding $\mathcal{P}(S)$ from the collection of filters is not a generally accepted practice. Some authors include both $\mathcal{P}(S)$ and $\emptyset$ but distinguish them from the so-called *proper filters.*

Remark 7.1.6 If S is any nonempty set and $\mathcal{X} \subseteq \mathcal{P}(S)$, we may ask whether we can find a filter $\mathcal{U}$ over S such that $\mathcal{X} \subseteq \mathcal{U}$. If there is such a filter, then we might just as well look at $\bigcap \{\mathcal{U} \subseteq \mathcal{P}(S) \, ; \, \mathcal{U}$ is a filter and $\mathcal{X} \subseteq \mathcal{U}\}$. By doing this, we realize that there is a *smallest filter* containing $\mathcal{X}$, the so-called **filter generated by** $\mathcal{X}$, in which case $\mathcal{X}$ is said to generate this filter.

Remark 7.1.7 A further conclusion from the above is that a system $\mathcal{X}$ of subsets of S generates a filter iff $\mathcal{X}$ satisfies the **finite intersection property f.i.p.**, by which we mean that

$$\text{for any } \textit{finite} \text{ subset } \{U_1, \dots, U_n\} \subseteq \mathcal{X} \text{ we have } U_1 \cap \dots \cap U_n \neq \emptyset.$$

Satisfaction of the f.i.p. is a necessary and sufficient condition for being a subset of a filter. This is easily seen after showing that generating a filter from a system $\mathcal{X}$ means collecting all the supersets of finite intersections of elements of $\mathcal{X}$.

Remark 7.1.8 It is obvious that for any $U \subseteq S$ with $U \neq \emptyset$, $\{V \subseteq S \, ; \, U \subseteq V\}$ is a filter over S. This filter is clearly the filter generated by $\{U\}$. We may even find an ultrafilter by choosing, for some $p \in S$,

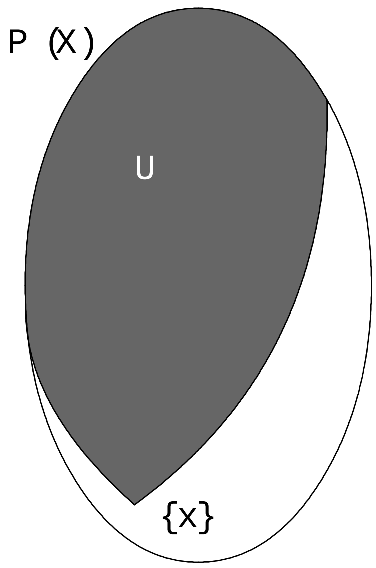
$$\mathcal{U} := \{U \subseteq S \, ; \, \{p\} \subseteq U\} = \{U \subseteq S \, ; \, p \in U\},$$

i.e. by taking the filter generated by $\mathcal{X} = \{\{p\}\}$. $\mathcal{U}$ thus defined is called the **ultrafilter fixed at** p.

Nevertheless, thinking of elements of $\mathcal{U}$ as *sufficiently large* (as we did in Remark 7.1.4) will not work with fixed ultrafilters, since *unrealistically small* subsets such as $\{p\}$ are contained in $\mathcal{U}$ as well. Of greater interest for our purposes are the **free ultrafilters**, i.e. ultrafilters $\mathcal{U}$ with

$$\bigcap \mathcal{U} = \bigcap \{U \subseteq S \, ; \, U \in \mathcal{U}\} = \emptyset.$$

(As an easy yet illuminating exercise, the reader is asked to verify that an ultrafilter $\mathcal{U}$ is fixed at some $p \in S$ iff $\bigcap \mathcal{U} \neq \emptyset$.)



1.

Figure 7.1: A fixed ultrafilter

Remark 7.1.9 $\mathcal{U}$ is an ultrafilter if and only if $\mathcal{U}$ is a *maximally proper filter*, i.e. iff $\mathcal{U}$ is a filter, $\mathcal{U} \neq \mathcal{P}(S)$, and for any filter $\mathcal{V}$ with $\mathcal{U} \subseteq \mathcal{V} \neq \mathcal{P}(S)$ we have $\mathcal{U} = \mathcal{V}$.

1. To understand this, note that if $\mathcal{U}$ is an ultrafilter, $\mathcal{V}$ a filter and $U \in \mathcal{V} \smallsetminus \mathcal{U}$, then $S \smallsetminus U \in \mathcal{U} \subseteq \mathcal{V}$, thus $\emptyset = U \cap (S \smallsetminus U) \in \mathcal{V}$, thus condition (iii) is violated for $\mathcal{V}$. For the other direction, if $\mathcal{U}$ is not an ultrafilter, then we find a $U \subseteq S$ such that both $U \notin \mathcal{U}$ and $(S \smallsetminus U) \notin \mathcal{U}$. But then at least one of the following holds:

$$\text{for all } V \in \mathcal{U}, U \cap V \neq \emptyset \qquad \text{or}$$

$$\text{for all } V \in \mathcal{U}, (S \smallsetminus U) \cap V \neq \emptyset.$$

   We conclude that $\mathcal{U} \cup \{U\}$ or $\mathcal{U} \cup \{S \smallsetminus U\}$ (or both) generate a filter properly extending $\mathcal{U}$.

Remark 7.1.10 As an exercise, the reader may prove that $\mathcal{U}$ is an ultrafilter over S if and only if $\mathcal{U}$ is a **prime filter** over S, i.e. iff, for any $U, V \subseteq S$, $U \cup V \in \mathcal{U}$ implies $U \in \mathcal{U}$ or $V \in \mathcal{U}$.

Remark 7.1.11 Closely related to the previous remark is the following observation (the proof of which is left as an exercise). Let $\mathcal{U}$ be an ultrafilter over some set S and $U \in \mathcal{U}$, and suppose that $U = U_1 \cup \ldots \cup U_n$ ($n \in \mathbb{N}$). Then, $U_i \in \mathcal{U}$ for some $i \in \{1, \ldots, n\}$.

1. If additionally $U_i \cap U_j = \emptyset$ for all $i, j \in \{1, \ldots, n\}$, $i \neq j$, then $U_i \in \mathcal{U}$ for exactly one $i \in \{1, \ldots, n\}$.

Remark 7.1.12 (This is intended for those readers who are already familiar with boolean algebras.) Regarding $\mathcal{P}(S)$ as a boolean algebra, we see that filters over S coincide with the inverse images of the largest element $\mathbf{1}$ under some boolean algebra-homomorphism η.
If, moreover, $\eta : \mathcal{P}(S) \longrightarrow \mathbf{B}_2$ where $\mathbf{B}_2$ is the two–element boolean algebra, then $\eta^{-1}(\mathbf{1}_{\mathbf{B}_2})$ is an ultrafilter. Thus,

1. *ultrafilters coincide with the inverse images of* $\mathbf{1}_{\mathbf{B}_2}$ *under boolean algebra homomorphisms with codomain* $\mathbf{B}_2$.

We have seen that filters, even ultrafilters, are easily constructed as upper closures under $\subseteq$ of some $U \in \mathcal{P}(S)$. Yet, as we have also mentioned, the *fixed* ultrafilters we thus obtain are not our main concern here, for reasons we will elaborate on later. The existence of *free* ultrafilters, on the other hand, is not obvious at all, as is shown by the (unavoidable!) use of Zorn's Lemma in the proof of the following Lemma.

Lemma 7.1.13 The following holds for any non–empty set S:

1. Any family $\mathcal{X} \subseteq \mathcal{P}(S)$ having the f.i.p. generates a (proper) filter.

2. Every proper filter $\mathcal{F}$ over S is contained in some ultrafilter $\mathcal{U} \supseteq \mathcal{F}$ over S.

Proof.

1. This is just an elaboration of our observations concerning the generation of filters. Suppose $\mathcal{X} \subseteq \mathcal{P}(S)$ has the f.i.p., and let

$$\mathcal{F} := \{U \subseteq S \,;\, X_1 \cap \ldots \cap X_n \subseteq U \text{ for some } X_1, \ldots, X_n \in \mathcal{X}\}.$$

We must check conditions (i) – (iii) for filters.

(i) If $U_1, U_2 \in \mathcal{F}$, then $X_1 \cap \ldots \cap X_n \subseteq U_1$ and $Y_1 \cap \ldots \cap Y_m \subseteq U_2$ for some $X_1, \ldots, X_n, Y_1, \ldots, Y_m \in \mathcal{X}$. It follows that

$$X_1 \cap \ldots \cap X_n \cap Y_1 \cap \ldots \cap Y_m \subseteq U_1 \cap U_2,$$

so $U_1 \cap U_2 \in \mathcal{F}$.

(ii) If $U \in \mathcal{F}$ and $U \subseteq V$, then

$$X_1 \cap \ldots \cap X_n \subseteq U \text{ for some } X_1, \ldots, X_n \in \mathcal{X},$$

but then also $X_1 \cap \ldots \cap X_n \subseteq V$, so $V \in \mathcal{F}$.

(iii) $\emptyset \notin \mathcal{F}$ by f.i.p., and $\mathcal{F} \neq \emptyset$ since $\mathcal{X} \subseteq \mathcal{F}$.

2. Let $\mathcal{F}$ be a proper filter over S and let

$$P := \{\mathcal{V} \subseteq \mathcal{P}(S)\ ;\ \mathcal{V} \text{ is a filter over over } S \text{ and } \mathcal{F} \subseteq \mathcal{V}\}\ .$$

$P \neq \emptyset$ since $\mathcal{F} \in P$. Now take any non–empty chain[1] $C \subseteq P$ and let $\mathcal{V}_0 := \bigcup C$. Then, we see that $\mathcal{V}_0$ is a filter over S. We verify this again by checking (i) – (iii) for filters:

(i) If $U_1, U_2 \in \mathcal{V}_0$, then $U_1 \in \mathcal{V}_1$ and $U_2 \in \mathcal{V}_2$ for some $\mathcal{V}_1, \mathcal{V}_2 \in C$. But then, since C is a chain, $U_1, U_2 \in \mathcal{V}_1$ or $U_1, U_2 \in \mathcal{V}_2$ and thus, since $\mathcal{V}_1$ and $\mathcal{V}_2$ are filters, $U_1 \cap U_2 \in \mathcal{V}_1$ or $U_1 \cap U_2 \in \mathcal{V}_2$, so $U_1 \cap U_2 \in \mathcal{V}_0$.

(ii) If $U \in \mathcal{V}_0$ and $U \subseteq V$, then $U \in \mathcal{V}$ for some $\mathcal{V} \in C$, but then $V \in \mathcal{V}$ since $\mathcal{V}$ is a filter, thus $V \in \mathcal{V}_0$.

(iii) $\emptyset \notin \mathcal{V}_0$ since otherwise $\emptyset \in \mathcal{V}$ for some $\mathcal{V} \in C$, contradicting the fact that all $\mathcal{V} \in C$ are filters; $\mathcal{V}_0 \neq \emptyset$ since $\mathcal{F} \neq \emptyset$ and $\mathcal{F} \subseteq \mathcal{V}_0$.

As we can see, $\mathcal{V}_0 \in P$ is an upper bound for C in P. Using Zorn's Lemma, we find that there is a maximal element $\mathcal{U}$ in P, i.e. $\mathcal{U}$ is a maximal filter containing $\mathcal{F}$. Using the remark above, we see that $\mathcal{U}$ is an ultrafilter extending $\mathcal{F}$.

■

Please note the use of Zorn's Lemma in the proof. Actually, the statement of Lemma 7.1.13 2 can be shown to be equivalent to Zorn's Lemma and thus also to the Axiom of Choice. Readers who are somewhat familiar with the history and axiomatics of Set Theory will remember that one of the disadvantages of using the Axiom of Choice in a proof of existence is the non–constructive nature of that proof. In the above situation, this means that we will work with ultrafilters to construct models without constructing the ultrafilters themselves. Thus the models whose existence we are proving will remain hidden from us by the veil of non–constructiveness. However, even their existence has many interesting implications, as we shall see later.

[1] I.e. for all $\mathcal{V}_1, \mathcal{V}_2 \in C$, $\mathcal{V}_1 \subseteq \mathcal{V}_2$ or $\mathcal{V}_2 \subseteq \mathcal{V}_1$. On this occasion, please recall *Zorn's Lemma* (cf. Section **??**) since this is the technique of proof used here.

Next, we want to rephrase the above Lemma 7.1.13 and reduce it to the most commonly used form. Combining clauses 1. and 2., we find the next lemma.

Lemma 7.1.14 (Existence of Ultrafilters)
Let S be any non–empty set. Then, any family $\mathcal{X} \subseteq \mathcal{P}(S)$ having the f.i.p. is contained in some ultrafilter $\mathcal{U} \supseteq \mathcal{X}$ over S.

A word of warning: In no way is the ultrafilter extending $\mathcal{X}$ required to be unique. This is clear from the observation that *any* filter contains S, thus $\mathcal{X} = \{S\}$ is contained in any fixed ultrafilter, of which we find the same number as there are elements in S, and with this we have not even accounted for the free ultrafilters.

We are now going to introduce an example for a "construction" of free ultrafilters, which is very popular since it gives rise to a great variety of useful examples of reduced products as will be shown later. For this, we will use the following definition.

Definition 7.1.15 Let S be any set. A subset $T \subseteq S$ is called **co–finite** iff $S \setminus T$ is finite. $\mathcal{P}_{\text{cof}}S := \{T \subseteq S\, ;\, T \text{ is co–finite}\}$.

Of course for finite sets S, $\mathcal{P}_{\text{cof}}S = \mathcal{P}(S)$. Infinite sets tend to be more interesting, as we can see in the following example.

Example 7.1.16 Let $\mathcal{F}$ be given by $\mathcal{F} = \mathcal{P}_{\text{cof}}\mathbb{N}$. Then, $\mathcal{F}$ has the f.i.p.; furthermore, $\mathcal{F}$ is even a filter, the so–called **Fréchet–Filter**, over $\mathbb{N}$. (To show that $\mathcal{F}$ is a filter, recall that, by DeMorgan's Laws,

$$T_1 \cap \ldots \cap T_n = \mathbb{N} \setminus ((\mathbb{N} \setminus T_1) \cup \ldots \cup (\mathbb{N} \setminus T_n)).$$

The rest should present no real problems and is left as an exercise.) Using Lemma 7.1.14 we find that there is an ultrafilter $\mathcal{U}$ over $\mathbb{N}$ with $\mathcal{P}_{\text{cof}}\mathbb{N} \subseteq \mathcal{U}$.
We note that for such a $\mathcal{U}$, $\bigcap \mathcal{U} = \emptyset$, since for any $n \in \mathbb{N}$, $\mathbb{N} \setminus \{n\} \in \mathcal{U}$, and already $\bigcap \{\mathbb{N} \setminus \{n\}\, ;\, n \in \mathbb{N}\} = \emptyset$, so any ultrafilter extending the Fréchet–Filter is a free ultrafilter.

Here is another nice example we will use later, e.g. when formulating and proving a semantical analog for the Compactness Theorem 2.4.5.

Example 7.1.17 For $X \neq \emptyset$, let $S := \{s \subseteq X\, ;\, s \text{ finite}\}$ and, for $x \in X$, $T_x := \{s \in S\, ;\, x \in s\}$.
Let $\mathcal{F} \subseteq \mathcal{P}(S)$ be given by $\mathcal{F} := \{T_x\, ;\, x \in X\}$. Then, $\mathcal{F}$ has the f.i.p., since

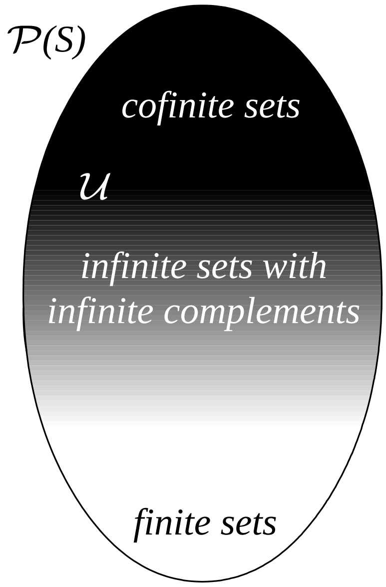


Figure 7.2: A free ultrafilter

clearly $\{x_1, \ldots, x_n\} \subseteq T_{x_i}$ $(i = 1, \ldots, n)$, thus

$$\{\mathrm{x}_1, \ldots, \mathrm{x}_\mathrm{n}\} \subseteq \mathrm{T}_{\mathrm{x}_1} \cap \ldots \cap \mathrm{T}_{\mathrm{x}_\mathrm{n}}.$$

(As an exercise: Can you tell whether $\mathcal{F}$ is a filter or not?)
Using Lemma 7.1.14 we once more find:
There is an ultrafilter $\mathcal{U}$ over S with $\mathcal{F} \subseteq \mathcal{U}$.
Is $\mathcal{U}$ fixed or free? Assume $\mathcal{U}$ is fixed, i.e for some $s \in S$, $s \in \bigcap \mathcal{F}$. This is equivalent to

$$s \in T_x \text{ for all } x \in X$$

and this again is the same as

$$x \in s \text{ for all } x \in X.$$

This means that $s \in \bigcap \mathcal{F}$ if and only if $s = X$. Since $s \in S$ is finite, the above leads to a contradiction if X is infinite. Hence we find, for infinite X, no $s \in \bigcap \mathcal{F} \supseteq \bigcap \mathcal{U}$ and thus clearly $\bigcap \mathcal{U} = \emptyset$. We conclude that if X is infinite, then any ultrafilter extending $\mathcal{F}$ is free.

7.2 Ultraproducts

We now have all the necessary tools to define ultraproducts. Recall from Section 1.5 the definition of direct and reduced products as well as those of their associated canonical projections.

Definition 7.2.1 Let $\mathcal{L}$ be a formal language and, for $s \in S$, let $\mathcal{A}_s$ be an $\mathcal{L}$–structure. Let $\mathcal{U}$ be an ultrafilter over S. The relation $\sim_{\mathcal{U}}$ on $\prod_{s\in S} |\mathcal{A}_s|$ is defined by

$$\langle a_s \,;\, s \in S\rangle \sim_{\mathcal{U}} \langle b_s \,;\, s \in S\rangle \text{ iff } \{s \in S \,;\, a_s = b_s\} \in \mathcal{U}$$

for any $\langle a_s \,;\, s \in S\rangle, \langle b_s \,;\, s \in S\rangle \in \prod_{s\in S} |\mathcal{A}_s|$.

The relation $\sim_{\mathcal{U}}$ is an equivalence relation. Reflexivity and symmetry are trivial to prove, and transitivity follows easily from the fact that $\mathcal{U}$ is closed under intersection. (Exercise: Why?)

Let us now consider the quotient $\prod_{s\in S} |\mathcal{A}_s| / \sim_{\mathcal{U}}$ and define on it interpretations for the function–, relation– and constant–symbols of $\mathcal{L}$. We will find a new $\mathcal{L}$–structure, the so called *ultraproduct of the family* $\langle \mathcal{A}_s \,;\, s \in S\rangle$ *under* $\mathcal{U}$, which we will denote by $\prod_{s\in S} \mathcal{A}/\mathcal{U}$. For further convenience we will write $\pi_{\mathcal{U}}$ instead of $\pi_{\sim_{\mathcal{U}}}$ for the canonical projection.

Definition 7.2.2 (Ultraproducts of $\mathcal{L}$–Structures)
The $\mathcal{L}$-structure $\mathcal{A} = \prod_{s\in S} \mathcal{A}_s/\mathcal{U}$, the *ultraproduct of the family* $\langle \mathcal{A}_s \,;\, s \in S\rangle$ *under* $\mathcal{U}$, is defined as follows:

1. **Universe:**
$$|\mathcal{A}| := \{\pi_{\mathcal{U}}(a) \,;\, a \in \prod_{s\in S} |\mathcal{A}_s|\} \,.$$

2. **Relations:** For a relation–symbol R_i, we define
$$R_i^{\mathcal{A}}(\langle \pi_{\mathcal{U}}(a_1), \ldots, \pi_{\mathcal{U}}(a_{\lambda(i)})\rangle)$$
iff
$$\{s \in S \,;\, R_i^{\mathcal{A}_s}(\langle \pi_s(a_1), \ldots, \pi_s(a_{\lambda(i)})\rangle)\} \in \mathcal{U}$$
for all $a_1, \ldots, a_{\lambda(i)} \in \prod_{s\in S} |\mathcal{A}_s|$.

3. **Functions:** For a function–symbol f_j, we define
$$f_j^{\mathcal{A}}(\pi_{\mathcal{U}}(a_1), \ldots, \pi_{\mathcal{U}}(a_{\mu(j)})) := \pi_{\mathcal{U}}(\langle f_j^{\mathcal{A}_s}(\pi_s(a_1), \ldots, \pi_s(a_{\mu(j)})) \,;\, s \in S\rangle)$$

for all $a_1, \ldots, a_{\mu(j)} \in \prod_{s\in S} |\mathcal{A}_s|$.

4. **Constants:** For a constant–symbol c_k, we define

$$c_k^{\mathcal{A}} := \pi_{\mathcal{U}}(\langle c_k^{\mathcal{A}_s} \; ; \; s \in S\rangle).$$

Exercise 7.2.3 Verify that this is a valid definition, i.e. show that

1. if $a_1 \sim_{\mathcal{U}} b_1, \ldots, a_{\lambda(i)} \sim_{\mathcal{U}} b_{\lambda(i)}$, then

$$R_i^{\mathcal{A}}(\pi_{\mathcal{U}}(a_1), \ldots, \pi_{\mathcal{U}}(a_{\lambda(i)})) \text{ iff } R_i^{\mathcal{A}}(\pi_{\mathcal{U}}(b_1), \ldots, \pi_{\mathcal{U}}(b_{\lambda(i)}));$$

2. if $a_1 \sim_{\mathcal{U}} b_1, \ldots, a_{\mu(j)} \sim_{\mathcal{U}} b_{\mu(j)}$, then

$$f_j^{\mathcal{A}}(\pi_{\mathcal{U}}(a_1), \ldots, \pi_{\mathcal{U}}(a_{\mu(j)})) = f_j^{\mathcal{A}}(\pi_{\mathcal{U}}(b_1), \ldots, \pi_{\mathcal{U}}(b_{\mu(j)})).$$

Simplifying our notation further, we will sometimes use the following notational conventions for $a \in \prod_{s\in S} |\mathcal{A}_s|_s$, $s \in S$, ultrafilters $\mathcal{U}$ and valuations h in $\prod_{s\in S} |\mathcal{A}_s|_s$:

$$a(s) := a_s := \pi_s(a) \quad \text{and} \quad a_{\mathcal{U}} := \pi_{\mathcal{U}}(a),$$

$$h_s := \pi_s \circ h \quad \text{and} \quad h_{\mathcal{U}} := \pi_{\mathcal{U}} \circ h.$$

In this simplified notation, the definition above would be written as follows:

1. $|\mathcal{A}| := \{a_{\mathcal{U}} \; ; \; a \in \prod_{s\in S} |\mathcal{A}_s|\}$
2. $f_j^{\mathcal{A}}(a_{1\mathcal{U}}, \ldots, a_{\mu(j)\mathcal{U}}) := \langle f_j^{\mathcal{A}_s}(a_{1s}, \ldots, a_{\mu(j)s}) \; ; \; s \in S\rangle_{\mathcal{U}}$
3. $\langle a_{1\mathcal{U}}, \ldots, a_{\lambda(i)\mathcal{U}}\rangle \in R_i^{\mathcal{A}}$ iff $\{s \in S \; ; \; \langle a_{1s}, \ldots, a_{\lambda(i)s}\rangle \in R_i^{\mathcal{A}_s}\} \in \mathcal{U}$
4. $c_k^{\mathcal{A}} := \langle c_k^{\mathcal{A}_s} \; ; \; s \in S\rangle_{\mathcal{U}}$

Concerning the valuation it is also worth noticing that

1. for any $s \in S$ and for any valuation h' into $\mathcal{A}_s$, there is a valuation h into $\prod_{s\in S} \mathcal{A}_s$ such that $h_s = h'$;
2. for any $s \in S$ and for any valuation h' into $\prod_{s\in S} \mathcal{A}_s / \mathcal{U}$, there is a valuation h into $\prod_{s\in S} \mathcal{A}_s$ such that $h_s = h'$.

One more special case must be mentioned before we can look at some results.

Definition 7.2.4 If for some set $S \neq \emptyset$, $\mathcal{A}_s = \mathcal{B}$ for all $s \in S$ (i.e. all the factors are the same), then we write $\mathcal{B}^S/\mathcal{U}$ for $\prod_{s\in S} \mathcal{A}_s / \mathcal{U}$ and call this the **ultrapower of $\mathcal{B}$ to S under $\mathcal{U}$**.

Example 7.2.5 Let us have a look at the direct product of linear orders. Let $\mathcal{A}$ be the two–element chain, i.e. $\mathcal{A} := \langle\{0,1\}, \leq\rangle$ where $0 \leq 1$. Now, looking at the direct power $\mathcal{B} := \mathcal{A}^{\mathbb{N}}$, it is evident that $\mathcal{B}$ is *not* a linear order, since e.g. $\langle 0,1,1,1,\ldots\rangle$ and $\langle 1,0,0,0,\ldots\rangle$ are not comparable under $\leq^{\mathcal{B}}$. On the other hand, if we let $\mathcal{U}$ be an ultrafilter over $\mathbb{N}$ and $\mathcal{C} := \mathcal{B}/\mathcal{U}$, everything turns out to be comparable again. Suppose we find $a, b \in |\mathcal{B}|$ such that $a_{\mathcal{U}} \leq^{\mathcal{C}} b_{\mathcal{U}}$ does *not* hold. Then, by the definition of the interpretation of relations in ultraproducts,

$$\{n \in \mathbb{N}\,;\, a_n \leq b_n\} \notin \mathcal{U},$$

but since $\mathcal{U}$ is an ultrafilter, this is the same as

$$\mathbb{N} \setminus \{n \in \mathbb{N}\,;\, a_n \leq b_n\} \in \mathcal{U},$$

and this in turn is equivalent to

$$\{n \in \mathbb{N}\,;\, a_n \not\leq b_n\} \in \mathcal{U},$$

and finally, since $\mathcal{A}$ is a linear order, this means

$$\{n \in \mathbb{N}\,;\, b_n \leq a_n\} \in \mathcal{U},$$

which, by the definition of $\leq^{\mathcal{C}}$, means $b_{\mathcal{U}} \leq^{\mathcal{C}} a_{\mathcal{U}}$. So, from the assumption that $a_{\mathcal{U}} \leq^{\mathcal{C}} b_{\mathcal{U}}$ does not hold, we proved that $b_{\mathcal{U}} \leq^{\mathcal{C}} a_{\mathcal{U}}$ holds. This is exactly the definition of "linear order". Thus, we found that our ultrapower of linear orders is, unlike the direct product, indeed a linear order.

Example 7.2.6 Reconsidering the previous example, we may wonder just how large this ultraproduct is. Set Theory tells us that the direct product has the same cardinality as $\mathcal{P}(\mathbb{N})$ and is thus uncountable. However, using an argument very similar to the one we applied to show that $\mathcal{C}$ is linearly ordered by $\leq^{\mathcal{C}}$, we can show that $\mathcal{C}$ is actually rather small: For $a \in |\mathcal{B}|$ arbitrary, we know (by the definition of ultrafilters) that exactly one of the sets $T_{a,0}$ and $T_{a,1}$ is in $\mathcal{U}$, where

$$T_{a,i} := \{n \in \mathbb{N}\,;\, a_n = i\} \quad (i \in \{0,1\}).$$

Thus,

$$\text{either } \mathrm{a}_{\mathcal{U}} = \langle 0,0,0,\ldots\rangle_{\mathcal{U}} \text{ or } \mathrm{a}_{\mathcal{U}} = \langle 1,1,1,\ldots\rangle_{\mathcal{U}}.$$

Also, clearly $\langle 0,0,0,\ldots\rangle_{\mathcal{U}} \neq \langle 1,1,1,\ldots\rangle_{\mathcal{U}}$. Hence, we have shown that the universe of $\mathcal{C}$ consists of exactly two elements. (The Theorem of Łoś 7.2.11 puts this result in a wider context. We will see that ultrapowers of finite structures are of the same cardinality as their sole factor.)

Example 7.2.7 What is the point of all this? We constructed an ultrapower and, after musing about this new structure for a few minutes, were able to prove that we have gained absolutely nothing. However, we constrained ourselves in two ways: first, by looking at an ultra*power*, and second, by taking a finite structure as the starting point of our venture. (Note that we did not make any assumptions about $\mathcal{U}$ being fixed or free.) This and the next example show that even dropping one of these limitations is enough to make the ultraproduct differ from its original structures.

For $\mathcal{A}$ take the set of natural numbers $\mathbb{N}$ together with the natural order $\leq$ and the function $+$. Again, let $S := \mathbb{N}$ and $\mathcal{U}$ be a free ultrafilter over S expanding the Fréchet–Filter. Let $\mathcal{B} := \mathcal{A}^{\mathbb{N}}$ and $\mathcal{C} := \mathcal{B}/\mathcal{U}$.

We notice that $\mathcal{C}$ is infinite (as is $\mathbb{N}$), since for $m, n \in \mathbb{N}, m \neq n$ we have $\langle m, m, m, \ldots \rangle_{\mathcal{U}} \neq \langle n, n, n, \ldots \rangle_{\mathcal{U}}$. Yet, something has indeed changed. Remember that the *Archimedean Property of* $\mathbb{N}$ expresses the fact that any natural number n can be reached by a consecutive sum of 1's with just enough summands:

$$n \leq \underbrace{1 + \ldots + 1}_{n \text{ summands}} .$$

This property is no longer valid for our ultrapower $\mathcal{C}$; i.e. for example for $a = \langle 0, 1, 2, \ldots \rangle_{\mathcal{U}}$, there is no way of writing a as a finite sum of 1's. This follows from the fact that

$$\text{for any } m \in \mathbb{N}, \quad \langle m, m, m, \ldots \rangle_{\mathcal{U}} \leq^{\mathcal{C}} \langle 0, 1, 2, \ldots \rangle_{\mathcal{U}},$$

which in itself follows from

$$\{s \in \mathbb{N}\,;\, \langle m, m, m \rangle_s \leq \langle 1, 2, 3, \rangle_s\} = \{m, m+1, m+2, \ldots\} \in \mathcal{U}$$

by co–finiteness, since $\{0, 1, \ldots, m-1\}$ is finite. Thus, you could say that $\mathcal{C}$ looks like $\mathbb{N}$ with additional infinitely large elements. However, please be aware that we did not give a first–order formula expressing the *Archimedean Property* or its consequence above.

Example 7.2.8 Now let, for $s \in \mathbb{N}$, $\mathcal{A}_s := \langle \{0, 1, 2, \ldots, s\} \rangle$; i.e. $\mathcal{A}$ is an $\mathcal{L}$–structure for the formal language $\mathcal{L}$ lacking any non–logical symbols. As before, our index set is $\mathbb{N}$ and the ultrafilter $\mathcal{U}$ is supposed to expand the Fréchet–Filter. Let $\mathcal{C}$ be the ultraproduct of $\langle \mathcal{A}_s\,;\, s \in \mathbb{N} \rangle$ under $\mathcal{U}$.

Now consider the following elements of $\mathcal{C}$:

$$\begin{aligned}
a_0 &:= \langle 0,0,0,0,0,\ldots\rangle_{\mathcal{U}}\\
a_1 &:= \langle 0,1,1,1,1,\ldots\rangle_{\mathcal{U}}\\
a_2 &:= \langle 0,1,2,2,2,\ldots\rangle_{\mathcal{U}}\\
&\ldots\\
a_n &:= \langle 0,1,2,\ldots,n-1,n,n,n\ldots\rangle_{\mathcal{U}}\\
&\ldots
\end{aligned}$$

(Exercise: Show that those are actually elements of the universe of $\mathcal{C}$.)
With an argument analogous to the one we used in the previous example to show that a is infinitely big, we can now show that all the a_i's are pairwise distinct, from which we conclude that we are dealing with an example of an ultraproduct of finite structures which is no longer finite. Since a finite structure is never elementary equivalent to an infinite one, we have here an example of an ultraproduct which is not elementary equivalent to any of its factors.

Examples 7.2.7 and 7.2.8 instantiate the central aspects and motivations which have led us to consider ultraproducts in the first place:

- By building the "infinitely large" element of the ultrapower of $\mathbb{N}$ we entered the realm of non–standard models of the axioms of the natural numbers. The proof of Łoś's Theorem 7.2.11 (cf. Appendix A) will make it clear that the difference between $\mathbb{N}$ and the non–standard structure provided by the ultrapower–construction is not a distinction that can be expressed in First–Order Logic; in other words, First–Order Logic cannot prevent the axioms of $\mathbb{N}$ from having models with infinitely large elements.

- We will later show that an ultraproduct of finite structures can be infinite and that elementary classes are necessarily closed under ultraproducts. These observations together yield the following generalization: Suppose in an elementary class $\mathbb{K}$, we find, for every $n \in \mathbb{N} \setminus \{0\}$, a *finite* structure $\mathcal{A}_n$ such that $\mathcal{A}_n$ has at least n elements; from this we can conclude that there is also an infinite structure in $\mathbb{K}$. In a more concrete example, this implies that finding an axiomatization of exactly the finite groups in terms of First–Order Logic is a task that is bound to fail.

Exercise 7.2.9 By using arguments similar to the ones used in the above examples, show that:

1. The direct product of a family of fields is not necessarily a field. (Hint: concentrate on the fact that in a field $a \cdot b = 0$ implies $a = 0$ or $b = 0$).

2. In an ultraproduct of a family of fields $a \cdot b = 0$ implies $a = 0$ or $b = 0$.

We will now turn our attention towards the main result of this chapter, the Main Theorem on Ultraproducts 7.2.11. Although its proof is deferred to Appendix A, a few introductory lemmata will be helpful to understand the the full extent of this result's implications.

For the relationship between interpretations of terms and the projections, we note the following Lemma.

Lemma 7.2.10 If $\langle \mathcal{A}_s\ ;\ s \in S\rangle$ is a family of $\mathcal{L}$–structures and $\mathcal{C} = \prod_{s\in S} \mathcal{A}_s\ /\mathcal{U}$ the ultraproduct of this family under the ultrafilter $\mathcal{U}$, then for any $\mathcal{L}$–term t and any valuation h into $\prod_{s\in S} \mathcal{A}_s$,

1. $t^{\mathcal{C}}[h_{\mathcal{U}}] = (t^{\mathcal{B}}[h])_{\mathcal{U}}$, and
2. $t^{\mathcal{A}_s}[h_s] = (t^{\mathcal{B}}[h])_s$ for any $s \in S$.

Proof. We proceed by structural induction over the definition of $\mathcal{L}$–terms, thereby we will deal with both statements simultaneously:

1. • If t is a variable x, then

$$\begin{aligned} t^{\mathcal{C}}[h_{\mathcal{U}}] &= h_{\mathcal{U}}(x) = \pi_{\mathcal{U}} \circ h(x) \\ &= (h(x))_{\mathcal{U}} = (t^{\mathcal{B}}[h])_{\mathcal{U}} \end{aligned}$$

and

$$t^{\mathcal{A}_s}[h_s] = h_s(x) = \pi_s \circ h(x) = (h(x))_s = (t^{\mathcal{B}}[h])_s.$$

• If t is a constant–symbol c_k, then

$$t^{\mathcal{C}}[h_{\mathcal{U}}] = c_k^{\mathcal{C}} = (c_k^{\mathcal{B}})_{\mathcal{U}} = (t^{\mathcal{B}}[h])_{\mathcal{U}}$$

and

$$t^{\mathcal{A}_s}[h_s] = c_k^{\mathcal{A}_s} = (c_k^{\mathcal{B}})_s = (t^{\mathcal{B}}[h])_s.$$

• If $t = f_j(t_1, \ldots, t_{\mu(j)})$ for a function–symbol f_j and $\mathcal{L}$–terms $t_1, \ldots, t_{\mu(j)}$, then

$$\begin{aligned} t^{\mathcal{C}}[h_{\mathcal{U}}] &= f_j^{\mathcal{C}}(t_1^{\mathcal{C}}[h_{\mathcal{U}}], \ldots, t_{\mu(j)}^{\mathcal{C}}[h_{\mathcal{U}}]) \\ &= f_j^{\mathcal{C}}((t_1^{\mathcal{B}}[h])_{\mathcal{U}}, \ldots, (t_{\mu(j)}^{\mathcal{B}}[h])_{\mathcal{U}}) \\ &\qquad \text{(by the induction hypothesis)} \\ &= (f_j^{\mathcal{B}}(t_1^{\mathcal{B}}[h], \ldots, t_{\mu(j)}^{\mathcal{B}}[h]))_{\mathcal{U}} \\ &\qquad \text{(by the definition of } f_j^{\mathcal{C}}) \\ &= (t^{\mathcal{B}}[h])_{\mathcal{U}} \end{aligned}$$

and

$$\begin{aligned} t^{\mathcal{A}_s}[h_s] &= f_j^{\mathcal{A}_s}(t_1^{\mathcal{A}_s}[h_s], \ldots, t_{\mu(j)}^{\mathcal{A}_s}[h_s]) \\ &= f_j^{\mathcal{A}_s}((t_1^{\mathcal{B}}[h])_s, \ldots, (t_{\mu(j)}^{\mathcal{B}}[h])_s) \\ &\qquad \text{(by the induction hypothesis)} \\ &= (f_j^{\mathcal{B}}(t_1^{\mathcal{B}}[h], \ldots, t_{\mu(j)}^{\mathcal{B}}[h]))_s \\ &\qquad \text{(by the definition of } f_j^{\mathcal{A}_s}) \\ &= (t^{\mathcal{B}}[h])_s. \end{aligned}$$

Please note that we took advantage of a universal property of the projections π_s and the direct product $\prod_{s\in S} A_s$:

$$\langle \pi_s(a)\ ;\ s \in S\rangle = a \text{ for any } a \in \prod_{s\in S} A_s\ .$$

■

We are now going to state the Main Theorem on Ultraproducts. Its proof is deferred to Appendix A, because it is rather technical and does not lead to further enlightment with respect to the theorem or its applications.

Theorem 7.2.11 (Łoś, Main Theorem on Ultraproducts)
For a formal language $\mathcal{L}$ let $\langle \mathcal{A}_s\ ;\ s \in S\rangle$ be a family of $\mathcal{L}$–structures and $\mathcal{U}$ be an ultrafilter over S. For the sake of readability, let $\mathcal{B} := \prod_{s\in S} \mathcal{A}_s$ be the direct product and $\mathcal{A} := \mathcal{B}/\mathcal{U}$ the ultraproduct of the family $\langle \mathcal{A}_s\ ;\ s \in S\rangle$ under $\mathcal{U}$. Then, the following two statements hold:

1. For any valuation h into $\mathcal{B}$ and for any $\mathcal{L}$–formula φ,

$$\mathcal{A} \models \varphi[h_{\mathcal{U}}] \text{ iff } \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s]\} \in \mathcal{U}.$$

2. For any $\mathcal{L}$–sentence α,

$$\mathcal{A} \models \alpha \text{ iff } \{s \in S\ ;\ \mathcal{A}_s \models \alpha\} \in \mathcal{U}.$$

Proof. Cf. Appendix A.

■

When applied to free ultraproducts, Łoś's Theorem states that a formula holds if it is satisfied in *sufficiently many* factor–structures. Note that for a fixed ultrafilter, say $\mathcal{U} = \{U \subseteq S\ ;\ p \in U\}$, the second part of Łos's Theorem reads as

$$\mathcal{C} \models \varphi[h_{\mathcal{U}}] \text{ iff } \mathcal{A}_s \models \varphi[h_s].$$

Figure 7.3: Jerzi Łoś (1920-1998)

A more elegant formulation can be found for the special case of ultrapowers: Ultrapowers are elementary equivalent to their factor.

Corollary 7.2.12 (Ultrapowers) For any $\mathcal{L}$–structure $\mathcal{A}$, any non–empty set S and any ultrafilter $\mathcal{U}$ over S,

$$\mathcal{A}^S/\mathcal{U} \equiv \mathcal{A}.$$

Proof. If you notice that for any $\mathcal{L}$–formula φ, $T_h^{\mathcal{A}^S/\mathcal{U}}(\varphi)$ is either S or $\emptyset$, then you will not come across any difficulties completing the proof. ∎

7.3 The Compactness Theorem Revisited

We are now in a position to present a semantic analogue to the Compactness Theorem of First–Order Logic 2.4.5. The following corollary presents a way of "constructing"[2] a model for an infinite set of sentences from the models of its finite subsets.

Corollary 7.3.1 Let Σ be a set of $\mathcal{L}$–sentences such that, for any finite $\Theta \subseteq \Sigma$, there is a model $\mathcal{A}_\Theta$ of Θ. Then, there is an ultraproduct $\mathcal{C}$ of the family $\langle \mathcal{A}_\Theta\ ;\ \Theta \subseteq \Sigma, \Theta \text{ finite}\rangle$ under some ultrafilter $\mathcal{U}$ such that $\mathcal{C} \models \Sigma$.

Proof. We proceed in small steps that we deliberately do not develop in full detail to leave room for the readers do do some work of their own. Note that the index set S for the ultraproduct is chosen to be the set of all finite subsets of Σ, thus an element $s \in S$ is a set, and an element of an ultrafilter over S will be a set of sets.

- First, we define for $\alpha \in \Sigma$ the set

$$S_\alpha := \{s \in S\ ;\ \alpha \in s\}\ .$$

[2] The quotes are used to express the author's unease with calling a description based heavily on the Axiom of Choice a "construction".

Thus, S_α is the set of finite subsets of Σ containing α. It follows that, for $\alpha_1, \ldots, \alpha_n \in \Sigma$,

$$\{\alpha_1, \ldots, \alpha_n\} \in S_{\alpha_1} \cap \ldots \cap S_{\alpha_n}.$$

- Now we collect all these S_α's in the set $\mathcal{F} := \{S_\alpha\,;\, \alpha \in \Sigma\}$. Then, $\mathcal{F}$ has the f.i.p. (we proved this in Example 7.1.17) and is thus a subset of some ultrafilter $\mathcal{U}$ over S.
- For $\alpha \in \Sigma$ and $s \in S_\alpha$, we have $\alpha \in s$ and thus $\mathcal{A}_s \models \alpha$.
- From this we find $S_\alpha \subseteq \{s \in S\,;\, \mathcal{A}_s \models \alpha\} \in \mathcal{U}$.
- Finally, by defining $\mathcal{C} := \prod_{s\in S} \mathcal{A}_s\, /\mathcal{U}$, we see that for any $\alpha \in \Sigma$, $\mathcal{C} \models \alpha$.

∎

7.4 The Upward Löwenheim–Skolem Theorem Revisited

As was promised in Section 5.4, we will now give an alternative proof of the Upward Löwenheim-Skolem Theorem, which makes use of the technique introduced in this Chapter.

Please recall from Example 7.1.17, that for every set X, if $S := \{s \subseteq X; s \text{ finite}\}$ and, for $x \in X$, $T_x := \{s \in S; x \in s\}$, then there is an ultrafilter over S extending $\{T_x; x \in X\}$.

Theorem 7.4.1 Let $\mathcal{L}$ be a formal language and $\mathcal{A} \in \mathrm{Str}\,\mathcal{L}$ be infinite. Then, for any $\kappa \geq \parallel \mathcal{L} \parallel$ there is an $\mathcal{L}$–structure $\mathcal{B}$ with $\mathcal{B} \equiv \mathcal{A}$ and $\mathrm{card}\,\mathcal{B} = \kappa$.

Proof. Let $\Sigma := \mathrm{Th}\,\mathcal{A}$, and let $\mathcal{A}_0$ be an $\mathcal{L}$–structure with $\mathcal{A}_0 \equiv \mathcal{A}$ and $\mathrm{card}\,\mathcal{A}_0 \leq \parallel \mathcal{L} \parallel$. ($\mathcal{A}_0$ exists by the Downward Löwenheim–Skolem Theorem 5.3.1.) Let $C := \{c_\alpha; \alpha < \kappa\}$ be a set of distinct constant symbols not already in $\mathcal{L}$, and let $\mathcal{L}'$ be $\mathcal{L}$ with these new constant symbols added. (From the notation it should be clear that the set C has cardinality κ). Moreover, let $\Gamma \subseteq \mathrm{Sen}\,\mathcal{L}'$ be given by

$$\Gamma := \{\neg c_\alpha \doteq c_\beta; \alpha < \beta < \kappa\}$$

and $\Sigma' := \Sigma \cup \Gamma$. Now define $S := \{s \subseteq S; s \text{ finite}\}$.
Since any finite subset of Σ' can contain only finitely many of the inequalities in Γ, and these finitely many inequalities can be satisfied simply by choosing different interpretations for the constant symbols involved (and arbitrary interpretations for those constant symbols that are not in the inequalities), it

follows that for any $s \in S$, there is an $\mathcal{L}'$–structure $\mathcal{A}_s$, which has the same carrier as $\mathcal{A}$ but carries additional interpretations for the constant symbols which appear in Γ but do not belong to $\mathcal{L}$, such that $\mathcal{A}_s$ is a model for s. Setting $T_\alpha := \{s \in S; \alpha \in s\}$ for $\alpha \in \Sigma'$ and letting $\mathcal{B}$ be the ultrapower

$$\mathcal{B} := \prod_{s \in S} \mathcal{A}_s / \mathcal{U}$$

where $\mathcal{U}$ is an ultrafilter over S extending the set $\mathcal{F} := \{T_\alpha; \alpha \in \Sigma'\}$, we see that $\mathcal{B} \models \Sigma'$.
Indeed, if $\alpha \in \Sigma'$, then $\mathcal{A}_s \models \alpha$ for all $s \in T_\alpha$, and $T_\alpha \in \mathcal{U}$, hence

$$\{s \in S; \mathcal{A}_s \models \alpha\} \subseteq T_\alpha \in \mathcal{U}$$

and, therefore, $\{s \in S; \mathcal{A}_s \models \alpha\} \in \mathcal{U}$. From Theorem 7.2.11 we know that this implies $\mathcal{B} \models \alpha$. Because $\alpha \in \Sigma'$ was arbitrary, $\mathcal{B}$ is a model for Σ' and thus has at least κ elements (the interpretations of the new constant symbols; cf. Lemma 5.4.1). If $\mathcal{B}$ has too many elements (i.e. if $\operatorname{card} \mathcal{B} > \kappa$), we apply the Downward Löwenheim–Skolem Theorem 5.3.1 to find $\mathcal{B}'$ of the desired cardinality. ■

Exercise 7.4.2 In the proof of Theorem 7.4.1, explain how the $\mathcal{L}'$–structures $\mathcal{A}_s$ have to be defined (w.r.t. the interpretations of the new constant symbols). Explain in more detail why they exist.

Chapter 8

The Semantical Characterization of Elementary Classes

Elementary classes were defined in Definition 4.1.6 using syntactical notions. In this chapter we are looking for a more direct characterization in the sense that we are trying to avoid using any reference to theories.

8.1 Ultraproducts in Elementary Classes

A first step is to show that elementary classes are closed under ultraproducts, or, to put it in the form of a negative result, that a class of $\mathcal{L}$–structures that is not closed under ultraproducts is never axiomatizable by a set of first–order sentences.

Remember that we have already verified that elementary classes are closed under elementary equivalence. We now ask if the converse is also true, i.e. if every class of $\mathcal{L}$–structures which is closed under elementary equivalence is also an elementary class. Looking at the following very basic example will not reveal the answer directly, but it will give us a hint on where to look for a counter–example.

Let $\mathbb{K}$ be the class of all finite sets, which are considered as a class of $\mathcal{L}$-structures for the (unique!) formal language without any extra non–logical symbols. According to the definition of elementary classes, one way of showing that $\mathbb{K}$ is elementary would be to provide an axiomatization for $\mathbb{K}$. Yet, as hard as we try, we eventually fail to find an appropriate set of axioms. This in itself is clearly not a proof of $\mathbb{K}$ not being elementary, but it reveals a basic asymmetry between being a model of a set Σ of sentences and not being a model of Σ. The asymmetry we are hinting at is not really an asymmetry as long as we require Σ to be finite. For infinite sets Σ however, $\mathbb{K}$ being the class of models for Σ means that *every* structure in $\mathbb{K}$ satisfies *every* sentence in Σ, whereas for $\mathrm{Str}\,\mathcal{L} \setminus \mathbb{K}$, a similar formulation would be that the structures in $\mathrm{Str}\,\mathcal{L} \setminus \mathbb{K}$ are violating *at least one* sentence in Σ.

Before drowning in abstract elaborations, we better present a model class which exemplifies these observations:

Let $Set_{\geq\omega}$ be the class of infinite sets. From First–Order Logic, we remember that for any $n \in \mathbb{N}$ there is a sentence α_n which is satisfied in a structure $\mathcal{A}$ if and only if the universe of $\mathcal{A}$ has at least n elements. (This holds true regardless of the language $\mathcal{L}$ under consideration; we may thus assume $\mathcal{L}$ to be as above, a formal language without any non–logical symbols.) Being an *infinite* structure is then equivalent to being a model of the set $\Sigma = \{\alpha_n \,;\, n \in \mathbb{N}\}$, whereby we see that $Set_{\geq\omega}$ is indeed an elementary class.[1] The complement of $Set_{\geq\omega}$ contains all finite structures. Being a finite structure means *not* satisfying all the sentences in Σ, i.e. violating at least one of them. Hence, we are in the rather uncomfortable position where one class of structures is well captured by the notion of satisfaction of a set of sentences, whereas for the complementary class, this feat does not work, since we lack the same elegant way of expression. Satisfaction of a set Σ of sentences expresses the simultaneous satisfaction of *each* sentence in the set (which can be viewed as an "infinite conjunction"), whereas the failure to satisfy Σ amounts to violating *at least one* sentence in Σ. What we need is a formalization of *failure* which is on a par with the "infinite conjunction". But this would have to be some sort of "infinite disjunction" which, unfortunately, we do not have.[2]

We will show later that, for finite sets of sentences, the class of non–models is equally well describable. This will lead us to the notion of basic–elementary classes.

The first steps towards a semantical rephrasing of the termin "elementary class" are the following observations. First, remember that every elementary class is closed under elementary equivalence $\equiv$ (cf. Lemma 4.1.12). We also mentioned that the converse is false, i.e. there are classes of structures which are closed under elementary equivalence, yet they are not elementary. For the time being, however, we are not in the position to give such an example.

[1] It may be confusing that the notion of "infinite", i.e. *not* finite, is the positive notion in the context of this example, while finite will be the negative counterpart. This stands in a refreshing contrast to all the finitist approaches to logic.

[2] This lack of symmetry between validity and its contrary is rooted already in the notion of deducibility, where we deal only with deducible sentences, as for non–deducibility, we are rather helpless before we are equipped with completeness–theorems. Also, compactness is formulated only for the positive case.

Our second observation is expressed in the following lemma.

Lemma 8.1.1 Every elementary class is closed under ultraproducts.

Proof. Let $\{\mathcal{A}_s\ ;\ s \in S\} \subseteq \mathrm{Mod}\,\Sigma$ and $\mathcal{B} := \prod_{s \in S} \mathcal{A}_s\ /\mathcal{U}$ be an ultraproduct under the ultrafilter $\mathcal{U}$ over the index set $S \neq \emptyset$. Then, for any $\alpha \in \Sigma$, we have $\{s \in S; \mathcal{A}_s \models \alpha\} = S \in \mathcal{U}$ and thus $\mathcal{B} \models \alpha$. We conclude $\mathcal{B} \in \mathrm{Mod}\,\Sigma$. ■

Thus, being closed under elementary equivalence and ultraproducts is a necessary condition for a class being elementary. Is this condition sufficient?

Assume we would like to find a proof for this. Our next step would be to show that a class $\mathbb{K}$ of $\mathcal{L}$–structures that is closed under elementary equivalence and ultraproducts is elementary, that is to say, it satisfies $\mathbb{K} = \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$. From 4.1.4, we already know that $\mathbb{K} \subseteq \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$, thus we "only" have to show that $\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K} \subseteq \mathbb{K}$. Let us assume that $\mathbb{K}$ is closed under elementary equivalence and ultraproducts and let $\mathcal{B} \in \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$. How can we tell whether $\mathcal{B} \in \mathbb{K}$ or not? Using assumptions we made on $\mathbb{K}$, a good try would be to show that $\mathcal{B}$ is elementary equivalent to the ultraproduct of a family of structures we already know to be in $\mathbb{K}$. But in order to do so, we have to know more about the structures in $\mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$.

The next lemma shows a few similarities to the semantical formulation of compactness. It claims that finite parts of theories of models in the smallest elementary class extending $\mathbb{K}$ are already modeled in $\mathbb{K}$.

Lemma 8.1.2 Let $\mathcal{B} \in \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$. Then, for any finite $\Theta \subseteq \mathrm{Th}\,\mathcal{B}$, there is $\mathcal{A}_\Theta \in \mathbb{K}$ with $\mathcal{A}_\Theta \models \Theta$.

Proof. First, note that the assumption $\mathcal{B} \in \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$ implies that $\mathrm{Th}\,\mathbb{K}$ is consistent and thus $\mathbb{K} \neq \emptyset$.
Now, Θ being finite means that $\Theta = \emptyset$ or $\Theta = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathrm{Th}\,\mathbb{K}$ for some $n \in \mathbb{N}, n \geq 1$.
If $\Theta = \emptyset$, then any $\mathcal{A} \in \mathbb{K}$ is a model for Θ.
If $\Theta \neq \emptyset$, we proceed by *reductio ad absurdum*, i.e. we assume that for no $\mathcal{A} \in \mathbb{K}$, $\mathcal{A} \models \Theta$, and from this deduce a contradiction as follows:
If for no $\mathcal{A} \in \mathbb{K}$, $\mathcal{A} \models \Theta$, then

$$\text{for } \textit{every}\ \mathcal{A} \in \mathbb{K}, \mathcal{A} \models \neg(\alpha_1 \wedge \ldots \wedge \alpha_n),$$

so

$$\neg(\alpha_1 \wedge \ldots \wedge \alpha_n) \in \mathrm{Th}\,\mathbb{K}$$

and thus, since $\mathcal{B} \in \mathrm{Mod}\,\mathrm{Th}\,\mathbb{K}$,

$$\mathcal{B} \models \neg(\alpha_1 \wedge \ldots \wedge \alpha_n).$$

On the other hand, since $\Theta \subseteq \operatorname{Th}\mathcal{B}$ and $\operatorname{Th}\mathcal{B}$ is deductively closed, we have

$$\mathcal{B} \models \alpha_1 \wedge \ldots \wedge \alpha_n,$$

which is a contradiction. ■

From here we follow a route similar to the one described in the proof of the semantical formulation of compactness and try to find an ultraproduct built upon these models of finite subsets of the theory of $\mathcal{B}$, hoping that this ultraproduct is elementary equivalent to $\mathcal{B}$.

Please recall the result from Example 7.1.17, which is indeed a proof for the following lemma.

Lemma 8.1.3 Let $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$, $S := \{\Theta \subseteq \Sigma\,;\, \Theta \text{ finite}\}$. For $\alpha \in \Sigma$, let $T_\alpha := \{\Theta \in S\,;\, \alpha \in \Theta\}$. Then, there is an ultrafilter $\mathcal{U} \supseteq \{T_\alpha\,;\, \alpha \in \Sigma\}$.

Proof. Follows immediately from Example 7.1.17. ■

Now we are ready to prove the following lemma.

Lemma 8.1.4 If $\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$ is closed under elementary equivalence and ultraproducts, then $\operatorname{Mod}\operatorname{Th}\mathbb{K} \subseteq \mathbb{K}$.

Proof. Let $\mathcal{B} \in \operatorname{Mod}\operatorname{Th}\mathbb{K}$. According to 8.1.2 we find that for any *finite* $\Theta \subseteq \operatorname{Th}\mathcal{B}$, there is an $\mathcal{A}_\Theta \in \mathbb{K}$ with $\mathcal{A} \models \Theta$.
Now, let $S := \{\Theta \subseteq \operatorname{Th}\mathcal{B}\,;\, \Theta \text{ finite}\}$, for $\alpha \in \operatorname{Th}\mathcal{B}$ let $T_\alpha := \{\Theta \in S\,;\, \alpha \in \Theta\}$ and finally $\mathcal{F} := \{T_\alpha\,;\, \alpha \in \operatorname{Th}\mathcal{B}\}$.
Using 8.1.1 (and the Axiom of Choice!) we obtain a family $\langle \mathcal{A}_\Theta\,;\, \Theta \in S\rangle$ such that $\mathcal{A}_\Theta \models \Theta$ for all $\Theta \in S$, and by Lemma 8.1.3, we know that there is an ultrafilter $\mathcal{U}$ over S whith $\mathcal{F} \subseteq \mathcal{U}$.
Thus, we set $\mathcal{C} := \prod_{\Theta \in S} \mathcal{A}_\Theta \,/\mathcal{U}$ to be the ultraproduct of $\langle \mathcal{A}_\Theta\,;\, \Theta \in S\rangle$ under $\mathcal{U}$.
Take any $\alpha \in \operatorname{Th}\mathcal{B}$. Then, for $\Theta \in S$ with $\alpha \in \Theta$ we have $\mathcal{A}_\Theta \models \alpha$, so

$$\mathcal{A}_\Theta \models \alpha \text{ for all } \Theta \in T_\alpha,$$

whence $T_\alpha \subseteq \{\Theta \in S\,;\, \mathcal{A}_\Theta \models \alpha\}$, and since $T_\alpha \in \mathcal{U}$, we conclude that

$$\{\Theta \in S\,;\, \mathcal{A}_\Theta \models \alpha\} \in \mathcal{U}$$

und thus $\mathcal{C} \models \alpha$. Thus, $\operatorname{Th}\mathcal{B} \subseteq \operatorname{Th}\mathcal{C}$, and by Corollary 6.1.12, we conclude that $\mathcal{B} \equiv \mathcal{C}$.
Thus, we have found an ultraproduct $\mathcal{C}$ of structures $\mathcal{A}_\Theta \in \mathbb{K}$ which is elementary equivalent to $\mathcal{B}$. However, since $\mathbb{K}$ is closed under ultraproducts and elementary equivalence, we conclude that $\mathcal{B} \in \mathbb{K}$. ■

The main result of this section simply follows by putting everything together.

Theorem 8.1.5 (Elementary Classes) For $\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$, the following are equivalent:

(i) $\mathbb{K}$ is an elementary class.

(ii) $\mathbb{K}$ is closed under elementary equivalence and under ultraproducts.

Proof. If $\mathbb{K}$ is an elementary class, then, by Lemma 4.1.12, $\mathbb{K}$ is closed under elementary equivalence and, by Lemma 8.1.1, it is also closed under ultraproducts.
If $\mathbb{K}$ is closed under elementary equivalence and ultraproducts, then, by 8.1.4, $\operatorname{Mod}\operatorname{Th}\mathbb{K} \subseteq \mathbb{K}$ and, by 4.1.4, $\mathbb{K} \subseteq \operatorname{Mod}\operatorname{Th}\mathbb{K}$, thus $\mathbb{K} = \operatorname{Mod}\operatorname{Th}\mathbb{K}$. ■

This is as good a semantic characterization of elementary classes as we can find with the means provided in this module. Actually, to entirely discard all syntactic notions in the characterization of elementary classes we must find a semantic description of *elementary equivalence*. Below we will state a result (without proof) which provides the necessary and sufficient conditions for a class to be elementary in purely semantical terms relying on the notions of ultraproduct and isomorphism. For the time being, we have to make do with ultraproducts and elementary equivalence, with the latter still firmly rooted in syntax.

Notation 8.1.6 For a class $\mathbb{K}$ of $\mathcal{L}$–structures, let

$$\mathbb{K}_{\text{fin}} := \{\mathcal{A} \in \mathbb{K} \,;\, |\mathcal{A}| \text{ is finite}\},$$

and consequently

$$\mathbb{K}_{\text{inf}} := \{\mathcal{A} \in \mathbb{K} \,;\, |\mathcal{A}| \text{ is infinite}\} \,.$$

Example 8.1.7 Let $\mathcal{L}$ be the trivial language containing no non–logical symbols[3], so $\mathcal{L}$–structures are sets with no extra relations, functions or constants. For $n \in \mathbb{N}$ let $\mathcal{A} := \{0, \ldots, n\}$, so every $\mathcal{A}_n$ is an $\mathcal{L}$–structure with exactly $n+1$ elements. Let $\mathcal{U}$ be an ultrafilter over $\mathbb{N}$ and $\mathcal{C} := \prod_{n\in\mathbb{N}} \mathcal{A}_n \,/\mathcal{U}$.
For each $n \in \mathbb{N}$ let α_n be an $\mathcal{L}$–sentence holding in an $\mathcal{L}$–structure $\mathcal{A}$ if and only if $|\mathcal{A}|$ has exactly $n+1$ elements. (Exercise: Find examples of such sentences!) Then, for each $n \in \mathbb{N}$,

$$\{s \in \mathbb{N} \,;\, \mathcal{A}_s \models \alpha_n\} = \{n\} \notin \mathcal{U}$$

[3]Remember that the *non–logical symbols* are given by the index sets I, J and K and the arity–functions λ and μ which characterize $\mathcal{L}$. If a formal language $\mathcal{L}$ contains no non–logical symbols, these index sets are all empty and consequently $\lambda = \mu = \emptyset$ as well. This justifies the formulation "*the* trivial language" instead of "*a* trivial language".

and thus $\mathcal{C} \not\models \alpha_n$. It follows that $|\mathcal{C}|$ is infinite (since $|\mathcal{C}|$ is obviously not empty). Using these observations we can conclude that $\operatorname{Str}\mathcal{L}_{\text{fin}}$ is not closed under ultraproducts and thus not elementary.

Exercise 8.1.8 Analyzing the above argumentation, find a way to show that for any language $\mathcal{L}$ and for any $\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$ one finds that:

If for any $n \in \mathbb{N}$ there is a $\mathcal{A} \in \mathbb{K}_{\text{fin}}$ with card $|\mathcal{A}| \geq n$, then $\mathbb{K}_{\text{fin}}$ is not elementary.

Please note that by constructing an infinite model as an ultraproduct of a family of structures whose cardinalities are finite but without upper bound in $\mathbb{N}$ we implicitely proved that

> there is no set of sentences Σ such that (1) Σ has only finite models but (2) there is no upper bound for the cardinalities of models of Σ.

Also, by using the upward direction of the Löwenheim–Skolem Theorems, Theorem 5.4.3, we may conclude that

> if there is no finite upper bound for the sizes of *finite* structures in some elementary class, then there is no upper bound at all to the cardinalities of the structures;

i.e.

> if, in a given elementary class $\mathbb{K}$, for every $n \in \mathbb{N}$, there is a *finite* structure $\mathcal{A} \in \mathbb{K}$ with card $|\mathcal{A}| \geq n$, then there is, for any cardinal κ, a $\mathcal{B} \in \mathbb{K}$ with $|\mathcal{B}| \geq \kappa$.

Example 8.1.9 Let $\mathcal{L} = \{+, -, \cdot, 0, 1\}$ be the language of rings and fields and let $\mathbb{F}$ be the class of fields with characteristics different from 0. Then, $\mathbb{F}$ is not elementary, as shows the following argument (which exhibits several analogies to the above example):
First, we remember that the characteristic of a field is either a prime number or 0, that $\mathbb{Z}_p$ is a finite field for $p \in \mathbb{N}$ prime and that clearly the characteristic of $\mathbb{Z}_p$ is p.
Then, we set $S := \{p \in \mathbb{N}\,;\, p \text{ prime}\}$. S is infinite and thus there is a free ultrafilter $\mathcal{U}$ over S containing every co–finite subset of S. Now, let $\mathcal{C} := \prod_{p \in S} \mathbb{Z}_p \,/\mathcal{U}$. Since *being a field* is expressible in terms of (finitely many) $\mathcal{L}$–sentences, $\mathcal{C}$ is a field by Łos's Theorem. Yet, by the very same theorem, the characteristic of $\mathcal{C}$ cannot be a prime and thus must be 0.

Exercise 8.1.10 Work out the details to Example 8.1.9.[4]

We conclude that

> The class of fields having characteristic different from 0 is not closed under ultraproducts and thus not elementary.

If we recall that being an elementary class means being axiomatizable by a set of $\mathcal{L}$–sentences, we see that the above examples provide classes of models which are not fully describable in First–Order Logic.

Example 8.1.11
Let $\mathcal{L}$ be the language equipped with a binary function symbol $+$, a binary relation symbol $\leq$ and the constant symbols 0 and 1. Consider the $\mathcal{L}$-structure $\mathbb{N} := \langle \mathbb{N}, +, \leq, 0, 1 \rangle$ with the obvious interpretations of the symbols and let $\mathbb{K} := \operatorname{Mod} \operatorname{Th} \mathbb{N}$. Then, $\mathbb{K}$ is elementary by definition and thus closed under ultraproducts. However, as we saw in 7.2.7, the ultrapower $\mathbb{N}^{\mathbb{N}}/\mathcal{U}$ under a free ultrafilter $\mathcal{U}$ does not have the *Archimedean Property* of $\mathbb{N}$ and has elements exhibiting the behaviour of "infiniteness". Yet, clearly $\mathbb{N}^{\mathbb{N}}/\mathcal{U}$ is a model of $\operatorname{Th} \mathbb{N}$. This is what is sometimes expressed by the following statement:
Every first–order axiomatization[5] of $\mathbb{N}$ has non–standard models.

To sum up, so far we used ultraproducts in two ways: One was to show that some classes are not elementary, and the other was to show that some elementary classes contain non-standard structures. A third application of ultraproducts now follows. Ultraproducts can be used to characterize the complement class of an elementary class.

8.2 Ultraproducts in Basic–Elementary Classes

Basic–elementary classes provide (partial) answers to quite a few questions you may or may not have asked yourself.

- Since an elementary class is just the class of models of some *arbitrary* set of sentences, what happens if we restrict ourselves to *finite* sets of sentences?

[4] Note that for a given p "being of characteristic p" is expressible as an $\mathcal{L}$–sentence. Also, note that in order to express "having characteristic 0" in first–order logic, we need an infinite set of $\mathcal{L}$–sentences, but this we will only be able to proof once we introduce the notion of basic–elementary classes.

[5] Another "detail" should be mentioned here. Clearly such an axiomatization of $\mathbb{N}$ (by first–order sentences) consists of infinitely many sentences. As Gödel proved in the early 1930s, such an axiomatization is inaccessible to a systematic approach in the sense that it is never recursive. So, although we very much believe that there are axiomatizations (e.g. $\operatorname{Th} \mathbb{N}$ itself), we will never fully grasp the form of such an axiomatization (provided Church was right when he claimed that "recursive" and "computable" are one and the same — but that is another story and part of the theory of recursive functions).

- What do maximally proper elementary classes look like?
- An elementary class $\mathbb{K}$ is a class of $\mathcal{L}$–structures axiomatized by a set of $\mathcal{L}$–sentences. Is there an equally elegant description for $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$?

At the end of this section the reader will hopefully have found answers to these questions. As it turns out the best way to access these problems is by starting with the first question.

Definition 8.2.1 A class $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$ is called a **basic–elementary class** if $\mathbb{K} = \mathrm{Mod}\{\alpha\}$ for some $\mathcal{L}$–sentence α.

Basic–elementary classes are obviously elementary. Moreover, the restriction that $\mathbb{K}$ be axiomatized by a *single* sentence can be weakened, as we can see in the following Lemma.

Lemma 8.2.2 $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$ is basic–elementary iff $\mathbb{K} = \mathrm{Mod}\,\Sigma$ for a *finite* $\Sigma \subseteq \mathrm{Sen}\,\mathcal{L}$.

Proof. This is a simple consequence of

$$\mathrm{Mod}\ \{\alpha_1, \ldots, \alpha_n\} = \mathrm{Mod}\{\alpha_1 \wedge \ldots \wedge \alpha_n\}.$$

■

Example 8.2.3 There are a lot of examples for this in "everyday math experience", e.g. the class of fields, rings, groups etc. since they are each characterized by a finite set of axioms.

Clearly, $\emptyset$ and $\mathrm{Str}\,\mathcal{L}$ are basic–elementary as well. (As an exercise: Can you tell why?)

Now we want to take a closer look at the complementary classes of basic–elementary classes.

Lemma 8.2.4 If $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$ is basic–elementary, then so is $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$.

Proof. Let $\mathbb{K} = \mathrm{Mod}\{\alpha\}$, α an $\mathcal{L}$–sentence. Then, for any $\mathcal{A} \in \mathrm{Str}\,\mathcal{L}$,

$$\begin{aligned} \mathcal{A} \in \mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K} \quad & \text{iff} \quad \mathcal{A} \notin \mathbb{K} \\ & \text{iff} \quad \mathcal{A} \not\models \alpha \\ & \text{iff} \quad \mathcal{A} \models \neg\alpha \\ & \text{iff} \quad \mathcal{A} \in \mathrm{Mod}\{\neg\alpha\}. \end{aligned}$$

Thus, $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ is basic–elementary. ■

So, if a class of $\mathcal{L}$–structures is axiomatized by a single sentence, its complement class is axiomatized by the negation of this sentence. Using finite set of sentences instead of a singleton set, this last sentence generalizes to the following statement:

> If a class of $\mathcal{L}$–structures is axiomatized by a *finite* set of sentences, then the complementary class is axiomatized by ... (Exercise: Complete this statement!)

Returning to merely elementary classes we note that basic–elementary classes are elementary classes with elementary complement–classes. Does the converse hold as well? Theorem 8.2.5 confirms this.

Theorem 8.2.5 (Basic–Elementary Classes)
$\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$ is a basic–elementary class iff both $\mathbb{K}$ and $\operatorname{Str}\mathcal{L} \smallsetminus \mathbb{K}$ are elementary classes.

Proof. By Lemma 8.2.4, if $\mathbb{K}$ is basic–elementary, then so is $\operatorname{Str}\mathcal{L} \smallsetminus \mathbb{K}$. Since basic–elementary classes are elementary, $\mathbb{K}$ and $\operatorname{Str}\mathcal{L}$ are both elementary.

So it remains to prove the converse.

Assume $\mathbb{K}$ and $\overline{\mathbb{K}} := \operatorname{Str}\mathcal{L} \smallsetminus \mathbb{K}$ are both elementary classes. If $\mathbb{K} = \emptyset$ or $\mathbb{K} = \operatorname{Str}\mathcal{L}$, then $\mathbb{K}$ is basic–elementary as we have seen in the examples.

So, without loss of generality, assume $\mathbb{K} \neq \emptyset$ and $\operatorname{Str}\mathcal{L} \smallsetminus \mathbb{K} \neq \emptyset$. Let Σ and $\overline{\Sigma}$ be the axiomatizations of $\mathbb{K}$ and $\operatorname{Str}\mathcal{L} \smallsetminus \mathbb{K}$, respectively, i.e. $\mathbb{K} = \operatorname{Mod}\Sigma$ and $\overline{\mathbb{K}} = \operatorname{Mod}\overline{\Sigma}$.

Since $\emptyset = \operatorname{Mod}\Sigma \cap \operatorname{Mod}\overline{\Sigma} = \operatorname{Mod}(\Sigma \cup \overline{\Sigma})$, completeness implies that $\Sigma \cup \overline{\Sigma}$ is inconsistent. By applying the Compactness Theorem 2.4.5 we find a finite $\Sigma_0 \subseteq \Sigma \cup \overline{\Sigma}$ which is already inconsistent.

Now let $\Delta := \Sigma_0 \smallsetminus \Sigma$.

Since Σ_0 is inconsistent, $\Sigma_0 \subseteq \Sigma$ would imply Σ inconsistent and thus $\mathbb{K} = \operatorname{Mod}\Sigma = \emptyset$. So we conclude $\Delta \neq \emptyset$, $\Delta = \{\delta_1, \dots, \delta_n\}$.

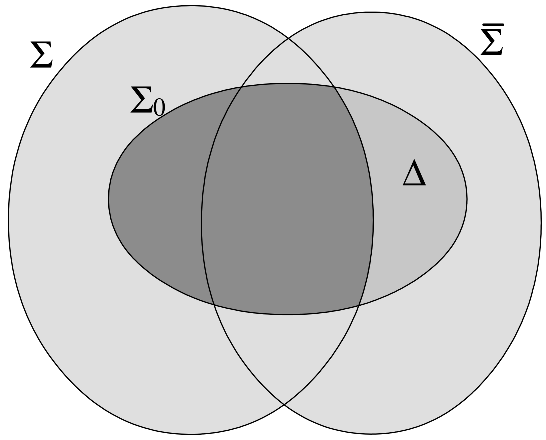
We claim that $\mathcal{A} \in \operatorname{Mod}\Sigma$ iff $\mathcal{A} \not\models \Delta$.

For sufficiency, let $\mathcal{A} \in \operatorname{Mod}\Sigma = \mathbb{K}$ and assume $\mathcal{A} \models \Delta$. Then, $\mathcal{A} \models \Sigma \cup \Delta$, thus $\mathcal{A} \in \operatorname{Mod}(\Sigma \cup \Delta) \subseteq \operatorname{Mod}\Sigma_0 = \emptyset$, a contradiction.

Conversely assume $\mathcal{A} \not\models \Delta$. Then, $\mathcal{A} \not\models \overline{\Sigma}$, since $\Delta \subseteq \overline{\Sigma}$. But this means $\mathcal{A} \notin \operatorname{Mod}\overline{\Sigma}$, so $\mathcal{A} \in \operatorname{Str}\mathcal{L} \smallsetminus \operatorname{Mod}\overline{\Sigma} = \operatorname{Mod}\Sigma$.

Thus, we conclude that $\mathbb{K} = \operatorname{Str}\mathcal{L} \smallsetminus \operatorname{Mod}\Delta$ is the complement of a basic–elementary class, thus, by Lemma 8.2.4, $\mathbb{K}$ is basic–elementary. ■

Hence, if we were really able to provide a purely semantical description of elementary classes (which we are not, since we have not discarded “elementary equivalence” yet), we could equally well describe finitely axiomatizable classes of structures by semantical notions. However, the above will do for the moment.



Without mentioning elementary classes, we can formulate the result from Theorem 8.2.5 in the following corollary.

Corollary 8.2.6 $\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$ is a basic–elementary class iff both $\mathbb{K}$ and $\operatorname{Str}\mathcal{L} \setminus \mathbb{K}$ are both

- closed under elementary equivalence $\equiv$ and
- closed under ultraproducts.

In the proof of Theorem 8.2.5 (Basic–Elementary Classes) we implicitely applied a direct consequence of the Compactness Theorem 2.4.5 which deserves being stated in its own right.

Lemma 8.2.7 (The Covering–Lemma:)
Assume $\Sigma, \Theta \subseteq \operatorname{Sen}\mathcal{L}$ with

$$\operatorname{Mod}\Sigma \subseteq \bigcup_{\alpha\in\Theta} \operatorname{Mod}\alpha.$$

($\bigcup_{\alpha\in\Theta} \operatorname{Mod}\alpha$ is a *covering* of $\operatorname{Mod}\Sigma$.) Then there are $\alpha_1, \ldots, \alpha_n \in \Theta$ such that

$$\operatorname{Mod}\Sigma \subseteq \operatorname{Mod}\alpha_1 \cup \ldots \cup \operatorname{Mod}\alpha_{\mathrm{n}}\,.$$

Exercise 8.2.8 Prove the Covering Lemma.
(Hint: Note that $\operatorname{Mod}\Sigma \subseteq \bigcup_{\alpha\in\Theta} \operatorname{Mod}\alpha$ iff $\operatorname{Mod}\Sigma \cap \bigcap_{\alpha\in\Theta} \operatorname{Mod}\neg\alpha = \emptyset$.)

Example 8.2.9 For $\mathcal{L}$ the trivial language, we saw in Example 8.1.7 that $\operatorname{Str}\mathcal{L}_{\text{fin}}$ is not closed under ultraproducts and thus is not elementary. So $\operatorname{Str}\mathcal{L}_{\text{inf}} = \operatorname{Str}\mathcal{L} \setminus \operatorname{Str}\mathcal{L}_{\text{fin}}$ is clearly not basic–elementary. Still, $\operatorname{Str}\mathcal{L}_{\text{inf}} =$

$\operatorname{Str}\mathcal{L} \setminus \operatorname{Str}\mathcal{L}_{\text{fin}}$ is elementary, as you will have no trouble showing by finding an appropriate set of sentences Σ which is satisfied in a $\mathcal{L}$–structure if and only if the structure has an infinite universe. (Exercise: Do so!)
Thus we observe for the trivial language $\mathcal{L}$ that the class of finite $\mathcal{L}$–structures cannot be axiomatized at all. The class of infinite $\mathcal{L}$–structures can be axiomatized, but not by a finite set of $\mathcal{L}$–sentences.
Please note, we did not claim that there are no theories with only finite models. What we did say is that given a set of sentences whose model–class contains exclusively finite models, there is an $n \in \mathbb{N}$ such that every model in this class has at most n elements. (Exercise: We are quite sure you will have no trouble finding such a set of sentences which has only finite models.)

Example 8.2.10 For $\mathcal{L} = \{+, -, \cdot, 0, 1\}$, the language of rings and fields, we saw in Example 8.1.9 that the class $\mathbb{F}$ of fields with characteristic different from 0 is not elementary. Again, we can easily confirm that $\operatorname{Str}\mathcal{L} \setminus \mathbb{F}$ is elementary. So we conclude that although e.g. the class of *all* fields is finitely axiomatizable the subclass of fields with characteristic 0 is *not finitely* axiomatizable.

We now turn our attention to the question about maximally proper elementary classes. The first result is obvious.

Lemma 8.2.11 Any elementary class is the (class–)intersection of basic–elementary classes.

Proof. This is a simple consequence of

$$\operatorname{Mod}\Sigma = \bigcap \{\operatorname{Mod}\varphi \,;\, \varphi \in \Sigma\} .$$

■

From this we conclude the following characterization of basic–elementary classes and provide a partial answer to the second of the above questions: "What do maximally proper elementary classes look like?"

Corollary 8.2.12 If $\mathbb{K} \subseteq \operatorname{Str}\mathcal{L}$ is an elementary class which is maximal among the proper, elementary subclasses of $\operatorname{Str}\mathcal{L}$ (i.e. $\mathbb{K}_0$ elementary and $\mathbb{K} \subseteq \mathbb{K}_0 \neq \operatorname{Str}\mathcal{L}$ implies $\mathbb{K} = \mathbb{K}_0$), then $\mathbb{K}$ is basic–elementary.

Proof. We prove the contraposition. If $\mathbb{K}$ is elementary but not basic–elementary, then $\emptyset \neq \mathbb{K} \neq \operatorname{Str}\mathcal{L}$, so $\operatorname{Th}\mathbb{K} \not\subseteq \operatorname{Th}\operatorname{Str}\mathcal{L}$, since otherwise $\mathbb{K} = \operatorname{Mod}\operatorname{Th}\mathbb{K} \supseteq \operatorname{Mod}\operatorname{Th}\operatorname{Str}\mathcal{L} = \operatorname{Str}\mathcal{L}$.
Now take any $\alpha \in \operatorname{Th}\mathbb{K} \setminus \operatorname{Th}\operatorname{Str}\mathcal{L}$. Then,

$$\mathbb{K} = \operatorname{Mod}\operatorname{Th}\mathbb{K} \subseteq \operatorname{Mod}\alpha \text{ (as } \{\alpha\} \subseteq \operatorname{Th}\mathbb{K}).$$

Moreover, $\operatorname{Mod}\alpha \neq \operatorname{Str}\mathcal{L}$ since otherwise $\alpha \in \operatorname{Th}\operatorname{Mod}\alpha = \operatorname{Th}\operatorname{Str}\mathcal{L}$, which contradicts the choice of α. Thus, $\operatorname{Mod}\alpha$ is a basic–elementary class $\mathbb{K}_0$ with $\mathbb{K} \subsetneq \mathbb{K}_0 \subsetneq \operatorname{Str}\mathcal{L}$. ∎

Remark 8.2.13 Generally the converse of Corollary 8.2.12 is not true, as the following example shows.
Let $\mathcal{L}$ be the language with one unary function–symbol s and one constant–symbol 0. For $n \in \mathbb{N}$, $n \geq 1$, let $t_n \in \operatorname{Tm}\mathcal{L}$ be defined by

$$t_1 :\equiv s(0) \quad \text{and} \quad t_{n+1} :\equiv s(t_n).$$

Let $\Sigma := \{\alpha_n \,;\, n \in \mathbb{N}\}$ where $\alpha_n :\equiv \neg t_n \doteq 0$ and let $\beta :\equiv \forall x \neg s(x) \doteq 0$. For $n \in \mathbb{N}$ let $\mathcal{A}_n$ be the $\mathcal{L}$–structure with universe $\mathbb{Z}_n$ and $s^{\mathcal{A}_n}(m) := m + 1$. (This justifies the choice of s for the function symbol as an abbreviation of "successor".) Let $\mathcal{A}$ be the $\mathcal{L}$–structure with universe $\{k \in \mathbb{Z} \,;\, k \geq -1\}$ and again the successor function as the interpretation of s. Then, we claim that

1. $\beta \vdash \Sigma$, thus $\operatorname{Mod}\beta \subseteq \operatorname{Mod}\Sigma$;
2. there is an ultraproduct of $\langle \mathcal{A}_n \,;\, n \in \mathbb{N}\rangle$ which is a model for Σ while each $\mathcal{A}_n$ is not;
3. $\operatorname{Mod}\Sigma$ is not basic–elementary.

Thus $\operatorname{Mod}\beta$ is a basic–elementary class which is properly contained in the elementary but not basic–elementary class $\operatorname{Mod}\Sigma$.

Exercise 8.2.14 Fill in the details of the proofs of the claims stated in the above remark.

We have seen that the basic–elementary classes are exactly those elementary classes whose complement–class is elementary as well. Moreover, maximal elementary classes are basic–elementary.

Now we draw one more conclusion by combining our knowledge about basic–elementary classes with the Downward Löwenheim–Skolem Theorem 5.3.1:

Theorem 8.2.15 In every non–empty basic–elementary class there is a countable structure.

Proof. (Sketchy) If $\mathbb{K}$ is basic–elementary, then $\mathbb{K} = \operatorname{Mod}\alpha$ for some $\mathcal{L}$–sentence α. Since sentences are finite concepts, only finitely many non–logical symbols from $\mathcal{L}$ occur in α. Let $\mathcal{L}_0$ be the sub–language of $\mathcal{L}$ comprising exactly these symbols. Then α is an $\mathcal{L}_0$–sentence as well, and clearly every $\mathcal{L}$–structure

$\mathcal{A}$ may be easily made into an $\mathcal{L}_0$–structure simply by forgetting the interpretations of the symbols from $\mathcal{L}$ which do not belong to $\mathcal{L}_0$.Conversely, every $\mathcal{L}_0$–structure can be viewed as an $\mathcal{L}$–structure if we add arbitrary interpretations for the non–logical symbols, which does not affect the satisfaction of α or the cardinality of the structure. Moreover, every $\mathcal{L}$–structure is a model of α iff it is so as an $\mathcal{L}_0$–structure. From the Downward Löwenheim–Skolem Theorem 5.3.1 we know that there is a countable $\mathcal{L}_0$–structure which is a model for α, but this model can be extended to a model in $\mathbb{K}$. ∎

8.3 Discarding $\equiv$

Our original task of expressing closure properties of semantical classes without using syntactical concepts is not yet completely accomplished since we make extensive use of elementary equivalence. In this section, we finally want to show a way of using semantical concepts exclusively for working with elementary classes. However, we will not be able to give the proof of the main result, since this lies beyond the scope of this lecture.

Still, half the work is already done, since we showed in Theorem 4.4.6 that for any two $\mathcal{L}$–structures $\mathcal{A}$ and $\mathcal{B}$, if $\mathcal{A}$ and $\mathcal{B}$ are isomorphic, then they are elementary equivalent:

$$\text{if } \mathcal{A} \cong \mathcal{B}, \text{ then } \mathcal{A} \equiv \mathcal{B}.$$

The following main result we are now going to state should give you an idea of the importance of ultraproducts in model theory.

Theorem 8.3.1 (Elementary Equivalence via Isomorphism) Let $\mathcal{A}$ and $\mathcal{B}$ be two $\mathcal{L}$–structures. Then the following are equivalent:

(i) $\mathcal{A} \equiv \mathcal{B}$.

(ii) For some index sets I and J and ultrafilters $\mathcal{U}$ over I and $\mathcal{V}$ over J, $\mathcal{A}^I/\mathcal{U} \cong \mathcal{B}^J/\mathcal{V}$.

This means that two structures are elementary equivalent iff they have isomorphic ultrapowers.

Proof. To show that (ii) implies (i) is not difficult and is thus left as an exercise. The converse, however, is rather complex and the reader is referred to the literature. ∎

We use Theorem 8.3.1 to finally prove the following corollary, the objective of our endeavour.

Corollary 8.3.2 For $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$, the following statements are equivalent

(i) $\mathbb{K}$ is an elementary class;

(ii) $\mathbb{K}$ is closed under isomorphisms and ultraproducts, and $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ is closed under ultrapowers.

Proof. Both directions rely on Theorem 8.1.5 by which $\mathbb{K}$ *is elementary iff* $\mathbb{K}$ *is closed under elementary equivalence and ultraproducts.*

(i) implies (ii): If $\mathbb{K}$ is an elementary class, then $\mathbb{K}$ is closed under ultraproducts (by Theorem 8.1.5). Moreover, if $\mathcal{B} \cong \mathcal{A} \in \mathbb{K}$, then $\mathcal{B} \equiv \mathcal{A}$ by Theorem 4.4.6, so $\mathcal{B} \in \mathbb{K}$ (by Theorem 8.1.5). Thus, $\mathbb{K}$ is closed under isomorphisms. Moreover, if $\mathcal{A}^I/\mathcal{U} \in \mathbb{K}$ for some ultrapower $\mathcal{A}^I/\mathcal{U}$ of $\mathcal{A}$, then $\mathcal{A} \in \mathbb{K}$, since $\mathcal{A}^I/\mathcal{U} \equiv \mathcal{A}$ and $\mathbb{K}$ is closed under $\equiv$, so contraposition shows $\mathcal{A}^I/\mathcal{U} \notin \mathbb{K}$ whenever $\mathcal{A} \notin \mathbb{K}$, i.e. $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ is closed under ultrapowers.

(ii) implies (i): Assume $\mathbb{K}$ is closed under isomorphisms and ultraproducts, and $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ is closed under ultrapowers. By Theorem 8.1.5, all that remains to be shown is that $\mathbb{K}$ is closed under elementary equivalence. Assume $\mathcal{B} \equiv \mathcal{A} \in \mathbb{K}$. By Theorem 8.3.1 there are isomorphic ultrapowers $\mathcal{A}^I/\mathcal{U}$ of $\mathcal{A}$ and $\mathcal{B}^J/\mathcal{V}$ of $\mathcal{B}$. $\mathcal{A}^I/\mathcal{U} \in \mathbb{K}$ by the premise, so $\mathcal{B}^J/\mathcal{V} \in \mathbb{K}$ since $\mathbb{K}$ is closed under $\cong$ by the premise as well. Now, if $\mathcal{B} \notin \mathbb{K}$, then $\mathcal{B} \in \mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$, but then also $\mathcal{B}^J/\mathcal{V} \in \mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ by the premise. This is a contradiction. We conclude $\mathcal{B} \in \mathbb{K}$, i.e. $\mathbb{K}$ is elementary according to Theorem 8.1.5. ■

For basic–elementary classes, the criterion is even simpler.

Corollary 8.3.3 For $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$, the following statements are equivalent

(i) $\mathbb{K}$ is a basic–elementary class.

(ii) Both $\mathbb{K}$ and $\mathrm{Str}\,\mathcal{L} \smallsetminus \mathbb{K}$ are closed under isomorphisms and ultraproducts.

Proof. As an exercise combine Corollary 8.3.2 and Theorem 8.2.5 to prove the claim. ■

We find that (basic) elementary classes can be described using only the semantical notions of ultraproducts, ultrapowers and isomorphisms.

Chapter 9

Universal Algebra

The word "algebra" carries different meanings in mathematics: First of all, it stands for one of the large fields into which mathematics is commonly divided, at the same level as geometry or analysis. Then, it also names a very specific sort of mathematical structure, namely vector spaces with a multiplication defined for their vectors; a prime example is the set of complex numbers, thought of as a vector space over the reals.

Starting with this chapter, we will use the term "algebra" for a concept situated at an intermediate level of generality. We will use it for a set equipped with some specified operations. Since Model Theory is the main subject of this lecture, we shall root the notion of a *(universal) algebra* within the realm of structures (and languages). Later we will gradually abandon the Model Theoretic approach and treat algebras as a fundamental notion.

9.1 Algebras

Looking at structures for a formal language the way we did, we roughly had four components characterizing a structure: Its universe, its constants, its functions and its relations. The exertions with the Löwenheim–Skolem Theorems showed that relations play an entirely different role than the rest of the semantical components making up a structure. To be the universe of a substructure a subset has to comply to certain closure–conditions concerning the constants and functions, the relations however are simply *imposed* on the smaller structure by restriction. This aspect is even more emphasized when we deal with structures lacking any relations. These are called *(universal) algebras*. At first glance, the definition looks quite different from the definition of a structure, but, as we will see, these divergences are marginal and cannot really obscure the common origin of the two concepts.

Definition 9.1.1 A **type** t consists of a family $t = \langle r_s ; s \in S\rangle$ of nonnegative integers r_s, together with a family (over the same index set S) $\langle f_s ; s \in S\rangle$ of **operation symbols**. r_s is the **arity** $\mathrm{ar}(f_s)$ of f_s.

There is nothing remarkable about this definition, except for the fact that we did not exclude $S = \emptyset$, nor $r_s = 0$. The significance of this will become clear after the next definition.

Definition 9.1.2 If $t = \langle r_s ; s \in S\rangle$ is a type, then a **(universal) algebra** $\mathbf{A} = \langle A; f_s^{\mathbf{A}}\rangle_{s\in S}$ **of type** t consists of a set A, the **universe** of $\mathbf{A}$, and a family $\langle f_s^{\mathbf{A}} ; s \in S\rangle$ of functions $f_s^{\mathbf{A}} : A^{r_s} \longrightarrow A$, the so called **fundamental operations** of $\mathbf{A}$. Algebras of the same type are called **similar**.

Since we are concerned exclusively with universal algebras in this module, the attribute "universal" will be dropped in most cases. The same goes for "fundamental" as far the operations of $\mathbf{A}$ are concerned. Moreover, if the algebra $\mathbf{A}$ is clear from the context, the operation symbol f_i and the operation $f_i^{\mathbf{A}}$ will be used interchangeably.

Many of the popular examples presented in this chapter will be of finitary character, that is, the index set S of the type (and thus the type itself) will be finite. If this is the case, say $t := \langle r_1, \ldots, r_n\rangle$, we will denote algebras $\mathbf{A}$ of type t in the form $\mathbf{A} = \langle \mathbf{A}; f_1^{\mathbf{A}}, \ldots, f_n^{\mathbf{A}}\rangle$. Some of our key concerns for this lecture are examples and case studies, and these examples typically deal with algebras having very few fundamental operations.

In the light of the introductory remarks above, we see that the type fixes a formal language $\mathcal{L}$ such that an algebra $\mathbf{A}$ of this type is nothing less than an $\mathcal{L}$–structure. This justifies adopting many of the conventions and notations for structures to algebra. We will, for example, call an algebra finite if and only if its universe is finite. Denotationally, we distinguish algebras from structures in that we denote algebras by boldface roman letters $\mathbf{A}$, $\mathbf{B}$, whereas structures are denoted by calligraphic letters $\mathcal{A}$, $\mathcal{B}$.

Of course, languages $\mathcal{L}$ for algebras are devoid of relation symbols. Consequently, we will call such a language a **functional** language. Thus, given any functional language $\mathcal{L}$ we find that all $\mathcal{L}$–structures are algebras. On the other hand, any type gives rise to some functional language. This language, although not uniquely determined, is still fixed in the sense that there is a strict correspondence of function– and constant–symbols and the arities involved of any two such languages. Consequently, we will address such a language as *the* **underlying language** of a type or of an algebra or of a class of algebras of common type.

The special role of **nullary operations** (i.e. operations having arity 0) will become clear in the following remarks.

Remark 9.1.3 The definition just given looks *very* general — which is quite fitting for the notion of a *universal* algebra. However, it is the product of several deliberate choices:

1. By not excluding $n = 0$, we allow a set A with no operations at all to be an algebra.

2. The set A may be empty according to our definition. This is handy in most cases, but may require some care as exhibited below and clearly stands in contrast to the definition of $\mathcal{L}$–structures.

3. By allowing non–negative integers as arities, we include the possibility $r_i = 0$ for some i. What is an operation of arity 0? In Set Theory , A^{r_i} is constructed as the set of all maps from $\{0, \ldots, r_i - 1\}$ into A. If $r_i = 0$, this becomes the set of all maps from $\emptyset$ into A. Thinking of maps as sets of ordered pairs, there is exactly one such map, namely $\emptyset$ (not depending on A being empty or not!); in other words. $A^0 = \{\emptyset\}$.

   Consequently, an operation of arity 0 is a map $f : \{\emptyset\} \longrightarrow A$. Now if $A \neq \emptyset$, f is completely determined by $f(\emptyset) \in A$. Summing up, operations of arity 0 may be identified with elements of A, called **constants** or **designated elements** in this context — provided that A is not empty. If $A = \emptyset$, there exist *no* maps from $\{\emptyset\}$ into A, so an empty algebra cannot have any nullary operations. (In fact, in an empty algebra only very few operations are possible. Which one?)

4. Restricting arities to integers excludes operations taking infinitely many "inputs". This restriction to so-called **finitary operations** is standard practice and does not affect the topics we will discuss in this module in any way.

5. Operations as defined in 9.1.2 are defined for **every** tuple of elements of A (of the correct length). This excludes so-called **partial algebras** where maps $f_s : D \longrightarrow A$ are admissible as operations for arbitrary subsets $D \subseteq A^{r_s}$. The theory of partial algebras is well beyond the scope of this lecture; the interested reader is referred to further literature.

We will now consider a preliminary batch of examples in order to sketch the scope of Definition 9.1.2. With one notable exception, most of the algebraic structures encountered in any undergraduate curriculum fit neatly in the frame of Definition 9.1.2. In familiar situations we use the conventional infix notation in place of both f_s and f_s^{A}.

Example 9.1.4 1. Any *group* G is an algebra $\mathbf{G}$ of type $(2,1,0)$ with $\mathbf{G} = \langle G; \cdot, ^{-1}, e\rangle$, $\cdot$ denoting multiplication, $^{-1}$ inverses and e the neutral element.

2. Another algebra $\mathbf{S}$ of type $(2,1,0)$ is obtained by choosing the power set $P(X)$ of some fixed set X as universe S, set intersection $\cap$ as operation of arity 2, set complement ^{c} with respect to X as operation of arity 1 and $\emptyset$ as constant, thus $\mathbf{S} = \langle S; \cap, ^c, \emptyset\rangle$.

   Comparing these algebras with groups, we see that similar algebras need not be very similar in the nontechnical sense of the word.

3. Among all algebras of type $(2,1,0)$, groups may be characterized by the familiar *group laws*

   (A) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

   (N) $x \cdot e = e \cdot x = x$

   (I) $x \cdot x^{-1} = x^{-1} \cdot x = e$

   requiring that $\cdot$ be associative, that e acts as neutral element and x^{-1} as inverse element for x with respect to $\cdot$. Such "laws" specify a subclass of a given class of algebras of a given type, and they present an axiomatization of this class as a *(basic–)elementary class* (cf. Section **??**).

4. Any group $\mathbf{G}$ may also be viewed as an algebra of type (3) with a ternary operation m defined on G by $m(x,y,z) := x \cdot y^{-1} \cdot z$ where $\cdot$ of course denotes the original multiplication operation which comes with $\mathbf{G}$. The question arises naturally whether one may find group laws formulated exclusively in terms of m which characterize groups among all algebras of type (3). (Exercise: Answer this question.)

5. Any (unitary) *ring* $\mathbf{R} = \langle R; +, -, \cdot, 0, 1\rangle$ is an algebra $\mathbf{R}$ of type $(2,1,2,0,0)$, where $+$ denotes addition, $-$ additive inverses, $\cdot$ multiplication, 0 the additive and 1 the multiplicative neutral element.

6. A *combinatory algebra* is an algebra $\mathbf{X} = \langle X; \cdot, \mathsf{K}, \mathsf{S}\rangle$ of type $(2,0,0)$ satisfying $(\mathsf{K} \cdot x) \cdot y = y$ and $((\mathsf{S} \cdot x) \cdot y) \cdot z = (x \cdot z) \cdot (y \cdot z)$. Since the binary operation $\cdot$ in combinatory algebras need not be associative nor commutative, we will use such algebras mainly to produce unfamiliar examples or counterexamples. They originate from logic and provide a more algebraic counterpart to λ–calculus. (For more on the subject of combinatory algebras and λ–calculus the reader is referred to the literature.)

7. The one familiar algebraic structure missing from our list of examples is that of a vector space. While vector addition is an ordinary binary operation on V, everything becomes a little less finitary than in the above examples if the vector space we try to present as an algebra is a vector space over an infinite field F. Then we need a unary operation $f_x : V \longrightarrow V$ given by $f_x(v) := xv$, for any $x \in F$ to express scalar multiplication of the vector $v \in V$ by the scalar $x \in F$.

The next example is rather special, first, in the sense that it mixes different topics (syntax and semantics, algebras and languages), and second, in that it presents a rather elegant way of demonstrating an algebraic access to logic and model theory.

Example 9.1.5 If $\mathcal{L}$ is a formal language, we get an algebra $\mathbf{T}_{\mathcal{L}}$ by choosing as the universe the set $\mathrm{Tm}\,\mathcal{L}$ of all $\mathcal{L}$–terms, and as operations $f_j^{\mathbf{T}}$ the process of building a composite term $f_j(t_1, \ldots, t_{\mu(j)})$ from the terms $t_1, \ldots, t_{\mu(j)}$. So $f_j^{\mathbf{T}}(t_1, \ldots, t_{\mu(j)}) := f_j(t_1, \ldots, t_{\mu(j)})$, and the resulting algebra could (misusing the notation only very slightly) be written as

$$\mathbf{T} := \langle \mathrm{Tm}\,\mathcal{L}; f_j, c_k \rangle_{j \in J, k \in K}.$$

In the same vein, consider the set $\mathrm{Fml}\,\mathcal{L}$ of $\mathcal{L}$–formulae. Again, we could look at the process of composing formulae to more complex ones as a fundamental operation on $\mathrm{Fml}\,\mathcal{L}$, thus regarding $\mathrm{Fml}\,\mathcal{L}$ as an algebra $\mathbf{F}$. $\mathbf{F}$ would then, at least for the approach to formal languages chosen in this lecture, have one binary operation ($\wedge$), one unary operation ($\neg$) and countably infinite unary operations ($\forall v_n$ for all $n \in \mathbb{N}$).

The approach to algebras treated in these two examples relates closely to the term–structures for building syntactical models. Consequently, algebras thus defined are often called **term–algebras** and they share an important *universal* property: They are *freely generated* by some set of generators. Some more details on this subject are to follow in Section **??**.

9.2 Homomorphisms

Homomorphisms, isomorphisms and related notions were already introduced in the context of structures in chapter 4.4. We are now in the somewhat easier situation where we can disregard relations, which may or may not be preserved by a potential homomorphism. Definition 4.4.1 thus simplifies to the following Definition:

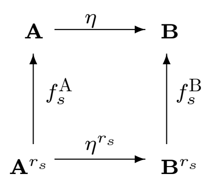
Definition 9.2.1 Let $\mathbf{A} = \langle A; f_s^{\mathbf{A}}\rangle_{s\in S}$, $\mathbf{B} = \langle B; f_s^{\mathbf{B}}\rangle_{s\in S}$ be two similar algebras. A map $\eta : A \longrightarrow B$ is a **homomorphism** from $\mathbf{A}$ into $\mathbf{B}$ iff for every $s \in S$ and every r_s-tuple $a_1, \dots, a_{r_s}$ of elements of A

$$\eta(f_s^{\mathbf{A}}(a_1, \dots, a_{r_s})) = f_s^{\mathbf{B}}(\eta(a_1), \dots, \eta(a_{r_s})).$$

$\mathbf{A}$ is called the **source** (or domain) and $\mathbf{B}$ the **target** (or co–domain) of η. If η is a surjective mapping, $\mathbf{B}$ is called a **homomorphic image** of $\mathbf{A}$. We write $\mathrm{Hom}(\mathbf{A}, \mathbf{B})$ for the set of all homomorphisms from $\mathbf{A}$ into $\mathbf{B}$.

In particular, a constant of $\mathbf{A}$ will be mapped to the corresponding constant of $\mathbf{B}$ under any homomorphism. (Exercise: Write a detailed proof of this claim, using 9.1.3.) Recalling the notation η^{r_s} from Section 4.4, the above situation (at least for the case $r_s > 0$) can be written as

$$\eta \circ f_s^{\mathrm{A}} = f_s^{\mathrm{B}} \circ (\eta^{r_s}).$$



The next Proposition is simply a reformulation of Lemma 4.4.2 in the context of algebras.

Proposition 9.2.2 If $\mathbf{A}$ and $\mathbf{B}$ are two algebras of the same type with underlying language $\mathcal{L}$, then $\eta : \mathcal{A} \longrightarrow \mathcal{B}$ is a homomorphism iff for any $\mathcal{L}$–term t and any valuation h into $\mathbf{A}$, $\eta(t^{\mathbf{A}}[h]) = t^{\mathbf{B}}[\eta \circ h]$.

If $\mathcal{L}$ is a formal language and $t(x_1, \dots, x_n)$ an $\mathcal{L}$–term containing exactly the free variables $x_1, \dots, x_n$, then, for any $\mathcal{L}$–structure, t uniquely determines an n–ary function $t^{\mathcal{A}} : |\mathcal{A}|^n \longrightarrow |\mathcal{A}|$ by

$$t^{\mathcal{A}}(a_1, \dots, a_n) := t^{\mathcal{A}}[h\binom{x_1}{a_1} \dots \binom{x_n}{a_n}]$$

for any $a_1, \dots, a_n \in |\mathcal{A}|$, h an arbitrary valuation into $\mathcal{A}$. An arbitrary function $f : |\mathcal{A}|^n \longrightarrow |\mathcal{A}|$ which is (as a set of ordered pairs) equal to a function stemming from a term in the above sense is called a **term–function** (on $\mathcal{A}$). Thus, the above proposition says that a map is a homomorphism between algebras if and only if it is compatible with all the term–functions.

For the sake of completeness it must be mentioned that we could, with some difficulty, have defined term–functions inductively. However, we feel that the meaning of the above is clear.

From Definition 4.4.3 and the remarks thereafter, we recall the special kinds of homomorphisms and their properties. It follows that in the context of algebras (i.e. in the absence of relations) a homomorphism $\eta : \mathbf{A} \longrightarrow \mathbf{B}$ is an **isomorphism** if and only if η is injective and surjective (i.e., bijective). As for $\mathcal{L}$–structures in general, isomorphisms delimit the degree of resolution adopted by Universal Algebra in studying its objects. Accordingly, a property of algebras is called an **algebraic property** if it is preserved under any isomorphism. As an example, the property of being a commutative group is algebraic while the property of being a group of permutations of $\{1, \dots, n\}$ is not, since any group of permutations of $\{1, \dots, n\}$ is isomorphic to some group of permutations of any set with n elements.

By definition, isomorphic algebras are always similar. The converse is not true. (Exercise: Find a simple example of two similar non–isomorphic algebras.)

Some more specializations deserve mentioning:

A homomorphism from $\mathbf{A}$ into itself is called an **endomorphism** of $\mathbf{A}$, and an isomorphism from $\mathbf{A}$ onto itself an **automorphism** of $\mathbf{A}$. The set of all endomorphisms $\mathrm{End}(\mathbf{A})$ of $\mathbf{A}$ is the universe of the algebra

$$\mathrm{End}(\mathbf{A}) = \langle \mathrm{End}(\mathbf{A}); \circ, \mathrm{id}_{\mathbf{A}} \rangle$$

of type $(2, 0)$ where $\circ$ denotes composition of maps as usual and $\mathrm{id}_{\mathbf{A}}$ is of course the identity map of A. (For those among the readers with some algebraic background: $\mathrm{End}(\mathbf{A})$ is a monoid.) Similarly, the set $\mathrm{Aut}(\mathbf{A})$ of all automorphisms of $\mathbf{A}$ is the universe of the algebra

$$\mathrm{Aut}(\mathbf{A}) = \langle \mathrm{Aut}(\mathbf{A}); \circ, {}^{-1}, \mathrm{id}_{\mathbf{A}} \rangle$$

of type $(2, 1, 0)$ where $^{-1}$ denotes inverses of maps ($\mathrm{Aut}(\mathbf{A})$ is a group).

Example 9.2.3

1. As an exercise: Verify for groups, rings and vector spaces that the respective definitions of homomorphisms match with the definition just given in this section.

2. Regarding the term–algebra for some formal language $\mathcal{L}$, the reader is invited to verify that any valuation h into an $\mathcal{L}$–structure $\mathcal{A}$ determines an $\mathcal{L}$–homomorphism $\eta_h : \mathbf{T}_{\mathcal{L}} \longrightarrow \mathcal{A}$ by $\eta_h(t) := t^{\mathcal{A}}[h]$. Also verify that if the language is functional (i.e. if it contains no relation–symbols) we are actually facing homomorphisms in the sense of the above definition.

9.3 Subuniverses and Subalgebras

Given a group $\mathbf{G} = \langle G; \cdot, ^{-1}, e\rangle$, a subgroup of $\mathbf{G}$ is, roughly said, any subset of G which is closed under the operations $\cdot$ and $^{-1}$ and contains e. Again, there is a straightforward generalization to the setting of arbitrary algebras, as we have already seen for substructures. Since we do not need to worry about relations (remember that they were defined on substructures simply as restrictions, without any further constraints), the definition of a subalgebra as a substructure of an algebra is straightforward. Recall the definition of subuniverse from Definition **??**.

Definition 9.3.1 Let $\mathbf{A} = \langle A; f_s^{\mathbf{A}}\rangle_{s\in S}$ be an algebra. An algebra $\mathbf{B} = \langle B; f_s^{\mathbf{B}}\rangle_{s\in S}$ of the same type as $\mathbf{A}$ is a **subalgebra** of $\mathbf{A}$ iff B is a subuniverse of $\mathbf{A}$ and $f_s^{\mathbf{B}}$ is the restriction to B of $f_s^{\mathbf{A}}$ for all $s \in S$. Given any subuniverse $B \subseteq A$, the canonical subalgebra living on B will be denoted by $\mathbf{B}$. We write $\mathbf{Sub}\,\mathbf{A}$ for the collection of all subuniverses of $\mathbf{A}$.

Example 9.3.2

1. Consider an arbitrary group G with operations $\cdot$ and $^{-1}$, and neutral element e. Considered as an algebra $\mathbf{G}$ of type $(2, 1, 0)$, $\mathbf{G} = \langle G; \cdot, ^{-1}, e\rangle$, the subalgebras of $\mathbf{G}$ are just the ordinary subgroups of G. If $\mathbf{G} = \langle G; \cdot\rangle$, the subuniverses are the subsets closed under $\cdot$, including $\emptyset$. Viewed as an algebra of type (3) with the operation m as defined in 9.1.4.4, the subalgebras are the co–sets[1] (left or right) of any ordinary subgroup of G and $\emptyset$, endowed with the restriction of m as its only operation. (Exercise: Prove this statement.)

2. A subalgebra of the term–algebra $\mathbf{T}_{\mathcal{L}}$ as defined in 9.1.5 can easily be found by restriction to *closed terms*, i.e. the terms containing no variables.

We list some straightforward generalizations of facts well-known from the group, ring or vector space setting. The proof is left as an exercise.

Proposition 9.3.3 Let $\mathbf{A}$ and $\mathbf{B}$ be algebras and $\eta : \mathbf{A} \longrightarrow \mathbf{B}$ any homomorphism.

1. If $S \subseteq A$ is a subuniverse of $\mathbf{A}$, then $\eta[S] = \{\eta(s)\ ;\ s \in S\}$ is a subuniverse of $\mathbf{B}$. Extended notation established in 9.3.1, the corresponding subalgebra of $\mathbf{B}$ will be written $\eta[\mathbf{S}]$. $\eta[\mathbf{S}]$ is clearly a homomorphic image of $\mathbf{S}$ in the sense of 9.2.1.

[1] If S is a subset of a group G, then a **left co–set** of S in G is a set of the form $g \cdot S := \{g \cdot s\ ;\ s \in S\}$ for some $g \in G$. Similarly, a **right co–set** of S in G is of the form $S \cdot g := \{s \cdot g\ ;\ s \in S\}$.

2. If $T \subseteq B$ is a subuniverse of $\mathbf{B}$, then $\eta^{-1}[T] = \{a \in A : h(a) \in T\}$ is a subuniverse of $\mathbf{A}$. Again, $\eta^{-1}[\mathbf{T}]$ stands for the corresponding subalgebra of $\mathbf{A}$.

3. If $\{S_k ; k \in K\}$ is a chain[2] of subuniverses of $\mathbf{A}$, then their set union $\bigcup_{k \in K} S_k$ is a subuniverse of $\mathbf{A}$. More generally, if $\mathcal{D} = \{S_k ; k \in K\}$ is a set of subuniverses of $\mathbf{A}$ directed by $\subseteq$ (i.e. for $S_{k_1}, S_{k_2} \in \mathcal{D}$ there is always an $S \in \mathcal{D}$ such that $S_{k_1} \cup S_{k_2} \subseteq S$), then $\bigcup \mathcal{D}$ is a subuniverse of $\mathbf{A}$.

If $\langle S_k ; k \in K \rangle$ is any family of subuniverses of an algebra $\mathbf{A}$, then the set intersection $\bigcap_{k \in K} S_k$ is a subuniverse of $\mathbf{A}$. This even applies to the empty family of subuniverses whose intersection is A itself. We conclude with the following proposition.

Proposition 9.3.4 **Sub A** is a closure system and thus a complete lattice (under $\subseteq$).

The tricky part is to be more specific about the supremum of a family of subuniverses. Examples where suprema do not coincide with the set union are easily found (Exercise!), consequently we have to ask ourselves how these suprema are to be defined. From Section **??** we know that for any $X \subseteq A$ there is a smallest sub*universe* of $\mathbf{A}$ extending X, the substructure generated by X, $\mathbf{A}[X]$. Clearly, by restricting the fundamental operations of $\mathbf{A}$ to $\mathbf{A}[X]$, we obtain a sub*algebra* of $\mathbf{A}$, the **subalgebra generated by** X **in** $\mathbf{A}$, and the process of *generating* $\mathbf{A}[X]$ *from below* by closing X under all the fundamental operations as described in **??** applies to the setting of algebras as well, so

$$\mathbf{A}[X] = \bigcup_{n \in \mathbb{N}} G^n[X]$$

where G is the set of fundamental operations of $\mathbf{A}$, including the nullary ones. (For the notation, see Definition **??**.)

It follows easily from Proposition 9.3.4 that a subset S of an algebra[3] $\mathbf{A}$ is a subuniverse of $\mathbf{A}$ if and only if $S = \mathbf{A}[S]$.

Example 9.3.5 Consider the group $\langle \mathbb{Z}; +, -, 0 \rangle$, and let $X := \{2\}$. Since in this setting subalgebras are subgroups, $\mathbb{Z}[X]$ is the set of even integers, as we learned from Group Theory. Bottom-up construction yields $G^0 = \{2\}$, $G^1 = \{-2, 0, 2, 4\}$, $G^2 = \{-4, -2, 0, 2, 4, 6, 8\}$, ... It is a tedious but still instructing

[2] Remember that $\{S_k ; k \in K\}$ is a chain if, for $k_1, k_2 \in K$, either $S_{k_1} \subseteq S_{k_2}$ or $S_{k_2} \subseteq S_{k_1}$.

[3] "subset of an algebra" is a clear but not entirely correct abbreviation of "a subset S of the the universe A of an algebra $\mathbf{A}$".

exercise to write out the details for the calculation $G^3[X] = G[G^2[X]]$. (Cf. Definition **??** for details.)

Call a subuniverse S **finitely generated** if and only if $S = \mathbf{A}[X]$ for some *finite* subset X; especially, an algebra $\mathbf{A}$ is finitely generated iff its universe A is finitely generated (as a subuniverse).

The following Proposition is another consequence of 9.3.3.

Proposition 9.3.6 For any algebra $\mathbf{A}$,

$$\mathbf{A}[X] = \bigcup \{\mathbf{A}[Y] \, ; \, Y \subseteq X \text{ and } Y \text{ finite}\} .$$

Proof. Let $\mathcal{D} := \{\mathbf{A}[Y] \, ; \, Y \subseteq X \text{ and } Y \text{ finite}\}$.

If $Y \subseteq X$, then $\mathbf{A}[Y] \subseteq \mathbf{A}[X]$, so $\bigcup \mathcal{D} \subseteq \mathbf{A}[X]$.

On the other hand, since $\mathcal{D}$ is directed (cf. Proposition 9.3.33.), $\bigcup \mathcal{D}$ is a subuniverse of $\mathbf{A}$; moreover $X \subseteq \bigcup \mathcal{D}$, so $\mathbf{A}[X] \subseteq \mathcal{D}$. ■

In particular, the universe of any algebra is the union of its finitely generated subuniverses. A related notion is the following: $\mathbf{A}$ is called **locally finite** if and only if every finitely generated subalgebra of $\mathbf{A}$ is finite (note that $\mathbf{A}$ itself need not be finitely generated).

Example 9.3.7 Consider $\mathbf{A} = \langle \mathbb{Z}; +, -, 0 \rangle$ as the additive group with neutral element 0. Then, $\mathbf{A}$ is finitely generated since $\mathbf{A} = \mathbf{A}[\{1\}]$, but $\mathbf{A}$ is not locally finite for exactly this reason. On the other hand, if we let $\mathbf{B}$ be $\mathbb{Z}_2^{\mathbb{N}}$, the product of countably infinite many two–element groups with operations defined componentwise, we see that $\mathbf{B}$ is locally finite but not finitely generated. (The verification is left as an exercise.)

Clearly every finite algebra is locally finite. Also, if an infinite algebra is locally finite, it is *never* finitely generated. The proofs are left as an exercise.

We conclude this section by looking at *extremal* subalgebras. The largest subalgebra of any algebra $\mathbf{A}$ is, trivially, $\mathbf{A}$ itself. Consequently, our next step will be to look for coatoms (cf. Definition 3.1.9) in the lattice $\mathbf{Sub}\,\mathbf{A}$.

Definition 9.3.8 $\mathbf{B}$ is a **maximal subalgebra** of $\mathbf{A}$ iff $B \neq A$ and for any subuniverse S of $\mathbf{A}$, $B \subseteq S \subseteq A$ implies $S = B$ or $S = A$.

Maximal subalgebras need not exist in a given algebra $\mathbf{A}$. (Exercise: Prove that the additive group of rational numbers has no maximal subgroups.)

The following digression is intended to spice up the discussion with some nontrivial flavor. Given any algebra $\mathbf{A}$, let the **Frattini algebra** $\Phi(\mathbf{A})$ be the intersection of all maximal subalgebras of $\mathbf{A}$. (If $\mathbf{A}$ has no maximal subalgebras,

$\Phi(\mathbf{A})$ is reduced to the intersection of the empty family of subalgebras of $\mathbf{A}$, which is $\mathbf{A}$ itself.) An element $a \in A$ is called a **non–generator** if and only if it can be dropped from any set generating $\mathbf{A}$, more precisely, if and only if $\mathbf{A}[X]= A$ implies $\mathbf{A}[X \setminus \{a\}]= A$ for any $X \subseteq A$. These notions are connected by the following Proposition.

Proposition 9.3.9 For any algebra $\mathbf{A}$, the universe of $\Phi(\mathbf{A})$ coincides with the set of all non–generators of $\mathbf{A}$.

Proof. We show that $a \in A$ fails to be a non–generator exactly if a lies outside some maximal subalgebra $\mathbf{M}$ of $\mathbf{A}$. Suppose a is not a non–generator. Then, there exists $X \subseteq A$ such that $\mathbf{A}[X]\neq A$ but $\mathbf{A}[X \cup \{a\}]= A$. Let $\mathcal{S}$ be the collection of all subuniverses of $\mathbf{A}$ containing X but excluding a. $\mathcal{S} \neq \emptyset$ since $\mathbf{A}[X]\in \mathcal{S}$. Let C be a chain in $\mathcal{S}$ and let $S = \bigcup C$. Then S is a subuniverse by 9.3.3.3 and $a \notin S$, so $S \in \mathcal{S}$. By applying Zorn's Lemma (cf. Section **??** or Appendix B) we conclude that $\mathcal{S}$ contains a subuniverse M maximal with respect to $\subseteq$, and that the subalgebra $\mathbf{M}$ with universe M is a maximal subalgebra of $\mathbf{A}$. Indeed, if a subalgebra $\mathbf{B}$ properly includes M, then $X \subseteq B$ and $a \in B$, hence $\mathbf{B} = \mathbf{A}$. Since $\mathbf{M}$ is therefore a maximal subalgebra not containing a, a is not contained in the intersection $\Phi(\mathbf{A})$ of all maximal subalgebras of $\mathbf{A}$. This proves that any non–generator lies outside the Frattini algebra. For the other direction, assume $\mathbf{M}$ is a maximal subalgebra of $\mathbf{A}$ and $a \notin M$. Then, $\mathbf{A}[M \cup \{a\}]= A$ while $\mathbf{A}[M]= M \neq A$, thus a is not a non–generator. ∎

The smallest subalgebra of $\mathbf{A}$ is the subalgebra with universe $\mathbf{A}[\emptyset]$. (Proof: exercise.) As mentioned earlier in this section, if the type of $\mathbf{A}$ does not include any nullary operations, this is just the empty algebra of this type. If there are nullary operations, i.e. constants, we have

$$\mathbf{A}[\emptyset]=\mathbf{A}[\{c\,;\, c \text{ is a constant}\}]\ .$$

(Exercise: Prove this statement!) This is familiar from Ring Theory, where the notion of the *characteristic* of a (commutative, with 1) domain D is defined as the cardinality of the subring generated in D by the constants 0 and 1. In the same way as above for maximal subalgebras, we may define here **minimal subalgebras**, as opposed to smallest subalgebras, as the atoms (cf. Definition 3.1.9) in the lattice $\mathbf{Sub}\,\mathbf{A}$: Again, they need not exist in a given algebra $\mathbf{A}$. (Exercise: The additive group of rational numbers has no minimal subgroups, as has the additive group of the integers.)

9.4 Direct Products

In Section 7.2, we were first acquainted with products of families of $\mathcal{L}$–structures and saw that, from the viewpoint of First Order Logic , products are not an appropriate tool to build new models, since satisfaction of axioms is violated. Algebras on the other hand will prove to be much more resistant to the hostile effects of building products, since they lack relations. By restricting the classes of algebras under consideration to those axiomatized by *equations* we will later see that satisfaction of these axioms is preserved under products.

From 7.1.1 recall the definition of the direct product $\prod_{s\in S}\mathcal{A}_s$ of a family $\langle\mathcal{A}_s\ ;\ s\in S\rangle$. It is clear that in the case of algebras, this definition is simplified since we do not have to deal with relations. Note that if $S=\emptyset$, i.e. if we consider an empty family of algebras (or structures), their direct product is just $\{\emptyset\}$ – a fact we used already in Remark 9.1.3.3 in order to explain what nullary operations are.

Definition 9.4.1 Let $\mathbf{A}_k$ $(k\in K)$ be similar algebras. Then the **direct product** $\prod_{k\in K}\mathbf{A}_k$ is the algebra $\mathbf{A}$ of the same type with universe $\prod_{k\in K}A_k$ and fundamental operations $f_i^{\mathbf{A}}$ given by

$$f_i^{\mathbf{A}}(a_1,\ldots,a_{r_i}) := (\ldots, f_i^{\mathbf{A}_k}(a_1(k),\ldots,a_{r_i}(k)),\ldots)$$

for any maps $a_1,\ldots,a_{r_i}$ in $\prod_{k\in K}A_k$. If $\mathbf{A}_k\cong\mathbf{A}$ for some algebra $\mathbf{A}$ and all $k\in K$, we write $\mathbf{A}^K$ instead of $\prod_{k\in K}\mathbf{A}_k$ and call $\mathbf{A}^K$ a **direct power** of $\mathbf{A}$.

We leave it to the readers to convince themselves that their favourite example of a direct product construction from classical algebra indeed falls under the scope of Definition 7.1.1. Note the case $K=\emptyset$: Here the direct product degenerates into the (unique) one-element algebra of the type considered. Constructions based on the direct product will play a major role in the following.

It is straightforward (i.e. an easy exercise) to check that the projections π_k associated with the notion of direct products are surjective homomorphisms from $\prod_{k\in K}\mathbf{A}_k$ onto $\mathbf{A}_k$ for any collection. Almost as simple is the following observation.

Lemma 9.4.2 If $\mathbf{B}$ and $\mathbf{A}_k$ $(k\in K)$ are similar algebras, and $g_k:\mathbf{B}\longrightarrow\mathbf{A}_k$ is a surjective homorphism for each $k\in K$, then there is a uniquely determined surjective homomorphism $g:\mathbf{B}\longrightarrow\prod_{k\in K}\mathbf{A}_k$ satisfying $g_k=\pi_k\circ g$ for all $k\in K$.

Proof. Indeed, g defined by $g(b):=\langle g_k(b)\ ;\ k\in K\rangle$ has the required properties. (Exercise: Develop the details.) ∎

More interesting is the fact that direct products of algebras can actually be characterized by the conclusion of Lemma 9.4.2: Assume $\mathbf{C}$ and $\mathbf{A}_k$ are similar algebras and $h_k : \mathbf{C} \longrightarrow A_k$ is a surjective homorphisms for every $k \in K$. Then we call the pair $\langle \mathbf{C}, \langle h_k \,;\, k \in K\rangle\rangle$ the **categorical product** of the family $\langle \mathbf{A}_k \,;\, k \in K\rangle$ if and only if, for any algebra $\mathbf{B}$ of the same type and any family $\langle g_k \,;\, k \in K\rangle$ of surjective homomorphisms $g_k : \mathbf{B} \longrightarrow \mathbf{A}_k$, there is a surjective homomorphism $g : \mathbf{B} \longrightarrow \mathbf{C}$ satifying $g_k = h_k \circ g$ for all $k \in K$. Thus, by the above remarks, the direct product is a categorical product. Interestingly, even the converse holds up to isomorphism, as the following Proposition shows.

Proposition 9.4.3 For any categorical product $\langle \mathbf{C}, \langle h_k \,;\, k \in K\rangle\rangle$ of a family $\langle \mathbf{A}_k \,;\, k \in K\rangle$ of similar algebras, $\mathbf{C} \cong \prod_{k\in K} \mathbf{A}_k$.

Proof. Suppose $\langle \mathbf{C}, \langle h_k \,;\, k \in K\rangle\rangle$ is a categorical product of the family $\langle \mathbf{A}_k \,;\, k \in K\rangle$. By 9.4.2 we get a surjective homomorphism $g : \prod_{k\in K} \mathbf{A}_k \longrightarrow \mathbf{C}$ with $h_k \circ g = \pi_k$. On the other hand, we have seen above by explicit construction that there is a onto homomorphism $h : \mathbf{C} \longrightarrow \prod_{k\in K} \mathbf{A}_k$ such that $\pi_k \circ h = h_k$ for all k. Hence $\pi_k \circ h \circ g = \pi_k$, which shows that $h \circ g : \prod_{k\in K} \mathbf{A}_k \longrightarrow \prod_{k\in K} \mathbf{A}_k$ must be the identity map. Thus we have shown that h and g are mutually inverse isomorphisms and we are done. ∎

In Category Theory, the definition of categorical products indeed *defines* products in a general sense, since categorically isomorphism stands for equality.

We conclude this section with the following fact which neatly connects all the basic concepts we have studied so far.

Proposition 9.4.4 Let $\mathbf{A}$ and $\mathbf{B}$ be similar algebras. Then $h : A \longrightarrow B$ is a homomorphism from $\mathbf{A}$ into $\mathbf{B}$ iff $\{\langle a, h(a)\rangle \,;\, a \in A\}$ is a subuniverse of $\mathbf{A} \times \mathbf{B}$.

Proof. This is left as an exercise we do not want to withhold. ∎

Chapter 10

Congruences

Up to now, the constructions and notions we presented were mere translations from Model Theoretic concepts to Universal Algebra. We obtained them by generalizing well-known constructions, thus producing something new from something old in a straightforward manner. Examples for this process are Homomorphic images (Def. 9.2.1 together with 9.3.3.1), subalgebras (Def. 9.3.1), and direct products (Def. 9.4.1). Conspicuously missing is another construction, which — although intimately connected with the formation of homomorphic images — emphasizes the functional character of Universal Algebras, namely the process of dividing the universe of an algebra into nonempty pieces and giving the collection of sets thus obtained the structure of an algebra of the same type (cf. Section 1.4). In Classical Algebra, say Group Theory, this process amounts to partitions of a group G into co–sets of some normal subgroup $N \subseteq G$. However, at this point, we need the more general concept of a *congruence relation*.

10.1 Congruences and quotient algebras

Suppose we are given an algebra $\mathbf{A}$ and an equivalence relation ϑ on the universe A of $\mathbf{A}$. We want to turn the quotient set A/ϑ into an algebra of the same type as $\mathbf{A}$ in such a way that the canonical map π_ϑ becomes a homomorphism. This requirement completely determines the fundamental operations $f_i^{\mathbf{A}/\vartheta}$. Indeed, we must have (by Definition 9.2.1)

$$f_i^{\mathbf{A}/\vartheta}(\pi_\vartheta(a_1), \ldots, \pi_\vartheta(a_{r_i})) = \pi_\vartheta(f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i})),$$

or in the notation established in Section 1.4,

$$f_i^{\mathbf{A}/\vartheta}([a_1], \ldots, [a_{r_i}]) = [f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i})],$$

for all fundamental operations f_i and all $a_1, \ldots, a_{r_i} \in A$. Now, if $a_j, a'_j \in A$ with $a_j \vartheta a'_j$ for $j \in \{1, \ldots, r_i\}$, then $[a_j] = [a'_j]$ and thus we *should* have $f_i^{\mathbf{A}/\vartheta}([a_1], \ldots, [a_{r_i}]) = f_i^{\mathbf{A}/\vartheta}([a'_1], \ldots, [a'_{r_i}])$. However, there is no reason why $f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i})$ and $f_i^{\mathbf{A}}(a'_1, \ldots, a'_{r_i})$ should be in the same ϑ-class, thus $f_i^{\mathbf{A}/\vartheta}$ as specified is not necessarily well–defined. This points out the restriction we have to impose on ϑ.

Definition 10.1.1 Let $\mathbf{A}$ be an algebra and ϑ an equivalence relation on the universe A of $\mathbf{A}$.

1. ϑ is **compatible** with the fundamental operation $f_i^{\mathbf{A}}$ iff

   $$a_j \vartheta a'_j \text{ for } j \in \{1, \ldots, r_i\} \text{ implies } f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i}) \vartheta f_i^{\mathbf{A}}(a'_1, \ldots, a'_{r_i})$$

   for any choice of $a_i, a'_j \in A$.

2. ϑ is a **congruence (relation)** on $\mathbf{A}$ iff ϑ is compatible with all fundamental operations of $\mathbf{A}$.

3. Given a congruence ϑ on $\mathbf{A}$, the quotient set A/ϑ equipped with operations $f_i^{\mathbf{A}/\vartheta}$ defined by

   $$f_i\mathbf{A}/\vartheta([a_1], \ldots, [a_{r_i}]) = [f_i(a_1, \ldots, a_{r_i})]$$

   is called the **quotient algebra** of $\mathbf{A}$ relative to ϑ. $\mathbf{Con}\,\mathbf{A}$ denotes the set of all congruence relations on $\mathbf{A}$.

Note that any equivalence relation is compatible with all nullary operations (cf. Section 9.1) defined on its carrier set. The discussion preceding Definition 10.1.1 shows that (i) the canonical map π_ϑ can be turned into a homomorphism if and only if ϑ is a congruence on $\mathbf{A}$, and (ii) the algebraic structure imposed on A by this requirement is uniquely determined.

To include but a few simple examples, we note that Δ_A and ∇_A (cf. Example 1.4.2) are congruences for any algebra $\mathbf{A}$, so they are clearly the smallest and the largest congruence on $\mathbf{A}$, respectively. As a less trivial example, define a relation ϑ on $\mathbb{Q}$ by $a\vartheta b$ if and only if $a - b \in \mathbb{Z}$. It is easy to check that ϑ is an equivalence and that ϑ is compatible with $+$ and $-$ but not with $\cdot$, thus ϑ is a congruence on the additive group $\langle \mathbb{Q}; +, -, 0\rangle$ but not on the ring $\langle \mathbb{Q}; +, -, \cdot, 0, 1\rangle$.

Lemma 10.1.2 If ϑ is a congruence on $\mathbf{A}$, then $\pi_\vartheta : \mathbf{A} \longrightarrow \mathbf{A}/\vartheta$ is a surjective homomorphism.

Proof. Exercise! ■

For equivalences we have seen that there was a mutual correspondence between the kernels of maps and the equivalence relations. Transposed to the context of congruences, this reads as in the following proposition.

Proposition 10.1.3 For any algebra $\mathbf{A}$, the congruences on $\mathbf{A}$ are precisely the kernels of the homomorphisms with source $\mathbf{A}$.

Proof. Another exercise. ■

Two different homomorphisms with source $\mathbf{A}$ may have the same kernel even if both of them are surjective. For example, all automorphisms of $\mathbf{A}$ (i.e. isomorphisms $\eta : \mathbf{A} \longrightarrow \mathbf{A}$) have the same kernel $\Delta_A = \{\langle a, a\rangle \,;\, a \in A\}$. However, the targets of surjective homomorphisms with the same kernel are isomorphic.

Proposition 10.1.4 Assume $g_1 : \mathbf{A} \longrightarrow \mathbf{B}_1$ and $g_2 : \mathbf{A} \longrightarrow \mathbf{B}_2$ are surjective homomorphisms and $\ker g_1 = \ker g_2$. Then $\mathbf{B}_1 \cong \mathbf{B}_2$.

Proof. Define $\eta : \mathbf{B}_1 \longrightarrow \mathbf{B}_2$ by $\eta(g_1(a)) := g_2(a)$. Then, η is well-defined, injective and surjective, since

$$g_1(a) = g_1(a') \text{ iff } g_2(a) = g_2(a'),$$

for any $a, a' \in A$, and g_1, g_2 are both surjective.

Now, for any $g_1(a_1), \ldots, g_1(a_{r_i}) \in \mathbf{B}_1$ we have

$$\begin{aligned}
\eta(f_i^{\mathbf{B}_1}(g_1(a_1), \ldots, g_1(a_{r_i}))) &= \eta(g_1(f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i}))) \\
&\quad \text{(since } g_1 \text{ is a homomorphism)} \\
&= g_2(f_i^{\mathbf{A}}(a_1, \ldots, a_{r_i})) \\
&\quad \text{(by definition of } \eta\text{)} \\
&= f_i^{\mathbf{B}_2}(g_2(a_1), \ldots, g_2(a_{r_i})) \\
&\quad \text{(since } g_2 \text{ is a homomorphism)} \\
&= f_i^{\mathbf{B}_2}(\eta(g_1(a_1)), \ldots, \eta(g_1(a_{r_i}))) \\
&\quad \text{(by definition of } \eta\text{).}
\end{aligned}$$

This shows that η is a homomorphism. ■

The most important consequence of Proposition 10.1.4 is covered by the following theorem.

Theorem 10.1.5 (The Homomorphism Theorem) Let $\eta : \mathbf{A} \longrightarrow \mathbf{B}$ be a surjective homomorphism from an algebra $\mathbf{A}$ to an algebra $\mathbf{B}$. Then, $\mathbf{B}$ and $\mathbf{A} / \ker \eta$ are isomorphic. In other words, homomorphic images and quotient algebras of $\mathbf{A}$ are identical (up to isomorphism).

Proof. Let $\kappa = \ker \eta$, then η and the projection $\pi_\kappa : \mathbf{A} \longrightarrow \mathbf{A} / \kappa$ are both surjective and have the same kernel. Consequently, our claim follows by Proposition 10.1.4. ■

Theorem 10.1.5 may be viewed as the *target end dual* of Proposition 10.1.3. If η in 10.1.5 is not surjective, the conclusion reads $\eta[\mathbf{A}] \cong \mathbf{A} / \ker \eta$. (Exercise: Modify the proof of Theorem 10.1.5 to obtain a proof of this statement.)

Given two congruences ϑ and ρ on some algebra $\mathbf{A}$, each of the congruences carries over to the quotient of $\mathbf{A}$ under the other in a natural way.

Definition 10.1.6 If $\mathbf{A}$ is an algebra and $\vartheta, \rho \in \mathbf{Con}\,\mathbf{A}$, then ϑ/ρ is the binary relation on A defined by

$$\vartheta/\rho := \{\langle [a]_\rho, [b]_\rho \rangle \in (\mathbf{A} / \rho)^2 \; ; \; a\vartheta b\} \; .$$

ϑ/ρ is a binary relation on A, and some simple calculations show that $\vartheta/\rho \in \mathbf{Con}\,\mathbf{A}$. (Proof: Exercise.)

Example 10.1.7 Consider the group $\mathbb{Z} = \langle \mathbb{Z}; +, 0\rangle$ as an algebra of type $(2, 0)$. For

$$\vartheta := \{\langle m, n\rangle \in \mathbb{Z}^2 \; ; \; m = n \mod 3\}$$

and

$$\rho := \{\langle m, n\rangle \in \mathbb{Z}^2 \; ; \; m = n \mod 4\},$$

we get $\mathbb{Z} / \rho = \mathbb{Z}_4$.

$[0]_\rho = [4]_\rho = [8]_\rho$, so since $4\vartheta 1$, we get $[0]_\rho \vartheta [1]_\rho$; since $8\vartheta 2$, we get $[0]_\rho \vartheta [2]_\rho$; and since $0\vartheta 3$, we get $[0]_\rho \vartheta [3]_\rho$. Thus we conclude that $\mathbf{A} / \vartheta/\rho$ is the trivial group with one element, $[0]_\rho$.

In general, there is no nice overall behaviour of ϑ/ρ expressible in terms ϑ and ρ. However, there are exceptions, as can be seen in the following Theorem.

Theorem 10.1.8 If ϑ and ρ are congruences on the algebra $\mathbf{A}$ satisfying $\rho \subseteq \vartheta$, then

$$(\mathbf{A} / \rho)/(\vartheta/\rho) \cong \mathbf{A} / \vartheta.$$

Proof. The isomorphism $\eta : (\mathbf{A} / \rho)/(\vartheta/\rho) \longrightarrow \mathbf{A} / \vartheta$ is given by

$$\eta([[a]_\rho]_{\eta/\rho}) := [a]_\vartheta.$$

The student is kindly invited to verify that η is well–defined and an isomorphism. ∎

Going back to Example 10.1.7, we see that for ϑ and ρ as defined we have $\rho \not\subseteq \vartheta$. As an alternative, consider $\rho := \{\langle m, n\rangle \in \mathbb{Z}^2 \,;\, m = n \mod 6\}$. Then $\rho \subseteq \vartheta$ and indeed $(\mathbb{Z}/\rho)/(\vartheta/\rho) \cong \mathbb{Z}/\vartheta = \mathbb{Z}_\mathbf{2}$.

The restriction of congruences to subalgebras is worth a few thoughts as well.

Definition 10.1.9 If $\mathbf{A}$ is an algebra, $\vartheta \in \mathbf{Con\,A}$ and $B \subseteq A$, then we define $B^\vartheta \subseteq A$ by

$$B^\vartheta := \{a \in A \,;\, a\vartheta b \text{ for some } b \in B\} .$$

Exercise 10.1.10 Which of the following statements is true?

1. For any algebra $\mathbf{A}$, any $\vartheta \in \mathbf{Con\,A}$ and any subset $B \subseteq A$, B^ϑ is a subuniverse of $\mathbf{A}$.

2. For any algebra $\mathbf{A}$ and any $\vartheta \in \mathbf{Con\,A}$, the assignment $B \mapsto B^\vartheta$ defines a closure operator on A.

We will need a special case of the first statement to formulate the next main result.

Lemma 10.1.11 Let $\mathbf{A}$ be an algebra and $\vartheta \in \mathbf{Con\,A}$. Then, for any *subuniverse* $B \subseteq A$, B^ϑ is a subuniverse[1] of $\mathbf{A}$.

Proof. If f is a fundamental operation of $\mathbf{A}$ and $a_1, \ldots, a_r \in B^\vartheta$, then there are $b_1, \ldots, b_r \in B$ with $a_1\vartheta b_1, \ldots, a_r\vartheta b_r$. Since B is a subuniverse, we have

$$f(b_1, \ldots, b_r) \in B,$$

and ϑ being a congruence, we have

$$f(a_1, \ldots, a_r)\vartheta f(b_1, \ldots, b_r).$$

We conclude that $f(a_1, \ldots, a_r) \in B^\vartheta$ by definition of B^ϑ. ∎

From the point of view of quotients, we obtain nothing new when we switch from ϑ to $\vartheta \cap B^2$ and from B to B^ϑ.

Theorem 10.1.12 If $\mathbf{A}$ is an algebra, $\vartheta \in \mathbf{Con\,A}$ and $\mathbf{B} \in \mathbf{Sub\,A}$, then

$$\mathbf{B}/(\vartheta \cap B^2) \cong B^\vartheta/(\vartheta \cap {B^\vartheta}^2).$$

[1] For the sake of completeness we must mention that we also write B^ϑ for the subalgebra of $\mathbf{A}$ whose subuniverse is B^ϑ.

Proof. Define $\eta : \mathbf{B}/(\vartheta \cap B^2) \longrightarrow B^\vartheta/(\vartheta \cap {B^\vartheta}^2)$ by

$$\eta([b]_{\vartheta\cap B^2}) := [b]_{\vartheta\cap {B^\vartheta}^2}.$$

To check that η is well–defined and an isomorphism is left as an exercise. ■

Since there usually is an abundance of congruences and equivalences on any given algebra **A**, the need arises to impose some structure onto the sets **Eq A** and **Con A**. The direct approach is to treat the relations as sets of ordered pairs and compare them using $\subseteq$. Thus, if ϑ and ρ are congruences on some algebra **A**, we say that ϑ is **finer**[2] than ρ if and only if $\vartheta \subseteq \rho$. If ϑ is finer than ρ, then ρ is said to be **coarser** than ϑ.

The sets **Con A** of congruences and **Eq A** of equivalences on an algebra A display interesting features when considered as ordered by $\subseteq$. It is easy to see that the intersection of a family of congruences (as sets!) is again a congruence, and the same holds for equivalences. On the other hand, the union of a family of congruences need not even be an equivalence. As we saw in 3.2.3, arbitrary intersections (i.e. infima of arbitrary families) give rise to arbitrary suprema, but in this case (as for **Sub A**), suprema are not identical with set theoretical unions.

Proposition 10.1.13 For any algebra **A**, the sets **Eq A** and **Con A**, ordered by $\subseteq$, are both complete lattices.

Proof. See 11.1.5, or even better, try it yourself. ■

For more details on the complete lattice of congruences, see 10.2 below.

We conclude this section with an example intended to show that our definitions indeed generalize notions well-known in Algebra.

Example 10.1.14 Let $\mathbf{G} = \langle G; \cdot, ^{-1}, e\rangle$ be any group, and $N \subseteq G$ a normal subgroup[3]. We write ab instead of $a \cdot b$ and aN for $\{an \,;\, n \in N\}$ and the like to keep notation familiar. Then we define a binary relation ϑ_N on G by $a\vartheta_N b$ if and only if $aN = bN$ for $a, b \in G$. ϑ_N is clearly an equivalence. (Exercise!) Suppose $a\vartheta_N a'$ and $b\vartheta_N b'$. Then, $abN = aNb = a'Nb = a'bN = a'b'N$ and $a^{-1}N = (N^{-1}a)^{-1} = (Na)^{-1} = (aN)^{-1} = (a'N)^{-1} = \ldots = a'^{-1}N$, so ϑ_N is compatible with $\cdot$ and $^{-1}$. Thus ϑ_N is a congruence. Now $aN = bN$ if and only if $ab^{-1} \in N$, hence $a\vartheta_N b$ if and only if $a \in Nb = bN$. Since $b = be \in N$, this shows that the ϑ_N-class of any $b \in G$ is just bN. In other words, $\mathbf{G}/\vartheta_N \cong \mathbf{G}/N$ canonical.

[2] Despite this being mere nomenclature, it still exhibits the intended use of congruences to "build blocks". A congruence thus is the finer the smaller its block are, i.e. the fewer elements are related by it.

[3] Recall that a subgroup N of a group G is a **normal subgroup** if and only if $gHg^{-1} = H$ for all $g \in G$, which is equivalent to $gH = Hg$ for all $\in G$

Starting with an arbitrary congruence ϑ on $\mathbf{G}$, it is easy to verify that $[e]_\vartheta =: N$ is in fact a normal subgroup of $\mathbf{G}$ and that $\vartheta = \vartheta_N$. Thus, congruences on $\mathbf{G}$ and normal subgroups of $\mathbf{G}$ correspond bijectively. Accordingly, group theorists work with normal subgroups rather than with congruences. Note, however, that this correspondence hinges on – among other facts – the presence of a neutral element e in $\mathbf{G}$. Consequently, we cannot expect to replace congruences by the consideration of special kinds of subalgebras in our general setting.

10.2 The lattice of congruences

As we have seen in Proposition 10.1.13, the set $\mathbf{Con}\,\mathbf{A}$ of congruences on an algebra $\mathbf{A}$ is a (complete) lattice when ordered by $\subseteq$. However, we have also seen that the suprema in this lattice are more complicated to describe than the infima. In this section, we want to spice up the results with some more details.

Definition 10.2.1 If $R_1, R_2 \subseteq A^2$ are binary relations on some set A, then the **relational product** $R_1 \circ R_2$ is defined by

$$R_1 \circ R_2 := \{\langle a, b\rangle \in A^2 \text{ ; for some } c, \langle a, c\rangle \in R_1 \text{ and } \langle c, b\rangle \in R_2\} .$$

Note that the product $f \circ g$ of two maps is *not* a special case of the relational product as defined above, since $\circ$ is not read in the same direction. This is, of course, inconsistent notation, but it serves the convenience of the reader, because it mirrors the natural direction of reading in both cases.

The relational product in general might display some completely unreasonable behaviors, in the sense that the product of two rather large relations may even be empty. The situation changes to the better when we focus on congruences and equivalences.

Exercise 10.2.2

1. Is the relational product commutative? Is it associative?
2. What if we restrict ourselves to equivalences on some set A?
3. Given two equivalences ϑ, ρ on some set A, is the relational product $\vartheta \circ \rho$ always an equivalence?
4. Show that for $\vartheta \in \mathbf{Eq}\ \mathbf{A}$, $\vartheta \circ \vartheta = \vartheta$.

In order to iterate the relational product over three and more factors, we have to agree on a reading in which we may omit the brackets. Thus, we *define* for binary relations $R_1, \ldots, R_n$ the product $R_1 \circ \ldots \circ R_n$ by

$$R_1 \circ \ldots \circ R_n := R_1 \circ (R_2 \circ (\ldots (R_{n-1} \circ R_n) \ldots)).$$

Another notion not totally foreign, but still unexpected in connection with relations, is the *inverse* of a relation.

Definition 10.2.3 If $R \subseteq A^2$ is a binary relation on a set A, then R^{-1} denotes the **(relational) inverse** of R and is defined by

$$R^{-1} := \{ \langle b, a \rangle \in A^2 \, ; \, \langle a, b \rangle \in R \} \, .$$

The notion *inverse* demands for a reference to an operation, which in this case is, of course, the relational product $\circ$. Unfortunately, things are not as straightforward as with inverses in, say, groups. Clearly, the inverse of a relation on A is always defined (other than with functions), and it is again a relation on A. However, we do not have a nice *cancellation property* as in the case of groups where the product of an element and its inverse results in the neutral element of the binary product. On any set A, Δ_A is the unit element with respect to $\circ$ (why?), but we do not have $R \circ R^{-1} = \Delta_A$ in general (why not?). The next exercise states the most direct consequences.

Exercise 10.2.4 Show that for $\vartheta_1, \vartheta_2 \subseteq A^2$, we have

(i) $(\vartheta_1 \circ \vartheta_2)^{-1} = \vartheta_2^{-1} \circ \vartheta_1^{-1}$;

(ii) $\vartheta_1 \subseteq \vartheta_2$ iff $\vartheta_1^{-1} \subseteq \vartheta_2^{-1}$.

For equivalences, fortunately, matters are much simpler.

Exercise 10.2.5 Show that for $\vartheta \in \mathbf{Eq\ A}$, we have $\vartheta^{-1} = \vartheta$.

The title of the current section promised that we will deal with the lattice of congruences, therefore we should have a closer look at the inner structure of this lattice. We already know from 10.1.13 that $\mathbf{Con\,A}$ is a complete lattice, and we know how to compute the infimum of a given set of congruences on $\mathbf{A}$. We also know that the supremum is generally *not* the same as the set union. Of course, in some cases we might be lucky as the following proposition shows.

Proposition 10.2.6 If Θ is a **directed** set of congruences, i.e. if $\vartheta_1, \vartheta_2 \in \Theta$ implies $\vartheta_1, \vartheta_2 \subseteq \rho$ for some $\rho \in \Theta$, then $\operatorname{Sup} \Theta = \bigcup \Theta$.

Proof. It suffices to show that $\bigcup \Theta$ is a congruence. Since all the $\vartheta \in \Theta$ are reflexive, $\bigcup \Theta$ is also reflexive. The same argument applies to symmetry. To show transitivity, assume $\langle a,b\rangle,\langle b,c\rangle\in\bigcup\Theta$. Then there are $\vartheta_1,\vartheta_2\in\Theta$ with $\langle a,b\rangle\in\vartheta_1$ and $\langle b,c\rangle\in\vartheta_2$. By directedness, we find $\rho\in\Theta$ with $\vartheta_1\subseteq\rho\supseteq\vartheta_2$, hence $\langle a,b\rangle,\langle b,c\rangle\in\rho$, and since ρ is a congruence and therefore transitive, we find $\langle a,c\rangle\in\rho\subseteq\bigcup\Theta$.

For compatibility, assume that $\langle a_1,b_1\rangle,\ldots,\langle a_n,b_n\rangle\in\bigcup\Theta$ and that f is an n–ary fundamental operation of the algebra carrying the congruences in Θ. We have to show that

$$\langle f(a_1,\ldots,a_n),f(b_1,\ldots,b_n)\rangle\in\bigcup\Theta.$$

From $\langle a_1,b_1\rangle,\ldots,\langle a_n,b_n\rangle\in\bigcup\Theta$ we conclude that there are $\vartheta_1,\ldots,\vartheta_n\in\Theta$ with $\langle a_i,b_i\rangle\in\vartheta_i$. Applying directedness of Θ $n-1$ times, we find $\rho\in\Theta$ such that $\vartheta_1,\ldots,\vartheta_n\subseteq\rho$, hence $\langle a_1,b_1\rangle,\ldots,\langle a_n,b_n\rangle\in\rho$. Finally, since ρ is a congruence,

$$\langle f(a_1,\ldots,a_n),f(b_1,\ldots,b_n)\rangle\in\rho\subseteq\bigcup\Theta.$$

■

Exercise 10.2.7

1. Does the converse also hold? That is, does $\operatorname{Sup}\Theta = \bigcup \Theta$ imply that Θ is directed?

2. Does Proposition 10.2.6 hold for equivalences as well?

Luckily there are other characterizations of the supremum of a set of congruences, which work in a more general setting.

Proposition 10.2.8 If Θ is a set of congruences on an algebra $\mathbf{A}$, then

$$\operatorname{Sup}\Theta = \bigcup \{\vartheta_0 \circ \ldots \circ \vartheta_n \ ; \ n \in \mathbb{N}, \vartheta_0, \ldots, \vartheta_n \in \Theta\} \ .$$

Proof. We shall only outline the procedure and leave the details to the reader:

Let $\Xi := \{\vartheta_0 \circ \ldots \circ \vartheta_n \ ; \ n \in \mathbb{N}, \vartheta_0, \ldots, \vartheta_n \in \Theta\}$. Then the following two statements hold.

1. $\bigcup \Xi$ is a congruence, and $\vartheta \subseteq \bigcup \Xi$, for any $\vartheta \in \Theta$. Thus, $\operatorname{Sup}\Theta \subseteq \bigcup \Xi$.

2. $\vartheta_0 \circ \ldots \circ \vartheta_n \subseteq \operatorname{Sup}\Theta$, for all $n \in \mathbb{N}$ and all $\vartheta_0, \ldots, \vartheta_n \in \Theta$. Thus, $\bigcup \Xi \subseteq \operatorname{Sup}\Theta$.

■

In other words, the supremum of congruences is the union of the (finite) iterated product of these congruences.

Note that the supremum of a set of congruences is expressed by finite products: $\langle a, b\rangle \in \operatorname{Sup}\Theta$ if and only if, for some $n \in \mathbb{N}$, some $\vartheta_0, \ldots, \vartheta_n \in \Theta$ and some $a_0, \ldots, a_{n+1} \in A$,

$$a = a_0 \vartheta_0 a_1 \vartheta_1 a_2 \vartheta_2 \ldots a_n \vartheta_n a_{n+1} = b.$$

Exercise 10.2.9 Does Proposition 10.2.8 remain valid if "congruence" is replaced with "equivalence"?

Using the characterization given in Proposition 10.2.8, we find the following result.

Proposition 10.2.10 For $\vartheta_1, \vartheta_2 \in \mathbf{Con\,A}$, the following are equivalent:

(i) $\vartheta_1 \circ \vartheta_2 = \vartheta_2 \circ \vartheta_1$;

(ii) $\operatorname{Sup}\{\vartheta_1, \vartheta_2\} = \vartheta_1 \circ \vartheta_2$;

(iii) $\vartheta_1 \circ \vartheta_2 \subseteq \vartheta_2 \circ \vartheta_1$.

Proof. Exercise. ■

If one (and thus any) of the clauses of Proposition 10.2.10 holds, we say that ϑ_1 and ϑ_2 are **permutable**. An algebra **A** such that any two congruences in **Con A** are permutable is called **congruence–permutable**, and consequently a class $\mathbb{K}$ of algebras is called **congruence–permutable** if any algebra in $\mathbb{K}$ is congruence–permutable.

Finally, we would like to show that the congruence–lattice of a quotient is isomorphic to a sublattice of the original congruence–lattice. We first need the notion of an *interval* in a lattice.

Definition 10.2.11 If $\langle L, \leq\rangle$ is a lattice, then a subset $I \subseteq L$ is called a **(closed) interval** in **L** if, for some $a_1, a_2 \in L$, $I = \{b \in L\,;\, a_1 \leq b \leq a_2\}$. In this case we write $I = [a_1, a_2]$.

Clearly, the closed intervals in $\mathbb{R}$ are intervals in the sense of the above definition.

The following Lemma is worth some attention.

Lemma 10.2.12 For any lattice $\mathbf{L} = \langle L, \leq\rangle$ and any $a, a', b, b' \in L$, the following two statements hold.

1. $[a, b] = [a', b']$ iff $a = a'$ and $b = b'$.

2. $[a, b]$ is a sublattice[4] of **L**.

Proof. Exercise. ■

Theorem 10.2.13 (Correspondence Theorem) If **A** is an algebra and $\vartheta \in$ **Con A**, then the lattices **Con A** $/\vartheta$ and $[\vartheta, \nabla_A]$ are isomorphic.

Proof. The isomorphism η is given by $\eta(\rho) := \rho/\vartheta$. The details of this proof are left as an exercise. ■

10.3 Generating congruences

When considering the group $\langle \mathbb{Z}; +, 0\rangle$, our experiences in Algebra tell us that congruences on $\mathbb{Z}$ correspond to calculations *modulo* some n. More to the point, for any congruence ϑ on $\mathbb{Z}$ there is a number n such that, for any $k, l \in \mathbb{Z}$, $k\vartheta l$ if and only if $k = l \mod n$. If we divide $\mathbb{Z}$ by this congruence, we find the (quotient–)group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.

[4] With this we clandestinely switched from the order–theoretic aspect of **L** to the algebraic one, since here *sublattice* stands for *subalgebra*.

There is another way of looking at this: Starting from $\mathbb{Z}$, we wonder what would be the consequences if we considered n to be the same as 0. Clearly we would then conclude that $n+1=1$, $n+2=2,\ldots,2n=0,2n+1=1,\ldots$ and so on. Finally we would end up with the very structure we called *quotient group* a little earlier.

As we can see in the following Definition, there is even an algebraic way of formulating this little game of "I wonder what would happen if n were 0".

Definition 10.3.1 If $\mathbf{A}$ is an algebra and $R \subseteq A^2$, then

$$\theta(R) := \bigcap \{\vartheta \in \mathbf{Con\,A}\ ;\ R \subseteq \vartheta\}$$

is called the **congruence generated by R in $\mathbf{A}$**.

If $R = \{\langle a_1, b_1\rangle, \ldots, \langle a_n, b_n\rangle\}$ is finite, then we simply write

$$\theta(\langle a_1, b_1\rangle, \ldots, \langle a_n, b_n\rangle)$$

for $\theta(R)$. Moreover, if $S \subseteq A$, then we write $\theta(S)$ for $\theta(S^2)$.

As seen in Proposition 10.1.13, $\mathbf{Con\,A}$ is a complete lattice with meet corresponding to set–intersection. Therefore, congruences generated by some R are indeed *congruences*.

Corollary 10.3.2 *For any algebra $\mathbf{A}$ and any $R \subseteq A^2$, $\theta(R)$ is a congruence on $\mathbf{A}$. In fact, it is the least congruence ϑ on $\mathbf{A}$ such that for any $\langle a, b\rangle \in R$, $a\vartheta b$.*

Proposition 10.3.3 *If $\mathbf{A}$ is an algebra and $\Theta \subseteq \mathbf{Con\,A}$, then*

$$\mathrm{Sup}\,\Theta = \theta(\bigcup \Theta)\,.$$

Proof. Exercise. ∎

In other words, the supremum of congruences is the congruence generated by the set–union of the respective congruences.

Complete lattices stand in direct correspondence to closure systems (cf. Definition 3.2.4), and thus the generation of congruences defines a closure operator (see Definition 3.2.5). The proof of this is left as an exercise.

We started this section by examining congruences in the group $\mathbb{Z}$. This may be somewhat misleading since almost every congruence on $\mathbb{Z}$ is of the form $\theta(\langle a, b\rangle)$. (The proof of this claim and to find its exceptions is left as an exercise.) Congruences generated by a single pair of elements of the carrier (cf. Definition 1.4.1) even have their own names.

Definition 10.3.4 A **principal congruence** on an algebra $\mathbf{A}$ is a congruence $\theta(\langle a, b, \rangle)$ generated by two elements.

As mentioned at the end of the previous section, our aim is to develop another *compactness–result* (cf. Compactness Theorem 2.4.5) in the context of congruences.

Proposition 10.3.5 Let $\mathbf{A}$ be an algebra and $\vartheta \in \mathbf{Con\,A}$. Then the following statement holds.

$$\begin{aligned}\vartheta = \text{Sup } \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} &= \bigcup \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \\ &= \bigcup \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\}\ .\end{aligned}$$

Proof. We proceed in five steps which will yield the desired equations.

(1) $\vartheta \subseteq \bigcup \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\}$,
since $\langle a, b\rangle \in \theta(\langle a, b\rangle)$;

(2) $\bigcup \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \subseteq \text{Sup } \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\}$,
since $\bigcup \Theta \subseteq \text{Sup}\,\Theta$ for all sets Θ of congruences on some algebra;

(3) $\text{Sup } \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \subseteq \text{Sup } \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\}$,
since $\theta(\langle a, b\rangle) = \theta(\{a, b\})$ and therefore $\{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \subseteq \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\}$;

(4) $\text{Sup } \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\} = \bigcup \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\}$,
since $\{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\}$ is directed (cf. Proposition 10.2.6);

(5) $\bigcup \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\} \subseteq \vartheta$,
since $\theta(R) \subseteq \vartheta$ for all $R \subseteq \vartheta$.

Putting everything together, we find

$$\begin{aligned}\vartheta &\subseteq \bigcup \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \\ &\subseteq \text{Sup } \{\theta(\langle a, b\rangle)\ ;\ a\vartheta b\} \\ &\subseteq \text{Sup } \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\} \\ &= \bigcup \{\theta(R)\ ;\ R \subseteq \vartheta, R \text{ finite}\} \\ &\subseteq \vartheta.\end{aligned}$$

■

Thus, a congruence is the supremum of its principal sub–congruences.

Exercise 10.3.6 Which of the equations in Proposition 10.3.5 hold if *congruence* is replaced by *equivalence* (and consequently *generation of congruences* by *generation of equivalences*)?

We like to put this result in context with closure systems. The last term in the equation above states that the closure operator assigning to each relation the smallest closure operator containing it is what we call an **algebraic closure operator**.

Definition 10.3.7 A closure operator C on a set X is called **algebraic** if for any $S \subseteq X$, $C(S) = \bigcup \{C(Z)\,;\, Z \subseteq S, Z \text{ finite}\}$. A lattice is called an **algebraic lattice** if it is order–isomorphic (i.e. isomorphic as ordered sets) to the closure system associated to an algebraic closure operator.

As a direct consequence of being order–isomorphic to a closure system, every algebraic lattice is clearly complete. An alternative characterization of algebraic complete lattices uses the notion of *compact* elements of a complete lattice.

Definition 10.3.8 An element a of a complete lattice $\mathbf{L} = \langle L, \leq\rangle$ is called **compact** if, for any subset $S \subseteq A$, the following statement holds.

$$\text{Whenever } a \leq \operatorname{Sup} S, \text{ then for some } \textit{finite } S_0 \subseteq S, a \leq \operatorname{Sup} S_0\,.$$

Here we find Compactness again (cf. Proposition 10.3.5 and Theorem 2.4.5). The setting is closely related to the notion of compact (subsets of) topological spaces, where — as you might know if you have ever dealt with set–theoretical topology — compactness stands for the property that every covering by open sets has a finite sub–covering. Examples of compact closed subspaces of the real line $\mathbb{R}$ as a topological space are the closed intervals $[\alpha, \beta]$.

Compact elements a in complete lattices can always be reached in a finite number of steps in the following sense. If for some *chain* $C \subseteq A$ we have $\operatorname{Sup} C \geq a$, then $c \geq a$ for some $c \in C$. (Exercise: prove this!)

As an exercise, so as to become used to the notion of compact elements, you might like to show the following statements.

- Every complete lattice has at least one compact element.
- Finite lattices consist of compact elements exclusively.
- The compact elements in the complete lattice $\langle \mathcal{P}(X), \subseteq\rangle$ are exactly the finite $Y \subseteq X$.

Trivially, every element in a complete lattice is an upper bound of the compact elements that lie below it:

$$a \geq \operatorname{Sup}\ \{c \leq a\,;\, c \text{ compact}\}\,.$$

In an algebraic lattice, a is even the *least* upper bound.

Proposition 10.3.9 The following statement is true for any complete lattice $\mathbf{L} = \langle L, \leq \rangle$.

$$\mathbf{L} \text{ is algebraic iff for all } a \in A,\ a = \text{Sup } \{c \leq a\ ;\ c \text{ compact}\}\,.$$

As suggested by the nomenclature, examples are primarily found in the field of algebra.

Proposition 10.3.10 **Con A**, **Eq A** and **Sub A** are algebraic lattices for any algebra **A**.

Proof. For **Con A**, the claim follows directly from Proposition 10.3.5.

For **Eq A**, we first prove the analogue of Proposition 10.3.5 for equivalences (which should have been done in Exercise 10.3.6). The rest is simple.

For **Sub A**, Proposition 9.3.6 is exactly what we need. ∎

Examples of non–algebraic complete lattices, on the other hand, seem to be less natural than their algebraic cousin.

Exercise 10.3.11 Find an example of a non–algebraic complete lattice.

The relationship between algebraic lattices and algebras is a very close one, as can be seen in the next result. Its complete proof exceeds the scope of this lecture.)

Theorem 10.3.12 (Birkhoff and Frink)
Let **L** be a complete lattice.

$$\mathbf{L} \text{ is algebraic iff } \mathbf{L} \text{ is order–isomorphic to } \mathbf{Sub\,A} \text{ for some algebra } \mathbf{A}\,.$$

Proof. Since we know that **Sub A** is algebraic for any algebra **A**, and we know that *being an algebraic complete lattice* is preserved under order–isomorphisms (proof: exercise), the statement that every lattice order–isomorphic to the lattice of sub–algebras of some algebra **A** is algebraic, is quite obvious.

The tough part, for which we refer to the literature, is to show the converse. Here we have to construct an algebra whose lattice of sub–algebras is order–isomorphic to **L**. ∎

Chapter 11

Orders as Algebras

As mentioned in Chapter 3, semilattices and lattices give rise to an *algebraic* way of looking at order–relations by treating Sup and Inf as binary operations. The aim of this chapter is to study fundamental properties of ordered sets which allow for this kind of algebraization of their internal structure, thereby providing examples for universal algebras.

11.1 (Semi–)Lattices as Algebras

Definition 11.1.1 1. A binary operation $\cdot$ on some set S is called a **semilattice–operation** iff it is

- **associative**, i.e. $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ for all $x, y, z \in S$,
- **commutative**, i.e. $x \cdot y = y \cdot x$ for all $x, y \in S$ and
- **idempotent**, i.e. $x \cdot x = x$ for all $x \in S$.

2. A **semilattice** is an algebra $\mathbf{S} = \langle S; \cdot \rangle$ of type 2 such that $\cdot$ is a semilattice–operation.

While 3.1.12 provided the relational aspect of semilattices, the above definition reflects their algebraic side. As we shall see, these two aspects are freely interchangeable.

Proposition 11.1.2 Every Sup–semilattice (every Inf–semilattice) $\langle S, \leq \rangle$ is a semilattice. Indeed, Sup (Inf) restricted to 2-element subsets of S and considered as a binary operation on S is obviously a semilattice operation.
Conversely, every semilattice operation $\cdot$ on a set S may be used to define two orders $\leq_s$ and $\leq_i$ on S by setting for any $x, y \in S$ $x \leq_s y$ if and only if $y = x \cdot y$ and $x \leq_i y$ if and only if $x = x \cdot y$, respectively. Then $\langle S, \leq_s \rangle$ is a Sup–semilattice

and $\langle S, \leq_i \rangle$ an Inf –semilattice. The orders $\leq_s$ and $\leq_i$ are dual to one another. Moreover, these transformations are mutually inverse, that is, the binary operation Sup derived from $\leq_s$ coincides with $\cdot$ for any semilattice $\langle S; \cdot \rangle$, and $\leq_s$ derived from binary Sup in any Sup –semilattice $\langle S, \leq \rangle$ coincides with $\leq$ (and analogously for Inf –semilattices and $\leq_i$).

It is customary to write $\sqcup$ for the semilattice operation corresponding to binary Sup and to call the associated algebra $\mathbf{S} = \langle S; \sqcup \rangle$ a **join–semilattice**; similarly, if the operation corresponds to binary Inf , it is written as $\sqcap$ and the algebra $\mathbf{S} = \langle S; \sqcap \rangle$ is called a **meet–semilattice**. The choice whether a given semilattice should be regarded as a join–semilattice or a meet–semilattice amounts to specifying which one of two orders $\leq_i$, $\leq_s$ we wish to impose on the semilattice's carrier. However, picking $\sqcap$ to denote the semilattice operation always means that the associated order is $\leq_i$ respective to which $\sqcap$ is just binary Sup (and similarly for $\sqcup$ and $\leq_s$).

Since we showed that the two ways of looking at semilattices (order vs. algebra) are interchangeable in the above sense, it is natural to wonder whether this works for lattices as well. Thus, first we need the *algebraic* notion of a lattice.

Definition 11.1.3 An algebra $\mathbf{L} = \langle L; \sqcup, \sqcap \rangle$ of type $(2, 2)$ is a **lattice** iff both $\langle L; \sqcup \rangle$ and $\langle L; \sqcap \rangle$ are semilattices and the following **absorption identities** hold for any $x, y \in L$:

$$x \sqcup (x \sqcap y) \quad = \quad x \qquad (*)$$

and

$$x \sqcap (x \sqcup y) \quad = \quad x \qquad (**)$$

Proposition 11.1.4 If $\langle L, \leq \rangle$ is a lattice (as a poset), then $\langle L; \mathrm{Sup}_{\leq}, \mathrm{Inf}_{\leq} \rangle$ is a lattice (as an algebra).
Conversely, if $\langle L; \sqcup, \sqcap \rangle$ is a lattice (as an algebra), then the two order–relations $\leq_\sqcup$ and $\leq_\sqcap$ as defined in Proposition 11.1.2 are identical and $\langle L, \leq_\sqcup \rangle$ is a lattice (as a poset). Moreover, the transformation is mutually connected by

$$\langle L, \leq_{\sqcup_{\leq}} \rangle = \langle L, \leq \rangle$$

and

$$\langle L; \sqcup_{\leq_\sqcup}, \sqcap_{\leq_\sqcup} \rangle = \langle L; \sqcup, \sqcap \rangle \,.$$

Proof. Exercise. ■

Example 11.1.5

1. Recall from Definition 1.4.1 that **Eq X** denotes the set of all equivalence–relations on the set X. Consider the order $(\mathbf{Eq\ X}, \leq)$ (X any set) given by $\alpha \leq \beta$ iff $\alpha \subseteq \beta$ for any $\alpha, \beta \in \mathbf{Eq\ X}$ (cf. Proposition 10.1.13). We have $\alpha \leq \beta$ iff $\langle x_1, x_2\rangle \in \alpha$ implies $\langle x_1, x_2\rangle \in \beta$ iff every α-block $[x]_\alpha$ is contained as a subset in a (uniquely determined) β-block $[x]_\beta$ (verify!). We will show that Inf and Sup of $\{\alpha, \beta\}$ always exist, this means that $(\mathbf{Eq\ X}, \leq)$ is a lattice under the the corresponding operations; however, these operations are not so closely tied to the set operations $\cup$ or $\cap$.

   The proof of the existence of Inf is easy: If $\gamma \in \mathbf{Eq\ X}$, $\gamma \leq \alpha$ and $\gamma \leq \beta$, then every γ-block C is contained in some α-block A and in some β-block B, thus $C \subseteq A \cap B$. The collection of all sets $A \cap B$ with A ranging over α-blocks and B ranging over β-blocks is a partition of X (verify!) and so determines an equivalence $\mu \in \mathbf{Eq\ X}$. Clearly $\mu \leq \alpha$, $\mu \leq \beta$ and $\gamma \leq \mu$, so μ indeed is the Inf of α and β in $(\mathbf{Eq\ X}, \leq)$.

   To show the existence of Sup, suppose $\alpha \leq \gamma$, $\beta \leq \gamma$ and consider an α-block A and a β-block B such that $A \cap B \neq \emptyset$. The unique γ-block C containing A then has a nonempty intersection with B and must thus coincide with the unique γ-block C' containing B. Iterating the argument, consider a sequence of blocks $H_1 \ldots, H_m$ from either α or β such that $H_i \cap H_{i+1} \neq \emptyset$ for $1 \leq i < n$: If $x_1 \in H_1$, $x_m \in H_m$ and $x_1 \in C$ for some γ-block C, then also $x_m \in C$. Now define a binary relation κ on X by $x_1 \kappa x_m$ iff a sequence $H_1, \ldots, H_m$ exists as above with $x_1 \in H_1$ and $x_m \in H_m$. It is straightforward to check that κ is an equivalence relation (do so!); moreover, every κ-block is contained in a unique γ-block by construction. It follows that $\kappa \leq \gamma$. On the other hand, $\alpha \leq \kappa$ and $\beta \leq \kappa$ are immediate (consider sequences consisting of just one block from either α or β), so κ is indeed the Sup of α and β in $(\mathrm{Eq}, \leq)$.

2. Continue the preceding example and assume that α and β are actually congruences on some algebra **A**. We leave it as an exercise to the reader to show that Inf and Sup of α and β computed as equivalences as above are indeed congruences again. It follows that $\langle \mathbf{Con\,A}; \sqcap, \sqcup\rangle$ is a lattice with operations $\alpha \sqcap \beta = \mathrm{Inf}\,\{\alpha, \beta\}$ and $\alpha \sqcup \beta = \mathrm{Sup}\,\{\alpha, \beta\}$.

3. Consider the algebra $\mathbf{N} = \langle \mathbb{N}; \sqcap, \sqcup\rangle$ with $m \sqcap n =$ g.c.d of m and n, $m \sqcup n = \max\{m, n\}$. Both operations are semilattice operations, and $(**)$ of Definition 11.1.3 is satisfied since g.c.d.$(m, n) \leq m, n$. However, $(*)$ fails as e.g. $2 \sqcap (2 \sqcup 3) = 2 \sqcap 3 = 1$. Thus **N** is not a lattice. This example also shows that the two absorption identities (*) and (**) are independent. Clearly, by taking l.c.m. and min as operations, we get another example

of two semilattice operations satisfying (∗∗) but failing (∗) of Definition 11.1.3.

As has become apparent in the preceding examples, it often depends on the context whether the order relation or its two operations constitute the natural way to discuss a specific lattice. Accordingly, we will abuse notation sometimes and speak of the lattice $\mathbf{L} = (L, \leq)$ whenever $\leq$ is the order on L jointly induced by the operations $\sqcap$ and $\sqcup$ of the lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$.

If $\langle P_1, \leq_1 \rangle$ and $\langle P_2, \leq_2 \rangle$ are posets, a function $\eta : P_1 \longrightarrow P_2$ is said to **preserve the order** if

$$p \leq_1 q \text{ implies } \eta(p) \leq_2 \eta(q) \text{ for all } p, q \in P_1.$$

Funtions which preserve the order are also called **order–homomorphisms** (from P_1 to P_2). Note that this definition agrees with the definition of a $\mathcal{L}$–homomorphism (cf. Definition 4.4.1) for the formal language $\mathcal{L}$ having a binary relaion–symbol as its only non–logical symbol.

Exercise 11.1.6 Let $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ be a lattice and let $\leq$ be the order stemming from $\sqcap$ and $\sqcup$ in the sense of Proposition 11.1.4. Show that $x \leq x'$ and $y \leq y'$ imply $x \sqcap y \leq x' \sqcap y'$ and $x \sqcup y \leq x' \sqcup y'$, i.e.

$\sqcap : L^2 \longrightarrow L$ and $\sqcup : L^2 \longrightarrow L$ are order–preserving.

11.2 Distributive and Modular Lattices

There are some distinctive properties of lattices which are captured by *laws* resembling the *group laws* or *ring laws* familiar from classical algebra (see Example 9.1.4 3). Such laws are actually formulae of the underlying language of the class of algebras under consideration. To consider the collection of all algebras of the same type is most often pointless, since the algebraic properties and internal structure common to all algebras of a given type are far too general to be of interest. By imposing laws in the above sense, we limit the class of algebras we want to deal with, hopefully ending up with common properties worth studying.

Consider a lattice of sets $\mathbf{L} = \langle L; \cap, \cup \rangle$ as in Example 3.1.15 2. It is a beginner's exercise in Set Theory to show that for any three sets U, V and W the equations $U \cap (V \cup W) = (U \cap W) \cup (V \cap W)$ and $U \cup (V \cap W) = (U \cup W) \cap (V \cup W)$ always are satisfied. Abstracting to an arbitrary lattice $\mathbf{L}$, we adopt the following definition.

Definition 11.2.1 A lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is called **distributive** iff it satisfies the two equations

$$x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z) \qquad (D_\sqcap)$$

and

$$x \sqcup (y \sqcap z) = (x \sqcup y) \sqcap (x \sqcup z) \qquad (D_\sqcup)$$

for all $x, y, z \in L$.

It follows that every lattice of sets $\langle L; \cap, \cup \rangle$ is distributive. Most interestigly, the converse is also true; however, the proof of this highly nontrivial fact lies beyond the scope of this lecture.

There are many equivalent statements characterizing distributive lattices. We only mention here that $D_\sqcap$ and $D_\sqcup$ imply each other in the sense that whenever either one is satisfied in a lattice for *all* x, y, z, then so is the other. (Exercise: Prove this claim.)

Exercise 11.2.2 Find out which of the following finite lattices are distributive: $\mathbf{2}$, $\mathbf{N_5}$, $\mathbf{M_3}$, $\mathbf{B}_3$ (cf. Example 3.1.16).

Consider the collection $\mathcal{S}$ of all subspaces of a vector space $\mathbf{V}$. Since the intersection of any family of subspaces is a subspace again, $\mathcal{S}$ is a closure system on the carrier V of $\mathbf{V}$. Consequently, $\langle S, \subseteq \rangle$ is a complete lattice. In general, it is simple to find subspaces A, B and C of V such that $A \cap (B \cup C) \neq (A \cap B) \cup (A \cap C)$ (consider three pairwise different 1-dimensional subspaces!), so we conclude that such a lattice of subspaces is not distributive in general. However, the equation for distributivity is true when restricted to the case $A \supseteq C$. This is the motivating example for the next definition.

Definition 11.2.3 A lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is called **modular** iff it satisfies the implication

$$x \geq z \rightarrow x \sqcap (y \sqcup z) = (x \sqcap y) \sqcup (x \sqcap z) \qquad (M)$$

for all $x, y, z \in L$.

So $\mathbf{L}$ is modular if and only if it satisfies ($\mathrm{D}_\sqcap$) in special cases; hence, every distributive lattice is modular. The converse is not true, as examples in the following Exercise show.

Exercise 11.2.4 Find out which of the following finite lattices are modular: $\mathbf{2}$, $\mathbf{N_5}$, $\mathbf{M_3}$, $\mathbf{B}_3$ (cf. Example 3.1.16 and Exercise 11.2.2).

Again, there are many different statements characterizing modular lattices. Here, we only note the following.

Exercise 11.2.5 Show that a lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is modular iff

$$((x \sqcap z) \sqcup y) \sqcap z = (x \sqcap z) \sqcup (y \sqcap z)$$

for all $x, y, z \in L$.

11.3 Complemented Lattices

Given any set X, $\langle \mathcal{P}(X), \subseteq \rangle$ is a distributive lattice $\mathcal{P}(\mathbf{X}) = \langle \mathcal{P}(X); \cap, \cup \rangle$, as we have seen above, with $\mathbf{0}_{\mathcal{P}(\mathbf{X})} = \emptyset$ and $\mathbf{1}_{\mathcal{P}(\mathbf{X})} = X$. Moreover, for each $A \in \mathcal{P}(X)$ there exists $B \in \mathcal{P}(X)$ such that $A \cup B = X$ and $A \cap B = \emptyset$. We abstract this situation in the following definition.

Definition 11.3.1 A lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is **bounded** iff it contains a least element $\mathbf{0}_{\mathbf{L}}$ and a greatest element $\mathbf{1}_{\mathbf{L}}$. If $\mathbf{L}$ is bounded and $x \in L$, an element $y \in L$ is a **complement** of x iff $y \sqcup x = \mathbf{1}$ and $y \sqcap x = \mathbf{0}$. A bounded lattice $\mathbf{L}$ is **complented** iff *every* $x \in L$ has a complement.

Do not be misled to think that — as is the case for $\mathcal{P}(X)$ — complements need to be unique. For example, N_5 is complemented, but there is an element which has more than one complement (which one?). The reason for this ambiguity lies in the non–distributivity of N_5.

Lemma 11.3.2 In a distributive lattice any element has at most one complement.

Proof. Exercise. ■

However, distributivity in itself does not guarantee the existence of complements.

Exercise 11.3.3 Show that in a chain, only the top and bottom elements have complements, so chains considered as lattices are not complemented if they have more than two elements.

Lattices that are both distributive and complemented present an important special case. For historical reasons, they are called **Boolean lattices**. So all lattices $\mathcal{P}(\mathbf{X})$ as considered above are Boolean. The converse, however, is not true, not even in a weaker form, i.e. not all Boolean lattices are isomorphic[1]

[1] A **order–isomorphism** is an order–preserving, bijective function with order–preserving inverse function. Two lattices L and L' are **isomorphic** if there is an order–isomorphism $\eta : L \longrightarrow L'$. This is easily shown to be equivalent to *being isomorphic as algebras*.

Figure 11.1: George Boole (1815-1864)

to a lattice $\mathcal{P}(\mathbf{X})$ for some set X. Consider, for example, the lattice with all finite and cofinite subsets of some *infinite* set X as its carrier set and with the operations of set union and intersection. (Exercise: Verify that this defines indeed a Boolean lattice.)

In a Boolean lattice $\mathbf{L}$ the assignment of a (unique) complement to each element may be regarded as a unary operation on L. Writing $'$ for this operation, and adding the bounds $\mathbf{0}$ and $\mathbf{1}$ as constants to the algebra, we turn $\mathbf{L}$ into an algebra $\mathbf{B}$ of type $(2, 2, 1, 0, 0)$ by setting $\mathbf{B} = \langle L; \sqcap, \sqcup, ', \mathbf{0}, \mathbf{1}\rangle$. Such algebras are called **Boolean algebras**; they were first considered in 1854 by George Boole in his investigations in propositional calculus. They are, together with the permutation groups of roots of polynomials as studied by Galois, among the first *abstract* algebras in our sense. Again, the power set lattices above provide standard examples. To keep notation clean, we write $\mathbf{B}(\mathbf{X})$ for the Boolean algebra $\mathbf{B}(\mathbf{X}) = \langle \mathcal{P}(X); \cap, \cup, ^c, \emptyset, X\rangle$ where ^{c} stands for set complementation with respect to X, and just $\mathbf{B}_n$ if X has n elements. Again, not every Boolean algebra is of this sort as the finite-cofinite example above shows. Finally, BA denotes the class of all Boolean algebras.

Now consider a topological space $\mathcal{X} = \langle X, \mathcal{O}\rangle$, that is, a set X together with a collection $\mathcal{O}$ of subsets of X which is closed under *finite* set intersections and arbitrary set unions and furthermore contains $\emptyset$ and X ($\mathcal{O}$ is the collection of **open** sets of the space $\mathcal{X}$). $\mathcal{O}$ is easily seen to be a bounded distributive lattice under the operations $\cap$ and $\cup$. As the standard examples like the real line $\mathbb{R}$ or the real plane $\mathbb{R}^2$ show, an *open* set $U \in \mathcal{O}$ has very rarely a complement within $\mathcal{O}$, that is, an open set $V \in \mathcal{O}$ such that $U \cup V = X$ and $U \cap V = \emptyset$. (As an exercise, look for open sets with an open complement in $\mathbb{R}$ or $\mathbb{R}^2$.)

In a *distributive* bounded lattice $\langle \mathbf{L}, \leq\rangle$, the complement x' of some $x \in L$ has the property that it is the largest $z \in L$ such that $z \sqcap x = 0$. Indeed, if $z \sqcap x = \mathbf{0}$, then

$$z = z \sqcap \mathbf{1} = z \sqcap (x \sqcup x') = (z \sqcap x) \sqcup (z \sqcap x') = \mathbf{0} \sqcup (z \sqcap x') = z \sqcap x'$$

and thus $z \leq x'$ in the canonical order of **L**. Returning to the lattices $\mathcal{O}$ of open sets, it is not difficult to see that there is always a *largest open* set which is disjoint from some given open set U: The set of all *interior* points of $X \setminus U$. We abstract the weaker property, not confining ourselves to the distributive case.

Definition 11.3.4 Let $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ be a lattice with bottom element $\mathbf{0_L}$, and $x, x^* \in L$. Then x^* is called a **pseudocomplement** of x iff

(i) $x \sqcap x^* = 0$ and

(ii) $x \sqcap z = \mathbf{0}$ implies $z \leq x^*$ for all $z \in L$.

L is called **pseudocomplemented** iff every $x \in L$ has a pseudocomplemented.

The definition of pseudocomplements requires only the existence of a least element $\mathbf{0}$ in **L**. However, if $a, b \in L$ have pseudocomplements a^*, b^* in L, respectively, and $a \leq b$, then $b^* \leq a^*$ (Exercise: Prove this!); so a pseudocomplemented lattice will always have a greatest element, namely $\mathbf{0}^*$.

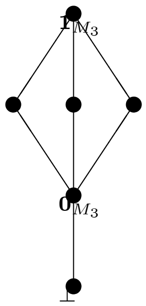
Every complemented lattice is pseudocomplemented but not vice versa (as the lattices $\mathcal{O}$ of open sets show); also every chain, considered as a lattice, is pseudocomplemented but not complemented whenever it has at least three elements (Exercise: Why?).

As for complements in lattices, the assignment of the (unique) pseudocomplement to each element may be regarded as a unary operation on L. Writing * for this operation, we make **L** an algebra **A** of type $(2, 2, 1, 0, 0)$ by setting $\mathbf{A} = \langle L; \sqcap, \sqcup, ^*, \mathbf{0}, \mathbf{1} \rangle$ where $\mathbf{0}$ and $\mathbf{1}$ are constants denoting the bounds of **L**. We will call such algebras **lattices with pseudocomplementation**. It is no loss of generality to include $\mathbf{1}$ among the constants since is definable by means of * and $\mathbf{0}$.

The most interesting case occurs if the lattice with pseudocomplementation is *distributive* as a lattice. Such algebras are commonly called **p–algebras** and the corresponding class is denoted by PALG. The open set lattices $\mathcal{O}$ are typical examples of p–algebras. p–algebras arise in logic much in the same way as Boolean algebras. While the latter serve as algebraic models of *classical* propositional calculus, p-algebras model (a fragment of) *intuitionistic* propositional calculus; see [**?**] for a thorough study of this connection. A more recent application of p–algebras within Computer Science is to use them as models for so-called *rough sets*, see e.g. [**?**] and the references given there. Of course, every Boolean algebra is a p–algebra. p–algebras will serve as a prime source of examples to illustrate many of the universal algebraic concepts addressed throughout this lecture.

Nondistributive lattices which are pseudocomplemented but not complemented exist in abundance. Indeed, pick your favorite bounded nondistributive lattice **L** and add a new bottom element $\bot$; the resulting ordered set is a lattice where exactly $\mathbf{1}_L$ and $\bot$ have complements, and where $\bot$ serves as a pseudocomplement of everything except itself.

Exercise 11.3.5 Check the claim from the previous paragraph for $\mathbf{M}_3$.



$\mathbf{M3}_\bot$

As a consequence, the class of algebras arising from such lattices by adding pseudocomplementation as a fundamental operation is too diverse to admit a meaningful structure theory (with a few exceptions).

Note that in Definition 11.3.4 only the meet operation $\sqcap$ is used, so pseudocomplements may meaningfully be defined in any meet semilattice with a zero. As we will see, such algebras have many interesting features, especially when compared to p–algebras.

Definition 11.3.6 A meet–semilattice $\mathbf{S} = \langle S; \sqcap \rangle$ with least element $\mathbf{0_S}$ is **pseudocomplemented** iff every $x \in S$ has a pseudocomplement in the sense of Def. **??**.

We will exclusively consider pseudocomplemented semilattices in the type $(2, 1, 0, 0)$, that is, with pseudocomplementation as a fundamental operation as well as $\mathbf{0}$ and $\mathbf{1} = \mathbf{0}^*$. We call such algebras **p–semilattices** and write PCS for the their class. Here is a p–semilattice which is not a lattice:

Example 11.3.7 Let $\mathcal{U}$ be the collection of open subsets of $\mathbb{R}$ which are contained in the open interval $(-2, 2)$ but do not contain $(-1, 1)$, together with all intervals of the form $(-2-1/n, 2+1/n)$, and $\emptyset$ and $\mathbb{R}$. Ordered by set inclusion, $\mathcal{U}$ is a meet semilattice with $\cap$ as Inf but not a lattice since, e.g., $(-1.5, 0)$ and $(0, 1.5)$ have no Sup within $\mathcal{U}$ (why?).

The pseudocomplement of any nonempty set in $\mathcal{U}$ is $\emptyset$, which itself has $\mathbb{R}$ as its pseudocomplement.

Our discussion of complements would not be complete without mentioning the numerous variants of complementation properties which are considered in the lattice theoretic literature. There we might encounter dual pseudocomplements (defined in terms of join and greatest element), relative complements and pseudocomplements (with respect to an interval $[a, b]$ within a lattice), or operations which obey some but not all of the identities valid for complementation in a distributive lattice. These variants lead us to algebras known as Heyting algebras, Post algebras and De Morgan algebras, just to mention a few. [**?**] is a good reference.

11.4 Order and First–Order Logic

Taking a closer look at the definition of complete lattices, we realize that it is formulated using notions that are not expressible in first–order logic: *Every* subset must have an infimum and a supremum. We therefore might argue that the definition uses second–order concepts and is thus not restateable in first–order languages. This argument is clearly short–sighted in that it draws a general conclusion from our failure to find an appropriate formalization. From the results of Chapter 8, we know what it means for a class of structures not to be axiomatizable in first–order logic. Thus in order to show that the class of complete lattices cannot be captured using first–order concepts exclusively, we use the semantic way and show that this class is not closed under ultraproducts.

As a formal language we use $\mathcal{L}_{\leq}$, the formal language having but the binary relation–symbol $\leq$ as a non–logical symbol. An $\mathcal{L}_{\leq}$–structure is thus basically a set with a binary relation defined on it. However, since we want to consider lattices in general and complete lattices specifically, we note the following proposition.

Proposition 11.4.1 The class of partially ordered sets is basic–elementary, as is the class of lattices.

Proof. To affirm a class of being (basic–)elementary is easier than to refute it, since for the former we only need to present a (finite) set of axioms which does the job. As an exercise, find an appropriate set of $\mathcal{L}_{\leq}$–sentences Σ_L such that $\operatorname{Mod}\Sigma_L$ is exactly the class of lattices. ■

The ultraproduct we are going to construct next may look funny as a complete lattice, yet it is suitable for our purpose.

Let $\hat{\mathbb{N}}$ be the set $\mathbb{N}$ enriched by a new element we will denote by $\top$. Using the (canonical) order–relation $\leq$ on $\mathbb{N}$, we define the binary relation $\leq_{\hat{\mathbb{N}}}$ on $\hat{\mathbb{N}}$

by setting $\top$ to be the (new) greatest element[2], i.e.

$$x \leq_{\hat{\mathbb{N}}} y \text{ iff } [x, y \in \mathbb{N} \text{ and } x \leq y] \text{ or } y = \top$$

for all $x, y \in \hat{\mathbb{N}}$. Since $\leq_{\hat{\mathbb{N}}}$ is an extension of $\leq$, we risk no confusion in simply writing $\leq$ for both relations.

Exercise 11.4.2 Show that $\leq$ as defined above is an order on $\hat{\mathbb{N}}$ and that $\langle \hat{\mathbb{N}}, \leq \rangle$ is a complete lattice.

Let $\mathcal{U}$ be an ultrafilter over $\mathbb{N}$ containing all cofinite subsets of $\mathbb{N}$, and let $\mathcal{A} := \hat{\mathbb{N}}^{\mathbb{N}}/\mathcal{U}$. The proofs of the following observations are left as exercises for the reader.

- $\mathcal{A}$ is a lattice with greatest element $\top_{\mathcal{A}} := \langle \top, \top, \top, \ldots \rangle_{\mathcal{U}}$ and least element $0_{\mathcal{A}} := \langle 0, 0, 0, \ldots \rangle_{\mathcal{U}}$.
- All elements of $\mathcal{A}$ except $\top_{\mathcal{A}}$ and $0_{\mathcal{A}}$ have an upper and a lower cover (cf. Section 3.1).
- $\langle 0, 1, 2, \ldots \rangle_{\mathcal{U}}$ is an upper bound of $\overline{\mathbb{N}} := \{\overline{n}\,;\, n \in \mathbb{N}\}$ in $\mathcal{A}$, where $\overline{n} := \langle n, n, n, \ldots \rangle_{\mathcal{U}}$.
- If $b \in |\mathcal{A}|$ is an upper bound of $\overline{\mathbb{N}}$, then so is the lower cover of b. (Hint: If such a lower cover c were not an upper bound of $\overline{\mathbb{N}}$, then for some $\overline{n} \in \overline{\mathbb{N}}$, $\overline{n} \leq c \leq \overline{n+1}$, from which we conclude $c = \overline{n}$ or $c = \overline{n+1}$. But then b, being the upper cover of c, could not be an upper bound of $\overline{\mathbb{N}}$.)
- $\overline{\mathbb{N}}$ does not have a supremum in $\mathcal{A}$ (and the set of upper bounds of $\overline{\mathbb{N}}$ has no infimum).

So we find a subset of $\mathcal{A}$ having no supremum, thus $\mathcal{A}$ is not a complete lattice. Since $\mathcal{A}$ is an ultraproduct of complete lattices, we have a proof for the following proposition.

Proposition 11.4.3 The class of complete lattices is not an elementary class.

11.5 Order vs. Algebra

In some cases, ordered structures allow for algebraization of their order–relation. Although the relational and the algebraic aspect are mutually dual, there are

[2] From Set Theory you probably remember the set $\omega + 1$, the ordinal successor of the first infinite ordinal ω. Actually, our set $\hat{\mathbb{N}}$ is exactly of ordinal type $\omega + 1$, or in other words, the ordered sets $\langle \hat{\mathbb{N}}, \leq \rangle$ and $\langle \omega + 1, \in \rangle$ are isomorphic.

differences if we consider classes of ordered sets as a whole, especially if we include structure–preserving maps in our considerations.

From Section 11.1, recall the notions of order–preserving maps, order–homomorphisms and order–isomorphism. If $\langle L_1, \leq_1\rangle$ and $\langle L_2, \leq_2\rangle$ are lattices, then a map $\eta : L_1 \longrightarrow L_2$ is a **lattice–homomorphism** if η preserves Sup and Inf , i.e. if

$$\eta(\operatorname{Sup} x, y) = \operatorname{Sup} \eta(x), \eta(y)$$

and

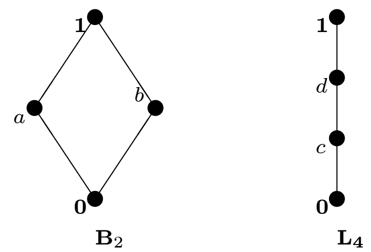
$$\eta(\operatorname{Inf} x, y) = \operatorname{Inf} \eta(x), \eta(y)$$

for all $x, y \in L_1$, where the Sup 's and Inf 's are taken in the appropriate lattice L_1 or L_2. Of course, if we consider just one of the two identities above, we get a definition of a **semilattice–homomorphism**.

A **lattice–isomorphism** is a bijective lattice–homomorphism whose inverse map is again a lattice–homomorphism.

Please note that the definition of lattice–homomorphisms and –isomorphisms agree with the definition of homomorphisms and isomorphisms for algebras of type $(2, 2)$, of which lattices are concrete instances, ans similarly for semilattice–homomorphisms.

Now consider the lattices $\mathbf{B}_2$, the four–element boolean algebra, and $\mathbf{L_4}$, the four–element chain:



It is easily seen that there is an order–homomorphism from $\mathbf{B}_2$ to $\mathbf{L_4}$, e.g. given by $\mathbf{0} \mapsto \mathbf{0}$, $\mathbf{1} \mapsto \mathbf{1}$, $a \mapsto c$ and $b \mapsto d$. However, it also rather obvious that this map is not a lattice–homomorphism. The reason for this discrepancy lies in the asymmetry of the constraints on mappings to preserve relations on one hand and to preserve functions on the other hand. Relational homomorphisms such as order–homomorphisms are only preserving in one direction, e.g. *if* $x \leq y$, *then* $\eta(x) \leq \eta(y)$. There is nothing said about the case where x and y are incomparable. $\eta(x)$ may be incomparable to $\eta(y)$, but not necessarily. On the other hand, when operations such as Sup are considered, *every* pair of elements x, y has to satisfy $\operatorname{Sup} \eta(x), \eta(y) = \eta(\operatorname{Sup} x, y)$ for η to preserve Sup , so the constraint is not a conditional one, but a generally formulated identity.

For those among you who feel uncomfortable with the above elaborations since they emphasize the weak point of relational homomorphisms in the sense that, although a and b are not in relation, their images c and d are comparable, we add another example which circumnavigates this point.

Example 11.5.1 Taking $\mathbf{L_4}$ from above, we define $\mathbf{L_4}_{\perp\top}$ to be the lattice resulted from adding a new least element $\perp$ and a new greatest element $\top$ to $\mathbf{L_4}$. Then, the map from $\mathbf{L_4}$ to $\mathbf{L_4}_{\perp\top}$ given by $\mathbf{0} \mapsto \perp$, $\mathbf{1} \mapsto \top$, $a \mapsto a$ and $b \mapsto b$ is an order–homomorphism but not a lattice–homomorphism. (Proof: Exercise.)

So we notice that

Proposition 11.5.2 If $\mathbf{A}$ and $\mathbf{B}$ are lattices, then every lattice–homomorphism $\eta : \mathbf{A} \longrightarrow \mathbf{B}$ is a order–homomorphism. The converse is not true in general.

Proof. For the first part, we note that in a lattice, $a \leq b$ is equivalent to $\text{Inf}\, a, b = a$. The rest is simple.

The second part was done by mentioning the above examples. ■

As we can see, *lattice–homomorphisms contain more structural information than order–homomorphisms.*

11.6 Distributivity via Sublattices

Distributivity and modularity of lattices is easily formulated using equations (or identities). It is clear that an arbitrary lattice can be tested for distributivity by checking if the laws of distributivity hold for any triple of elements of the lattice. However, there is a better, more universal way for this.

If $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is a lattice, then a **sub–lattice** of $\mathbf{L}$ is a subalgebra of $\mathbf{L}$ (viewed as an algebra of type $(2, 2)$).

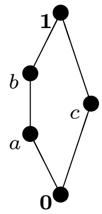
Lemma 11.6.1 For any lattice $\mathbf{L}$, the following two statements hold:

1. If $\mathbf{L}$ is modular, then so is any sub–lattice of $\mathbf{L}$.

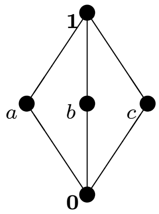
2. If $\mathbf{L}$ is distributive, then so is any sub–lattice of $\mathbf{L}$.

Proof. Trivial. ■

If you succeeded in exercises 11.2.2 and 11.2.4, you will know by now that $\mathbf{M_3}$ and $\mathbf{N_5}$ are examples of non–distributive lattices and that $\mathbf{N_5}$ is an example of a non–modular lattice.



N₅



M₃

The rather surprising fact is that they are also the *prototypes* for the respective case in the sense of the following two Lemmata.

Lemma 11.6.2 If a lattice **L** is non–modular, then there is a sub–lattice of **L** which is isomorphic to $\mathbf{N_5}$.

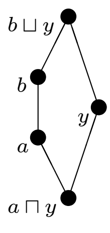
Proof. Assume $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is non–modular. Then, there are $x, y, z \in L$, $x \geq z$ such that $x \sqcap (y \sqcup z) \neq (x \sqcap y) \sqcup (x \sqcap z)$. Since $x \geq x \sqcap y$ and $y \sqcup z \geq z = x \sqcap z$, we conclude $x \sqcap (y \sqcup z) > (x \sqcap y) \sqcup z$. Setting $a := (x \sqcap y) \sqcup z$ and $b := x \sqcap (y \sqcup z)$ we thus have $x \geq b > a \geq z$.

We now claim that a and y are incomparable as are b and y. To see this, assume $y \geq a$. Then, $y \geq z$, and we have

$$\begin{aligned} b = x \sqcap (y \sqcup z) &= x \sqcap y \text{ (since } y \geq z) \\ &\leq (x \sqcap y) \sqcup z = a \end{aligned}$$

i.e. $b \leq a$, contradicting $b > a$. Dual arguments show that $y \leq b$ leads to a contradiction. Thus, we conclude that $y \not\geq a$ and $y \not\leq b$, from which we also conclude (using $b > a$) $y \not\geq b$ and $y \not\leq a$. Putting these last four facts together we have proved the claim.

Since $a < b$, $a \| y$ and $b \| y$ imply that a, b, y, $a \sqcap y$ and $b \sqcup y$ are pairwise distinct, we have the following situation:



So it remains to show that $b \sqcap y = a \sqcap y$ and $a \sqcup y = b \sqcup y$. Since $x \sqcap y \leq (x \sqcap y) \sqcup z = a \leq b$, we have

$$x \sqcap y = x \sqcap y \sqcap y \leq a \sqcap y \leq b \sqcap y,$$

and since $b \leq x$, we also have

$$b \sqcap y \leq x \sqcap y,$$

and thus

$$x \sqcap y = a \sqcap y = b \sqcap y.$$

Similarly for $a \sqcup y = b \sqcup y$.

It follows that $\{a \sqcap y, b \sqcup y, a, b, y\}$ is a sublattice isomorphic to $\mathbf{N_5}$. ■

Lemma 11.6.3 If a *modular* lattice $\mathbf{L}$ is non–distributive, then there is a sub–lattice of $\mathbf{L}$ which is isomorphic to $\mathbf{M_3}$.

To summarize, if $\mathbf{L}$ is non–distributive, then $\mathbf{L}$ has a sublattice which is isomorphic either to $\mathbf{N_5}$ (in which case $\mathbf{L}$ is not even modular) or to $\mathbf{M_3}$ (in which case $\mathbf{L}$ might still be modular).

We are thus left with the following alternative way of characterizing modularity and distributivity of lattices.

Theorem 11.6.4 Let $\mathbf{L}$ be a lattice. Then

1. $\mathbf{L}$ is modular
   iff
   $\mathbf{L}$ contains no sublattice isomorphic to $\mathbf{N_5}$
   iff
   $\mathbf{N_5}$ is not isomorphically embeddable into $\mathbf{L}$.

2. $\mathbf{L}$ is distributive
   iff
   $\mathbf{L}$ contains no sublattice isomorphic to either $\mathbf{N_5}$ or $\mathbf{M_3}$
   iff
   neither $\mathbf{N_5}$ nor $\mathbf{M_3}$ is isomorphically embeddable into $\mathbf{L}$.

Appendix A

A Proof for the Theorem of Łoś

In this section, we will give a proof of the Theorem of Łoś, the main theorem on ultraproducts 7.2.11. For the notation, please recall Section 7.2, especially the discussion following Exercise 7.2.3. Thus, for $a \in \prod_{s \in S} |\mathcal{A}_s|$, $s \in S$, ultrafilters $\mathcal{U}$ and valuations h in $\prod_{s \in S} |\mathcal{A}_s|$:

$$a(s) := a_s := \pi_s(a) \quad \text{and} \quad a_{\mathcal{U}} := \pi_{\mathcal{U}}(a),$$

$$h_s := \pi_s \circ h \quad \text{and} \quad h_{\mathcal{U}} := \pi_{\mathcal{U}} \circ h.$$

Theorem of Łoś 7.2.11 (Main Theorem on Ultraproducts).
For a formal language $\mathcal{L}$ let $\langle \mathcal{A}_s \,;\, s \in S \rangle$ be a family of $\mathcal{L}$–structures and $\mathcal{U}$ be an ultrafilter over S. For the sake of readability, let $\mathcal{B} := \prod_{s \in S} \mathcal{A}_s$ be the direct product and $\mathcal{A} := \mathcal{B}/\mathcal{U}$ the ultraproduct of the family $\langle \mathcal{A}_s \,;\, s \in S \rangle$ under $\mathcal{U}$. Then, the following holds:

1. For any valuation h into $\mathcal{B}$ and for any $\mathcal{L}$–formula φ,

$$\mathcal{A} \models \varphi[h_{\mathcal{U}}] \text{ iff } \{s \in S \,;\, \mathcal{A}_s \models \varphi[h_s]\} \in \mathcal{U}.$$

2. For any $\mathcal{L}$–sentence α,

$$\mathcal{A} \models \alpha \text{ iff } \{s \in S \,;\, \mathcal{A}_s \models \alpha\} \in \mathcal{U}.$$

The proof of Theorem 7.2.11 rests on two rather technical lemmata, which we are going to state and prove beforehand. In order to keep notation readable, we introduce the following abbreviation.

For $S, \mathcal{U}$, $\langle \mathcal{A}_s; s \in S\rangle$, φ and h as in the preamble and Clause 1 of the theorem, define $T(\varphi, h) \subseteq S$ by

$$T(\varphi, h) := \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s]\}\ .$$

Thus, $T(\varphi, h)$ is the set of indices s for which $\mathcal{A}_s$ is a model for φ under the (projected) valuation h_s. Using this convention, Clause 1 of the Theorem of Łos may be written as

$$\mathcal{A} \models \varphi[h_{\mathcal{U}}] \text{ iff } T(\varphi, h) \in \mathcal{U}.$$

The following lemma states that logic operators correspond to set theoretic operations on the subsets of S in a quite natural way.

Lemma A.1 For some formal language $\mathcal{L}$ let $\langle \mathcal{A}_s\ ;\ s \in S\rangle$ be a family of $\mathcal{L}$–structures and $\mathcal{B} := \prod_{s \in S} \mathcal{A}_s$. Then, for any valuation h into $\mathcal{B}$ and any $\mathcal{L}$–formulae φ, ψ and any variable x the following statements hold:

1. $T(\varphi \wedge \psi, h) = T(\varphi, h) \cap T(\psi, h)$,
2. $T(\neg\varphi, h) = S \setminus T(\varphi, h)$,
3. $T(\forall x \varphi, h) = \bigcap \{T(\varphi, h\binom{x}{b})\ ;\ b \in |\mathcal{B}|\}$.

Proof.

1.
$$\begin{aligned} T(\varphi \wedge \psi, h) &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi \wedge \psi[h_s]\} \\ &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s] \text{ and } \mathcal{A}_s \models \psi[h_s]\} \\ &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s]\} \cap \{s \in S\ ;\ \mathcal{A}_s \models \psi[h_s]\} \\ &= T(\varphi, h) \cap T(\psi, h) \end{aligned}$$

2.
$$\begin{aligned} T(\neg\varphi, h) &= \{s \in S\ ;\ \mathcal{A}_s \models \neg\varphi[h_s]\} \\ &= \{s \in S\ ;\ \mathcal{A}_s \not\models \varphi[h_s]\} \\ &= S \setminus \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s]\} \\ &= S \setminus T(\varphi, h) \end{aligned}$$

3.
$$\begin{aligned} T(\forall x \varphi, h) &= \{s \in S\ ;\ \mathcal{A}_s \models \forall x \varphi[h_s]\} \\ &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s\tbinom{x}{a}] \text{ for all } a \in |\mathcal{A}_s|\} \\ &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h_s\tbinom{x}{b_s}] \text{ for all } b \in |\mathcal{B}|\} \\ &\quad \text{(since } \pi_s \text{ is onto)} \\ &= \{s \in S\ ;\ \mathcal{A}_s \models \varphi[h\tbinom{x}{b}_s] \text{ for all } b \in |\mathcal{B}|\} \\ &\quad \text{(since } h_s\tbinom{x}{b_s} = (\pi_s \circ h)\tbinom{x}{b_s} = \pi_s \circ h\tbinom{x}{b} = h\tbinom{x}{b}_s \\ &= \bigcap \{T(\varphi, h\tbinom{x}{b})\ ;\ b \in |\mathcal{B}|\} \end{aligned}$$

■

The next lemma will be used when dealing with universally quantified formulae in the proof of the main theorem. Together with Clause 3 of Lemma A.1 it will show that

$$T(\forall x\psi, h) \in \mathcal{U} \text{ iff } T(\psi, h\binom{x}{b}) \in \mathcal{U} \text{ for all } b \in |\mathcal{B}|,$$

i.e. the set of indices s such that $\mathcal{A}_s \models \forall x\varphi[h_s]$ is *not* in the ultrafilter $\mathcal{U}$ if and only if there is a $b_0 \in \mathcal{B}$ (a *counter–example to* $\forall x\varphi[h_s]$*!*) such that $T(\psi, h\binom{x}{b_0}) \notin \mathcal{U}$.

Lemma A.2 For some formal language $\mathcal{L}$ let $\langle \mathcal{A}_s\ ;\ s \in S\rangle$ be a family of $\mathcal{L}$–structures, $\mathcal{U}$ an ultrafilter over S, and $\mathcal{B} := \prod_{s\in S} \mathcal{A}_s$. Then, for any valuation h into $\mathcal{B}$, any $\mathcal{L}$–formula ψ, and any variable x, the following holds:

$$\bigcap \{T(\psi, h\tbinom{x}{b})\ ;\ b \in |\mathcal{B}|\} \in \mathcal{U} \text{ iff } T(\psi, h\tbinom{x}{b}) \in \mathcal{U}, \text{ for all } b \in |\mathcal{B}|.$$

Proof. To show that the l.h.s. implies the r.h.s. is easy, since

$$\bigcap \{T(\psi, h\tbinom{x}{b})\ ;\ b \in |\mathcal{B}|\} \subseteq T(\psi, h\tbinom{x}{b}),$$

for any $b \in |\mathcal{B}|$. Thus, if $\bigcap \{T(\psi, h\binom{x}{b})\ ;\ b \in |\mathcal{B}|\} \in \mathcal{U}$, then clearly also $T(\psi, h\binom{x}{b})$ for all $b \in |\mathcal{B}|$, since $\mathcal{U}$ is an ultrafilter and therefore is closed under supersets.
To show the other direction, we proceed by contraposition and assume that $\bigcap \{T(\psi, h\binom{x}{b})\ ;\ b \in |\mathcal{B}|\} \notin \mathcal{U}$. We have to find some $b_0 \in |\mathcal{B}|$ such that $T(\psi, h\binom{x}{b_0}) \notin \mathcal{U}$.
$\bigcap \{T(\psi, h\binom{x}{b})\ ;\ b \in |\mathcal{B}|\} \notin \mathcal{U}$ iff $S \setminus \bigcap \{T(\psi, h\binom{x}{b})\ ;\ b \in |\mathcal{B}|\} \in \mathcal{U}$ (since $\mathcal{U}$ is an ultrafilter) iff $V := \{s \in S\ ;\ \text{there is } b_0 \in |\mathcal{B}| \text{ with } s \notin T(\psi, h\binom{x}{b_0})\} \in \mathcal{U}$.
Therefore, $s \in V$ iff $\mathcal{A}_s \not\models \psi[h\binom{x}{b_s}]$ for some $b \in |\mathcal{B}|$.
Now, for $s \in S$ define $D_s \subseteq |\mathcal{A}_s|$ by

$$D_s := \begin{cases} \mathcal{A}_s, & \text{iff } \mathcal{A}_s \models \psi[h_s\binom{x}{a}] \text{ for all } a \in |\mathcal{A}_s| \\ \{a \in |\mathcal{A}_s|\ ;\ \mathcal{A}_s \not\models \psi[h_s\binom{x}{a}]\}, & \text{else.} \end{cases}$$

(So, D_s contains either all the counter-examples, if there are any, or else all the examples for ψ in $\mathcal{A}_s$.)
Clearly $D_s \neq \emptyset$ for all $s \in S$, and (using the Axiom of Choice! Cf. Appendix B) we may conclude that $\prod_{s\in S} D_s \neq \emptyset$ as well.
Claim: For any $b \in \prod_{s\in S} D_s$,

$$T(\psi, h\tbinom{x}{b}) \notin \mathcal{U}.$$

To see this, take any $b \in \prod_{s\in S} D_s$ and assume $s \in T(\psi, h\binom{x}{b})$. Then,

$$\mathcal{A}_s \models \psi[h\tbinom{x}{b}_s]$$

by the definition of $T(\psi, [h\binom{x}{b}])$, which is equivalent to

$$\mathcal{A}_s \models \psi[h_s\tbinom{x}{b_s}],$$

so clearly

$$\mathcal{A}_s \models \psi[h_s\tbinom{x}{a}] \text{ for some } a \in |\mathcal{A}_s|, \text{ namely for } a = b_s.$$

By the definition of D_s we conclude that

$$D_s = |\mathcal{A}_s|,$$

hence,

$$\mathcal{A}_s \models \psi[h\tbinom{x}{a}] \text{ for } \textit{all } a \in |\mathcal{A}|,$$

but this is in turn equivalent to

$$\mathcal{A}_s \models \psi[h_s\tbinom{x}{b_s}] \text{ for all } b \in |\mathcal{B}|$$

and thus,

$$s \notin V.$$

Since $s \in T(\psi, h\binom{x}{b})$ was arbitrarily chosen, we conclude that

$$T(\psi, h\tbinom{x}{b}) \subseteq S \setminus V,$$

and $S \setminus V \notin \mathcal{U}$, hence,

$$T(\psi, h\tbinom{x}{b}) \notin \mathcal{U}.$$

Thus, choosing any $b \in \prod_{s\in S} D_s$ as b_0, we have $T(\psi, h\binom{x}{b_0}) \notin \mathcal{U}$. ■

Using these lemmata, it is not very difficult to prove the Theorem of Łos.

Proof (of the Theorem of Łos 7.2.11). Using Noetherian Induction on the structure of φ, we will show a somewhat stronger statement than Clause 1. We show that

$$\text{for } \textit{all} \text{ valuations } h \text{ into } \mathcal{B}, \mathcal{A} := \mathcal{B}/\mathcal{U} \models \varphi[h_{\mathcal{U}}] \text{ iff } T(\varphi, h) \in \mathcal{U}.$$

- If $\varphi = t_1 \doteq t_2$ for $\mathcal{L}$–terms t_1, t_2:

$$
\begin{array}{lll}
\mathcal{A} \models \varphi[h_{\mathcal{U}}] & \text{iff} & t_1^{\mathcal{A}}[h_{\mathcal{U}}] = t_2^{\mathcal{A}}[h_{\mathcal{U}}] \\
& \text{iff} & (t_1^{\mathcal{B}}[h])_{\mathcal{U}} = (t_2^{\mathcal{B}}[h])_{\mathcal{U}} \\
& \text{iff} & \{s \in S \,;\, (t_1^{\mathcal{B}}[h])_s = (t_2^{\mathcal{B}}[h])_s\} \in \mathcal{U} \\
& \text{iff} & \{s \in S \,;\, t_1^{\mathcal{A}}[h_s] = t_2^{\mathcal{A}}[h_s]\} = T(\varphi, h) \in \mathcal{U}.
\end{array}
$$

- If $\varphi = R_i(t_1, \dots, t_{\lambda_i})$ for $\mathcal{L}$–terms $t_1, \dots, t_{\lambda_i}$ and a relation symbol R_i:

$$
\begin{array}{lll}
\mathcal{A} \models \varphi[h_{\mathcal{U}}] & \text{iff} & \langle t_1^{\mathcal{A}}[h_{\mathcal{U}}] \,;\, t_{\lambda_i}^{\mathcal{A}}[h_{\mathcal{U}}]\rangle \in R_i^{\mathcal{A}} \\
& \text{iff} & \langle t_1^{\mathcal{B}}[h]_{\mathcal{U}} \,;\, t_{\lambda_i}^{\mathcal{A}}[h]_{\mathcal{U}}\rangle \in R_i^{\mathcal{A}} \\
& \text{iff} & \{s \in S \,;\, \langle (t_1^{\mathcal{B}}[h])_s \,;\, (t_{\lambda_i}^{\mathcal{B}}[h])_s\rangle \in R_i^{\mathcal{A}_s}\} \in \mathcal{U} \\
& \text{iff} & \{s \in S \,;\, \langle t_1^{\mathcal{A}_s}[h_s] \,;\, t_{\lambda_i}^{\mathcal{A}_s}[h_s]\rangle \in R_i^{\mathcal{A}_s}\} \in \mathcal{U} \\
& \text{iff} & \{s \in S \,;\, \mathcal{A}_s \models \varphi[h_s]\} \in \mathcal{U}.
\end{array}
$$

- If $\varphi = \neg\psi$ for an $\mathcal{L}$–formula ψ:

$$
\begin{array}{lll}
\mathcal{A} \models \varphi[h_{\mathcal{U}}] & \text{iff} & \mathcal{A} \not\models \psi[h_{\mathcal{U}}] \\
& \text{iff} & T(\psi, h) \notin \mathcal{U} \text{ (by the induction hypothesis)} \\
& \text{iff} & S \setminus T(\psi, h) = T(\varphi, h) \in \mathcal{U} \text{ (using Lemma A.1).}
\end{array}
$$

- If $\varphi = \psi \wedge \vartheta$ for $\mathcal{L}$–formulae ψ, ϑ:

$$
\begin{array}{lll}
\mathcal{A} \models \varphi[h_{\mathcal{U}}] & \text{iff} & \mathcal{A} \models \psi[h_{\mathcal{U}}] \text{ and } \mathcal{A} \models \vartheta[h_{\mathcal{U}}] \\
& \text{iff} & T(\psi, h) \in \mathcal{U} \text{ and } T(\vartheta, h) \in \mathcal{U} \\
& & \quad \text{(by the induction hypothesis)} \\
& \text{iff} & T(\psi, h) \cap T(\vartheta, h) = T(\varphi, h) \in \mathcal{U} \\
& & \quad \text{(using Lemma A.1).}
\end{array}
$$

- If $\varphi = \forall x\psi$ for an $\mathcal{L}$–formula ψ and a variable x:

$$
\begin{array}{lll}
\mathcal{A} \models \varphi[h_{\mathcal{U}}] & \text{iff} & \mathcal{A} \models \psi[h_{\mathcal{U}}\tbinom{x}{a}] \text{ for all } a \in |\mathcal{A}| \\
& \text{iff} & \mathcal{A} \models \psi[h\tbinom{x}{b}_{\mathcal{U}}] \text{ for all } b \in |\mathcal{B}| \\
& \text{iff} & T(\psi, h\tbinom{x}{b}) \in \mathcal{U} \text{ for all } b \in |\mathcal{B}| \\
& & \text{(by the inductive hypothesis)} \\
& \text{iff} & \bigcap \{T(\psi, h\tbinom{x}{b}) \,;\, b \in |\mathcal{B}|\} = T(\forall x\psi, h) = T(\varphi, h) \in \mathcal{U} \\
& & \text{(by Lemmata A.1 and A.2).}
\end{array}
$$

This proves Clause 1. For Clause 2 we note that for any $\mathcal{L}$–*sentence* α and any valuations h, h' into $\mathcal{B}$,

$$T(\alpha, h) = T(\alpha, h').$$

Therefore,

$$
\begin{array}{lll}
\mathcal{A} \models \alpha & \text{iff} & \mathcal{A} \models \alpha[h] \text{ for all valuations } h \text{ into } \mathcal{B} \\
& \text{iff} & T(\alpha, h) \in \mathcal{U} \text{ for all valuations } h \text{ into } \mathcal{B} \text{ (by } i.) \\
& \text{iff} & \{s \in S \,;\, \mathcal{A}_s \models \alpha[h_s]\} \in \mathcal{U} \text{ for some valuation } h \text{ into } \mathcal{B} \\
& & \text{(by the above remark)} \\
& \text{iff} & \{s \in S \,;\, \mathcal{A}_s \models \alpha\} \in \mathcal{U} \text{ (since } \alpha \text{ is a sentence),}
\end{array}
$$

which completes the proof of Clause 2 of the Theorem of Łos. ■

Appendix B

A Quick Introduction to Set Theory

What appears in this Module as an appendix intended to grant self-containedness should actually be put at the very beginning since Set Theory lies in the center of almost every mathematical field of work. It is both the foundation of the majority of the theories and subject of mathematical studies in itself.

In the 19th and the beginning of the 20th century, the increasing need for formal foundation and axiomatic description of Mathematics led to the Axiomatic Set Theory. Although we will not choose this approach to Set Theory, it is an interesting fact that the foundation of Logic and Model Theory may be regarded as rooted in the very same fields for which it should provide the foundation. So we will hint at the possibility to consider Set Theory as a formal theory in the sense of this module, formalized in a very simple formal language of first–order logic.

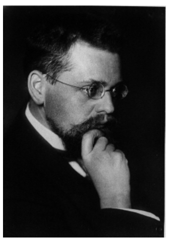


Figure B.1: Ernst Zermelo (1871-1953)

B.1 What are Sets?

The aim of this section is to provide the framework wherein the set theoretical notions essential for the understanding of this module will be defined.

The question which serves as a title for this section will not be answered. The disappointment should not be too big, since by now we should have gotten used to the fact that formal theories (and as such we must regard Set Theory) do not describe the entities entirely, but merely provide a sets of "rules" or "laws" which must hold for these entities. So what we will provide is a more or less informal description of these rules, which in principle (and please keep this in mind!) could be formalized as first–order axioms.

The entities considered in Set Theory are exclusively sets and classes. In fact, the main *raison d'être* for the axioms of Set Theory is to state that certain operations preformed on sets produce still sets. Sets may contain elements, which themselves are also sets (there is nothing else, remember). If a set contains no elements, it is called empty. In fact, since Set Theory focuses on content and not on structure, the elements of (entities contained in) a set are regarded as presented all at once, so there is no order involved. This leads to the fact that appart from the content, there is no way to distinguish sets. To be more precise,

Sets are equal if and only in they contain the same elements.

It is well–known that the fact that some set A belongs to (is contained in, is an element of) another set B is symbolized by

$$A \in B.$$

So the above formulation of the criteriom for identitfication of sets may be formalized as

$$A = B \iff \forall x(x \in A \iff x \in B).$$

This is one of the axioms of Set Theory, the Axiom of Extensionality. Obviously, the above formalization is in a formal language of first–order logic with one non–logical symbol, the binary relation symbol $\in$.

The Axiom of Extensionality has a direct consequence for sets without elements, namely that there exists at most one empty set. It is common practise to denote this empty set, whose existence will be guaranteed by other axioms, by $\emptyset$.

This is about everything we can say about what sets are. We now rather concentrate on how they behave and how we may build new sets from given ones. To this end, Set Theory provides the following set of axioms. To simplify notation, we use the generally accepted symbols and abbreviations:

- $A \in B$ denotes the fact that A is an element of B;
- $A \subseteq B$ ("A is subset of B") denotes the fact that every element of A is also an element of B;
- We will write $A \cup B$ to denote the union of A and B, i.e. the set which contains exactly all the element in A or B; similarly, $\bigcup A$ denotes the union of the elements of A, i.e. the set which contains exactly all the elements contained in some element of A;
- $A \cap B$ to denote the intersection of A and B, i.e. the set which contains exactly the elements that are both in A and in B; similarly, $\bigcap A$ denotes the intersection of the elements of A, i.e. the set which contains exactly the elements contained in all elements of A
- A and B are called disjoint if $A \cap B = \emptyset$;
- $\emptyset$ for the set containing no elements;
- $\mathcal{P}(A)$ for the powerset of A, i.e. the set of all subsets of A;

The standard universe of sets as we make use of in our mathematical studies is in fact grounded on a rather simple set of axioms. These axioms together formalize the *Set Theory of Zermelo and Fränkel with Choice*, ZFC for short. It is only one of a big variety of theories intended to serve as a foundation of the mathematical universe, but it is also the most generally accepted.

Axiom of Pairing For any sets x and y, there is a set containing both x and y as elements.

Axiom of Union For any set x, there is a set z such that for any $y \in x$, $y \subseteq z$.

eparation (Comprehension) For any set x and any formula φ in the formal language of Set Theory having at most one free variable, there is a set, denoted by $\{y \in x \,;\, \varphi(y)\}$, which contains all the elements of x which satisfy φ.

Axiom of Regularity Any non–empty set x contains an element which is disjoint to x.

Axiom of Infinity There is a set x such that $\emptyset \in x$ and for any y, if $y \in x$, then $y \cup \{y\} \in x$.

Power set Axiom For any set x there is a set containing all the subsets of x as elements;

The Axiom of Choice Given a set x of nonempty pairwise disjoint sets, there exists a set that contains exactly one element of each set in x; alternatively, the cartesian product of a non–empty family of non–empty sets is non–empty.

Figure B.2: Kazimierz Kuratowski (1896-1980)

This set of axioms is clearly infinite, but it is also countable (a notion we will only be able to define on the basis of these axioms ...) and it is of a simple form, or more precisely, *decidable*. By this we mean that there is a rather simple procedure to decide whether any given sequence of symbols from the language of Set Theory is one of the axioms. But nevertheless the theorems of this axiom–system provide (under the correct and intended interpretation) astonishing results. Let us have a look at how the universe of sets is populated with increasingly more complicated entities by use of these axioms.

That there is a set at all is a direct consequence of the Axiom of Infinity. Constructions such as direct products, relations and functions are equally made possible by these axioms, as we want to show in the next few paragraphs.

From the Axiom of Pairing, we may conclude that the bf ordered pair $\langle x, y \rangle$ exists for any two sets x and y, where $\langle x, y \rangle$ is given by the following definition originating from Kuratowsky:

$$\langle x, y \rangle := \{\{x\}, \{x, y\}\}.$$

The main purpose behind this seemingly rather arbitrary definition is to bring back some order in the otherwise order– and structure–less notion of sets. In an ordered pair, we may uniquely identify the first and the second component. This is formalized by the following result, whose proof is left as an exercise:

$$\langle x_1, y_1 \rangle = \langle x_2, y_2 \rangle \text{ if and only if } x_1 = x_2 \text{ and } y_1 = y_2.$$

Using ordered pairs, we may continue to define set–like objects which have more structure than mere sets: Relations and functions. A **relation**, as should be remembered, is a *set of ordered pairs*. To be more precise, a relation between the sets A and B is a set of ordered pairs $\langle a, b \rangle$ with $a \in A$ and $b \in B$. A **function** f from a set A to a set B, on the other hand, is a relation between A and B such that for any $a \in A$ there is *exactly one* $b \in B$ such that $\langle a, b \rangle \in f$.

The usual conventions for functions and relations apply, but to mention them all would turn this appendix into a book or module of its own. The important point is this: Functions and relations between sets can be shown to exist according to the axioms of ZFC, but not all of them are describable by nice formulas. As to *how many* functions and relations exist between two sets, is subject to additional axioms we will not discuss here. But nicely describable ones are easily proved to exist on the basis of our axioms.

Having defined relations and functions, we may continue towards tuples and families, direct products and direct powers. We may also show that the existence of structures in the sense of Model Theory, algebras and ordered sets can be taken for granted in our set–theoretically embedded universe.

But there are entities which provably are no longer sets since they are too big or too general in nature. The entities we are talking about are called **(proper) classes**, reflecting the idea that sets are a specialization of the notion of classes. Thus all sets are classes, but not vice versa. The most prominent example of a proper class is the class of all sets V. In fact, this class is to blame for the arising of axiomatic Set Theory early in the last century, when Bertrand Russell observed that the acceptance of V as a set leads to the following antinomy: If V is a set, then so is the subset R of all elements of V which are not member of itself,

$$B := \{\in V \,;\, x \notin x\} \,.$$

There are two cases to consider: (1) If $B \in B$, then $B \notin B$ by the definition of B, so the assumption $B \in B$ leads to a contradiction. So we conclude (2) $B \notin B$. However, this is equally contradictive, since then (again by the definition of B) $B \in B$. So we arrive at the contradiction

$$B \in B \text{ if and only if } B \notin B.$$

The way out of this antinomy is to exclude the class V from the collection of sets, since then B no longer is provably a set. To be more precise, we are only allowed to accept sub–classes as sets if they are sub–classes of sets which are separated from their parent–set by a formula in the language of axiomatic Set Theory, which is exactly the content of the Axiom Scheme of Separation.

From the fact that the class of all sets V is a proper class, we find many other examples of proper classes. Especially, as has been mentioned in Section 4.1, any class $\mathcal{L}$–structures is a proper class. Also, any class of models (apart from the empty class!) of a set Σ of $\mathcal{L}$–sentences is a proper class. The proof is left as an exercise.

Finite sets such as $\emptyset$, $\{\{\{\emptyset\}, \emptyset\}\}$ are no real challenge to our imagination. But matters tend to get a little more confusing and less intuitive when we start dealing with infinite sets. But to be able to distinguish between finite and

Figure B.3: Abraham Fraenkel (1891-1965)

infinite sets at all, we must have a measure for the size of sets, a measure of *counting the elements* provided purely within the set–theoretical framework.

B.2 Ordinals, Cardinals

The concept of cardinality is an abstraction of the numbering and counting of our everyday experience. Counting is a finite process by which we attach a natural number (whatever that is) to a certain collection of entities or objects. In Set Theory, the concept of cardinals (or cardinal numbers) is abstracting the process of counting the elements.

Before we can define what a cardinal is, we need to know what an ordinal is:

Definition B.2.1 If A is a set and R is a binary relation on A, R is said to **well–order** A (or "R is a well–ordering on A", or $\langle A, R\rangle$ is a well–order, or A is well–ordered by R) if R is reflexive, anti–symmetric, transitive, total and every non–empty subset of A contains a minimal element with respect to R. Thus, using infix–notation, $\langle A, R\rangle$ is a well–order if and only if

- $R \subseteq A \times A$;
- for all $x \in A$, xRx;
- for all $x, y \in A$, xRy and yRx imply $x = y$;
- for all $x, y, z \in A$, xRy and $yRz \Longrightarrow xRz$;
- for all $x, y \in A$, xRy or yRx;
- for all $B \subseteq A$, $B \neq \emptyset$ there is a $x \in B$ such that for no $y \in B$, yRx.

Thus, a well–ordering is an total, well–founded[1] order in the sense of Chapter 3.

[1] We use *well–founded* here rather than *noetherian* to establish some connection to the Axiom of Regularity which is sometimes also called the *Axiom of Foundation*.

One of the key features of Ordinal Numbers will be that they are transitive sets well–ordered by $\subseteq$, so this is a good moment to introduce the notion of a transitive set:

Definition B.2.2 A set A is called **transitive** if every element of A is also a subset of A, i.e. if $x \in A, y \in x \Longrightarrow y \in A$.

Definition B.2.3 A set A is an **Ordinal (Number)** iff the following properties hold;

(i) if $b \in A$ then $b \subseteq A$;

(ii) for all $a, b \in A$, either $a \in b$, $b \in a$ or $a = b$;

(iii) for all $B \subseteq A$, $B \neq \emptyset$, there is a $b \in B$ such that $b \cap B = \emptyset$.

The following observations provide an abundance of examples for ordinals: $\emptyset$ is an ordinal. If ξ is an ordinal, then so is $\succ(\xi) := \xi \cup \{\xi\}$. If Γ is a set of ordinals, then $\bigcup \Gamma$ and $\bigcap \Gamma$ are both ordinals.

The **finite ordinals** are defined by induction:

- $\emptyset$ is a finite ordinal;
- if n is a finite ordinal, then so is $\succ(\xi)$;
- no other sets are finite ordinals.

The finite ordinals mirror the process of counting as mentioned above. Therefore it is usual to denote finite ordinals by numbers (or *numerals*) according to the inductive rules $0 := \emptyset$ and $n + 1 :=\succ(n)$. Via this definition, the natural numbers are represented in Set Theory, and with a little additional work, we are able to represent the whole arithmetic in the framework of Set Theory. But this is not the aim of this chapter, our goal lies in the direction of infinity.

It is not very difficult to see that the set of finite ordinals is again an ordinal. Of course this new ordinal (we call it ω) is not a finite ordinal; as a matter of fact, it is the smallest infinite ordinal in the sense that any ordinal which is not finite contains ω as an element.

Appendix C

Exercises

C.1 Chapter ??

Exercise C.1.1 Show that, for a family $\langle X_i \,;\, i \in I\rangle$ $(I \neq \emptyset)$ and $\pi_i : \prod_{i\in I} X_i \longrightarrow X_i$ the canonical projection onto the ith component, there is a very natural bijective correspondence between $\prod_{i\in I} X_i \,/\, \ker \pi_i$ and X_i.

Exercise C.1.2 Show that for any filter $\mathcal{F}$ over the set I, the relation $\sim_{\mathcal{F}}$ defined on the direct product $\prod_{i\in I} X_i$ by

$$\langle x_i \,;\, i \in I\rangle \sim_{\mathcal{F}} \langle y_i \,;\, i \in I\rangle \ \text{ iff } \ \{i \in I \,;\, x_i = y_i\} \in \mathcal{F}$$

is an equivalence.

Exercise C.1.3 Show that $\mathcal{P}_{\mathrm{cof}}S$ is a filter on S.

Exercise C.1.4 Show that the **canonical embedding** $\iota : X \longrightarrow X^S/\mathcal{F}$ defined by $\iota(x) := [\langle x \,;\, s \in S\rangle]$ is 1–1 for any filter $\mathcal{F}$.

C.2 Chapter 3

Exercise C.2.1 Show that in an ordered set,

(a) greatest and least elements of are unique, provided they exist;

(b) any greatest element of some subset S is a maximum of S, and any least element of S is a minimum of S;

(c) maxima and minima need not exist for a given subset, and even if they exist, they need not be unique.

Exercise C.2.2 Show that in an ordered set, least upper bounds and greatest lower bounds of subsets are unique, whenever they exist.

Exercise C.2.3 Prove the statements in Example 3.1.10; especially, show that there are no covers in $(\mathbb{R}, \rho)$, $\operatorname{Inf}\{1/n; n \in \mathbb{N}\} = 0$ and that in Example 2, the atoms are exactly the prime numbers while there are no coatoms.

Exercise C.2.4 Why do the special cases $\operatorname{Sup}\emptyset$ and $\operatorname{Inf}\emptyset$ coincide with $\perp_{\mathbf{X}}$ and $\top_{\mathbf{X}}$ respectively, provided they exist?

Exercise C.2.5 Let $U \neq \emptyset$ be any set and pick a nonempty proper subset U_0 of U. Define $S := \{Z \subseteq U\,;\, U_0 \not\subseteq Z\}$. Show that with the order $\leq$ given by $Z_1 \leq Z_2$ iff $Z_1 \supseteq Z_2$, $\langle S, \leq\rangle$ is a Sup–semilattice but not an Inf-semilattice; moreover, calculate the semilattice operation for $\langle S, \leq\rangle$.

Exercise C.2.6 Show that for any lattice $\langle L, \leq\rangle$ and all $x, y \in L$,

$$\begin{aligned} \operatorname{Sup}\{x, \operatorname{Inf}\{x,y\}\} &= x; \text{ and} \\ \operatorname{Inf}\{x, \operatorname{Sup}\{x,y\}\} &= x. \end{aligned}$$

Exercise C.2.7 Show that for the divisibility order δ on $\mathbb{N}$, $\operatorname{Inf}\{m,n\} =$ g.c.d. of m and n and $\operatorname{Sup}\{m,n\} =$ l.c.m. of m and n.

Exercise C.2.8 Show that the power set $\mathcal{P}(X)$ of any set X is a complete lattice under the order of set–inclusion.

Exercise C.2.9 Show that every complete lattice L has a least element $\perp_{\mathbf{L}}$ and a greatest element $\top_{\mathbf{L}}$.

Exercise C.2.10 Show that $\mathbb{N}$ with the order of divisibility is a complete lattice.

Exercise C.2.11 Show that for any closure system $\mathcal{C}$ and any closure operator C, $C_{\mathcal{C}}$ is a closure operator and that $\mathcal{C}_C$ is a closure system, and that moreover

$$C_{\mathcal{C}_C} = C \text{ and } \mathcal{C}_{C_{\mathcal{C}}} = \mathcal{C}.$$

C.3 Chapter 4

Exercise C.3.1 Show that Mod Th and Th Mod are closure operators.

Exercise C.3.2 Show that for any $\mathcal{A}, \mathcal{B} \in \operatorname{Str} \mathcal{L}$ and $\varphi \operatorname{Sen} \mathcal{L}$,

$$\mathcal{A} \equiv \mathcal{B} \text{ iff } [\mathcal{A} \models \varphi \text{ iff } \mathcal{B} \models \varphi].$$

Exercise C.3.3 Show that if $\Sigma \vdash \varphi$ with $\varphi \notin \mathcal{L}(\Sigma)$, then $\vdash \varphi$.

Exercise C.3.4 Let $\mathcal{L}$ be a formal language. Show that

(a) $\operatorname{card} \operatorname{Tm} \mathcal{L} \leq \operatorname{card} \operatorname{Fml} \mathcal{L} = \operatorname{card} \operatorname{Sen} \mathcal{L}$;

(b) If $\mathcal{L}$ is countable, then so are $\operatorname{Tm} \mathcal{L}$, $\operatorname{Fml} \mathcal{L}$ and $\operatorname{Sen} \mathcal{L}$;

(c) if $\operatorname{card} \mathcal{L}$ is infinite, then $\operatorname{card} \operatorname{Tm} \mathcal{L} \leq \operatorname{card} \mathcal{L} = \operatorname{card} \operatorname{Fml} \mathcal{L}$;

(d) for any set X, if $\operatorname{card} \mathcal{L} \leq \operatorname{card} X$, then $\operatorname{card} \operatorname{Tm} \mathcal{L}_X = \operatorname{card} \operatorname{Fml} \mathcal{L} = \operatorname{card} X$.

Exercise C.3.5 Let $\mathcal{L}$ be the language having $\leq$ as only non–logical symbol, and consider the two $\mathcal{L}$–structures $\mathcal{A} := \langle \{(0,1),(1,0)\}, \leq \rangle$ and $\mathcal{B} := \langle \{(0,0),(1,1)\}, \leq \rangle$, where in both cases $\leq$ is the point–wise ordering. Show that there is a $\mathcal{L}$–homomorphism that is both injective and surjective, but $\mathcal{A}$ and $\mathcal{B}$ are not isomorphic.

Exercise C.3.6 Show that for a homomorphism $\eta : \mathcal{A} \longrightarrow \mathcal{B}$, the following are equivalent:

(i) η is a $\mathcal{L}$–isomorphism;

(ii) η is injective and surjective and η^{-1} is a $\mathcal{L}$–homomorphism;

(iii) η is injective and surjective and

$$R_i^{\mathcal{A}}(a_1, \ldots, a_{\lambda(i)}) \text{ iff } R_i^{\mathcal{B}}(\eta(a_1), \ldots, \eta(a_{\lambda(i)}))$$

for all relation–symbols R_i of $\mathcal{L}$ and all $a_1, \ldots, a_{\lambda(i)} \in |\mathcal{A}|$.

Exercise C.3.7 Show that for any isomorphism $\eta : \mathcal{A} \longrightarrow \mathcal{B}$, the inverse map η^{-1} is an isomorphism from $\mathcal{B}$ to $\mathcal{A}$

Exercise C.3.8 Treat the cases $\varphi \equiv R_i(t_1, \ldots, t_n)$ and $\varphi \equiv \vartheta \wedge \psi$ in the proof of Theorem 4.4.6.

Exercise C.3.9 Write out the details in the proof of Lemma 4.4.7.

C.4 Chapter 5

Exercise C.4.1 Show that equipotency, as introduced in Section 5.1 using bijections, is an equivalence relation.

Exercise C.4.2 Since A is regarded as of smaller cardinality than B if there is a injective mapping from A to B, what kind of mappings from B to A would in an equally plausible way constitute the fact that B is larger than A?

Exercise C.4.3 How do we get to the contradiction in the Diagonal argument?

Exercise C.4.4 Show that for $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$, $\varphi \in \operatorname{Fml}\mathcal{L}$ and c a new constant–symbol not in $\mathcal{L}$,

$$\text{if, in } \mathcal{L} \cup \{c\}, \Sigma \vdash \varphi\,(x/c) \text{ then, in } \mathcal{L},\ \Sigma \vdash \varphi.$$

Exercise C.4.5 Show that for $\Sigma \subseteq \operatorname{Sen}\mathcal{L}$ consistent,

(a) if $\varphi(x) \in \operatorname{Prop}\mathcal{L}$ and c is a new constant–symbol not in $\mathcal{L}$, then $\Sigma \cup \{\exists\varphi(x) \rightarrow \varphi(c)\}$ is a consistent set of $\mathcal{L} \cup \{c\}$–sentences;

(b) if $\varphi_1(x_1), \ldots, \varphi_n(x_n) \in \operatorname{Prop}\mathcal{L}$ and $c_1, \ldots, c_n$ are pairwise distinct, new constant–symbols not in $\mathcal{L}$, then

$$\Sigma \cup \{\exists\varphi_1(x_1) \rightarrow \varphi_1(c_1), \ldots, \exists\varphi_n(x_n) \rightarrow \varphi_n(c_n)\}$$

is a consistent set of $\mathcal{L} \cup \{c_1, \ldots, c_n\}$–sentences;

(c) if S is any set, $\{\varphi_s(x_s)\ ;\ s \in S\} \subseteq \operatorname{Prop}\mathcal{L}$ and $\{c_s\ ;\ s \in S\}$ a set of pairwise distinct new constant–symbols not in $\mathcal{L}$, then

$$\Sigma \cup \{\exists\varphi_s(x_s) \rightarrow \varphi_s(c_s)\ ;\ s \in S\}$$

is a consistent set of $\mathcal{L} \cup \{c_s\ ;\ s \in S\}$–sentences.

Exercise C.4.6 Show that for $\{\Sigma_s\ ;\ s \in S\}$ a set of consistent sets of $\mathcal{L}$–sentences linearly ordered by $\subseteq$, $\bigcup_{s\in S} \Sigma_s$ is consistent.

Exercise C.4.7 Show that a sentence $\alpha \in \operatorname{Sen}\mathcal{L}$ having a model has a countable model.

Exercise C.4.8 Show that a (consistent) theory having only uncountable models cannot be axiomatized by countably many axioms or in a countable language.

Exercise C.4.9 Which of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, $\mathbb{N}^2$ are countable?

Exercise C.4.10 Let $C := \{c_\beta\ ;\ \beta < \kappa\}$ be a set of cardinality κ of (pairwise distinct) constant–symbols. Let $\Gamma := \{\neg c \doteq c'\ ;\ c, c' \in C, c \neq c'\}$. Show that for any language $\mathcal{L}$ and any $\mathcal{L}_C$–structure $\mathcal{A}$, if $\mathcal{A} \models \Gamma$, then card $|\mathcal{A}| \geq \kappa$.

Exercise C.4.11 Analyze the proof of Proposition 5.4.2 to find out why the model of Σ has to be infinite.

Exercise C.4.12 Show that there is a set Σ of $\mathcal{L}$–sentences (for an adequate language $\mathcal{L}$) such that $\mathcal{A}$ is a model for Σ iff $\mathcal{A}$ has either exactly one or infinitely many elements.
(Hint: Consider $\Sigma := \{(\forall x x \doteq x) \vee \neg c \doteq c'\ ;\ c, c' \in C, c \neq c'\}$ for adequate C.)

Exercise C.4.13 In the proof of Theorem 7.4.1, explain how the $\mathcal{L}'$–structures $\mathcal{A}_s$ have to be defined (w.r.t. the interpretations of the new constant symbols). Explain in more detail why they exist.

C.5 Chapter ??

Exercise C.5.1 Show that the assignment $a'_n \mapsto b'_n$ from the proof of Proposition 6.2.2 does indeed define an isomorphism.

Exercise C.5.2 In the construction of the proof to Proposition 6.2.2, run the first few steps with your favorite enumerations of $\mathbb{Q}$ and $(0,1) \cap \mathbb{Q}$. (If you do not have a favorite enumeration of $(0,1) \cap \mathbb{Q}$, take the enumeration of $\mathbb{Q}$ and restrict it to the interval $(0,1)$.)

Exercise C.5.3 Show that $\{\zeta \in \mathbb{R}\ ;\ 0 < \zeta < 1\} \cup \{q \in \mathbb{Q}\ ;\ 1 \leq q < 2\}$ is a dense order without endpoints under the usual order.

Exercise C.5.4 Is it possible to apply the construction of the proof to Proposition 6.2.2 to show that, under the usual order,

(a) $\mathbb{R} \setminus \{0\}$ and $\mathbb{R}$ are not isomorphic;

(b) $\langle A \cap \mathbb{Q}, \leq \rangle$ and $\mathbb{Q}$ are not isomorphic?

C.6 Chapter 6

Exercise C.6.1 Show that for any language $\mathcal{L}$, $\mathrm{Sen}\,\mathcal{L}$ is the only inconsistent $\mathcal{L}$–theory.

Exercise C.6.2 Show that any complete set of $\mathcal{L}$–sentences is a $\mathcal{L}$–theory.

Exercise C.6.3 Show that for $\mathcal{L}$–structures $\mathcal{A}, \mathcal{B}$,

$$\mathrm{Th}\,\mathcal{A} \subseteq \mathrm{Th}\,\mathcal{B} \text{ iff } \mathcal{A} \equiv \mathcal{B}.$$

Exercise C.6.4 Show that any elementary class $\mathbb{K}$ is closed under elementary equivalence.

C.7 Chapter 7

Exercise C.7.1 Is $\mathbb{Z}_2 \times \mathbb{Z}_2$ – with operations defined by component — a field?

Exercise C.7.2 Show that an ultrafilter $\mathcal{U}$ over some set S is fixed at some $p \in S$ iff $\bigcap \mathcal{U} \neq \emptyset$.

Exercise C.7.3 Show that $\mathcal{U} \subseteq \mathcal{P}(S)$ is an ultrafilter over S iff $\mathcal{U}$ is a **prime filter** over S, i.e. iff, for any $U, V \subseteq S$, $U \cup V \in \mathcal{U}$ implies $U \in \mathcal{U}$ or $V \in \mathcal{U}$.

Exercise C.7.4 Show that if $\mathcal{U}$ is an ultrafilter over some set S and $U \in \mathcal{U}$ with $U = U_1 \cup \ldots \cup U_n$, then $U_i \in \mathcal{U}$ for some $i \in \{1, \ldots, n\}$.

Exercise C.7.5 Show that $\mathcal{P}_{\mathrm{cof}}\mathbb{N}$ is a filter over $\mathbb{N}$.

Exercise C.7.6 Let $X \neq \emptyset$, $S := \{s \subseteq X \,;\, s \text{ finite}\}$ and $T_x := \{s \in S \,;\, x \in s\}$ (for all $x \in X$). Is $\{T_x \,;\, x \in X\}$ a filter or not?

Exercise C.7.7 Show that for any ultrapower $\mathcal{A}^S/\mathcal{U}$ of a $\mathcal{L}$–structure $\mathcal{A}$, the function $\eta : \mathcal{A} \longrightarrow \mathcal{A}^S/\mathcal{U}$ given by $\eta(a) := \langle a \,;\, s \in S\rangle_{\mathcal{U}}$ is an embedding of $\mathcal{A}$ into $\mathcal{A}^S/\mathcal{U}$.

Exercise C.7.8 For $n \in \mathbb{N}$, find a $\mathcal{L}$–sentence holding in a $\mathcal{L}$–structure $\mathcal{A}$ iff $|\mathcal{A}|$ has exactly n elements.

C.8 Chapter 8

Exercise C.8.1 Use the argumentation of 8.1.7 to show that, for any language $\mathcal{L}$ and for any $\mathbb{K} \subseteq \mathrm{Str}\,\mathcal{L}$, if for any $n \in \mathbb{N}$ there is a $\mathcal{A} \in \mathbb{K}_{\mathrm{fin}}$ with card $|\mathcal{A}| >= n$, then $\mathbb{K}_{\mathrm{fin}}$ is not elementary.

C.9 Chapter 9

Exercise C.9.1 What operations are possible in an empty algebra?

Exercise C.9.2 Try to find group laws as in Example 9.1.4 which characterize groups among all algebras of type (3), or show that there are no such laws.

Exercise C.9.3 Using Remark 9.1.3, write a detailed proof for the fact that a constant of **A** will be mapped to the corresponding constant of **B** under any homomorphism $\eta \in \mathrm{Hom}(\mathbf{A}, \mathbf{B})$.

Exercise C.9.4 Find a simple example of two similar non–isomorphic algebras.

Exercise C.9.5 Verify for groups, rings and vector spaces that the respective definitions of homomorphisms match with the definition given in Definition 9.2.1, i.e. show that for example group homomorphisms coincide with homomorphisms of groups as universal algebras.

Exercise C.9.6 Show that any valuation h into an $\mathcal{L}$–structure $\mathcal{A}$ determines an $\mathcal{L}$–homomorphism $\eta_h : \mathbf{T}_\mathcal{L} \longrightarrow \mathcal{A}$ from the term–algebra $\mathbf{T}_\mathcal{L}$ into $\mathcal{A}$, given by $\eta_h(t) := t^\mathcal{A}[h]$. Also verify that if the language is functional, η_h is a homomorphism in the sense of Definition 9.2.1.

Exercise C.9.7 Show that if groups are considered as algebras of type (3) with the operation m as defined in 9.1.4.4, the subalgebras are exactly the co–sets (left or right) of ordinary subgroups of G and $\emptyset$.

Exercise C.9.8 Prove Proposition 9.3.3, i.e. show that for all algebras **A** and **B**any $\eta \in \mathrm{Hom}(\mathbf{A}, \mathbf{B})$,

1. if $S \in \mathbf{Sub}\,\mathbf{A}$, then $\eta[S] \in \mathbf{Sub}\,\mathbf{B}$;
2. if $T \in \mathbf{Sub}\,\mathbf{B}$, then $\eta^{-1}[T] \in \mathbf{Sub}\,\mathbf{A}$;
3. the union of a chain of subuniverses of **A** is a subuniverse of **A**, and the union of a directed system of subuniverse of **A** is a subuniverse of **A**.

Exercise C.9.9 Show that if $\langle S_k \,;\, k \in K \rangle$ is any family of subuniverses of an algebra **A**, then the set intersection $\bigcap_{k \in K} S_k$ is a subuniverse of **A**.

Exercise C.9.10 Find an algebra **A** and subuniverses **B** and **C** of **A** such that $\mathbf{B} \cup \mathbf{C}$ is *not* a subuniverse of **A**.

Exercise C.9.11 Consider the group $\langle \mathbb{Z}; +, -, 0\rangle$, and let $X := \{2\}$. Moreover let $G = \{+, -, 0\}$ be the set of fundamental operations of $\mathbb{Z}$.

Write out detailed calculations of $G^0[X], G^1[X], G^2[X]$ and $G^3[X]$ according to Definition **??**. Is there a general way to describe $G^n[X]$? And can you show directly (i.e. without using Lemma **??** or a similar result) that $\bigcup_{n\in\mathbb{N}} G^n[X]$ is the set of even integers?

Exercise C.9.12

1. Verify in detail that $\langle \mathbb{Z}; +, -, 0\rangle$ is not locally finite.
2. Verify that $\mathbb{Z}_2^{\mathbb{N}}$, the product of countably infinite many two–element groups with operations defined componentwise, is locally finite but not finitely generated.
3. Show that every finite algebra is locally finite.
4. Show that there is no infinite, locally finite, finitely generated algebra.

Exercise C.9.13 Show that the additive group of rational numbers has no maximal subgroups.

Exercise C.9.14 Show that

1. In every algebra, the smallest subalgebra is the subalgebra generated by $\emptyset$.
2. If there are nullary operations, $\mathbf{A}[\emptyset]$=$\mathbf{A}[\{c\,;\ c$ is a constant$\}]$.

Exercise C.9.15 Show that both $\langle \mathbb{Q}; +, -, 0\rangle$ and $\langle \mathbb{Z}; +, -, 0\rangle$ have no minimal subgroups.

Exercise C.9.16 Show that the following constructions fall into the scope of the definition of a direct product of (universal) algebras:

1. The direct product of groups.
2. The direct product of rings.
3. The direct product of fields.
4. The direct product of vector–spaces.

Exercise C.9.17 Show that the projections π_k associated with the notion of direct products are surjective homomorphisms from $\prod_{k\in K} \mathbf{A}_k$ onto $\mathbf{A}_k$ for any collection.

Exercise C.9.18 Let $\mathbf{B}$ and $\mathbf{A}_k$ ($k \in K$) be similar algebras, and $g_k : \mathbf{B} \longrightarrow \mathbf{A}_k$ a surjective homorphism for each $k \in K$. Show that there is a uniquely determined surjective homomorphism $g : \mathbf{B} \longrightarrow \prod_{k \in K} \mathbf{A}_k$ satisfying $g_k = \pi_k \circ g$ for all $k \in K$.

Exercise C.9.19 Show that $h : A \longrightarrow B$ is a homomorphism from $\mathbf{A}$ into $\mathbf{B}$ iff $\{\langle a, h(a)\rangle \; ; \; a \in A\}$ is a subuniverse of $\mathbf{A} \times \mathbf{B}$.

C.10 Chapter 10

Exercise C.10.1 Define a relation ϑ on $\mathbb{Q}$ by $a\vartheta b$ if and only if $a - b \in \mathbb{Z}$. Show that

1. ϑ is an equivalence.
2. ϑ is compatible with $+$ and $-$.
3. ϑ is not compatible with $\cdot$.
4. ϑ is a congruence on the additive group $\langle \mathbb{Q}; +, -, 0\rangle$ but not on the ring $\langle \mathbb{Q}; +, -, \cdot, 0, 1\rangle$.

Exercise C.10.2 Show that if ϑ is a congruence on $\mathbf{A}$, then $\pi_\vartheta : \mathbf{A} \longrightarrow \mathbf{A} / \vartheta$ is a surjective homomorphism.

Exercise C.10.3 Prove Proposition 10.1.3: For any algebra $\mathbf{A}$, the congruences on $\mathbf{A}$ are precisely the kernels of the homomorphisms with source $\mathbf{A}$.

Exercise C.10.4 Modify the proof of Theorem 10.1.5 to show that for any $\eta : \mathbf{A} \longrightarrow \mathbf{B}$, $\eta[\mathbf{A}] \cong \mathbf{A} / \ker \eta$.

Exercise C.10.5 Show that for congruences $\vartheta, \rho \in \mathbf{Con}\,\mathbf{A}$, ϑ/ρ is a binary relation on A and $\vartheta/\rho \in \mathbf{Con}\,\mathbf{A}$.

Exercise C.10.6 For $\vartheta, \rho \in \mathbf{Con}\,\mathbf{A}$, show that $\eta : (\mathbf{A} / \rho)/(\vartheta/\rho) \longrightarrow \mathbf{A} / \vartheta$ with $\eta([[a]_\rho]_{\eta/\rho}) := [a]_\vartheta$ is well–defined and an isomorphism.

Exercise C.10.7 Which of the following statements is true?

1. For any algebra $\mathbf{A}$, any $\vartheta \in \mathbf{Con}\,\mathbf{A}$ and any subset $B \subseteq A$, B^ϑ is a subuniverse of $\mathbf{A}$.
2. For any algebra $\mathbf{A}$ and any $\vartheta \in \mathbf{Con}\,\mathbf{A}$, the assignment $B \mapsto B^\vartheta$ defines a closure operator on A.

Exercise C.10.8 Show that $\eta : \mathbf{B}/(\vartheta \cap B^2) \longrightarrow B^\vartheta/(\vartheta \cap {B^\vartheta}^2)$ with

$$\eta([b]_{\vartheta \cap B^2}) := [b]_{\vartheta \cap {B^\vartheta}^2}.$$

is well–defined and an isomorphism.

Exercise C.10.9 Show that for a normal subgroup N of a group G, the relation ϑ_N on G with $a\vartheta_N b$ if and only if $aN = bN$ is an equivalence relation.

Exercise C.10.10

1. Is the relational product commutative? Is it associative?
2. What if we restrict ourselves to equivalences on some set A?
3. Given two equivalences ϑ, ρ on some set A, is the relational product $\vartheta \circ \rho$ always an equivalence?
4. Show that for $\vartheta \in \mathbf{Eq}\ \mathbf{A}$, $\vartheta \circ \vartheta = \vartheta$.

Exercise C.10.11 Show that Δ_A is the unit element with respect to the relational product $\circ$ for any set A. On the other hand, find a set A and $R \subseteq A \times A$ such that $R \circ R^{-1} \neq \Delta_A$.

Exercise C.10.12 Show that for $\vartheta_1, \vartheta_2 \subseteq A^2$, we have

(a) $\vartheta_1 \circ \vartheta_2^{-1} = \vartheta_2^{-1} \circ \vartheta_1^{-1}$;

(b) $\vartheta_1 \subseteq \vartheta_2$ iff $\vartheta_1^{-1} \subseteq \vartheta_2^{-1}$.

Exercise C.10.13 Show that for $\vartheta \in \mathbf{Eq}\ \mathbf{A}$, we have $\vartheta^{-1} = \vartheta$.

Exercise C.10.14

1. Does $\operatorname{Sup}\Theta = \bigcup \Theta$ imply that Θ is directed?
2. Does $\operatorname{Sup}\Theta = \bigcup \Theta$ hold for *directed* sets Θ of *equivalences*?

Exercise C.10.15 Complete the proof of Proposition 10.2.8, i.e. show that for any set Θ of congruences on an algebra $\mathbf{A}$ and $\Xi := \{\vartheta_0 \circ \ldots \circ \vartheta_n \,;\, n \in \mathbb{N}, \vartheta_0, \ldots, \vartheta_n \in \Theta\}$,

1. $\vartheta \subseteq \bigcup \Xi$ for any $\vartheta \in \Theta$;
2. $\bigcup \Xi$ is a congruence;

3. $\vartheta_0 \circ \ldots \circ \vartheta_n \subseteq \operatorname{Sup}\Theta$ for all $n \in \mathbb{N}$ and all $\vartheta_0, \ldots, \vartheta_n \in \Theta$.

Exercise C.10.16 Show that for $\vartheta_1, \vartheta_2 \in \mathbf{Con\,A}$, the following are equivalent:

(i) $\vartheta_1 \circ \vartheta_2 = \vartheta_2 \circ \vartheta_1$;

(ii) $\operatorname{Sup}\{\vartheta_1, \vartheta_2\} = \vartheta_1 \circ \vartheta_2$;

(iii) $\vartheta_1 \circ \vartheta_2 \subseteq \vartheta_2 \circ \vartheta_1$.

Exercise C.10.17 Show that $\eta : [\vartheta, \nabla_A] \longrightarrow \mathbf{Con\,A}/\vartheta$ with $\eta(\rho) := \rho/\vartheta$ is a lattice–isomorphism.

Exercise C.10.18 Show that if $\mathbf{A}$ is an algebra and $\Theta \subseteq \mathbf{Con\,A}$, then $\operatorname{Sup}\Theta = \theta(\bigcup \Theta)$.

Exercise C.10.19 Show that all but finitely many congruences on $\mathbb{Z}$ are of the form $\theta(\langle a, b, \rangle)$ and find the exceptions.

Exercise C.10.20 Show that the generation of congruences defines a closure operator, i.e. show that

(i) $R \subseteq \theta(R)$,

(ii) $\theta(\theta(R)) = \theta(R)$ and

(iii) $R \subseteq S$ implies $\theta(R) \subseteq \theta(S)$

for all subsets R, S of some algebra.

Exercise C.10.21 Show that for any congruence ϑ, the set $\{\theta(R)\,;\ R \subseteq \vartheta, R \text{ finite}\}$ is directed.

Exercise C.10.22 Which of the equations in Proposition 10.3.5 hold if "congruence" is replaced by equivalence (and consequently generation of congruences by generation of equivalences)?

Exercise C.10.23 Let a be a compact element of a complete lattice. Show that if $\operatorname{Sup} C \geq a$ for a chain $C \subseteq A$, then $c \geq a$ for some $c \in C$.

Exercise C.10.24 Find proofs for the following statements:

(a) Every complete lattice has at least one compact element.

(b) Finite lattices consist of compact elements exclusively.

(c) The compact elements in the complete lattice $\langle \mathcal{P}(X), \subseteq \rangle$ are exactly the finite $Y \subseteq X$.

Exercise C.10.25 Find an example of a non–algebraic complete lattice.

C.11 Chapter 11

Exercise C.11.1 Prove Proposition 11.1.4, i.e. show that

1. If $\langle L, \leq \rangle$ is a lattice (as a poset), then $\langle L; \mathrm{Sup}_{\leq}, \mathrm{Inf}_{\leq} \rangle$ is a lattice (as an algebra).

2. If $\langle L; \sqcup, \sqcap \rangle$ is a lattice (as an algebra), then the two order–relations $\leq_{\sqcup}$ and $\leq_{\sqcap}$ as defined in Proposition 11.1.2 are identical and $\langle L, \leq_{\sqcup} \rangle$ is a lattice (as a poset).

3. The transformation is mutually connected by

$$\langle L, \leq_{\sqcup_{\leq}} \rangle = \langle L, \leq \rangle$$

and

$$\langle L; \sqcup_{\leq_{\sqcup}}, \sqcap_{\leq_{\sqcup}} \rangle = \langle L; \sqcup, \sqcap \rangle \, .$$

Exercise C.11.2 Show that if α and β are congruences on some algebra, then Inf and Sup of α and β computed as equivalences as in Example 11.1.5 1 are indeed congruences again.

Exercise C.11.3 Let $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ be a lattice and let $\leq$ be the order stemming from $\sqcap$ and $\sqcup$ in the sense of Proposition 11.1.4. Show that $x \leq x'$ and $y \leq y'$ imply $x \sqcap y \leq x' \sqcap y'$ and $x \sqcup y \leq x' \sqcup y'$, i.e.
$\sqcap : L^2 \longrightarrow L$ and $\sqcup : L^2 \longrightarrow L$ are order–preserving.

Exercise C.11.4 Show that any lattice satisfying $D_{\sqcap}$ (cf. Definition 11.2.1) also satisfies $D_{\sqcup}$ and vice versa.

Exercise C.11.5 Find out which of the following finite lattices are distributive: $\mathbf{2}$, $\mathbf{N_5}$, $\mathbf{M_3}$, $\mathbf{B}_3$ (cf. Example 3.1.16).

Exercise C.11.6 Find out which of the following finite lattices are modular: $\mathbf{2}$, $\mathbf{N_5}$, $\mathbf{M_3}$, $\mathbf{B}_3$ (cf. Example 3.1.16).

Exercise C.11.7 Show that a lattice $\mathbf{L} = \langle L; \sqcap, \sqcup \rangle$ is modular iff

$$((x \sqcap z) \sqcup y) \sqcap z = (x \sqcap z) \sqcup (y \sqcap z)$$

for all $x, y, z \in L$.

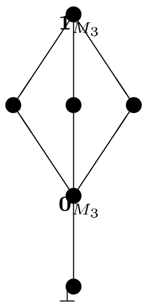
Exercise C.11.8 Show that in a distributive lattice any element has at most one complement.

Exercise C.11.9 Show that in a chain, only the top and bottom elements have complements, so chains considered as lattices are not complemented if they have more than two elements.

Exercise C.11.10 Show that for any infinite set X, the lattice with all finite and cofinite subsets of X as its carrier set and with the operations of set union and intersection is a Boolean lattice.

Exercise C.11.11 Which open subsets of $\mathbb{R}$ have an open complement in $\mathbb{R}$? And which open subsets of $\mathbb{R}^2$ have an open complement in $\mathbb{R}^2$?

Exercise C.11.12 Show that $\mathbf{M}_{3\perp}$ as depicted below is a lattice where exactly $\mathbf{1}_{M_3}$ and $\perp$ have complements, and where $\perp$ serves as a pseudocomplement of everything except itself.



$\mathbf{M}_{3\perp}$

Exercise C.11.13 Let $\mathcal{U}$ be the collection of open subsets of $\mathbb{R}$ which are contained in the open interval $(-2, 2)$ but do not contain $(-1, 1)$, together with all intervals of the form $(-2 - 1/n, 2 + 1/n)$, and $\emptyset$ and $\mathbb{R}$. Show that ordered by set inclusion, $\mathcal{U}$ is a meet semilattice with $\cap$ as Inf but not a lattice since, e.g., $(-1.5, 0)$ and $(0, 1.5)$ have no Sup within $\mathcal{U}$.

Exercise C.11.14 Let $\mathcal{L}_{\leq}$ be the formal language having the binary relation–symbol $\leq$ as its only non–logical symbol. Show that the class of lattices is basic–elementary by finding an appropriate set of $\mathcal{L}_{\leq}$–sentences Σ_L such that $\operatorname{Mod}\Sigma_L$ is exactly the class of lattices.

Exercise C.11.15 Let $\hat{\mathbb{N}}$ be the set $\mathbb{N}$ enriched by a new element $\top$, and set

$$x \leq_{\hat{\mathbb{N}}} y \text{ iff } [x, y \in \mathbb{N} \text{ and } x \leq y] \text{ or } y = \top$$

for all $x, y \in \hat{\mathbb{N}}$.

(a) Show that $\leq_{\hat{\mathbb{N}}}$ is an order on $\hat{\mathbb{N}}$ and that $\langle\hat{\mathbb{N}}, \leq_{\hat{\mathbb{N}}}\rangle$ is a complete lattice.

Let $\mathcal{U}$ be an ultrafilter over $\mathbb{N}$ containing all cofinite subsets of $\mathbb{N}$, and let $\mathcal{A} := \hat{\mathbb{N}}^{\mathbb{N}}/\mathcal{U}$. Show that

(b) $\mathcal{A}$ is a lattice with greatest element $\top_{\mathcal{A}} := \langle\top, \top, \top, \ldots\rangle_{\mathcal{U}}$ and least element $0_{\mathcal{A}} := \langle 0, 0, 0, \ldots\rangle_{\mathcal{U}}$.

(c) all elements of $\mathcal{A}$ except $\top_{\mathcal{A}}$ and $0_{\mathcal{A}}$ have an upper and a lower cover.

(d) $\langle 0, 1, 2, \ldots\rangle_{\mathcal{U}}$ is an upper bound of $\overline{\mathbb{N}} := \{\overline{n}\,;\, n \in \mathbb{N}\}$ in $\mathcal{A}$, where $\overline{n} := \langle n, n, n, \ldots\rangle_{\mathcal{U}}$.

- if $b \in |\mathcal{A}|$ is an upper bound of $\overline{\mathbb{N}}$, then so is the lower cover of b. (Hint: If such a lower cover c were not an upper bound of $\overline{\mathbb{N}}$, then for some $\overline{n} \in \overline{\mathbb{N}}$, $\overline{n} \leq c \leq \overline{n+1}$, from which we conclude $c = \overline{n}$ or $c = \overline{n+1}$. But then b, being the upper cover of c, could not be an upper bound of $\overline{\mathbb{N}}$.)

(e) $\overline{\mathbb{N}}$ does not have a supremum in $\mathcal{A}$ (and the set of upper bounds of $\overline{\mathbb{N}}$ has no infimum).

Exercise C.11.16 Let $\mathbf{L_4}$ be the four element chain, and let $\mathbf{L_4}_{\perp\top}$ be the lattice resulted from adding a new least element $\perp$ and a new greatest element $\top$ to $\mathbf{L_4}$. Show that the map from $\mathbf{L_4}$ to $\mathbf{L_4}_{\perp\top}$ given by $\mathbf{0} \mapsto \perp$, $\mathbf{1} \mapsto \top$, $a \mapsto a$ and $b \mapsto b$ is an order–homomorphism but not a lattice–homomorphism.

C.12 Chapter ??

C.13 Chapter ??

C.14 Chapter ??

C.15 Chapter B

Exercise C.15.1 Show that from the axiom of pairing, using First–order Logic, we can proof that the ordered pair of any two sets exists.

Exercise C.15.2 Show that for any sets x_1, x_2, y_1, y_2,

$$\langle x_1, y_1\rangle = \langle x_2, y_2\rangle \text{ if and only if } x_1 = x_2 \text{ and } y_1 = y_2.$$

Exercise C.15.3 Show that for any formal language $\mathcal{L}$ and any consistent set Σ of $\mathcal{L}$–sentences, $\mathrm{Mod}\,\Sigma$ is a proper class. (Hint: Take any model of Σ and any element a in the universe of this model. Replace a by an arbitrary set and show that the relations, functions etc. of the structure may be modified in such a way that the result is again a structure. The claim follows by the arbitrariness of the set and the Axiom of Union.)

Afterword

Index