# Introduction to
# CATEGORY THEORY
## and
# CATEGORICAL LOGIC

Thomas Streicher

SS 03 and WS 03/04

# Contents

# Introduction

The aim of this course is to give an introduction to the basic notions of Category Theory and Categorical Logic.

The first part on Category Theory should be of interest to a general mathematical audience with interest in algebra, geometry and topology where at least the language of category theory and some of its basic notions like limits, colimits and adjoint functors are indispensible nowadays. However, for following the lectures in a profitable way one should have already attended a course in basic algebra or topology because algebraic structures like groups, rings, modules etc. and topological spaces serve as the most important source of examples illustrating the abstract notions introduced in the course of the lectures.

The second part will be of interest to people who want to know about logic and how it can be modelled in categories. In particular, we will present *cartesian closed categories* where one can interpret typed $\lambda$-calculus, the basis of modern functional programming languages, and (elementary) toposes providing a most concise and simple notion of model for constructive higher order logic. Guiding examples for both notions will be presented *en detail*. Some knowledge about constructive logic would be helpful (as a motivating background) but is not necessary for following the presentation itself.

We conclude this most concise introduction with a list of suggestions for further reading.

# References

[Aw]   S. Awodey *Category Theory* OUP (2006).

[BW1] M.Barr, Ch. Wells *Toposes, Triples and Theories* Springer (1985).

[BW2] M.Barr, Ch. Wells *Category Theory for Computing Science* Prentice Hall (1990).

[Bor]  F. Borceux  *Handbook of Categorical Algebra*  3 vols., Cambridge University Press (1994).

[FS]   P.J. Freyd, A. Sčedrov *Categories, Allegories* North Holland (1990).

[Jac]    B. Jacobs *Categorical Logic and Type Theory* North Holland (1999).

[Joh]    P. T. Johnstone *Sketches of an Elephant. A Topos Theory Compendium.* 2 vols. OUP (2002).

[JM]    A. Joyal, I. Moerdijk *Algebraic Set Theory* CUP (1995).

[LaS]    J. Lambek, P. J. Scott *Introduction to Higher Order Categorical Logic* CUP (1986).

[LS]    F.W. Lawvere, S. Schanuel *Conceptual Mathematics* CUP (1997).

[LR]    F.W. Lawvere, R. Rosebrugh *Sets for Mathematics. A first introduction to categories.* CUP (2003).

[McL]    C. McLarty *Elementary Categories, Elementary Toposes* OUP (1995).

[ML]    S. MacLane *Categories for the Working Mathematician* Spinger (1971).

[MM]    S. MacLane, I. Moerdijk *Sheaves in Geometry and Logic. A First Introduction to Topos Theory.* Spinger (1992).

[PRZ]    M. La Palme Reyes, G. Reyes, H. Zolfaghari *Generic figures and their glueings. A constructive approach to functor categories.* Polimetrica (2004).

[PT]    P. Taylor *Practical Foundations* CUP (1999).

# Part I    CATEGORY THEORY

# 1 Categories

We first introduce our basic notion of structure, namely categories.

**Definition 1.1** *A* category $\mathbb{C}$ *is given by the following data*

- *a class* $\mathsf{Ob}(\mathbb{C})$ *of* objects *of* $\mathbb{C}$

- *a family* $\mathsf{Mor}(\mathbb{C})$ *associating with every pair* $A, B \in \mathsf{Ob}(\mathbb{C})$ *a class* $\mathsf{Mor}(\mathbb{C})(A, B)$ *of* morphisms *from $A$ to $B$*

- *for all* $A, B, C \in \mathsf{Ob}(\mathbb{C})$ *a mapping*

$$\circ_{A,B,C} : \mathsf{Mor}(\mathbb{C})(B, C) \times \mathsf{Mor}(\mathbb{C})(A, B) \to \mathsf{Mor}(\mathbb{C})(A, C)$$

  *called* composition

- *for all* $A \in \mathsf{Ob}(\mathbb{C})$ *a distinguished morphism*

$$\mathsf{id}_A \in \mathsf{Mor}(\mathbb{C})(A, A)$$

  *called* identity morphism *for $A$*

*required to satisfy the following conditions*

- *for all* $A, B, C, D \in \mathsf{Ob}(\mathbb{C})$ *and* $f \in \mathsf{Mor}(\mathbb{C})(A, B)$, $g \in \mathsf{Mor}(\mathbb{C})(B, C)$ *and* $h \in \mathsf{Mor}(\mathbb{C})(C, D)$ *it holds that*

$$\text{(Ass)} \quad h \circ (g \circ f) = (h \circ g) \circ f$$

  *standing as an abbreviation for the more explicit, but also more unreadable equation* $\circ_{A,C,D}(h, \circ_{A,B,C}(g, f)) = \circ_{A,B,D}(\circ_{B,C,D}(h, g), f)$

- *for all* $A, B, C \in \mathsf{Ob}(\mathbb{C})$ *and* $f \in \mathsf{Mor}(\mathbb{C})(A, B)$ *and* $g \in \mathsf{Mor}(\mathbb{C})(C, A)$ *it holds that*

$$\text{(Id)} \quad f \circ \mathsf{id}_A = f \quad and \quad \mathsf{id}_A \circ g = g$$

  *standing as an abbreviation for the more explicit, but also more unreadable equations* $\circ_{A,A,B}(f, \mathsf{id}_A) = f$ *and* $\circ_{C,A,A}(\mathsf{id}_A, g) = g.$ $\qquad \Diamond$

Notice that the identity morphisms are uniquely determined by $\circ$ and the requirement (Id). (Exercise!)

**Some remarks on notation.**

As already in Definition 1.1 we write simply $g \circ f$ instead of the more explicit $\circ_{A,B,C}(g, f)$ whenever $f \in \mathsf{Mor}(\mathbb{C})(A, B)$ and $g \in \mathsf{Mor}(\mathbb{C})(B, C)$. Instead of the somewhat clumsy $\mathsf{Mor}(\mathbb{C})(A, B)$ we often write simply $\mathbb{C}(A, B)$ and for $f \in \mathsf{Mor}(\mathbb{C})(A, B)$ we simply write $f : A \to B$ when $\mathbb{C}$ is clear from the context. Instead of $\mathsf{id}_A$ we often write $1_A$ or simply $A$. When the object $A$ is clear from the context we often write simply $\mathsf{id}$ or $1$ instead of $\mathsf{id}_A$ or $1_A$, respectively.

Next we consider some

**Examples of Categories**

(1) The category whose objects are sets, whose morphisms from $A$ to $B$ are the set-theoretic functions from $A$ to $B$ and where composition is given by $(g \circ f)(x) = g(f(x))$ is denoted as **Set**. Of course, in **Set** the identity morphism $\mathsf{id}_A$ sends every $x \in A$ to itself. For obvious reasons we call **Set** the *category of sets (and functions)*.

(2) We write **Set**$_*$ for the category of sets with a distinguished element (denoted by $*$) and functions preserving this distinguished point.

(3) We write **Mon** for the category of monoids and monoid homomorphisms. This makes sense as monoid homomorphisms are closed under composition and identity maps preserve the monoid structure.

(4) We write **Grp** for the full subcategory[1] of **Mon** whose objects are groups.

(5) We write **Ab** for the full subcategory of **Grp** whose objects are the abelian (i.e. commutative) groups.

(6) We write **Rng** for the category whose objects are rings and whose morphisms are ring homomorphisms and **CRng** for the full subcategory of **Rng** on *commutative* rings.

---

[1]$\mathbb{B}$ is a *subcategory* of $\mathbb{A}$ if $\mathsf{Ob}(\mathbb{B}) \subseteq \mathsf{Ob}(A)$, $\mathbb{B}(X, Y) \subseteq \mathbb{A}(X, Y)$ for all $X, Y \in \mathsf{Ob}(\mathbb{B})$ and composition and identities in $\mathbb{B}$ are inherited from $\mathbb{A}$ (by restriction). A subcategory $\mathbb{B}$ of $\mathbb{A}$ is called *full* if $\mathbb{B}(X, Y) = \mathbb{A}(X, Y)$ for all $X, Y \in \mathsf{Ob}(\mathbb{B})$.

(7) For a commutative ring $R$ we write $\mathbf{Mod}_R$ for the category of $R$-modules and their homomorphisms (if $R$ is a field $k$ then we write $\mathbf{Vect}_k$ instead of $\mathbf{Mod}_k$). The category of $R$-algebras and their homomorphisms is denoted as $\mathbf{Alg}_R$.

(8) We write $\mathbf{Sp}$ for the category of topological spaces and continuous maps.

(9) Identifying *homotopy equivalent* maps in $\mathbf{Sp}$ gives rise to the category $\mathbf{Sp}_h$.[2]

(10) Every monoid $M = (M, \cdot, 1)$ can be understood as a category with one object (usually denoted as $*$). Categories with one object are precisely the monoids.

(11) Every preorder $P = (P, \leq)$ (i.e. where $\leq$ is a reflexive and transitive binary relation on $P$) can be considered as a category whose objects are the elements of $P$ and whose morphisms from $x$ to $y$ are given by the set $\{* \mid x \leq y\}$. Categories arising this way are those categories $\mathbb{C}$ where $\mathbb{C}(X, Y)$ contains at most one element for all $X, Y \in \mathsf{Ob}(\mathbb{C})$ and they are called *posetal*. $\diamond$

When "inverting the direction of arrows" in a given category this gives rise to the so-called "dual" or "opposite" category $\mathbb{C}^{\mathsf{op}}$ which in general is quite different from $\mathbb{C}$.

**Definition 1.2** *Let $\mathbb{C}$ be a category. Then its* dual *or* opposite category $\mathbb{C}^{\mathsf{op}}$ *is given by* $\mathsf{Ob}(\mathbb{C}^{\mathsf{op}}) = \mathsf{Ob}(\mathbb{C})$, $\mathbb{C}^{\mathsf{op}}(A, B) = \mathbb{C}(B, A)$ *and* $\circ_{A,B,C}^{\mathbb{C}^{\mathsf{op}}}(g, f) = \circ_{C,B,A}^{\mathbb{C}}(f, g)$. $\diamond$

Obviously, for every object $A$ the morphism $\mathsf{id}_A$ is the identity morphism for $A$ also in $\mathbb{C}^{\mathsf{op}}$.

Next we consider some properties of morphisms generalising the notions injective, surjective and bijective known from $\mathbf{Set}$ to arbitrary categories.

---

[2]Continuous maps $f_0, f_1 : X \to Y$ are called *homotopy equivalent* (notation $f_0 \sim f_1$) iff there is a continuous map $f : [0, 1] \times X \to Y$ with $f_i(x) = f(i, x)$ for all $x \in X$ and $i \in \{0, 1\}$. One easily checks that $f_0 \sim f_1$ and $g_0 \sim g_1$ implies $g_0 \circ f_0 \sim g_1 \circ f_1$ (whenever the composition is defined), i.e. composition respects homotopy equivalence. This explains why identifying homotopy equivalent continuous maps gives rise to a category.

**Definition 1.3** *Let $\mathbb{C}$ be a category and $f : A \to B$ be a morphism in $\mathbb{C}$.*
*The morphism $f$ is called a* monomorphism *or* monic *iff for all $g, h : C \to A$ from $f \circ g = f \circ h$ it follows that $g = h$.*
*The morphism $f$ is called an* epimorphism *or* epic *iff for all $g, h : B \to C$ from $g \circ f = h \circ f$ it follows that $g = h$.*
*The morphism $f$ is called an* isomorphism *iff there exists a morphism $g : B \to A$ with $g \circ f = \text{id}_A$ and $f \circ g = \text{id}_B$. Such a morphism $g$ is unique provided it exists in which case it is denoted as $f^{-1}$.* $\diamond$

It is easy to show that an isomorphism is both monic and epic (Exercise!). However, the converse is not true in general: consider for example the inclusion $\mathbb{Z} \hookrightarrow \mathbb{Q}$ which is a morphism in **Rng** (and, of course, also **CRng**) which is epic and monic, but obviously not an isomorphism.(Exercise!) Categories where monic and epic implies isomorphism are called *balanced*.

**Exercises**

1. Show that in **Set** and **Ab** a morphism $f : A \to B$ is monic iff $f$ is one-to-one, is epic iff $f$ is onto, is an isomorphism iff $f$ is bijective.

2. Do the above equivalences hold also in **Sp**?

3. Find a category which is not posetal but where all morphisms are monic.

# 2 Functors and Natural Transformations

This section is devoted to the concept of a functor, i.e. structure preserving map between categories, inspired by and generalising both monoid homomorphism and monotonic maps.

**Definition 2.1** *Let $\mathbb{A}$ and $\mathbb{B}$ be categories. A (covariant) functor $F$ from $\mathbb{A}$ to $\mathbb{B}$ (notation $F : \mathbb{A} \to \mathbb{B}$) is given by a function*

$$F_{\mathsf{Ob}} : \mathsf{Ob}(\mathbb{A}) \to \mathsf{Ob}(\mathbb{B})$$

*(called "object part" of $F$) together with a family of functions*

$$F_{\mathsf{Mor}} = \big(F_{A,B} : \mathbb{A}(A, B) \to \mathbb{B}(F(A), F(B))\big)_{A,B \in \mathsf{Ob}(\mathbb{A})}$$

*(called "morphism part" of $F$) satisfying the following requirements*

(1) *$F_{A,A}(\mathsf{id}_A) = \mathsf{id}_{F_{\mathsf{Ob}}(A)}$ for all $A \in \mathsf{Ob}(\mathbb{A})$*

(2) *$F_{A,C}(g \circ f) = F_{B,C}(g) \circ F_{A,B}(f)$ for all $f : A \to B$ and $g : B \to C$.*

*A* contravariant functor from $\mathbb{A}$ to $\mathbb{B}$ *is a covariant functor from $\mathbb{A}^{\mathsf{op}}$ to $\mathbb{B}$.* ◇

Most of the time we suppress the indices of $F$ as they could be reconstructed from the context without any pain. Under this convention for example conditions (1) and (2) of Definition 2.1 are formulated as

$$F(\mathsf{id}_A) = \mathsf{id}_{F(A)} \qquad F(g \circ f) = F(g) \circ F(f)$$

which certainly is more readable.
Very often one needs functors with more than one argument. These are subsumed by Definition 2.1 taking $\mathbb{A}$ as $\mathbb{A}_1 \times \cdots \times \mathbb{A}_n$, i.e. a (cartesian) product of categories, whose straightforward definition we leave to the reader as an exercise(!).
Actually, functors are ubiquitous in mathematics as can be seen from the following

**Examples of Functors**

(1) Functors between categories with precisely one object correspond to monoid homomorphisms.

(2) Functors between posetal categories correspond to monotone functions between preorders.

(3) If $\mathbb{C}$ is a group or a monoid then functors from $\mathbb{C}$ to $\mathbf{Vect}_k$, the category of vector spaces over field $k$, are called *linear representations* of $\mathbb{C}$.

(4) The powerset function $A \mapsto \mathcal{P}(A)$ is the object part of two different functors from $\mathbf{Set}$ to $\mathbf{Set}$, a covariant one and a contravariant one: let $f : A \to B$ be a morphism in $\mathbf{Set}$ then the covariant power set functor sends $X \in \mathcal{P}(A)$ to $f[X] = \{f(x) \mid x \in X\}$ and the contravariant power set functor sends $Y \in \mathcal{P}(B)$ to $f^{-1}[Y] = \{x \in A \mid f(x) \in Y\}$.

(5) For every category $\mathbb{C}$ we may consider its "hom-functor" $\mathsf{Hom}_{\mathbb{C}} : \mathbb{C}^{\mathsf{op}} \times \mathbb{C} \to \mathbf{Set}$ whose object part is given by

$$\mathsf{Hom}_{\mathbb{C}}(A, B) = \mathbb{C}(A, B)$$

and whose morphism part is given by

$$\mathsf{Hom}_{\mathbb{C}}(f, g) : \mathsf{Hom}_{\mathbb{C}}(A, B) \to \mathsf{Hom}_{\mathbb{C}}(A', B') : h \mapsto g \circ h \circ f$$

for $f : A' \to A$ and $g : B \to B'$. Notice that the hom-functor is contravariant in its first argument and covariant in its second argument. Fixing the first or second argument of $\mathsf{Hom}_{\mathbb{C}}$ gives rise to the functors

$$\mathsf{Y}^*_{\mathbb{C}}(A) = \mathsf{Hom}_{\mathbb{C}}(A, -) : \mathbb{C} \to \mathbf{Set}$$

and

$$\mathsf{Y}_{\mathbb{C}}(B) = \mathsf{Hom}_{\mathbb{C}}(-, B) : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$$

where $\mathsf{Y}$ refers to the mathematician NOBUO YONEDA who first considered these functors.

(6) Let $F : \mathbf{Set} \to \mathbf{Mon}$ be the functor sending set $A$ to $A^*$, the free monoid over $A$ and $f : A \to B$ to the monoid homomorphism $F(f)(a_1 \ldots a_n) = f(a_1) \ldots f(a_n)$. There is also a functor $U : \mathbf{Mon} \to \mathbf{Set}$ in the reverse direction "forgetting" the monoid structure. Such "forgetful" functors $U$ can be defined not only for $\mathbf{Mon}$ but for arbitrary categories of (algebraic) structures. The associated "left adjoint" functor $F$ doesn't always exist. What "adjoint" really means will be clarified later and turns out as a central notion of category theory.

Notice that in the above definition of the hom-functor $\mathsf{Hom}_{\mathbb{C}}$ we have cheated a bit in assuming that $\mathsf{Hom}_{\mathbb{C}}(A, B) = \mathbb{C}(A, B)$ actually is a set for all objects $A$ and $B$ in $\mathbb{C}$. However, this assumption is valid for most categories one meets in practice and they are called *locally small*. A category $\mathbb{C}$ is called *small* if not only all $\mathbb{C}(A, B)$ are sets but also $\mathsf{Ob}(\mathbb{C})$ itself is a set. Typically, categories of structures like **Mon**, **Grp**, **Sp** and also **Set** are locally small, but not small.

Next we will show how to organise the collection of functors from $\mathbb{A}$ to $\mathbb{B}$ into a category itself by defining an appropriate notion of morphism between functors which is called *natural transformation*.

**Definition 2.2** *Let $\mathbb{A}$ and $\mathbb{B}$ be categories and $F$ and $G$ be functors from $\mathbb{A}$ to $\mathbb{B}$. A* natural transformation *from $F$ to $G$ is a family of morphisms*

$$\tau = (F(A) \to G(A))_{A \in \mathsf{Ob}(\mathbb{A})}$$

*such that*

$$
\begin{array}{ccc}
F(A) & \xrightarrow{\ \tau_A\ } & G(A) \\
{\scriptstyle F(f)} \downarrow & & \downarrow {\scriptstyle G(f)} \\
F(A') & \xrightarrow[\ \tau_{A'}\ ]{} & G(A')
\end{array}
$$

*commutes for all $f : A \to A'$ in $\mathbb{A}$. We write $\tau : F \dot\to G$ or $\tau : F \Rightarrow G$ if $\tau$ is a natural transformation from $F$ to $G$.*
*For natural transformations $\sigma : F \dot\to G$ and $\tau : G \dot\to H$ their (vertical) composition $\tau \circ \sigma$ is defined pointwise as*

$$(\tau \circ \sigma)_A = \tau_A \circ \sigma_A$$

*for $A \in \mathsf{Ob}(\mathbb{A})$. The identity natural transformation $\mathsf{id}_F : F \dot\to F$ on $F$ is given by $(\mathsf{id}_F)_A = \mathsf{id}_{F(A)}$.* $\diamondsuit$

Obviously, functors from $\mathbb{A}$ to $\mathbb{B}$ with natural transformations as morphisms organize into a category denoted by $\mathbb{B}^{\mathbb{A}}$ or $\mathsf{Func}(\mathbb{A}, \mathbb{B})$.

Actually, natural transformations are ubiquitous in mathematics as can be seen from the following

**Examples of Natural Transformations**

(1) Let $\mathcal{P} : \mathbf{Set} \to \mathbf{Set}$ be the covariant powerset functor on $\mathbf{Set}$. Then a natural transformation $\mathsf{Id} \dot{\to} \mathcal{P}$ is given by $\sigma_A : A \to \mathcal{P}(A) : a \mapsto \{a\}$ and a natural transformation from $\mathcal{P}^2 = \mathcal{P} \circ \mathcal{P}$ to $\mathcal{P}$ is given by $\bigcup_A : \mathcal{P}^2(A) \to \mathcal{P}(A) : \mathcal{X} \mapsto \bigcup \mathcal{X}$ where, as usual, $\bigcup \mathcal{X}$ is defined as $\{x \in A \mid \exists X \in \mathcal{X}.\ x \in X\}$.

(2) Let $I$ be a set. Define $F : \mathbf{Set} \to \mathbf{Set}$ as $F(X) = X^I \times I$ and $F(f)(g, i) = \langle f \circ g, i \rangle$. Then a natural transformation $\varepsilon : F \dot{\to} \mathsf{Id}$ is given by $\varepsilon_A : A^I \times I \to A : \langle g, i \rangle \mapsto g(i)$.

(3) Let $\mathrm{GL}_n(R)$ be the group of invertible $n \times n$-matrices with entries in $R$ where $R$ is a field or a commutative ring. For a ring homomorphism $h : R \to R'$ let $\mathrm{GL}_n(h) : \mathrm{GL}_n(R) \to \mathrm{GL}_n(R')$ be the group homomorphism sending $A = (a_{ij})$ to $\mathrm{GL}_n(h)(A) = (h(a_{ij}))$. Let $\mathsf{Inv}$ be the functor from $\mathbf{CRng}$ to $\mathbf{Grp}$ sending a commutative ring to the abelian group of its units. Then a natural transformation $\det : \mathrm{GL}_n \dot{\to} \mathsf{Inv}$ is given by sending each matrix to its determinant as indicated in the diagram

$$
\begin{array}{ccccc}
\mathrm{GL}_n(R) & \xrightarrow{\ \det_R\ } & \mathsf{Inv}(R) & \lhook\joinrel\longrightarrow & R \\
{\scriptstyle \mathrm{GL}_n(h)} \downarrow & & {\scriptstyle \mathsf{Inv}(h)} \downarrow & & \downarrow {\scriptstyle h} \\
\mathrm{GL}_n(R') & \xrightarrow[\ \det_{R'}\ ]{} & \mathsf{Inv}(R') & \lhook\joinrel\longrightarrow & R'
\end{array}
$$

(4) Let $R_1$ and $R_2$ be functors from a group $G$ to $\mathbf{Vect}_k$, i.e. representations of $G$. Then a natural transformation from $R_1$ to $R_2$ is simply a linear map $h : R_1(*) \to R_2(*)$ such that

$$
\begin{array}{ccc}
R_1(*) & \xrightarrow{\ h\ } & R_2(*) \\
{\scriptstyle R_1(f)} \downarrow & & \downarrow {\scriptstyle R_2(f)} \\
R_1(*) & \xrightarrow[\ h\ ]{} & R_2(*)
\end{array}
$$

for all $f \in G$. In representation theory such an $h$ is commonly called an *equivariant* map from $R_1$ to $R_2$.

If $\sigma : F\dot{\to}F' : \mathbb{A} \to \mathbb{B}$ and $\tau : G\dot{\to}G' : \mathbb{B} \to \mathbb{C}$ then their *horizontal* composition $\tau * \sigma : GF\dot{\to}G'F'$ is defined as

$$(\tau * \sigma)_A = G'\sigma_A \circ \tau_{FA} = \tau_{F'A} \circ G\sigma_A$$

where the second equality follows from $\tau : G\dot{\to}G'$. We leave it as an exercise(!) to show that

$$(\tau_2 \circ \tau_1) * (\sigma_2 \circ \sigma_1) = (\tau_2 * \sigma_2) \circ (\tau_1 * \sigma_1)$$

for $\sigma_1 : F\dot{\to}F'$ and $\sigma_2 : F'\dot{\to}F''$ in $\mathsf{Func}(\mathbb{A}, \mathbb{B})$ and $\tau_1 : G\dot{\to}G'$ and $\tau_2 : G'\dot{\to}G''$ in $\mathsf{Func}(\mathbb{B}, \mathbb{C})$. Moreover, one easily sees that $\mathsf{id}_G * \mathsf{id}_F = \mathsf{id}_{GF}$. Thus, horizontal composition gives rise to a functor $* : \mathsf{Func}(\mathbb{B}, \mathbb{C}) \times \mathsf{Func}(\mathbb{A}, \mathbb{B}) \to \mathsf{Func}(\mathbb{A}, \mathbb{C})$ for all categories $\mathbb{A}, \mathbb{B}, \mathbb{C}$. A category whose hom-sets themselves carry the structure of a category such that composition is functorial are called *2-categories*. The category $\mathbf{Cat}$ of categories and functors is a 2-category where $\mathbf{Cat}(\mathbb{A}, \mathbb{B})$ is the functor category $\mathsf{Func}(\mathbb{A}, \mathbb{B})$ and the morphism parts of the composition functors are given by horizontal composition $*$ which is functorial as we have seen above.

One readily checks that a natural transformation $\tau : F\dot{\to}G$ is an isomorphism iff all its components are isomorphisms (exercise!). We say that functors $F$ and $G$ from $\mathbb{A}$ to $\mathbb{B}$ are isomorphic iff there is an isomorphism from $F$ to $G$ in $\mathbf{Cat}(\mathbb{A}, \mathbb{B}) = \mathsf{Func}(\mathbb{A}, \mathbb{B})$.

# 3 Subcategories, Full and Faithful Functors, Equivalences

Though already shortly mentioned in section 1 we now "officially" give the definition of the notion of *subcategory*.

**Definition 3.1** *Let $\mathbb{C}$ be a category. A* subcategory $\mathbb{C}'$ *of $\mathbb{C}$ is given by*

$\mathsf{Ob}(\mathbb{C}') \subseteq \mathsf{Ob}(\mathbb{C})$ *and*

$\mathbb{C}'(A, B) \subseteq \mathbb{C}(A, B)$ *for all $A, B \in \mathsf{Ob}(\mathbb{C}')$*

*such that*

    i) $\mathsf{id}_A \in \mathbb{C}'(A, A)$ *for all $A \in \mathsf{Ob}(\mathbb{C}')$*

    ii) *whenever $f : A \to B$ and $g : B \to C$ are morphisms in $\mathbb{C}'$ then $g \circ f$ is a morphism in $\mathbb{C}'$*

*i.e. identities and composition are inherited from $\mathbb{C}$.*
*A subcategory $\mathbb{C}'$ of $\mathbb{C}$ is called* full *iff $\mathbb{C}'(A, B) = \mathbb{C}(A, B)$ for all $A, B \in \mathsf{Ob}(\mathbb{C}')$ and it is called* replete *iff, moreover, for every isomorphism $f : A \to B$ in $\mathbb{C}$ the object $B$ is in $\mathbb{C}'$ whenever $A$ is in $\mathbb{C}'$.* $\diamondsuit$

For every subcategory $\mathbb{C}'$ of $\mathbb{C}$ there is an obvious *inclusion functor*

$$\mathcal{I} : \mathbb{C}' \to \mathbb{C}$$

with $\mathcal{I}(A) = A$ and $\mathcal{I}(f) = f$ for all objects $A$ and morphisms $f$ in $\mathbb{C}'$.

**Definition 3.2** *Let $F : \mathbb{A} \to \mathbb{B}$ be a functor. The functor $F$ is called* faithful *iff for all $f, g : A \to B$ in $\mathbb{A}$ from $F(f) = F(g)$ it follows that $f = g$, i.e. iff all morphism parts of $F$ are one-to-one. The functor $F$ is called* full *iff for all $A, B \in \mathsf{Ob}(\mathbb{A})$ and $g : F(A) \to F(B)$ there is an $f : A \to B$ with $g = F(f)$, i.e. all morphism parts of $F$ are onto.* $\diamondsuit$

Obviously, a functor is full and faithful iff all its morphism parts are bijections.

**Lemma 3.1** *If a functor $F : \mathbb{A} \to \mathbb{B}$ is full and faithful then $F$ reflects isomorphisms, i.e. $f$ is an isomorphism whenever $F(f)$ is an isomorphism. However, in general faithful functors need not reflect isomorphisms.*

*Proof:* Simple exercise(!) left to the reader. $\qquad\qquad\qquad\qquad\square$

Notice that, in particular, for subcategories $\mathbb{C}'$ of $\mathbb{C}$ the inclusion functor $\mathcal{I} : \mathbb{C}' \to \mathbb{C}$ need not reflect isomorphisms though it always is faithful.[3]

**Definition 3.3** *A functor $F : \mathbb{A} \to \mathbb{B}$ is called an* equivalence (of categories) *iff $F$ is full and faithful and for every $B \in \mathsf{Ob}(\mathbb{B})$ there is an $A \in \mathbb{A}$ with $B \cong F(A)$. We write $\mathbb{A} \simeq \mathbb{B}$ iff there is an equivalence between $\mathbb{A}$ and $\mathbb{B}$.* $\Diamond$

Notice that being an equivalence is much weaker than being an isomorphism of categories. Nevertheless equivalence is the better notion because for categories isomorphism is the right notion of equality for objects.
We leave it as an exercise(!) to show that (using the Axiom of Choice) for every equivalence $F : \mathbb{A} \to \mathbb{B}$ there is a functor $G : \mathbb{B} \to \mathbb{A}$ such that $GF \cong \mathsf{Id}_{\mathbb{A}}$ and $FG \cong \mathsf{Id}_{\mathbb{B}}$, i.e. $F$ has a "quasi-inverse" $G$.

---

[3]For this reason P. Freyd in his book *Categories, Allegories* [FS] required faithful functors also to reflect isomorphisms! Accordingly, he required for subcategories $\mathbb{C}' \subseteq \mathbb{C}$ also that $f^{-1}$ is in $\mathbb{C}'$ whenever $f$ is in $\mathbb{C}'$.

# 4   Comma Categories and Slice Categories

At first look the notions introduced in this section may appear as somewhat artificial. However, as we shall see later they turn out as most important for *categorical logic*.

**Definition 4.1** *Let* $F : \mathbb{A} \to \mathbb{C}$ *and* $G : \mathbb{B} \to \mathbb{C}$ *be functors. The* comma category $F{\downarrow}G$ *is defined as follows. The objects of* $F{\downarrow}G$ *are triples* $(A, f, B)$ *where* $A \in \mathsf{Ob}(\mathbb{A})$, $B \in \mathsf{Ob}(\mathbb{B})$ *and* $f \in \mathbb{C}(F(A), G(B))$. *A morphism from* $(A, f, B)$ *to* $(A', f', B')$ *in* $F{\downarrow}G$ *is a pair* $(g : A{\to}A', h : B{\to}B')$ *such that*

$$
\begin{array}{ccc}
F(A) & \xrightarrow{\;F(g)\;} & F(A') \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f'} \\
G(B) & \xrightarrow[\;G(h)\;]{} & G(B')
\end{array}
$$

*commutes. Composition of morphisms in* $F{\downarrow}G$ *is componentwise, i.e.* $(g', h') \circ (g, h) = (g' \circ g, h' \circ h)$ *and* $\mathsf{id}_{(A,f,B)} = (\mathsf{id}_A, \mathsf{id}_B)$. $\diamond$

Special cases of particular interest arise if $\mathbb{A}$ or $\mathbb{B}$ coincide with $\mathbf{1}$, the trivial category with one object and one morphism, or if $F$ or $G$ are identity functors. If $G = \mathsf{Id}_{\mathbb{C}}$ we write $F{\downarrow}\mathbb{C}$ for $F{\downarrow}\mathsf{Id}_{\mathbb{C}}$, if $F = \mathsf{Id}_{\mathbb{C}}$ we write $\mathbb{C}{\downarrow}G$ for $\mathsf{Id}_{\mathbb{C}}{\downarrow}G$ and if $F = \mathsf{Id}_{\mathbb{C}} = G$ then we write $\mathbb{C}{\downarrow}\mathbb{C}$ for $\mathsf{Id}_{\mathbb{C}}{\downarrow}\mathsf{Id}_{\mathbb{C}}$. If $\mathbb{B} = \mathbf{1}$ and $G(*) = X$ we write $F{\downarrow}X$ for $F{\downarrow}G$ and if $\mathbb{A} = \mathbf{1}$ and $F(*) = Y$ we write $Y{\downarrow}G$ for $F{\downarrow}G$. If $F = \mathsf{Id}_{\mathbb{C}}$ and $G : \mathbf{1} \to \mathbb{C} : * \mapsto X$ we write $\mathbb{C}/X$ for $F{\downarrow}G$ and if $F : \mathbf{1} \to \mathbb{C} : * \mapsto Y$ and $G = \mathsf{Id}_{\mathbb{C}}$ we write $Y/\mathbb{C}$ for $F{\downarrow}G$. The comma category $\mathbb{C}/X$ is commonly called *slice (category of* $\mathbb{C}$*) over* $X$.
We leave it as an exercise[4] to show that for all sets $I$ we have $\mathbf{Set}/I \simeq \mathbf{Set}^I$ where $\mathbf{Set}^I$ is the $I$-fold product of $\mathbf{Set}$.
**Notice that for categories $\mathbb{C}$ different from Set it always makes sense to consider for $I \in \mathsf{Ob}(\mathbb{C})$ the slice category $\mathbb{C}/I$ whereas $\mathbb{C}^I$ is meaningless because $I$ is not a set in this case.**

---

[4]The idea is to send $A \in \mathbf{Set}^I$ to the map $\coprod_{i \in I} A_i \longrightarrow I : (i, a) \mapsto i$ where as usual $\coprod_{i \in I} A_i = \{(i, a) \mid i \in I, a \in A_i\}$.

# 5   Yoneda Lemma

The Yoneda lemma says that for every locally small category $\mathbb{C}$ there is a *full and faithful* functor $\mathsf{Y}_{\mathbb{C}} : \mathbb{C} \to \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$, the so-called *Yoneda functor*, which allows one to consider $\mathbb{C}$ (up to equivalence) as a full subcategory of $\widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$, the category of (set-valued) *presheaves* over $\mathbb{C}$. As we shall see later this result is very important because categories of the form $\widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ are very "set-like" in the sense that they can be considered as categories of "generalised sets" and, therefore, used as models for (constructive) logic and mathematics!

**Definition 5.1** (Yoneda embedding)
*Let $\mathbb{C}$ be a locally small category. For $I \in \mathsf{Ob}(\mathbb{C})$ the functor $\mathsf{Y}_{\mathbb{C}}(I) : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ is defined as $\mathbb{C}(-, I)$, i.e. sends $J \in \mathsf{Ob}(\mathbb{C})$ to the set $\mathbb{C}(J, I)$ and sends $g : K \to J$ to the function $\mathbb{C}(g, I) : \mathbb{C}(J, I) \to \mathbb{C}(K, I) : h \mapsto h \circ g$. For $f : I \to J$ let $\mathsf{Y}_{\mathbb{C}}(f)$ be the natural transformation from $\mathsf{Y}_{\mathbb{C}}(I)$ to $\mathsf{Y}_{\mathbb{C}}(J)$ whose component at $K \in \mathsf{Ob}(\mathbb{C})$ is given by $\mathsf{Y}_{\mathbb{C}}(f)_K : \mathbb{C}(K, I) \to \mathbb{C}(K, J) : h \mapsto f \circ h$. These data give rise to a functor $\mathsf{Y}_{\mathbb{C}} : \mathbb{C} \to \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ called the Yoneda embedding for $\mathbb{C}$.* $\diamond$

For Definition 5.1 to make sense one has to verify (1) that $\mathsf{Y}_{\mathbb{C}}(f)$ is actually a natural transformation, i.e. that

$$
\begin{array}{ccc}
\mathbb{C}(K, I) & \xrightarrow{\;\mathsf{Y}_{\mathbb{C}}(f)_K\;} & \mathbb{C}(K, J) \\
{\scriptstyle \mathbb{C}(g, I)}\Big\downarrow & & \Big\downarrow{\scriptstyle \mathbb{C}(g, J)} \\
\mathbb{C}(K', I) & \xrightarrow[\;\mathsf{Y}_{\mathbb{C}}(f)_{K'}\;]{} & \mathbb{C}(K', J)
\end{array}
$$

commutes for all $g : K' \to K$, and (2) that $\mathsf{Y}_{\mathbb{C}}$ is actually functorial, i.e. satisfies $\mathsf{Y}_{\mathbb{C}}(\mathsf{id}_I) = \mathsf{id}_{\mathsf{Y}_{\mathbb{C}}(I)}$ and $\mathsf{Y}_{\mathbb{C}}(g \circ f) = \mathsf{Y}_{\mathbb{C}}(g) \circ \mathsf{Y}_{\mathbb{C}}(f)$. We leave the straighforward verifications of (1) and (2) to the inclined reader as a simple exercise(!).

**Lemma 5.1** (Yoneda Lemma)
*Let $\mathbb{C}$ be a locally small category and $A : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$. Then for all $I \in \mathsf{Ob}(\mathbb{C})$*

(1)  *natural transformations $\sigma, \tau : \mathsf{Y}_{\mathbb{C}}(I) \to A$ are equal iff $\sigma_I(\mathsf{id}_I) = \tau_I(\mathsf{id}_I)$ and*

(2) *for every $a \in A(I)$ there exists a unique natural transformation $\tau^{(a)}$ :*
$\mathsf{Y}_\mathbb{C}(I) \to A$ *with* $\tau_I^{(a)}(\mathsf{id}_I) = a$.

*Accordingly, we have a natural isomorphism*

$$\iota : \widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(-), A) \xrightarrow{\cong} A$$

*given by* $\iota_I : \widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(I), A) \to A(I) : \tau \mapsto \tau_I(\mathsf{id}_I)$ *for* $I \in \mathbb{C}$.

*Proof:* First notice that for a natural transformation $\tau : \mathsf{Y}_\mathbb{C}(I) \to A$ it holds that

$$(\dagger) \qquad \tau_J(f) = A(f)(\tau_I(\mathsf{id}_I)) \quad \text{for all } f : J \to I$$

because by naturality of $\tau$ the diagram

$$
\begin{array}{ccc}
\mathbb{C}(I, I) & \xrightarrow{\tau_I} & A(I) \\
{\scriptstyle \mathsf{Y}_\mathbb{C}(I)(f)} \downarrow & & \downarrow {\scriptstyle A(f)} \\
\mathbb{C}(J, I) & \xrightarrow[\tau_J]{} & A(J)
\end{array}
$$

commutes and, therefore,

$$
\begin{aligned}
\tau_J(f) = \tau_J(\mathsf{Y}_\mathbb{C}(I)(f)(\mathsf{id}_I)) &= \\
&= (\tau_J \circ \mathsf{Y}_\mathbb{C}(I)(f))(\mathsf{id}_I) = (A(f) \circ \tau_I)(\mathsf{id}_I) = \\
&= A(f)(\tau_I(\mathsf{id}_I))
\end{aligned}
$$

as desired.

Now for showing claim (1) assume that $\tau_I(\mathsf{id}_I) = \sigma_I(\mathsf{id}_I)$. Thus, due to $(\dagger)$ we have $\tau_J(f) = A(f)(\tau_I(\mathsf{id}_I)) = A(f)(\sigma_I(\mathsf{id}_I)) = \sigma_J(f)$ for all $f : J \to I$ in $\mathbb{C}$, i.e. $\tau = \sigma$ as desired.

For claim (2) suppose $a \in A(I)$. The desired $\tau^{(a)} : \mathsf{Y}_\mathbb{C}(I) \to A$ is given by $\tau_J^{(a)}(f) = A(f)(a)$ for $f : J \to I$ in $\mathbb{C}$. Obviously, we have $\tau_I^{(a)}(\mathsf{id}_I) = A(\mathsf{id}_I)(a) = a$. That $\tau^{(a)}$ is indeed a natural transformation we leave as an exercise(!) to the reader.

It is now obvious that $\iota_I(\tau) = a$ iff $\tau = \tau^{(a)}$ from which it follows that $\iota_I$ is a bijection for all $I \in \mathbb{C}$. For naturality of $\iota$ suppose $f : J \to I$ in $\mathbb{C}$. We have to show that

$$
\begin{array}{ccc}
\widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(I), A) & \xrightarrow{\iota_I} & A(I) \\
{\scriptstyle \widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(f), A)} \downarrow & & \downarrow {\scriptstyle A(f)} \\
\widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(J), A) & \xrightarrow{\iota_J} & A(J)
\end{array}
$$

commutes. Suppose $\tau : \mathsf{Y}_\mathbb{C}(I) \to A$ then we have

$$
\begin{aligned}
(A(f) \circ \iota_I)(\tau) = A(f)(\iota_I(\tau)) = A(f)(\tau_I(\mathsf{id}_I)) &= (A(f) \circ \tau_I)(\mathsf{id}_I) = \\
= (\tau_J \circ \mathsf{Y}_\mathbb{C}(I)(f))(\mathsf{id}_I) &= \tau_J(f) = \tau_J(\mathsf{Y}_\mathbb{C}(f)_J(\mathsf{id}_J)) = \\
= (\tau \circ \mathsf{Y}_\mathbb{C}(f))_J(\mathsf{id}_J) &= \iota_J(\tau \circ \mathsf{Y}_\mathbb{C}(f)) = \\
&= \iota_J(\widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(f), A)(\tau)) = \\
= (\iota_J \circ \widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(f), A))(\tau)
\end{aligned}
$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

As the Yoneda Lemma 5.1 allows us to identify in a canonical way $\widehat{\mathbb{C}}(\mathsf{Y}_\mathbb{C}(I), A)$ with $A(I)$ for all presheaves $A \in \widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ we often write simply $a : \mathsf{Y}_\mathbb{C}(I) \to A$ for the natural transformation $\tau^{(a)}$ as constructed in the proof of the Yoneda lemma.

**Corollary 5.1** *For every locally small category $\mathbb{C}$ the Yoneda functor $\mathsf{Y}_\mathbb{C} : \mathbb{C} \to \widehat{\mathbb{C}}$ is full and faithful.*

*Proof:* We have to show that for all $I, J \in \mathsf{Ob}(\mathbb{C})$ and $\tau : \mathsf{Y}_\mathbb{C}(I) \to \mathsf{Y}_\mathbb{C}(J)$ there exists a unique morphism $f : I \to J$ in $\mathbb{C}$ with $\tau = \mathsf{Y}_\mathbb{C}(f)$. Notice that by the Yoneda Lemma 5.1 we have $\tau_K(g) = \mathsf{Y}_\mathbb{C}(J)(g)(\tau_I(\mathsf{id}_I)) = \tau_I(\mathsf{id}_I) \circ g$ for all $g : K \to I$. Thus, as for $f = \tau_I(\mathsf{id}_I)$ we have $\mathsf{Y}_\mathbb{C}(f)_K(g) = f \circ g$ it follows from the Yoneda Lemma 5.1 that $\tau = \mathsf{Y}_\mathbb{C}(f)$ and $f$ is unique with this property. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

As full and faithful functors reflect isomorphisms we get that $\mathsf{Y}_\mathbb{C}(I) \cong \mathsf{Y}_\mathbb{C}(J)$ implies $I \cong J$ and, therefore, the category $\mathbb{C}$ may be considered as a full subcategory of $\widehat{\mathbb{C}}$. Those $A : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ with $A \cong \mathsf{Y}_\mathbb{C}(I)$ for some $I \in \mathsf{Ob}(\mathbb{C})$ are called *representable*. We give this definition an "official" status as representability is a key concept used for defining most of the subsequent notions.

**Definition 5.2** *Let $\mathbb{C}$ be locally small. Then $A : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ is called* representable *iff $A \cong \mathsf{Y}_\mathbb{C}(I)$ for some $I \in \mathsf{Ob}(\mathbb{C})$.* $\qquad\qquad\Diamond$

Notice that by the Yoneda lemma for representable $A$ the representing object $I$ with $A \cong \mathsf{Y}_\mathbb{C}(I)$ is unique up to isomorphism.

# 6 Grothendieck universes : big *vs.* small

Generally in category theory one doesn't worry too much about its set-theoretic foundations because one thinks that categories (namely toposes) provide a better foundation for mathematics (see e.g. [LR]).

However, as we have seen in the previous section when discussing the Yoneda lemma *local smallness* is of great importance in category theory. From the definition of local smallness it is evident that we want to call a collection "small" iff it is a set. Typical non-small collections are the collection of all sets or the collection of all functors from $\mathbb{C}^{\mathsf{op}}$ to **Set**. Already in traditional algebra and topology to some extent one has to consider large collections when quantifying over *all* groups, *all* topological spaces etc. In this case the usual "excuse" is that in Zermelo-Fraenkel set theory ZF(C) one may quantify over the universe of all sets and, therefore, also over collections which can be expressed by a predicate in the language of set theory (i.e. a first order formula using only the binary predicates $=$ and $\in$). In Gödel-Bernays-vonNeumann class theory GBN one may even give an "ontological status" to such collections which usually are referred to as classes.[5] The drawback, however, is that classes don't enjoy good closure properties, e.g. there doesn't exist the class of all functions from class $X$ to class $Y$ etc. Thus, the desire to freely manipulate big collections has driven categorists to introduce the following notion of *universe*.

**Definition 6.1** *A Grothendieck universe is a set $U$ such that the following properties*

    (U1)    *if $x \in a \in U$ then $x \in U$*
    (U2)    *if $a, b \in U$ then $\{a, b\}$ and $a \times b$ are elements of $U$*
    (U3)    *if $a \in U$ then $\bigcup a$ and $\mathcal{P}(a)$ are elements of $U$*
    (U4)    *the set $\omega$ of all natural numbers is an element of $U$*
    (U5)    *if $f : a \to b$ is surjective with $a \in U$ and $b \subseteq U$ then $b \in U$*

*hold for $U$.* $\diamond$

Notice that the conditions (U1)–(U5) guarantee that $U$ with $\in$ restricted to

---

[5]For a first order formula $\phi(x)$ in the language of set theory one may form the class $\{x \mid \phi(x)\}$ consisting of all sets $a$ satisfying $\phi(a)$ (where a set is a class contained in some class (e.g. $\{x \mid x = x\}$) as an element).

$U \times U$ provides a model for ZF(C) (i.e. a "small inner model"[6] in the slang of set theory). Thus, the universe $U$ is closed under the usual set-theoretic operations.[7]

Notice, however, that ZF(C) cannot prove the existence of Grothendieck universes as otherwise ZF(C) could prove its own consistency in contradiction to Gödel's second incompleteness theorem. Nevertheless, one shouldn't be too afraid of inconsistencies when postulating Grothendieck universes as it is nothing but a *reflection principle* claiming that what we have axiomatized in ZF(C) "actually exists". In other words, if ZF(C) + the assumption of a Grothendieck universe were inconsistent then it were rather the fault of the former[8] than the fault of the latter.

If having accepted one Grothendieck universe why shouldn't one accept more? Accordingly, when really worrying about set-theoretic foundations for category theory one usually adds to ZF(C) the requirement that for every set $a$ there is a Grothendieck universe $U$ with $a \in U$. Accordingly, for every Grothendieck universe $U$ there is a Grothendieck universe $U'$ with $U \in U'$ and, therefore, also $U \subseteq U'$. Thus, this axiomatic setting guarantees (at least) the existence[9] of a hierarchy $U_0 \in U_1 \in \cdots \in U_n \in U_{n+1} \in \cdots$ of universes which is also *cumulative* in the sense that $U_0 \subseteq U_1 \subseteq \cdots \subseteq U_n \subseteq U_{n+1} \subseteq \cdots$ holds as well.

**In the rest of these notes we will not refer anymore to the set-theoretical sophistications discussed in this section but, instead, will use the notions "big" and "small" informally.**

---

[6]Actually, the notion of universe is somewhat stronger than that of a small inner model as the latter need not satisfy (U5) for arbirary functions $f$ but only for first order definable $f$ as ensured by the set theoretic replacement axiom.

[7]We suggest it as an exercise(!) to directly derive from conditions (U1)–(U5) that $\langle a, b \rangle = \{\{a\}, \{a, b\}\} \in U$ whenever $a, b \in U$.

[8]What makes ZF so incredibly strong is the mixture of infinity, powersets and the replacement axiom!

[9]Notice that in presence of the axiom of choice such a cumulative hierarchy can be constructed internally by iterating the choice function $u$ assigning to every set $a$ a Grothendieck universe $u(a)$ with $a \in u(a)$. Of course, nothing prevents us from iterating this function $u$ along arbitrary transfinite ordinals!

# 7 Limits and Colimits

A lot of important properties of categories can be formulated by requiring that *limits* or *colimits* of a certain kind do exist meaning that certain functors are representable.

Before introducing the general notion of limit and colimit for a diagram we study some particular instances of these notions which, however, together will guarantee that all (small) limits (or colimits) do exist.

**Definition 7.1** (cartesian products)
*Let $\mathbb{C}$ be a category. The (cartesian) product of a family $A = (A_i \mid i \in I)$ of objects in $\mathbb{C}$ is given by an object $P$ of $\mathbb{C}$ together with a family $\pi = (\pi_i : P \to A_i \mid i \in I)$ of morphisms in $\mathbb{C}$ such that for every object $C$ and every family $f = (f_i : C \to A_i \mid i \in I)$ there exists a* **unique** *morphism $g : C \to P$ with $f_i = \pi_i \circ g$ for all $i \in I$. This unique morphism $g$ will often be denoted as $\langle f_i \rangle_{i \in I}$ and is called* the mediating arrow *from $f$ to $\pi$. The object $P$ together with the family $\pi$ is called* product cone *over $A$ and $\pi_i : P \to A_i$ is called $i$-th projection.* $\diamond$

Next we will show that product cones are unique up to isomorphism.

**Lemma 7.1** *Suppose $\pi = (\pi_i : P \to A_i \mid i \in I)$ and $\pi' = (\pi'_i : P' \to A_i \mid i \in I)$ are product cones over $A = (A_i \mid i \in I)$. Then for the unique arrows $f : P \to P'$ and $g : P' \to P$ such that $\pi'_i \circ f = \pi_i$ and $\pi_i \circ g = \pi'_i$ for all $i \in I$ it holds that $g \circ f = \mathsf{id}_P$ and $f \circ g = \mathsf{id}_{P'}$, i.e. that the cones $\pi$ and $\pi'$ are canonically isomorphic.*

*Proof:* Due to our assumptions we have that for all $i \in I$ it holds that

$$\pi_i \circ g \circ f = \pi'_i \circ f = \pi_i$$

and, therefore, by uniqueness of mediating arrows $\mathsf{id}_P = g \circ f$ (because $\pi_i \circ \mathsf{id}_P = \pi_i$ for all $i \in I$). Similarly one shows that $f \circ g = \mathsf{id}_{P'}$. $\square$

Notice that for $I = \emptyset$ a product cone $(P, \pi)$ over $(A_i \mid i \in I)$ is already determined by the object $P$ (because $\pi$ has to be the empty family) which has to satisfy the condition that for every object $C \in \mathsf{Ob}(\mathbb{C})$ there is a unique morphism from $C$ to $P$. Such objects $P$ will be called *terminal* and are usually denoted as $T$ or $1$.

**Definition 7.2** (terminal objects)
*An object $T$ in $\mathbb{C}$ is called* terminal *iff for every $C \in \mathsf{Ob}(\mathbb{C})$ there is a unique morphism $C \to T$ in $\mathbb{C}$ (often denoted as $!_C$).* ◇

In **Set** a product cone for $A = (A_i \mid i \in I)$ is given by $P = \prod_{i \in I} A_i$ where

$$\prod_{i \in I} A_i = \{s : I \to \bigcup_{i \in I} A_i \mid \forall i \in I.\ s(i) \in A_i\}$$

and $\pi = (\pi_i \mid i \in I)$ with $\pi_i(s) = s(i)$.
Terminal objects in **Set** are those sets containing precisely one element. Accordingly, we often write 1 for a terminal object.
We recommend it as an exercise(!) to construct products in the categories **Mon**, **Sp** etc.

**Definition 7.3** (equalisers) *Let $f, g : A \to B$ in $\mathbb{C}$. An* equaliser *of $f$ and $g$ is a morphism $e : E \to A$ such that*

(i) *$f \circ e = g \circ e$ and*

(ii) *for every $h : C \to A$ with $f \circ h = g \circ h$ there exists a **unique** morphism $k : C \to E$ with $h = e \circ k$.* ◇

Requirement (ii) of the above definition is usually visualized by the following diagram



One easily checks that an equaliser $e$ necessarily has to be a monomorphism (exercise!). A mono which appears as equaliser of some parallel pair of morphisms will be called *regular*. In **Set** an equaliser of $f, g : A \to B$ is given by $E = \{x \in A \mid f(x) = g(x)\}$ and $e : E \hookrightarrow A : x \mapsto x$.

**Definition 7.4** (small/finitely complete)
*A category $\mathbb{C}$ has (small) limits iff $\mathbb{C}$ has products of (small) families of objects and $\mathbb{C}$ has equalisers.*
*A category $\mathbb{C}$ has finite limits iff $\mathbb{C}$ has finite products (i.e. products for finite families of objects) and $\mathbb{C}$ has equalisers.*
*A category $\mathbb{C}$ is small/finitely complete iff $\mathbb{C}$ has finite/small limits.* $\diamondsuit$

Obviously, a category $\mathbb{C}$ has finite limits iff $\mathbb{C}$ has a terminal object, binary products and equalisers (exercise!).
Next we consider the most important case of *pullbacks* which exist in all categories with finite limits. Moreover, we will see that categories with a terminal object and all pullbacks will also have all finite limits.

**Definition 7.5** (pullbacks)
*A pullback in a category $\mathbb{C}$ is a commuting square*

$$
\begin{array}{ccc}
P & \xrightarrow{\ p_2\ } & A_2 \\
{\scriptstyle p_1}\big\downarrow & & \big\downarrow{\scriptstyle f_2} \\
A_1 & \xrightarrow[\ f_1\ ]{} & I
\end{array}
$$

*in $\mathbb{C}$ such that for all $g_i : B \to A_i$ ($i{=}1,2$) with $f_1 \circ g_1 = f_2 \circ g_2$ there exists a unique $h : B \to P$ with $p_i \circ h = g_i$ for $i{=}1,2$ as in the diagram*

*To indicate that a square is a pullback we write*

$$
\begin{array}{ccc}
P & \xrightarrow{\ p_2\ } & A_2 \\
{\scriptstyle p_1}\downarrow & \lrcorner & \downarrow{\scriptstyle f_2} \\
A_1 & \xrightarrow[\ f_1\ ]{} & I
\end{array}
$$

*The pair $p_1, p_2$ is called* pullback cone over $f_1, f_2$. $\qquad\qquad\qquad \diamondsuit$

Next we give a useful characterisation of having finite limits.

**Theorem 7.1** *Let $\mathbb{C}$ be a category with a terminal object $1$. Then the following two conditions are equivalent*

(1) *$\mathbb{C}$ has pullbacks*

(2) *$\mathbb{C}$ has binary products and equalisers.*

*Proof:* $(1) \Rightarrow (2)$ : Suppose $\mathbb{C}$ has pullbacks.
First we show that $\mathbb{C}$ has binary products. For objects $A_1, A_2$ in $\mathbb{C}$ consider their pullback

$$
\begin{array}{ccc}
A_1 \times A_2 & \xrightarrow{\ \pi_2\ } & A_2 \\
{\scriptstyle \pi_1}\downarrow & \lrcorner & \downarrow \\
A_1 & \xrightarrow{\qquad} & 1
\end{array}
$$

then $\pi_1, \pi_2$ is a product cone over $A_1, A_2$ because for morphisms $g_1 : B \to A_1$ and $g_2 : B \to A_2$ it holds that $!_{A_1} \circ g_1 = !_{A_2} \circ g_2$ and, therefore, due to the definition of pullback there exists a unique $h : B \to A_1 \times A_2$ with $\pi_i \circ h = g_i$ for $i=1,2$ as required for a product cone. As $\mathbb{C}$ has binary products and a terminal object it has all finite products.
For showing that $\mathbb{C}$ has equalisers suppose that $f_1, f_2 : A \to B$. Then the equaliser of $f_1$ and $f_2$ is constructed via the following pullback

$$
\begin{array}{ccc}
E & \xrightarrow{\ f'\ } & B \\
{\scriptstyle e}\downarrow & \lrcorner & \downarrow{\scriptstyle \delta_B} \\
A & \xrightarrow[\ \langle f_1, f_2 \rangle\ ]{} & B \times B
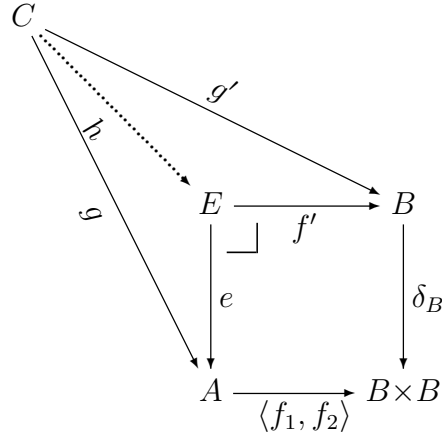\end{array}
$$

where $\delta_B = \langle \mathsf{id}_B, \mathsf{id}_B \rangle$. The morphism $e$ equalises $f_1$ and $f_2$ because we have

$$f_i \circ e = \pi_i \circ \langle f_1, f_2 \rangle \circ e = \pi_i \circ \delta_B \circ f' = \mathsf{id}_B \circ f' = f'$$

for $i{=}1, 2$ and, accordingly, that $f_1 \circ e = f_2 \circ e$. Now suppose $g : C \to A$ with $f_1 \circ g = f_2 \circ g$ for which morphism we write $g'$. As for $i{=}1, 2$ we have

$$\pi_i \circ \delta_B \circ g' = g' = f_i \circ g = \pi_i \circ \langle f_1, f_2 \rangle \circ g$$

and mediating arrows to product cones are unique we have that $\delta_B \circ g' = \langle f_1, f_2 \rangle \circ g$. Thus, there exists a unique arrow $h : C \to E$ with $e \circ h = g$ and $f' \circ h = g'$ as in



Now suppose that $h' : C \to E$ is a morphism with $e \circ h' = g$. Then we have also $f' \circ h' = g'$ because

$$f' \circ h' = \pi_1 \circ \delta_B \circ f' \circ h' = \pi_1 \circ \langle f_1, f_2 \rangle \circ e \circ h' =$$
$$= \pi_1 \circ \langle f_1, f_2 \rangle \circ e \circ h = \pi_1 \circ \delta_B \circ f' \circ h =$$
$$= f' \circ h$$

Thus, by uniqueness of mediating arrows to pullback cones it follows that $h' = h$. Thus, we have proved uniqueness of $h$ and $e$ is actually an equaliser of $f_1$ and $f_2$.

$(2) \Rightarrow (1)$ : Suppose $f_1 : A_1 \to I$ and $f_2 : A_2 \to I$. Let $\pi_1 : A_1 \times A_2 \to A_1$ and $\pi_2 : A_1 \times A_2 \to A_2$ be a product cone and $e : P \to A_1 \times A_2$ be an equaliser of $f_1 \circ \pi_1$ and $f_2 \circ \pi_2$. We define $p_i = \pi_i \circ e$ for $i{=}1, 2$ and show that it is a

pullback cone over $f_1, f_2$. First of all the diagram

$$\begin{array}{ccc} P & \xrightarrow{p_2} & A_2 \\ {\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle f_2} \\ A_1 & \xrightarrow[f_1]{} & I \end{array}$$

commutes as we have

$$f_1 \circ p_1 = f_1 \circ \pi_1 \circ e = f_2 \circ \pi_2 \circ e = f_2 \circ p_2$$

Now suppose $g_i : B \to A_i$ for $i{=}1, 2$ with $f_1{\circ}g_1 = f_2{\circ}g_2$. Then for $g = \langle g_1, g_2 \rangle$ we have

$$f_1 \circ \pi_1 \circ g = f_1 \circ g_1 = f_2 \circ g_2 = f_2 \circ \pi_2 \circ g$$

and, therefore, there exists a unique morphism $h : B \to P$ with $e \circ h = g$. Thus, we have $p_i \circ h = \pi_i \circ e \circ h = \pi_i \circ g = g_i$ for $i{=}1, 2$. The morphism $h$ is unique with this property as if $p_i \circ h' = g_i$ for $i{=}1, 2$ then $\pi_i \circ e \circ h' = g_i$ for $i{=}1, 2$ and, therefore, we have $e \circ h' = \langle g_1, g_2 \rangle = g$ from which it follows that $h = h'$. Thus, we have

$$\begin{array}{ccc} P & \xrightarrow{p_2} & A_2 \\ {\scriptstyle p_1}\downarrow \;\lrcorner & & \downarrow{\scriptstyle f_2} \\ A_1 & \xrightarrow[f_1]{} & I \end{array}$$

as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\;\;\square$

By inspection of the proof of Theorem 7.1 we get the following concrete construction of pullbacks in **Set**, namely

$$P = \{(x_1, x_2) \in A_1 {\times} A_2 \mid f_1(x_1) = f_2(x_2)\}$$

and $p_i(x_1, x_2) = x_i$ for $i{=}1, 2$.
From a logical point of view the pullback construction is important as it amounts to *reindexing a family of sets along a function*. Suppose $f : J \to I$ in **Set** and $g = \pi_1 : \coprod_{i \in I} A_i \to I$ for a family $(A_i \mid i \in I)$ of sets. Then we

have

$$\coprod_{j \in J} A_{f(j)} \xrightarrow{\ q\ } \coprod_{i \in I} A_i$$

$$\pi_1 \downarrow \qquad\qquad \downarrow \pi_1$$

$$J \xrightarrow{\ f\ } I$$

where $q(j,x) = (f(j),x)$ as follows from the obvious 1-1-correspondence between pairs $(j,x) \in \coprod_{j \in J} A_{f(j)}$ and tuples $(j,(i,x))$ with $i = f(j)$ and $x \in A_i$. Having seen how to compute pullbacks in **Set** we recommend it as an exercise(!) to construct pullbacks in the categories **Mon**, **Grp**, **Sp** etc.

By *dualisation* for every kind of limit there is a dual notion of *colimit* as summarised in the following table

| $\mathbb{C}$ | $\mathbb{C}^{\mathsf{op}}$ |
|---|---|
| sum | product |
| coequaliser | equaliser |
| pushout | pullback |
| initial object | terminal object |

We recommend it as an exercise(!) to show that **Set** has the above mentioned colimits. What about **Mon**, **Grp**, **Sp** etc.?

Because of their great importance for reasoning about pullbacks we prove the following lemmata.

**Lemma 7.2** *If $m : A' \to A$ is a monomorphism in category $\mathbb{C}$ then for every pullback diagram in $\mathbb{C}$*

$$B' \xrightarrow{\ f'\ } A'$$

$$n \downarrow \qquad\qquad \downarrow m$$

$$B \xrightarrow{\ f\ } A$$

*the arrow $n$ is a monomorphism, too.*

*Proof:* Suppose $g_1, g_2 : C \to B'$ with $n \circ g_1 = n \circ g_2$. Then we also have $f' \circ g_1 = f' \circ g_2$ because $m \circ f' \circ g_1 = f \circ n \circ g_1 = f \circ n \circ g_2 = m \circ f' \circ g_2$ and $m$ is a mono by assumption. Thus, it follows that $g_1 = g_2$ by uniqueness of mediating arrows. $\square$

**Lemma 7.3** *Whenever the following diagram*



*commutes in $\mathbb{C}$ and the right square is a pullback then the left square is a pullback if and only if the whole outer rectangle is a pullback.*

*Proof:* Suppose that the left square is a pullback. For showing that the outer rectangle is also a pullback suppose that $f \circ g \circ h = a \circ k$ for some arrows $h : B \to K$ and $k : B \to A$. Then as the right square is assumed to be a pullback there is a unique arrow $\alpha' : B \to A'$ with $g \circ h = a' \circ \alpha'$ and $k = f' \circ \alpha'$. Furthermore, due to the assumption that the left square is a pullback there exists a unique map $\alpha'' : B \to A''$ with $a'' \circ \alpha'' = h$ and $g' \circ \alpha'' = \alpha'$. The whole situation is summarised in the following diagram



Thus $\alpha''$ is a mediating arrow from $(h, k)$ to $(f \circ g, a)$. To show uniqueness of $\alpha''$ with this property assume that $\alpha : B \to A''$ with $a'' \circ \alpha = h$ and

$f' \circ g' \circ \alpha = k$. But then we have also that $a' \circ g' \circ \alpha = g \circ a'' \circ \alpha = g \circ h$ and, therefore, it follows that $\alpha' = g' \circ \alpha$ by uniqueness of mediating arrows. Thus, again by uniqueness of mediating arrows it follows that $\alpha = \alpha''$.

For the reverse direction suppose that the outer rectangle is a pullback. For showing that the left square is a pullback suppose that $g \circ h = a' \circ k$ for some maps $h : B \to K$ and $k : B \to A'$. Then we have also $f \circ g \circ h = f \circ a' \circ k = a \circ f' \circ k$. Thus, there exists a unique arrow $\alpha : B \to A''$ with $h = a'' \circ \alpha$ and $f' \circ k = f' \circ g' \circ \alpha$ as the outer rectangle is assumed to be a pullback. The situation is summarised in the following diagram
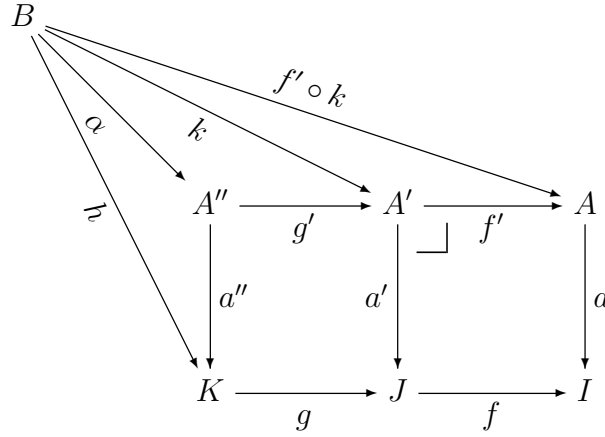


where $k = g' \circ \alpha$ follows from uniqueness of mediating arrows to the right pullback. Thus $\alpha$ is a mediating arrow from $(h, k)$ to $(a'', g')$. For uniqueness of $\alpha$ with this property suppose that $h = a'' \circ \beta$ and $k = g' \circ \beta$. Then also $f' \circ k = f' \circ g' \circ \beta$ and, therefore, we have $\alpha = \beta$ by uniqueness of mediating arrows to the outer pullback. □

Next we will define quite generally what is a limit for a diagram or a functor. Later we will show that a category has limits for all small diagrams iff it "has limits" in the sense of Definition 7.4.

But before we have to define precisely what we mean by a "diagram" in a category which in turn requires the notion of *graph* to be defined next.

**Definition 7.6** (graphs)
*A (directed) graph is a tuple $G = (G_0, G_1, d_0^G, d_1^G)$ where $G_0$ and $G_1$ are sets of* nodes *and* edges, *respectively, and $d_0^G, d_1^G : G_1 \to G_0$ are functions called* source *and* target, *respectively.*

*For graphs $G$ and $H$ a graph morphism from $G$ to $H$ is given by a pair $f = (f_0, f_1)$ where $f_i : G_i \to H_i$ for $i=0, 1$ satisfying*

$$d_i^H \circ f_1 = f_0 \circ d_i^G$$

*for $i=0, 1$.* $\diamondsuit$

Notice that graphs and graph morphisms give rise to a category **Graph** where composition of graph morphisms is defined componentwise, i.e. $(g \circ f)_i = g_i \circ f_i$ for $i=0, 1$, and $\mathsf{id}_G = (\mathsf{id}_{G_0}, \mathsf{id}_{G_1})$. Obviously, the category **Graph** is isomorphic to the functor category $\mathbf{Set}^{\mathbb{G}^{\mathrm{op}}}$ where $\mathbb{G}$ is the category

$$V \underset{d_1}{\overset{d_0}{\rightrightarrows}} E$$

with two objects $E$ ("edges") and $V$ ("vertices") whose only non-trivial arrows are $d_0$ and $d_1$ from $V$ to $E$.

Moreover, notice that every category $\mathbb{C}$ can be understood as a (possibly very large) graph by forgetting composition and identities.

Next we will define what is a diagram in a category $\mathbb{C}$ and what are natural transformations between diagrams "of the same shape".

**Definition 7.7** (diagrams)
*Let $\mathbb{C}$ be a category and $G = (G_0, G_1, d_0^G, d_1^G)$ be a graph. A diagram in $\mathbb{C}$ of shape $G$ is a graph morphism $D : G \to \mathbb{C}$ (where $\mathbb{C}$ is understood as a graph).*

*For diagrams $D, D' : G \to \mathbb{C}$ a natural transformation from $D$ to $D'$ is a family $\tau = (\tau_I : D(I) \to D'(I) \mid I \in G_0)$ such that for every edge $e : I \to J$ in $G$ (i.e. $d_0(e) = I$ and $d_1(e) = J$) the diagram*

$$
\begin{array}{ccc}
D(I) & \xrightarrow{\ \tau_I\ } & D'(I) \\
{\scriptstyle D(e)}\downarrow & & \downarrow{\scriptstyle D'(e)} \\
D(J) & \xrightarrow[\ \tau_J\ ]{} & D'(J)
\end{array}
$$

*commutes. We use the notation $\tau : D \to D'$ for stating that $\tau$ is a natural transformation from $D$ to $D'$.*

*Diagrams of shape $G$ and natural transformations between them form a category $\mathsf{Diag}(G, \mathbb{C})$ where composition is defined componentwise, i.e. $(\tau \circ \sigma)_I = \tau_I \circ \sigma_I$, and $(\mathsf{id}_D)_I = \mathsf{id}_{D(I)}$.* $\diamondsuit$

In order to formulate a notion of limit for diagrams of shape $G$ we need as an auxiliary notion a functor $\Delta_{\mathbb{C}}^G$ from $\mathbb{C}$ to $\mathsf{Diag}(G, \mathbb{C})$ assigning to every object $X$ in $\mathbb{C}$ "the constant diagram in $\mathbb{C}$ of shape $G$ with value $X$".

**Definition 7.8** *Let $\mathbb{C}$ be a category and $G$ a graph. Then the functor $\Delta_{\mathbb{C}}^G : \mathbb{C} \to \mathsf{Diag}(G, \mathbb{C})$ sends an object $X \in \mathsf{Ob}(\mathbb{C})$ to the diagram $\Delta_{\mathbb{C}}^G(X)$ with*

$$\Delta_{\mathbb{C}}^G(X)(I) = X \qquad \text{for } I \in G_0 \text{ and}$$
$$\Delta_{\mathbb{C}}^G(X)(e) = \mathsf{id}_X \qquad \text{for } e \in G_1$$

*and a morphism $f : X \to Y$ in $\mathbb{C}$ to the natural transformation $\Delta_{\mathbb{C}}^G(f) : \Delta_{\mathbb{C}}^G(X) \to \Delta_{\mathbb{C}}^G(Y)$ with $\Delta_{\mathbb{C}}^G(f)_I = f$ for $I \in G_0$.* ◇

Now we are ready to define what is a *limit cone* for a diagram $D : G \to \mathbb{C}$.

**Definition 7.9** (limit)
*Let $G$ be a graph and $D : G \to \mathbb{C}$ be a diagram of shape $G$ in the category $\mathbb{C}$. We define $\mathsf{Cone}(D)$ as the slice category $\Delta_{\mathbb{C}}^G {\downarrow} D$ where a morphism from cone $\tau' : \Delta(X') \to D$ to cone $\tau : \Delta(X) \to D$ is a $\mathbb{C}$-morphism $f : X' \to X$ making*

$$\Delta(X') \xrightarrow{\Delta(f)} \Delta(X)$$



*commute in $\mathsf{Diag}(G, \mathbb{C})$.*
*A* limit cone *for $D$ is a terminal object in $\mathsf{Cone}(D)$.* ◇

We recommend it as an exercise(!) to verify that $\pi : \Delta_{\mathbb{C}}^G(P) \to D$ is a limit cone for diagram $D : G \to \mathbb{C}$ if and only if the contravariant functor

$$\mathsf{Diag}(G, \mathbb{C})(\Delta(-), D) : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$$

is representable, i.e. there exists a natural isomorphism

$$\mathsf{Y}_{\mathbb{C}}(P) \xrightarrow[\cong]{\varphi} \mathsf{Diag}(G, \mathbb{C})(\Delta(-), D)$$

with $\varphi_P(\mathsf{id}_P) = \pi$.

We now will explicitate the above definition of limit which despite its conciseness and elegance might be a little bit difficult to grasp when seeing it the first time. First recall that a natural transformation $\tau : \Delta(X) \to D$ is nothing but a family of $\mathbb{C}$-morphisms $(\tau_I : X \to D(I) \mid I \in G_0)$ such that

$$
\begin{array}{ccc}
 & X & \\
\tau_I \swarrow & & \searrow \tau_J \\
D(I) & \xrightarrow{\quad D(e) \quad} & D(J)
\end{array}
$$

commutes for all edges $e : I \to J$ in $G_1$. (These commuting triangles are thought of as the "side faces of the cone $\tau$" explaining the terminology "cone".) Now for cones $\tau : \Delta(X) \to D$ and $\sigma : \Delta(Y) \to D$ over $D$ a morphism from $(Y, \sigma)$ to $(X, \tau)$ is nothing but a $\mathbb{C}$-morphism $f : Y \to X$ such that

$$
\begin{array}{ccc}
Y & \xrightarrow{\quad f \quad} & X \\
\sigma_I \searrow & & \swarrow \tau_I \\
 & D(I) &
\end{array}
$$

commutes for all $I \in G_0$.

From these considerations it should have become clear that products, equalisers and pullbacks are particular instances of the general notion of limit. Next we show that products and equalisers suffice for guaranteeing the existence of limits for arbitrary (small) diagrams.

**Theorem 7.2** *A category $\mathbb{C}$ has limits for all (small) diagrams if and only if $\mathbb{C}$ has (small) products and equalisers.*

*Proof:* Suppose $\mathbb{C}$ has (small) limits. Then, it has in particular limits for diagrams of the shape $\Delta(I)$, the discrete graph with $\Delta(I)_0 = I$ and $\Delta(I)_1 = \emptyset$, whenever $I$ is small. Thus $\mathbb{C}$ has products of $I$-indexed families of objects for all small $I$. Furthermore, limits for diagrams of shape

$$
\bullet \rightrightarrows \bullet
$$

provide equalisers as there is an obvious 1-1-correspondence between morphism $g : X \to A$ with $f_1 \circ g = f_2 \circ g$ and cones

$$
\begin{array}{c}
X \\
g \swarrow \quad \searrow h \\
A \underset{f_2}{\overset{f_1}{\rightrightarrows}} B
\end{array}
$$

(the latter being determined already by $g$ because $h = f_1 \circ g = f_2 \circ g$).
For the reverse direction assume that $\mathbb{C}$ has (small) products and equalisers. Suppose $D : G \to \mathbb{C}$ is a diagram with $G$ a small graph. We define

$$
A := \prod_{I \in G_0} D(I) \qquad \text{and} \qquad B := \prod_{e \in G_1} D(d_1(e))
$$

and $f, g : A \to B$ as the morphisms with

$$
\pi_e \circ f = \pi_{d_1(e)} \qquad \text{and} \qquad \pi_e \circ g = D(e) \circ \pi_{d_0(e)}
$$

for $e \in G_1$. Let $m : P \rightarrowtail A$ be an equalizer of $f$ and $g$. We will show that

$$
\mu_I = \pi_I \circ m \qquad (I \in G_0)
$$

is a limiting cone for $D$. First we show that $\mu = (\mu_I \mid I \in G_0)$ is a cone. For that purpose suppose $e : I \to J$ is an edge in $G$. Then we have

$$
D(e) \circ \mu_I = D(e) \circ \pi_I \circ m = \pi_e \circ g \circ m = \pi_e \circ f \circ m = \pi_J \circ m = \mu_J
$$

as desired. Suppose $\tau : \Delta(X) \to D$ is a cone over $D$. We have to show that there is a unique morphism $h : X \to P$ such that

$$
\mu_I \circ h = \tau_I
$$

for all $I \in G_0$. As $\mu_I = \pi_I \circ m$ the desired $h$ has to satisfy $\pi_I \circ m \circ h = \tau_I$ for all $I \in G_0$. As this requirement determines $m \circ h$ uniquely and $m$ is monic it follows that $h$ is determined uniquely. It remains to show the existence of $h$ with this property. Let $\alpha : X \to A$ be the unique arrow with $\pi_I \circ \alpha = \tau_I$. Thus, it suffices to show that $\alpha$ factors through $m$ via some $h$. This, however, is guaranteed if we can show that $\alpha$ equalises $f$ and $g$, i.e. $\pi_e \circ f \circ \alpha = \pi_e \circ g \circ \alpha$ for all $e \in G_1$. For this purpose suppose $e : I \to J$ in $G$. Then we have

$$
\pi_e \circ f \circ \alpha = \pi_J \circ \alpha = \tau_J = D(e) \circ \tau_I = D(e) \circ \pi_I \circ \alpha = \pi_e \circ g \circ \alpha
$$

as desired. $\qquad \square$

By inspection of the above proof we get that

**Corollary 7.1** *A category $\mathbb{C}$ has limits for all finite diagrams iff $\mathbb{C}$ has finite products and equalisers.*

As functors to $\mathbb{C}$ are particular cases of diagrams in $\mathbb{C}$ in the sense of Definition 7.7 it is clear that limits for diagrams subsume limits for functors. As products and equalisers are limits for functors it follows from Theorem 7.2 and Corollary 7.1 that a category has all small/finite limits iff it has limits for all functors $F : \mathbb{D} \to \mathbb{C}$ where $\mathbb{D}$ is small/finite.

# 8 Adjoint Functors

Because of its great importance for our treatment of adjoint functors we recall the notion of *representable presheaf* and give a simple characterisation of representability.

**Theorem 8.1** (characterisation of representability)
*Let $\mathbb{C}$ be a (small) category. Then*

(1) *a (contravariant) presheaf $F : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ is representable, i.e. iso-morphic to $\mathsf{Y}_{\mathbb{C}}(I) = \mathbb{C}(-, I)$ for some $I \in \mathsf{Ob}(\mathbb{C})$, iff there exists an $x \in F(I)$ such that for every $y \in F(J)$ there exists a unique morphism $u : J \to I$ in $\mathbb{C}$ with $y = F(u)(x)$*

(2) *a (covariant) presheaf $F : \mathbb{C} \to \mathbf{Set}$ is representable, i.e. isomorphic to $\mathbb{C}(I, -)$ for some $I \in \mathsf{Ob}(\mathbb{C})$, iff there exists an $x \in F(I)$ such that for all $y \in F(J)$ there exists a unique morphism $u : I \to J$ in $\mathbb{C}$ with $y = F(u)(x)$.*

*Proof:* We just prove (1) and leave the (analogous) argument for (2) to the reader as an exercise(!).
Suppose $F$ is representable, i.e. there exists a natural isomorphism $\varphi : \mathsf{Y}_{\mathbb{C}}(I) \overset{\cong}{\Rightarrow} F$. Then $x = \varphi_I(\mathsf{id}_I) \in F(I)$ has the desired property as for $y \in F(J)$ the morphism $\varphi_J^{-1}(y)$ is the unique arrow $u : J \to I$ with $y = F(u)(x)$ which can be seen as follows. We have

$$F(u)(x) = F(u)(\varphi_I(\mathsf{id}_I)) = \varphi_J(\mathsf{Y}_{\mathbb{C}}(I)(u)(\mathsf{id}_I)) = \varphi_J(u) = \varphi_J(\varphi_J^{-1}(y)) = y$$

and if $y = F(v)(x)$ then

$$\varphi_J^{-1}(y) = \varphi_J^{-1}(F(v)(x)) = \mathsf{Y}_{\mathbb{C}}(I)(v)(\varphi_I^{-1}(x)) = \mathsf{Y}_{\mathbb{C}}(I)(v)(\mathsf{id}_I) = \mathsf{id}_I \circ v = v \, .$$

For the reverse direction suppose that $x \in F(I)$ such that for every $y \in F(J)$ there exists a unique $u : J \to I$ with $y = F(u)(x)$. Then by the Yoneda lemma there exists a unique natural transformation $\varphi : \mathsf{Y}_{\mathbb{C}}(I) \to F$ with $\varphi_I(\mathsf{id}_I) = x$. From the proof of the Yoneda lemma we know that $\varphi_J(u) = F(u)(x)$. Thus, due to our assumption about $x$ we know that $\varphi_J$ is a bijection for all $J \in \mathsf{Ob}(\mathbb{C})$ and, therefore, the natural transformation $\varphi$ is an isomorphism in $\widehat{\mathbb{C}}$. $\qquad\square$

**Definition 8.1** (category of elements)
*Let $\mathbb{C}$ be a (small) category.*

(1) *For a (contravariant) presheaf $F : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ its category of elements $\mathbf{Elts}(F)$ is defined as follows: its objects are pairs $(I, x) \in \coprod_{I \in \mathbf{Ob}(\mathbb{C})} F(I)$, its morphisms from $(J, y)$ to $(I, x)$ are the $\mathbb{C}$-morphisms $u : J \to I$ with $y = F(u)(x)$ and composition and identities are inherited from $\mathbb{C}$.*

(2) *For a (covariant) presheaf $F : \mathbb{C} \to \mathbf{Set}$ its category of elements $\mathbf{Elts}(F)$ is defined as follows: its objects are pairs $(I, x) \in \coprod_{I \in \mathbf{Ob}(\mathbb{C})} F(I)$, its morphisms from $(I, x)$ to $(J, y)$ are the $\mathbb{C}$-morphisms $u : I \to J$ with $y = F(u)(x)$ and composition and identities are inherited from $\mathbb{C}$.* ◊*

Notice that for a contravariant presheaf $F : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ the category $\mathbf{Elts}(F)$ is isomorphic to the comma category $F^{\mathsf{op}}{\downarrow}1$ (where $F^{\mathsf{op}} : \mathbb{C} \to \mathbf{Set}^{\mathsf{op}}$). Similarly, for a covariant presheaf $F : \mathbb{C} \to \mathbf{Set}$ the category $\mathbf{Elts}(F)$ is isomorphic to $1{\downarrow}F$. Alternatively, for a contravariant presheaf $F : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ the category $\mathbf{Elts}(F)$ is isomorphic to $\mathsf{Y}_{\mathbb{C}}{\downarrow}F$ and for a covariant presheaf $F : \mathbb{C} \to \mathbf{Set}$ the category $\mathbf{Elts}(F)$ is isomorphic to $(\mathsf{Y}_{\mathbb{C}^{\mathsf{op}}}{\downarrow}F)^{\mathsf{op}}$. The verification of these claims we leave to the reader as an exercise(!).

Using the terminology of Definition 8.1 we can reformulate Theorem 8.1 quite elegantly as follows.

**Theorem 8.2** *Let $\mathbb{C}$ be a small category. Then $F : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ is representable iff $\mathbf{Elts}(F)$ has a terminal object and $F : \mathbb{C} \to \mathbf{Set}$ is representable iff $\mathbf{Elts}(F)$ has an initial object.*

*Proof:* We prove the claim just for contravariant $F$. For covariant $F$ the argument is analogous by duality.
By Theorem 8.1 $F$ is representable iff there exists an $x \in F(I)$ such that for every $y \in F(J)$ there is a unique arrow $u : J \to I$ with $y = F(u)(x)$, i.e. iff $(I, x)$ is terminal in $\mathbf{Elts}(F)$. Thus $F$ is representable iff $\mathbf{Elts}(F)$ has a terminal object. □

Based on the notion of representability we are ready to define the notion of *adjointness*.

**Definition 8.2** (left and right adjointable)

*A functor $F : \mathbb{A} \to \mathbb{B}$ is called* right adjointable *iff for every $B \in \mathsf{Ob}(\mathbb{B})$ the functor $\mathsf{Y}_\mathbb{B}(B) \circ F^{\mathsf{op}} = \mathbb{B}(F(-), B) : \mathbb{A}^{\mathsf{op}} \to \mathbf{Set}$ is representable.*

*A functor $U : \mathbb{B} \to \mathbb{A}$ is called* left adjointable *iff for all $A \in \mathsf{Ob}(\mathbb{A})$ the functor $\mathbb{A}(A, U(-)) : \mathbb{B} \to \mathbf{Set}$ is representable.* $\Diamond$

Obviously, due to Theorem 8.1 a functor $F : \mathbb{A} \to \mathbb{B}$ is right adjointable iff for every object $B$ in $\mathbb{B}$ there exists a morphism $\varepsilon_B : F(U_0 B) \to B$ such that for every object $A$ in $\mathbb{A}$ and $f : F(A) \to B$ there exists a unique morphism $g : A \to U_0 B$ with $f = \varepsilon_B \circ F(g)$ as indicated in the diagram

$$
\begin{array}{ccc}
U_0 B & F(U_0 B) \xrightarrow{\ \varepsilon_B\ } B \\
\uparrow g & \uparrow F(g) \quad \nearrow f \\
A & F(A)
\end{array}
$$

A typical example is the diagonal functor $\Delta : \mathbb{C} \to \mathbb{C} \times \mathbb{C}$ sending $X$ to $(X, X)$ and $f$ to $(f, f)$. That $\Delta$ is right adjointable means that for every $(A, B) \in \mathsf{Ob}(\mathbb{C} \times \mathbb{C})$ there exists an object $P$ in $\mathbb{C}$ together with $\varepsilon_{(A,B)} = (\pi_1, \pi_2) : (P, P) \to (A, B)$ such that for every $(f, g) : \Delta(C) \to (A, B)$, i.e. $f : C \to A$ and $g : C \to B$, there exists a unique morphism $h : C \to P$ with $(\pi_1, \pi_2) \circ \Delta(h) = (f, g)$, i.e. $\pi_1 \circ h = f$ and $\pi_2 \circ h = g$, as indicated in the diagram

$$
\begin{array}{ccc}
P & (P, P) \xrightarrow{\ (\pi_1, \pi_2)\ } (A, B) \\
\uparrow h & \uparrow (h, h) \quad \nearrow (f, g) \\
C & (C, C)
\end{array}
$$

It is easy to see (exercise!) that $\mathbb{C}$ has binary products if and only if the functor $\Delta : \mathbb{C} \to \mathbb{C} \times \mathbb{C}$ is right adjointable.

More generally, a category $\mathbb{C}$ has limits of diagrams of shape $G$ if and only if the functor

$$\Delta_\mathbb{C}^G : \mathbb{C} \to \mathsf{Diag}(G, \mathbb{C})$$

is right adjointable, i.e.

$$
\begin{array}{ccccc}
\lim D & & \Delta^G_{\mathbb{C}}(\lim D) & \xrightarrow{\ \pi\ } & D \\
\vdots & & \big\uparrow & & \nearrow \\
h & & \Delta^G_{\mathbb{C}}(h) & & \curvearrowleft \\
\vdots & & \big\uparrow & & \\
X & & \Delta^G_{\mathbb{C}}(X) & &
\end{array}
$$

A further example, particularly interesting from a "logical" point of view, is the following characterisation of function spaces. Let $A$ be a set. Then the functor $(-) \times A : \mathbf{Set} \to \mathbf{Set}$ (sending $f : Y \to X$ to $f \times A : Y \times A \to X \times A : (y, a) \mapsto (f(y), a)$) is right adjointable, i.e. there exists a set $B^A$ together with a map $\varepsilon : B^A \times A \to B$ such that for every $f : C \times A \to B$ there exists a unique map $g : C \to B^A$ with $\varepsilon \circ (g \times A) = f$. Naturally for $B^A$ one chooses the set of all functions from $A$ to $B$ and defines $\varepsilon : B^A \times A \to B$ as the *evaluation* map sending $(f, a)$ to $f(a)$. Now given $f : C \times A \to B$ the unique map $g : C \to B^A$ with $\varepsilon \circ (g \times A) = f$ is obtained by defining $g(c)(a) = f(c, a)$. One readily checks that $g$ is determined uniquely by this requirement. Accordingly, we may write $\lambda(f)$ for denoting this unique map $g$. The situation is summarized in the following diagram

$$
\begin{array}{ccccc}
B^A & & B^A \times A & \xrightarrow{\ \varepsilon\ } & B \\
\big\uparrow & & \big\uparrow & & \nearrow \\
\lambda(f) & & \lambda(f) \times A & & f \\
\big\uparrow & & \big\uparrow & & \\
C & & C \times A & &
\end{array}
$$

whose shape should be already quite familiar.

Next we consider examples for functors which are left adjointable. Let $U : \mathbf{Mon} \to \mathbf{Set}$ be the functor sending a monoid $M$ to its underlying set $U(M)$ and a monoid homomorphism $h : M \to M'$ to the function $h$. As $U$ forgets the monoid structure it is often called "forgetful functor". This forgetful functor $U$ is left adjointable as for every set $X$ the functor $\mathbf{Set}(X, U(-)) : \mathbf{Mon} \to \mathbf{Set}$ is representable via $\eta_X : X \to U(X^*)$ where $X^*$ is the monoid of *words over* $X$ whose binary operation is concatenation of words

$$
x_1 x_2 \ldots x_n \cdot y_1 y_2 \ldots y_m = x_1 x_2 \ldots x_n y_1 y_2 \ldots y_m
$$

and whose neutral element is given by the empty word (often denoted as $\varepsilon$). The map $\eta_X$ sends $x \in X$ to the word $x$ consisting just of the single "letter" $x$. It is easy to verify that for every monoid $M$ and every function $f : X \to U(M)$ there exists a unique monoid homomorphism $h : X^* \to M$ (sending $x_1 x_2 \ldots x_n$ to $h(x_1 x_2 \ldots x_n) = f(x_1) \cdot f(x_2) \cdot \ldots \cdot f(x_n)$ and the empty word to $h(\varepsilon) = 1_M$, the neutral element of the monoid $M$). The situation is summarized in the following diagram

$$
\begin{array}{ccccc}
X & \xrightarrow{\ \eta_X\ } & U(X^*) & & X^* \\
& \searrow{\scriptstyle f} & \big\downarrow{\scriptstyle U(h)} & & \vdots{\scriptstyle h} \\
& & U(M) & & M
\end{array}
$$

Actually, for every equationally[10] defined notion of algebraic structure (as e.g. group, ring, vector space (over a fixed scalar field $k$) etc. the forgetful functor to **Set** is left adjointable where for every set $X$ the map $\eta_X : X \to U(FX)$ is the inclusion of the set $X$ of *generators* into the underlying set of the *free algebraic structure $FX$ over $X$*.

**Definition 8.3** (adjunction)
*An* adjunction *is a tuple* $(F, U, \varphi)$ *where* $F : \mathbb{A} \to \mathbb{B}$ *and* $U : \mathbb{B} \to \mathbb{A}$ *are functors and* $\varphi$ *is a family*

$$
\varphi_{A,B} : \mathbb{B}(F(A), B) \xrightarrow{\cong} \mathbb{A}(A, U(B))
$$

*of bijections natural in $A$ and $B$, i.e. the diagram*

$$
\begin{array}{ccc}
\mathbb{B}(F(A), B) & \xrightarrow{\ \varphi_{A,B}\ } & \mathbb{A}(A, U(B)) \\
\big\downarrow{\scriptstyle \mathbb{B}(F(f), g)} & & \big\downarrow{\scriptstyle \mathbb{A}(f, U(g))} \\
\mathbb{B}(F(A'), B') & \xrightarrow[\varphi_{A',B'}]{} & \mathbb{A}(A', U(B'))
\end{array}
$$

---

[10]Notice, however, that fields are an exception! Actually, there is no free field over the empty set of generators. If there were a free field $F$ then it would have to be isomorphic to $\mathbb{Z}_2$ as the hypothetical unique morphism from $F$ to $\mathbb{Z}_2$ has to be one to one. However, there is no homomorphism from $\mathbb{Z}_2$ to $\mathbb{Q}$ and, therefore, there cannot exist a free field.

*commutes for all morphisms $f : A' \to A$ in $\mathbb{A}$ and $g : B \to B'$ in $\mathbb{B}$.*
*We write $F \dashv U$ iff there is a $\varphi$ such that $(F, U, \varphi)$ is an adjunction.*    $\Diamond$

Obviously, in the above definition the condition on $\varphi$ can be formulated more concisely as the requirement that $\varphi$ is a natural isomorphism from $\mathbb{B}(F(-_1), -_2)$ to $\mathbb{A}(-_1, U(-_2))$ in the functor category $\mathbf{Set}^{\mathbb{A}^{\mathrm{op}} \times \mathbb{B}}$.
Now if $F : \mathbb{A} \to \mathbb{B}$ is right adjointable then *using strong choice principles* one can augment $F$ to an adjunction $(F, U, \varphi)$ as follows. For every $B \in \mathsf{Ob}(\mathbb{B})$ choose a map $\varepsilon_B : F(U_0 B) \to B$ representing $\mathbb{B}(F(-), B)$. The object part of $U$ is defined by putting $U(B) = U_0 B$. The morphism part of $U$ is defined by sending every map $g : B \to B'$ in $\mathbb{B}$ to the unique arrow $U(g) : U_0 B \to U_0 B'$ making the diagram

$$
\begin{array}{ccc}
F(U_0 B) & \xrightarrow{\;\varepsilon_B\;} & B \\
{\scriptstyle F(U(g))}\big\downarrow & & \big\downarrow{\scriptstyle g} \\
F(U_0 B') & \xrightarrow[\;\varepsilon_{B'}\;]{} & B'
\end{array}
$$

commute. We recommend it as an exercise(!) to show that the so defined $U$ is actually a functor, i.e. preserves composition and identities. The natural isomorphism $\varphi$ is defined at $A \in \mathsf{Ob}(\mathbb{A})$ and $B \in \mathsf{Ob}(\mathbb{B})$ by sending an arrow $f : F(A) \to B$ to the unique arrow $\varphi_{A,B}(f) : A \to U(B)$ making the diagram

$$
\begin{array}{ccc}
FUB & \xrightarrow{\;\varepsilon_B\;} & B \\
{\scriptstyle F(\varphi_{A,B}(f))}\big\uparrow & \nearrow{\scriptstyle f} & \\
FA & &
\end{array}
$$

commute. As $\varepsilon_B$ represents $\mathbb{B}(F(-), B)$ the mapping $\varphi_{A,B}$ is a bijection between $\mathbb{B}(F(A), B)$ and $\mathbb{A}(A, U(B))$. We recommend it as an exercise(!) to verify that $\varphi$ is actually a natural transformation.
Analogously, every left adjointable functor $U : \mathbb{B} \to \mathbb{A}$ can be augmented to an adjunction $(F, U, \varphi)$ choosing for every object $A$ in $\mathbb{A}$ an arrow $\eta_A : A \to U(F_0 A)$ representing the presheaf $\mathbb{A}(A, U(-))$.

Moreover, for every adjunction $(F, U, \varphi)$ one can show that the functors $F$ and $U$ are right and left adjointable, respectively. An element representing $\mathbb{A}(A, U(-))$ is given by $\eta_A = \varphi_{A,FA}(\mathsf{id}_{FA})$, called *unit of the adjunction at $A$*, and an element representing $\mathbb{B}(F(-), B)$ is given by $\varepsilon_B = \varphi^{-1}_{UB,B}(\mathsf{id}_{UB})$, called *counit of the adjunction at $B$*. One can show (exercise!) that the so defined $\eta$ and $\varepsilon$ are natural tranformations, i.e.

$$\eta : \mathsf{Id}_{\mathbb{A}} \Rightarrow UF \qquad \text{and} \qquad \varepsilon : FU \Rightarrow \mathsf{Id}_{\mathbb{B}}$$

respectively, satisfying the so-called "triangle equalities"

$$\varepsilon_F \circ F\eta = \mathsf{id}_F \qquad \text{and} \qquad U\varepsilon \circ \eta_U = \mathsf{id}_U$$

i.e. the diagrams

$$
\begin{array}{ccc}
FA & \xrightarrow{F\eta_A} & FUFA \\
\| & & \downarrow{\varepsilon_{FA}} \\
& & FA
\end{array}
\qquad\qquad
\begin{array}{ccc}
UB & \xrightarrow{\eta_{UB}} & UFUB \\
\| & & \downarrow{U\varepsilon_B} \\
& & UB
\end{array}
$$

commute for all $A \in \mathsf{Ob}(\mathbb{A})$ and $B \in \mathsf{Ob}(\mathbb{B})$, respectively.[11]

Conversely, from natural transformations $\eta : \mathsf{Id}_{\mathbb{A}} \Rightarrow UF$ and $\varepsilon : FU \Rightarrow \mathsf{Id}_{\mathbb{B}}$ satisfying the triangle equalites $\varepsilon_F \circ F\eta = \mathsf{id}_F$ and $U\varepsilon \circ \eta_U = \mathsf{id}_U$ one can construct a natural isomorphism $\varphi : \mathbb{B}(F(-_1), -_2) \xrightarrow{\cong} \mathbb{A}(-_1, U(-_2))$ such that $(F, U, \varphi)$ is an adjunction with $\eta_A = \varphi_{A,FA}(\mathsf{id}_{FA})$ and $\varepsilon_B = \varphi^{-1}_{UB,B}(\mathsf{id}_{UB})$ simply by putting $\varphi_{A,B}(f) = U(f) \circ \eta_A$.

Before studying properties of adjunctions we observe that adjoints are unique up to isomorphism.

**Theorem 8.3** (uniqueness of adjoints)
*The right adjoints of a functor are all isomorphic and so are its left adjoints.*

*Proof:* Suppose $(F, U, \varphi)$ and $(F, U', \varphi')$ are adjunctions. Let $\varepsilon$ and $\varepsilon'$ be the corresponding counits of the adjunctions. We define a natural transformation

---

[11]e.g. the commutation of the first triangle follows from

$$\varphi_{A,FA}(\varepsilon_{FA} \circ F\eta_A) = \varphi_{UFA,FA}(\varepsilon_{FA}) \circ \eta_A = \mathsf{id}_{UFA} \circ \eta_{FA} = \eta_{FA} = \varphi_{A,FA}(\mathsf{id}_{FA})$$

because $\varphi_{A,FA}$ is one-to-one

$\iota : U \Rightarrow U'$ as follows: for $B \in \mathsf{Ob}(\mathbb{B})$ let $\iota_B : U(B) \to U'(B)$ be the unique arrow in $\mathbb{A}$ such that $\varepsilon'_B \circ F(\iota_B) = \varepsilon_B$. Notice that $\iota_B$ is an isomorphism whose inverse is given by the unique arrow $\iota_B^{-1}$ satisfying $\varepsilon_B \circ F(\iota_B^{-1}) = \varepsilon'_B$. That $\iota = (\iota_B)_{B \in \mathsf{Ob}(\mathbb{B})}$ is a natural transformation can be seen as follows. Suppose that $g : B \to B'$ then we have

$$\varepsilon'_{B'} \circ F(U'(g) \circ \iota_B) = \varepsilon'_{B'} \circ FU'g \circ F(\iota_B) = g \circ \varepsilon'_B \circ F(\iota_B) = g \circ \varepsilon_B$$

and

$$\varepsilon'_{B'} \circ F(\iota_{B'} \circ U(g)) = \varepsilon'_{B'} \circ F(\iota_{B'}) \circ FUg = \varepsilon_{B'} \circ FUg = g \circ \varepsilon_B$$

and, therefore, also $\varepsilon'_{B'} \circ F(U'(g) \circ \iota_B) = \varepsilon'_{B'} \circ F(\iota_{B'} \circ U(g))$ from which it follows that $U'(g) \circ \iota_B = \iota_{B'} \circ U(g)$ as desired.

Analogously, one proves that left adjoints of a functor are unique up to isomorphism. $\qquad\square$

**Theorem 8.4** *Right adjointable functors preserve colimits and left adjointable functors preserve limits.*

*Proof:* We just prove the first claim as the second one follows by duality. Thus, suppose that $F : \mathbb{A} \to \mathbb{B}$ is right adjointable, i.e. for all $B \in \mathsf{Ob}(\mathbb{B})$ there is an arrow $\varepsilon_B : FUB \to B$ such that for all $f : FA \to B$ in $\mathbb{B}$ there is a unique arrow $g : A \to UB$ in $\mathbb{A}$ such that $\varepsilon_B \circ Fg = f$. Now suppose $D : G \to \mathbb{A}$ is a diagram in $\mathbb{A}$ with colimiting cocone $\mu : D \Rightarrow \Delta(A)$. We will show that

$$F\mu = (F(\mu_I) : F(D(I)) \to F(A))_{I \in G_0}$$

is a colimiting cocone for $F \circ D$. First we show that $F\mu$ is a cocone. For that purpose suppose $u : I \to J$ in $G$. Then we have $F\mu_J \circ F(D(u)) = F\mu_I$ because $\mu_J \circ D(u) = \mu_I$ (as $\mu$ is a cocone over $D$) and $F$ preserves composition. Thus, it remains to show that $F\mu$ is a colimiting cocone for $F \circ D$. For that purpose suppose $\tau : F \circ D \Rightarrow \Delta(X)$. For $I \in G_0$ let $\sigma_I : D(I) \to UX$ be the unique arrow satisfying $\varepsilon_X \circ F(\sigma_I) = \tau_I$. We show that $\sigma : D \Rightarrow \Delta(UX)$. For $u : I \to J$ in $G$ we have

$$\varepsilon_X \circ F(\sigma_I) = \tau_I = \tau_J \circ F(D(u)) = \varepsilon_X \circ F(\sigma_J) \circ F(D(u)) = \varepsilon_X \circ F(\sigma_J \circ D(u))$$

and, therefore, also $\sigma_I = \sigma_J \circ D(u)$. Accordingly, there exists a unique arrow $h : A \to UX$ with $h \circ \mu_I = \sigma_I$ for all $I \in G_0$. Thus, for all $I \in G_0$ we have

$$\varepsilon_X \circ Fh \circ F\mu_I = \varepsilon_X \circ F\sigma_I = \tau_I$$

43

from which it follows that $k = \varepsilon_X \circ Fh$ is a mediating arrow from $F\mu$ to $\tau$. We have to show that $k$ is unique with this property. Suppose $k' : FA \to X$ with $k' \circ F\mu_I = \tau_I$ for all $I \in G_0$. Then there exists $h' : A \to UX$ with $\varepsilon_X \circ Fh' = k'$. Then for all $I \in G_0$ we have

$$\varepsilon_X \circ F(h' \circ \mu_I) = \varepsilon_X \circ Fh' \circ F\mu_I = k' \circ F\mu_I = \tau_I = \varepsilon_X \circ F(\sigma_I)$$

from which it follows that $h' \circ \mu_I = \sigma_I$ for all $I \in G_0$. Thus, we have $h = h'$ (as $\mu$ is a colimiting cocone) and, accordingly, also $k = \varepsilon_X \circ Fh = \varepsilon_X \circ Fh' = k'$ as desired. $\qquad\square$

Next we characterise some properties of a functor under the assumption that it has a right adjoint.

**Theorem 8.5** *Suppose $F \dashv U : \mathbb{B} \to \mathbb{A}$ is an adjunction with unit $\eta : \mathsf{Id} \Rightarrow UF$. Then*

(1) *$F$ is faithful iff all $\eta_A$ are monic*

(2) *$F$ is full iff all $\eta_A$ are split epic, i.e. $\eta_A \circ s = \mathsf{id}$ for some $s : UFA \to A$*

(3) *$F$ is full and faithful iff $\eta$ is a natural isomorphism.*

*Proof:*
*ad* (1) : For morphisms $f_1, f_2 : A' \to A$ in $\mathbb{A}$ we have

$$\eta_A \circ f_1 = \eta_A \circ f_2 \quad \text{iff} \quad UFf_1 \circ \eta_{A'} = UFf_2 \circ \eta_{A'} \quad \text{iff} \quad Ff_1 = Ff_2$$

from which it follows that $F$ is faithful iff all $\eta_A$ are monic.
*ad* (2) : If $F$ is full then for all $A \in \mathsf{Ob}(\mathbb{A})$ there is a map $s_A : UFA \to A$ with $F(s_A) = \varepsilon_{FA} : FUFA \to FA$. Thus, we have $\mathsf{id}_{UFA} = \varphi(\varepsilon_{FA}) = \varphi(\mathsf{id}_{FA} \circ F(s_A)) = \eta_A \circ s_A$ as desired.
For the reverse direction suppose $\eta_A \circ s_A = \mathsf{id}_{UFA}$ for all $A \in \mathsf{Ob}(A)$. Suppose $g : FA' \to FA$. We show that $g = Ff$ for $f = s_A \circ Ug \circ \eta_{A'}$. As $F \dashv U$ it is equivalent to show that $Ug \circ \eta_{A'} = UFf \circ \eta_{A'}$ which can be seen as follows

$$UFf \circ \eta_{A'} = \eta_A \circ f = \eta_A \circ s_A \circ Ug \circ \eta_{A'} = Ug \circ \eta_{A'}$$

*ad* (3) : immediate from (1) and (2). $\qquad\square$

Actually, claim (3) of Theorem 8.5 can be strengthened as follows.

**Lemma 8.1** *If $F \dashv U : \mathbb{B} \to \mathbb{A}$ and $UF$ is isomorphic to $\mathsf{Id}_{\mathbb{A}}$ then $\eta$ is a natural isomorphism, too.*

*Proof:* Let $\iota : \mathsf{Id} \Rightarrow UF$ be a natural isomorphism. Then $\eta$ is an isomorphism as for every $A$ the arrow $\eta_A$ is inverted by $j_A = \iota_A^{-1} \circ U\varepsilon_{FA} \circ UF\iota_A$ which can be seen as follows. We have

$$j_A \circ \eta_A = \iota_A^{-1} \circ U\varepsilon_{FA} \circ UF\iota_A \circ \eta_A = \iota_A^{-1} \circ U\varepsilon_{FA} \circ \eta_{UFA} \circ \iota_A = \iota_A^{-1} \circ \iota_A = \mathsf{id}_A$$

and from this it follows that

$$
\begin{aligned}
& \iota_A^{-1} \circ \eta_A \circ j_A \circ \iota_A = && \text{(naturality of } \eta\text{)} \\
& = \iota_A^{-1} \circ UFj_A \circ UF\iota_A \circ \eta_A = && \text{(naturality of } \iota^{-1}\text{)} \\
& = j_A \circ \iota_A \circ \iota_A^{-1} \circ \eta_A = \\
& = j_A \circ \eta_A = \mathsf{id}_A
\end{aligned}
$$

and therefore $\eta_A \circ j_A = \iota_A \circ \iota_A^{-1} = \mathsf{id}_{UFA}$. $\qquad\square$

We leave it as an exercise(!) to formulate and verify the dual analogues of Theorem 8.5 and Lemma 8.1.

# 9    Adjoint Functor Theorems

Already back in the 1960ies P.J.Freyd proved a couple of theorems providing criteria for the existence of adjoints under fairly general assumptions. The exposition of Freyd's Adjoint Functor Theorems given in this section essentially follows the presentation given in [ML].

We need the following two lemmas whose easy proof we leave as a straightforward exercise(!) to the reader.

**Lemma 9.1** *Let $\mathbb{C}$ be a category with small limits and $U : \mathbb{C} \to \mathbb{B}$ a functor preserving small limits. Then for every object $B$ in $\mathbb{B}$ the comma category $B{\downarrow}U$ has small limits.*

**Lemma 9.2** *Let $\mathbb{C}$ be a locally small category and $U : \mathbb{C} \to \mathbb{B}$ a functor. Then for every object $B$ in $\mathbb{B}$ the comma category $B{\downarrow}U$ is locally small, too.*

The key idea of the general Adjoint Functor Theorem (FAFT) is to establish the existence of a *weakly initial object* from which there follows the existence of an initial object under the assumptions of local smallness and small completeness.

**Definition 9.1** *An object $W$ of a category $\mathbb{C}$ is called* weakly initial *iff for every object $A$ of $\mathbb{C}$ there exists a morphism from $W$ to $A$.*                    $\Diamond$

**Lemma 9.3** *Let $\mathbb{C}$ be a locally small category having all small limits. Then $\mathbb{C}$ has an initial object if and only if $\mathbb{C}$ has a weakly initial object.*

*Proof:* Obviously, every initial object in $\mathbb{C}$ is also weakly initial in $\mathbb{C}$. Conversely, suppose that $W$ is weakly initial in $\mathbb{C}$. Let $i : I \rightarrowtail W$ be an equaliser of all endomorphisms of $W$ which exists because $\mathbb{C}(W, W)$ is small and $\mathbb{C}$ has all small limits. We show that $I$ is initial in $\mathbb{C}$. Let $A$ be an object in $\mathbb{C}$. Then there exists a map $f$ from $W$ to $A$ and thus $f \circ i : I \to A$. For uniqueness suppose that $f_1, f_2 : I \to A$. Let $e : E \rightarrowtail I$ be an equaliser of $f_1$ and $f_2$. As $W$ is weakly initial there exists a map $p : W \to E$. As $i$ equalises all endomorphisms of $W$ we have $i = i \circ e \circ p \circ i$. As $i$ is monic it follows that $\mathsf{id}_I = e \circ p \circ i$. Thus, we have also $e \circ p \circ i \circ e = e$ from which it follows that $p \circ i \circ e = \mathsf{id}_E$ as $e$ being an equaliser is monic. Thus $e$ is an isomorphism (with inverse $p \circ i$) from which it follows that $f_1 = f_2$ as desired.     $\square$

**Theorem 9.1** (General Adjoint Functor Theorem (FAFT)) *Let $\mathbb{C}$ be a locally small category with small limits. Then a functor $U : \mathbb{C} \to \mathbb{B}$ has a left adjoint iff $U$ preserves small limits and satisfies the following*

> **Solution Set Condition** *For every object $B$ of $\mathbb{B}$ there exists a small family $(f_i : B \to U(X_i))_{i \in I}$ such that for every map $g : B \to U(X)$ for some $i \in I$ there is a map $h : X_i \to X$ with $g = U(h) \circ f_i$.*

*Proof:* Due to Theorem 8.4 a functor $U$ preserves limits whenever it has a left adjoint $F$. Moreover, for every $B \in \mathbb{B}$ the unit $\eta_B : B \to UFB$ of $F \dashv U$ at $B$ is initial in $B{\downarrow}U$ from which it follows that the solution set condition holds.

Suppose that $U$ preserves small limits and the solution set condition holds. For $U$ having a left adjoint it suffices to show that for every $B \in \mathbb{B}$ the comma category $B{\downarrow}U$ has an initial object. This, however, follows from Lemma 9.3 because $B{\downarrow}U$ is locally small by Lemma 9.2, has all small limits by Lemma 9.1 and the solution set condition gives rise to a weakly initial object in $B{\downarrow}U$ by taking the product of the small family $(f_i)_{i \in I}$ in $B{\downarrow}U$. $\square$

Notice that the solution set condition is actually necessary as can be seen from the following counterexample. Let $\mathbb{C} = \mathbf{Ord}^{\mathrm{op}}$ where $\mathbf{Ord}$ is the large poset of (small) ordinals considered as a category. Let $U$ be the unique functor from $\mathbb{C}$ to the terminal category $\mathbf{1}$. Obviously, all assumptions of Theorem 9.1 are satisfied with the exception of the solution set condition. Now if $U$ had a left adjoint this would give rise to an initial object in $\mathbb{C}$ which does not exist as there is no greatest (small) ordinal.

Next we discuss a few applications of FAFT.

First we show that for every category $\mathbb{A}$ of equationally defined algebras (and all homomorphisms between them) the forgetful functor $U : \mathbb{A} \to \mathbf{Set}$ has a left adjoint $F$, i.e. that free $\mathbb{A}$-algebras always exist. Every equationally defined class $\mathbb{A}$ of algebras is locally small and has small limits which are preserved by $U$. Thus, the assumptions of FAFT are satisfied. The solution set condition is also valid for the following reason. Let $I$ be a set. Then for every $A \in \mathbb{A}$ every function $f : I \to A$ factors through the least subalgebra of $A$ generated by the image of $I$ under $f$. As up to isomorphism there is just a small collection of algebras in $\mathbb{A}$ which are generated by a subset of cardinality less or equal the cardinality of $I$ there is up to isomorphism just a small collection of maps $f : I \to A$ such that $A$ is the only subalgebra of $A$

containing the image of $I$ under $f$. Thus, the forgetful functor $U : \mathbb{A} \to \mathbf{Set}$ has a left adjoint $F$.

From universal algebra one knows that the map $\eta_I$ sending elements of $I$ to the corresponding generators in $U(F(I))$ is one-to-one provided $\mathbb{A}$ contains algebras of arbitrarily big size. This can be shown also by the following "abstract nonsense" argument without any inspection of the actual construction of the free algebra $F(I)$. Suppose $\eta_I(i) = \eta_I(j)$ for some $i, j \in I$ with $i \neq j$. Let $A$ be an algebra with $A \geq |I|$. Then there exists a function $f : I \to U(A)$ which is one-to-one. Thus, by adjointness there exists a homomorphism $h : F(I) \to A$ with $f = U(h) \circ \eta_I$ rendering $\eta_I(i) = \eta_I(j)$ impossible because we have $h(\eta_I(i)) = f(i) \neq f(j) = h(\eta_I(j))$.

Notice, however, that the forgetful functor from the category $\mathbf{CBool}$ of complete Boolean algebras to $\mathbf{Set}$ does not have a left adjoint because there exist[12] arbitrarily big complete Boolean algebras generated by a countable subset. As $\mathbf{CBool}$ is locally small and has small limits preserved by the forgetful functor $U$ this counterexample illustrates again the necessity of the solution set condition.

Next we consider the forgetful functor $U$ from $\mathbf{CompHaus}$ to $\mathbf{Set}$ where $\mathbf{CompHaus}$ is the full subcategory of $\mathbf{Sp}$ on compact Hausdorff spaces. FAFT is applicable because $\mathbf{CompHaus}$ is locally small and has small limits which are preserved by $U$. Let $I$ be a set. Then every map $f : I \to U(X)$ factors through $\overline{f[I]}$, the closure of the image of $X$ under $f$. Thus, for verifying the solution set condition it suffices to consider only maps $f : I \to U(X)$ with *dense* image. For any such map $f$ the size of $X$ is bounded by the size of $\mathcal{P}^2(I)$ for the following reason. Define $e : X \to \mathcal{P}^2(I)$ by sending $x \in X$ to the collection of all $J \subseteq I$ with $x \in \overline{f[J]}$. For showing that $e$ is one-to-one suppose $x_1$ and $x_2$ are distinct elements of $X$. Then there exist open disjoint sets $U_1$ and $U_2$ with $x_i \in U_i$ for $i = 1, 2$ from which it follows that $f^{-1}[U_1] \notin e(x_2)$ and $f^{-1}[U_2] \notin e(x_1)$. But for $i = 1, 2$ we have $x_i \in \overline{f[f^{-1}[U_i]]}$ since for every open neighbourhood $V$ of $x_i$ the sets $f[I]$ and $U_i \cap V$ have non-empty intersection (as $f[I]$ is dense in $X$ by assumption). Thus $f^{-1}[U_i] \in e(x_i)$ for $i = 1, 2$ and, therefore, since $f^{-1}[U_1] \notin e(x_2)$ and $f^{-1}[U_2] \notin e(x_1)$ it follows that $e(x_1) \neq e(x_2)$ as desired. Up to isomorphism there is just a small collection of compact Hausdorff spaces with cardinality less or equal $2^{2^{|I|}}$. As $\mathbf{CompHaus}$ is locally small up to isomorphism there

---

[12]as has been shown in R. Solovay *New proof of a theorem of Gaifman and Hales* Bull.Amer.Math.Soc.72, 1966, pp.282–284

is just a small collection of continuous maps in **CompHaus** that start from $I$ and have dense image. Thus, the solution set condition holds for $U$ from which it follows by FAFT that $U$ has a left adjoint. By an even simpler argument it can be shown (exercise!) that the forgetful functor from **Haus**, the category of Hausdorff spaces and continuous maps, into **Sp** has a left adjoint, i.e. that **Haus** forms a *full reflective subcategory* of **Sp**.

Next we prove the Special Adjoint Functor Theorem providing a criterion which is easier to check than the solution set condition but requires slightly stronger assumptions.

**Definition 9.2** *A family $(C_i)_{i \in I}$ of objects in a category $\mathbb{C}$ is called* cogenerating *iff maps $f, g : Y \to X$ in $\mathbb{C}$ are equal whenever for all $i \in I$ it holds that $h \circ f = h \circ g$ for all $h : X \to C_i$.* $\diamond$

**Lemma 9.4** *Let $\mathbb{C}$ be a locally small category with small limits and infima of arbitrary families of subobjects. If $\mathbb{C}$ has a cogenerating family $(C_i)_{i \in I}$ then $\mathbb{C}$ has an initial object.*

*Proof:* Let $0$ be the intersection, i.e. infimum, of *all* subobjects of $\prod_{i \in I} C_i$. Suppose $f, g : 0 \to X$. Then the equaliser $e : E \rightarrowtail 0$ of $f$ and $g$ is an isomorphism because $0$ is already the least subobject of $\prod_{i \in I} C_i$. Thus, it follows that $f = g$.

For initiality of $0$ it remains to show that for every object $X$ of $\mathbb{C}$ there exists some morphism from $0$ to $X$. Consider the pullback

$$
\begin{array}{ccc}
Y & \xrightarrow{\quad n \quad} & \prod_{i \in I} C_i \\
{\scriptstyle g}\Big\downarrow & \lrcorner & \Big\downarrow{\scriptstyle f} \\
X & \xrightarrow[\quad m \quad]{} & \prod_{i \in I} \prod_{h : X \to C_i} C_i
\end{array}
$$

where $\pi_{i,h} \circ m = h$ and $\pi_{i,h} \circ f = \pi_i$. As the family $(C_i)_{i \in I}$ is cogenerating for $\mathbb{C}$ it follows that $m$ is monic and, therefore, the map $n$ is monic, too, by Lemma 7.2. As $Y$ is a subobject of $\prod_{i \in I} C_i$ there is a map $k$ from $0$ to $Y$ and, therefore, we have $g \circ k : 0 \to X$ as desired. $\square$

**Theorem 9.2** (Special Adjoint Functor Theorem (SAFT)) *Let $\mathbb{C}$ be a locally small category with small limits, infima of arbitrary families of subobjects and a cogenerating family $(C_i)_{i \in I}$. Then for locally small categories $\mathbb{B}$ a functor $U : \mathbb{C} \to \mathbb{B}$ has a left adjoint iff $U$ preserves small limits and infima of arbitrary families of subobjects.*

*Proof:* Of course, if $U$ has a left adjoint then it preserves all limits, i.e. in particular small limits and arbitrary intersections of subobjects.
For showing the reverse direction suppose that $U$ preserves small limits and arbitrary intersections of subobjects. We will show that for all objects $B$ of $\mathbb{B}$ the comma category $B{\downarrow}U$ has an initial object from which it then follows that $U$ has a left adjoint. From Lemma 9.1 and 9.2 it follows that $B{\downarrow}U$ is locally small and has small limits. As $\mathbb{C}$ has and $U$ preserves arbitrary intersections of subobjects it follows that $B{\downarrow}U$ has arbitrary intersections of subobjects, too. As $\mathbb{B}$ is locally small the collection $\bigcup_{i \in I} \mathbb{B}(B, U(C_i))$ is small, too, and one easily shows that it provides a cogenerating family for $B{\downarrow}U$. Thus, it follows by Lemma 9.4 that $B{\downarrow}U$ has an initial object. $\square$

In most examples the extra assumption of subobjects being closed under arbitrary intersections is redundant as in these cases the categories under consideration are *well-powered* in the following sense.

**Definition 9.3** *A category $\mathbb{C}$ is* well-powered *iff for every object $A$ of $\mathbb{C}$ the poset $\mathsf{Sub}_{\mathbb{C}}(A)$ of subobjects of $A$ is small.* $\diamond$

Obviously, if a category $\mathbb{C}$ is well-powered and has small limits then $\mathbb{C}$ has also arbitrary intersections of subobjects. Moreover, if $U : \mathbb{C} \to \mathbb{B}$ preserves small limits then $U$ preserves[13] also arbitrary intersections of subobjects.
This gives rise to the following even more useful version of the Special Adjoint Functor Theorem (SAFT).

**Theorem 9.3** *Let $\mathbb{C}$ be a category which is locally small, well-powered and small complete. If $\mathbb{C}$ admits a small cogenerating family and $U : \mathbb{C} \to \mathbb{B}$ with $\mathbb{B}$ locally small then $U$ has a left adjoint iff $U$ preserves small limits.*

---

[13]From preservation of pullbacks it follows that monos are preserved, too.

*Proof:* The claim is immediate from Theorem 9.2 because if $\mathbb{C}$ is well-powered and $U$ preserves small limits then $U$ preserves also arbitrary intersections of subobjects. $\square$

As an illustration of the power of SAFT we show that the inclusion $U$ of the category **CompHaus** of compact Hausdorff spaces into the category **Sp** has a left adjoint $\beta$ called *Stone-Čech compactification*. Obviously, both **CompHaus** and **Sp** are locally small, have small limits and are well-powered. Due to Urysohn's Separation Lemma continuous maps to the space $[0,1]$ separate points in compact Hausdorff spaces. Thus, by Theorem 9.3 the forgetful functor $U$ has a left adjoint $\beta$. Unwinding the proof of SAFT one can see that for a space $X$ its reflection[14] to **CompHaus** is given by $\eta_X : X \to \beta(X)$ where $\beta(X)$ is the closure of the image of the map

$$\eta_X : X \to \widetilde{X} : x \mapsto (f \mapsto f(x))$$

with $\widetilde{X} = \prod_{f \in \mathbf{Sp}(X,[0,1])}[0,1]$ (which is compact by Tychonoff's Theorem).

---

[14]a left adjoint to a full and faithful functor is commonly called a *reflection*

# 10 Monads

Every adjunction $F \dashv U : \mathbb{B} \to \mathbb{A}$ induces a so-called *monad* $(T, \eta, \mu)$ on $\mathbb{A}$ where $T = U \circ F : \mathbb{A} \to \mathbb{A}$, $\eta : \mathsf{Id}_{\mathbb{A}} \to T$ is the unit of the adjunction and $\mu : T^2 \to T$ is given by $\mu_A = U \varepsilon_{FA}$. Using the triangular equalities for unit $\eta$ and counit $\varepsilon$ of the adjunction $F \dashv U$ it is a straightforward exercise(!) to show that $(T, \eta, \mu)$ is a *monad* in the sense of the following definition.

**Definition 10.1** (monad)
*A* monad *on a category* $\mathbb{C}$ *is given by a triple* $(T, \eta, \mu)$ *with* $T : \mathbb{C} \to \mathbb{C}$, $\eta : \mathsf{Id}_{\mathbb{C}} \to T$ *and* $\mu : T^2 \to T$ *satisfying the following equalities*

$$\mu \circ \eta_T = \mathsf{id}_T = \mu \circ T\eta \qquad and \qquad \mu \circ T\mu = \mu \circ \mu_T$$

*as illustrated by the diagrams*
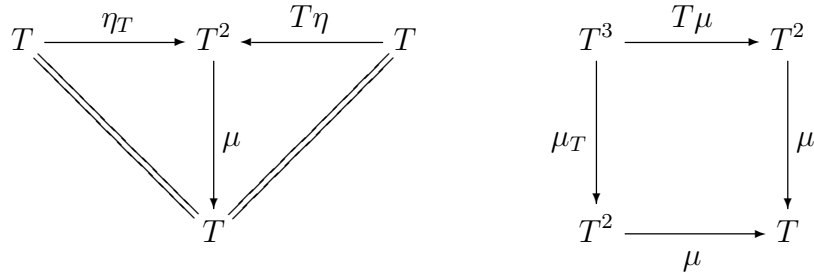


*The natural transforms* $\eta$ *and* $\mu$ *are called* unit *and* multiplication *of the monad, respectively.* $\diamond$

Of course, there arises the question to which extent every monad is induced by an adjunction. The answer will be positive but in most cases there is not a unique such adjunction even up to isomorphism. But we will show that there is a *minimal* and a *maximal* solution to this problem, the so-called *Kleisli category* $\mathbb{C}_T$ and the so-called *category* $\mathbb{C}^T$ *of Eilenberg-Moore algebras*, respectively.
But before we will give a few examples of monads which shall provide some intuition for this notion. Let $\mathbb{C}$ be some category of algebras (e.g. **Mon**, **Grp**, **Ab** etc.) and $U$ be the forgetful functor from $\mathbb{C}$ to **Set**. As seen in the previous section this forgetful functor $U$ has a right adjoint $F : \textbf{Set} \to \mathbb{C}$ sending $X$ to the free algebra over $X$. Then for $T = UF$ we have that $T(X)$ is the underlying set of $F(X)$, i.e. $T(X) = X^*$ in case of $\mathbb{C} = \textbf{Mon}$. The

function $\eta_X : X \to T(X)$ inserts the generators into the generated algebra, i.e. in case $\mathbb{C} = \mathbf{Mon}$ we have that $\eta_X(x)$ is the single letter word consisting of the letter $x$. More complex is $\mu_X : T^2(X) \to T(X)$ sending a term $t$ over $T(X)$ to the term $\mu_X(t)$ over $X$ by interpreting the generators occuring in $t$ as elements of $T(X)$. In case $\mathbb{C} = \mathbf{Mon}$ the operation $\mu$ sends $w_1 \ldots w_k$ in $X^{**}$ with $w_i = x_{i,1} \ldots x_{i,n_i}$ to $x_{1,1} \ldots x_{1,n_1} \ldots x_{k,1} \ldots x_{k,n_k}$ in $X^*$. Notice that in computer science this operation is called "flattening" (of lists).

Generally, in computer science monads play an important role because they allow one to capture so-called "computational effects" as for example non-termination. Consider the functor $L : \mathbf{Set} \to \mathbf{Set}$ (called *lifting*) sending $X$ to $L(X) = X+\{*\}$ where $*$ is a distinguished element (representing nontermination). For $f : X \to Y$ the map $L(f) : L(X) \to L(Y)$ behaves on $X$ like $f$ and sends $*$ to $*$. Let $\eta_X$ be the function including $X$ into $X+\{*\}$ and $\mu_X : L^2(X) \to L(X)$ be the function sending the "two different "$*$'s" of $L^2(X)$ to $*$ and elements of $X$ to themselves. One readily checks that this defines a monad on $\mathbf{Set}$ for which we will see later on that it captures non-termination (because the Kleisli category $\mathbf{Set}_L$ will turn out as isomorphic to the category of sets and partial functions).

Slightly more refined examples of "computational monads" are the following ones.

Fix a set $R$ of so-called "responses" and let $T : \mathbf{Set} \to \mathbf{Set}$ be the functor sending $X$ to $T(X) = R^{R^X}$ and $f : X \to Y$ to

$$T(f) = R^{R^f} : \phi \mapsto (p \mapsto \phi(p \circ f))$$

i.e. $T(f) = \lambda\phi{:}R^{R^X}.\lambda f{:}R^Y.\phi(\lambda x{:}X.p(f(x)))$ in the language of simply typed $\lambda$-calulus as familiar from functional programming. The unit $\eta_X : X \to R^{R^X}$ sends $x \in X$ to the map $\eta_X(x) : R^X \to R : p \mapsto p(x)$ and the multiplication $\mu_X$ is given by $R^{\eta_{R^X}}$, i.e.

$$\mu_X(\Phi) = \lambda p{:}R^X.\Phi(\eta_{R^X}(p)) = \lambda p{:}R^X.\Phi(\lambda\phi{:}R^{R^X}.\phi(p))$$

in $\lambda$-calculus notation. This monad is called "continuation monad" as a "computation" $\phi \in R^{R^X}$ is thought of as a "generalised point of $X$" described by its "response" $\phi(p)$ to an "observation" or "continuation" $p$. We leave it as an exercise(!) to show that the continuation monad is induced by the adjunction $R^{(\cdot)} \dashv (R^{(\cdot)})^{\mathsf{op}} : \mathbf{Set}^{\mathsf{op}} \to \mathbf{Set}$.

Another example is the so-called *state monad* which is described as follows. Fix a set $S$ of states and define $T(X) = (X \times S)^S$ whose elements $c$ are

thought of as "elements of $X$ with a side effect" because given a state $s \in S$ applying $c$ to it gives rise to a pair $c(s) = (x, s')$ where $x \in X$ is the resulting value and $s'$ is the (possibly) altered state after executing $c$ in state $s$. The unit of the state monad is given by $\eta_X(x)(s) = (x, s)$ sending $x \in X$ to the side-effect free computation $s \mapsto (x, s)$. Its multiplication $\mu$ is given by

$$\mu_X(C) = \lambda s{:}S.\pi_1(C(s))(\pi_2(C(s)))$$

where $C : S \to (X{\times}S)^S{\times}S$. We leave it as an exercise(!) to show that the state monad is induced by the adjunction $(\cdot){\times}S \dashv (\cdot)^S : \mathbf{Set} \to \mathbf{Set}$.

Notice that just for ease of exposition we have defined the above computational monads on **Set**. However, for the purposes of denotational semantics they should rather be defined on the cartesian closed category **PreDom** of predomains[15] and Scott continuous functions between them (which does contain **Set** as a full subcategory). Moreover, the monads $T : \mathbf{PreDom} \to \mathbf{PreDom}$ should be "enriched over **PreDom**" in the sense that the morphism parts of $T$ are themselves continuous functions in **PreDom**. Notice that this latter requirement makes sense for all cartesian closed categories $\mathbb{C}$. Such monads are usually called "strong monads" where the "enrichment" is often referred to as "the strength of the monad".

Next we will describe the construction of the Kleisli category $\mathbb{C}_T$ for a monad $T$ on $\mathbb{C}$.

**Theorem 10.1** (Kleisli category)
*Let $(T, \eta, \mu)$ be a monad on a category $\mathbb{C}$. Then we may construct a category $\mathbb{C}_T$, the so-called* Kleisli category *for $T$, with the same objects as $\mathbb{C}$ but with $\mathbb{C}_T(A, B) = \mathbb{C}(A, TB)$ where composition in $\mathbb{C}_T$ is given by*

$$g \circ_{\mathbb{C}_T} f = \mu_C \circ Tg \circ f$$

*as illustrated in the diagram*



---

[15] A predomain is a partial order with suprema of $\omega$-chains. A Scott continuous function between predomains is a monotonic mapping preserving suprema of $\omega$-chains.

*where $g^\dagger = \mu_C \circ Tg$ is called the "lifting of g". In $\mathbb{C}_T$ the identity on A is given by $\eta_A : A \to TA$ in $\mathbb{C}$.*

*Proof:* Straightforward equational reasoning using the defining equations for a monad. The details are left as an exercise(!) to the reader. $\square$

**Remark** Notice that one may equivalently define monads in terms of lifting $(\cdot)^\dagger$ (instead of multiplication) in the following way. One postulates a function $T : \mathsf{Ob}(\mathbb{C}) \to \mathsf{Ob}(\mathbb{C})$ together with maps $\eta_A : A \to TA$ and a lifting operation

$$(\cdot)^\dagger : \mathbb{C}(A, TB) \to \mathbb{C}(TA, TB)$$

for all $A, B \in \mathsf{Ob}(\mathbb{C})$ satisfying the requirements

(1) $\eta_A{}^\dagger = \mathsf{id}_{TA}$

(2) $f^\dagger \circ \eta_A = f$

(3) $(g^\dagger \circ f)^\dagger = g^\dagger \circ f^\dagger$

for all $f : A \to TB$ and $g : B \to TC$ in $\mathbb{C}$. One easily checks (exercise!) that requirements (1), (2) and (3) are satisfied for the lifting operation $(\cdot)^\dagger$ as defined in the proof of Theorem 10.1. Similarly one checks that from data $T$, $\eta$ and $(\cdot)^\dagger$ satisfying conditions (1), (2) and (3) one can define a monad by putting $T(f) = (\eta_B \circ f)^\dagger$ for $f : A \to B$ in $\mathbb{C}$ and $\mu_A = \mathsf{id}_{TA}{}^\dagger$. We leave it as an exercise(!) to show that these two processes are inverse to each other.

Next we will show that every monad $(T, \eta, \mu)$ on $\mathbb{C}$ is induced by an adjunction $F_T \dashv U_T : \mathbb{C}_T \to \mathbb{C}$.

**Theorem 10.2** *Let $(T, \eta, \mu)$ be a monad on $\mathbb{C}$. Then putting for $f : X \to Y$ in $\mathbb{C}$ and $h : X \to Y$ in $\mathbb{C}_T$*

$$F_T(X) = X \qquad and \qquad F_T(f) = \eta_Y \circ f$$
$$U_T(X) = TX \qquad and \qquad U_T(h) = \mu_Y \circ T(h)$$

*gives rise to an adjunction $F_T \dashv U_T : \mathbb{C}_T \to \mathbb{C}$ inducing the original monad $(T, \eta, \mu)$.*

*Proof:* It is a straightforward exercise(!) to verify that $F_T$ and $U_T$ are actually functors. The natural isomorphism $\varphi$ is given by

$$\varphi_{X,Y} : \mathbb{C}_T(F_T(X), Y) \to \mathbb{C}(X, U_T(Y)) : f \mapsto f .$$

We leave it as an exercise(!) to verify that $\varphi$ is actually natural. $\square$

**Theorem 10.3** *Let $F \dashv U : \mathbb{D} \to \mathbb{C}$ be an adjunction and $(T, \eta, \mu)$ the induced monad. Then there exists a unique functor $K : \mathbb{C}_T \to \mathbb{D}$ with $F = K \circ F_T$ and $U_T = U \circ K$ as illustrated in*



*Proof:* For $X \in \mathbb{C}$ we have $K(X) = K(F_T(X)) = F(X)$. For $h : X \to Y$ in $\mathbb{C}_T$, i.e. $h : X \to TY$ in $\mathbb{C}$, we have $U(K(h)) = U_T(h) = \mu_Y \circ T(h)$ from which it follows that $U(K(h)) \circ \eta_X = \mu_Y \circ T(h) \circ \eta_X = \mu_Y \circ \eta_{TY} \circ h = h$ and accordingly $K(h) = \varepsilon_{FY} \circ F(h)$ as follows from the universal property of $\eta_X$. We leave it as an exercise(!) to the reader to verify that the so defined $K$ is actually a functor from $\mathbb{C}_T$ to $\mathbb{D}$ and, moreover, satisfies the equalities $F = K \circ F_T$ and $U_T = U \circ K$. $\qquad \square$

Thus $F_T \dashv U_T$ may be considered as the *initial* factorization of the monad $T$ into an adjunction inducing it. Next we will construct a *terminal* such factorization for every monad $T$.

**Theorem 10.4** (Eilenberg-Moore algebras)
*Let $(T, \eta, \mu)$ be a monad on $\mathbb{C}$. Then an (Eilenberg-Moore) algebra w.r.t. $T$ (or $T$-algebra) is a pair $(A, \alpha)$ where $A$ is an object of $\mathbb{C}$ and $\alpha : TA \to A$ making the diagrams*



*commute. An* algebra morphism *from $(A, \alpha)$ to $(B, \beta)$ is a morphism $h :$*

$A \to B$ in $\mathbb{C}$ *making the diagram*

$$
\begin{array}{ccc}
TA & \xrightarrow{\;Th\;} & TB \\
\Big\downarrow{\alpha} & & \Big\downarrow{\beta} \\
A & \xrightarrow{\;h\;} & B
\end{array}
$$

*commute.*

*Eilenberg-Moore algebras and their morphisms give rise to a category $\mathbb{C}^T$ with composition and identities inherited from $\mathbb{C}$.*

*Proof:* Trivial verification left to the reader as an exercise!. $\qquad\square$

Notice that for every object $A$ in $\mathbb{C}$ from the defining equations for a monad it follows that $(TA, \mu_A)$ is an algebra w.r.t. $T$, namely the free algebra generated by $A$ as follows from the next theorem.

**Theorem 10.5** *Let $(T, \eta, \mu)$ be a monad on $\mathbb{C}$. Then putting*

$$
\begin{aligned}
F^T(X) &= (TX, \mu_X) & \text{and} && F^T(f) &= T(f) \\
U^T(A, \alpha) &= A & \text{and} && U^T(h) &= h
\end{aligned}
$$

*gives rise to an adjunction $F^T \dashv U^T : \mathbb{C}^T \to \mathbb{C}$ inducing the original monad $(T, \eta, \mu)$.*

*Proof:* We leave it as a straightforward exercise(!) to the reader to verify that $F^T$ and $U^T$ as defined above are actually functors.
For showing that $F^T \dashv U^T$ we construct the desired natural isomorphism $\varphi$ as follows: an $h \in \mathbb{C}^T(F^T(X), (A, \alpha))$ is sent to $\varphi(h) = h \circ \eta_X \in \mathbb{C}(X, U^T(A, \alpha))$. That the so defined $\varphi$ is actually a natural transformation is left to the reader as an exercise(!). The map $\varphi$ is one-to-one as if $h_1, h_2 : F^T(X) \to (A, \alpha)$ are algebra homorphisms with $\varphi(h_1) = \varphi(h_2)$, i.e. $h_1 \circ \eta_X = h_2 \circ \eta_X$, then we have

$$
h_1 = h_1 \circ \mu_X \circ T\eta_X = \alpha \circ Th_1 \circ T\eta_X = \alpha \circ Th_2 \circ T\eta_X = h_2 \circ \mu_X \circ T\eta_X = h_2
$$

For showing that $\varphi$ is onto suppose $f : X \to A = U^T(A, \alpha)$. Then $\alpha \circ T(f) : F^T(X) \to (A, \alpha)$ as can be seen from the following commuting diagram

$$
\begin{array}{ccccc}
T^2 X & \xrightarrow{\ T^2 f\ } & T^2 A & \xrightarrow{\ T\alpha\ } & TA \\
\Big\downarrow{\mu_X} & & \Big\downarrow{\mu_A} & & \Big\downarrow{\alpha} \\
TX & \xrightarrow[\ Tf\ ]{} & TA & \xrightarrow[\ \alpha\ ]{} & A
\end{array}
$$

because $T(\alpha \circ T(f)) = T\alpha \circ T^2 f$ and $\varphi(\alpha \circ T(f)) = \alpha \circ Tf \circ \eta_X = \alpha \circ \eta_A \circ f = f$ as desired.

Obviously, the unit of $F^T \dashv U^T$ is the original $\eta$. For $(A, \alpha)$ we have $\varepsilon_{(A,\alpha)} = \alpha$ and thus $U^T \varepsilon_{F^T X} = \mu_X$. Thus $F^T \dashv U^T$ induces the original monad. $\quad\square$

**Theorem 10.6** *Let $(T, \eta, \mu)$ be the monad on $\mathbb{C}$ induced by an adjunction $F \dashv U : \mathbb{D} \to \mathbb{C}$. Then there exists a unique functor $L : \mathbb{D} \to \mathbb{C}^T$ making the diagram*

$$
\begin{array}{ccccc}
\mathbb{C} & \xrightarrow{\ F^T\ } & \mathbb{C}^T & \xrightarrow{\ U^T\ } & \mathbb{C} \\
 & {\scriptstyle F}\searrow & \Big\uparrow{L} & \nearrow{\scriptstyle U} & \\
 & & \mathbb{D} & &
\end{array}
$$

*commute.*

*Proof:* We determine the functor $L$ from the requirements (1) $F^T = L \circ F$ and (2) $U = U^T \circ L$. From (2) it follows that $L(g) = U(g)$ for all morphisms $g$ in $\mathbb{D}$ and that $L(A)$ is of the form $(U(A), \alpha_A)$. From (1) it follows that $L(F(X)) = F^T(X) = (TX, \mu_X)$. For determining the structure map $\alpha_A$ of $L(A)$ consider the algebra morphism $L(\varepsilon_A) : L(F(UA)) \to L(A)$ where $\varepsilon_A : FUA \to A$ is the counit of $F \dashv U$ at $A$. As $U^T(L(\varepsilon_A)) = U(\varepsilon_A)$ and

$L(F(UA)) = F^T(UA) = (T(UA), \mu_{UA})$ we have

$$
\begin{array}{ccc}
UFUFUA & \xrightarrow{\;UFU\varepsilon_A\;} & UFUA \\
\downarrow{\scriptstyle \mu_{UA}} & & \downarrow{\scriptstyle \alpha_A} \\
UFUA & \xrightarrow[\;U\varepsilon_A\;]{} & UA
\end{array}
$$

and, therefore, it follows that

$$
\alpha_A = \alpha_A \circ UFU\varepsilon_A \circ UF\eta_{UA} = U\varepsilon_A \circ \mu_{UA} \circ UF\eta_{UA} = U\varepsilon_A
$$

because $\mu_{UA} = U\varepsilon_{FUA}$.

According to these consideration the object part of $L$ is given by

$$
L(A) = (UA, U\varepsilon_A)
$$

and its morphism part by $L(g) = U(g)$.

First we show that $(UA, U\varepsilon_A))$ is actually an algebra. That

$$
U\varepsilon_A \circ \eta_{UA} = \mathsf{id}_{UA}
$$

is just one of the triangle equalities of the adjunction $(F, U, \eta, \varepsilon)$. That

$$
\begin{array}{ccc}
UFUFUA & \xrightarrow{\;UFU\varepsilon_A\;} & UFUA \\
\downarrow{\scriptstyle U\varepsilon_{FUA}} & & \downarrow{\scriptstyle U\varepsilon_A} \\
UFUA & \xrightarrow[\;U\varepsilon_A\;]{} & UA
\end{array}
$$

follows from functoriality of $U$ and naturality of $\varepsilon : FU \Rightarrow \mathsf{Id}$.

That $L$ is functorial follows from functoriality of $U$ and $L(g) = U(g)$.

Requirement (1) holds as for $X \in \mathsf{Ob}(\mathbb{C})$ we have

$$
L(F(X)) = (UFX, U\varepsilon_{FX}) = (TX, \mu_X) = F^T(X)
$$

and for $f : X \to Y$ in $\mathbb{C}$ we have

$$
L(F(f)) = U(F(f)) = T(f) = F^T(f) \;.
$$

Requirement (2) holds as for $A \in \mathsf{Ob}(\mathbb{D})$ we have

$$U^T(L(A)) = U^T(UA, U\varepsilon_A) = UA$$

and for $g : A \to B$ in $\mathbb{D}$ we have

$$U^T(L(g)) = U^T(U(g)) = U(g) \ .$$

Thus, we have shown that $L$ is actually the unique functor from $\mathbb{D}$ to $\mathbb{C}^T$ with $L \circ F = F^T$ and $U^T \circ L = U$. $\qquad\qquad\square$

If $U : \mathbf{Mon} \to \mathbf{Set}$ is the forgetful functor and $F \dashv U$ its left adjoint then for the induced monad $T = UF$ one readily checks that $\mathbf{Set}^T$ is equivalent to the category of monoids and their homomorphisms and that $\mathbf{Set}_T$ is the full subcategory of free monoids. As not every monoid is free the comparison functor $\mathbf{Set}_T \hookrightarrow \mathbf{Set}^T$ is not essentially surjective.

**Beck's Monadicity Theorem**

There is a famous theorem of J. Beck characterising those adjunctions $F \dashv U : \mathbb{D} \to \mathbb{C}$ where the functor $L$ of Theorem 10.6 is an equivalence, i.e. those adjunctions whose right adjoint—up to equivalence—is a forgetful functor for a category of algebras. A functor $U : \mathbb{D} \to \mathbb{C}$ is called *monadic* iff it has a left adjoint $F$ such that the comparison functor $L : \mathbb{D} \to \mathbb{C}^T$ is an equivalence where $T = UF$ is the monad induced by $F \dashv U$. For formulating *Beck's Monadicity Theorem* we need the following terminology.

**Definition 10.2** *A pair $f, g : A \to B$ is* reflexive *iff they have a common section, i.e. an morphism $s : B \to A$ with $f \circ s = \mathsf{id}_B = g \circ s$.*
*A pair $f, g : A \to B$ is called* contractible *iff there exist a map $q : B \to C$ and sections $t$ and $s$ of $f$ and $q$ respectively making the diagram*

$$
\begin{array}{ccccc}
B & \xrightarrow{\ t\ } & A & \xrightarrow{\ f\ } & B \\
{\scriptstyle q}\downarrow & & {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle q} \\
Q & \xrightarrow[\ s\ ]{} & B & \xrightarrow[\ q\ ]{} & Q
\end{array}
$$

*commute.*
*If $U : \mathbb{D} \to \mathbb{C}$ then a pair $f, g : X \to Y$ is $U$-contractible iff $Uf, Ug : UX \to UY$ is a contractible pair in $\mathbb{D}$.* $\qquad\qquad\Diamond$

Notice that for a contractible pair $f, g : A \to B$ the map $q$ is necessarily a coequalizer of $f$ and $g$.

**Theorem 10.7** *A functor $U : \mathbb{C} \to \mathbb{D}$ is monadic iff the following conditions hold*

(1) *$U$ has a left adjoint*

(2) *$U$ reflects isomorphism*

(3) *$\mathbb{C}$ has coequalizers of reflexive $U$-contractible pairs and $U$ preserves them.*

*Proof:* For a proof of Beck's Theorem see e.g. section 3.3. of [BW1]. $\qquad\square$

# Part II    CATEGORICAL LOGIC

# 11 Cartesian Closed Categories and $\lambda$–Calculus

First we recall the definition of exponentials and introduce the notion of a cartesian closed category.

**Definition 11.1** *A category $\mathbb{C}$ is called* cartesian closed *iff $\mathbb{C}$ has finite products and exponentials, i.e. for all $A, B \in \mathsf{Ob}(\mathbb{C})$ there exists an object $B^A$ in $\mathbb{C}$ together with a morphism $\varepsilon : B^A \times A \to B$ such that for every $f : C \times A \to B$ there exist a unique map $\lambda(f) : C \to B^A$ with $f = \varepsilon \circ (\lambda(f) \times A)$ as indicated in the diagram*



*We write* ccc *as an abbreviation for "cartesian closed category".* $\Diamond$

Notice that $\mathbb{C}$ is a ccc iff it has finite products and for every $A, B \in \mathsf{Ob}(\mathbb{C})$ the presheaf $\mathbb{C}((-) \times A, B)$ is representable.

Recall that for maps $f : A \to B$, $g : C \to D$ in $\mathbb{C}$ the map $f \times g$ is defined as the unique morphism $h : A \times C \to B \times D$ making the diagram



commute, i.e. $f \times g = \langle f \circ \pi_1, g \circ \pi_2 \rangle$. Moreover, for $f_i : B \to A_i$ ($i{=}1, 2$) and $g : C \to B$ it holds that

$$\langle f_1, f_2 \rangle \circ g = \langle f_1 \circ g, f_2 \circ g \rangle$$

because $\pi_i \circ \langle f_1 \circ g, f_2 \circ g \rangle = f_i \circ g = \pi_i \circ \langle f_1, f_2 \rangle \circ g$ for $i{=}1, 2$.

The following claim is easy to show and will be used very often subsequently.

**Lemma 11.1** *In a cartesian closed category $\mathbb{C}$ we have*

$$\lambda(f) \circ g = \lambda(f \circ (g \times A))$$

*for morphism $f : C \times A \to B$ and $g : D \to C$ in $\mathbb{C}$.*

*Proof:* Left as an exercise! $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

The following are examples of ccc's

(1) the category **Set** of sets and functions

(2) the category **Pos** of posets and monotone functions

(3) the category **Cpo** of chain complete posets and (Scott) continuous functions.

Nonexamples are the categories of (abelian) groups, vectors spaces and topological spaces.[16]

## 11.1 Exponentials in Presheaf Categories

A very wide class of ccc's is provided by categories of presheaves, i.e. categories of the form $\widehat{\mathbb{C}}$ for some small category $\mathbb{C}$. Such categories have small limits (that are computed pointwise) and, thus, in particular all finite products. We will show next that $\widehat{\mathbb{C}}$ has also all exponentials.
The Yoneda lemma tells us how exponentials have to look like provided they exist. Let $A$ and $B$ be objects in $\widehat{\mathbb{C}}$. Then by Yoneda we have for all objects $I$ in $\mathbb{C}$ that
$$B^A(I) \cong \widehat{\mathbb{C}}(\mathsf{Y}_{\mathbb{C}}(I), B^A) \cong \widehat{\mathbb{C}}(\mathsf{Y}_{\mathbb{C}}(I) \times A, B)$$

where the second isomorphism is due to the assumption that $\widehat{\mathbb{C}}$ is cartesian closed. The morphism part of the presheaf $B^A$ can be read off from the

---

[16]In abelian groups and vector spaces $0 \cong 1$ and, therefore, there is precisely one morphism from $A$ to $B$ whenever $B^A$ exists. A proof that **Sp** is not cartesian closed can be found on pp.353-355 of vol.2 of [Bor] as proof of his Proposition 7.1.2.

following commuting diagram

$$
\begin{array}{ccccc}
\widehat{\mathbb{C}}(\mathsf{Y}(I){\times}A,B) & \xrightarrow{\ \cong\ } & \widehat{\mathbb{C}}(\mathsf{Y}(I),B^A) & \xrightarrow{\ \cong\ } & B^A(I) \\
{\scriptstyle \widehat{\mathbb{C}}(\mathsf{Y}(u){\times}A,B)}\big\downarrow & & {\scriptstyle \widehat{\mathbb{C}}(\mathsf{Y}(u),B^A)}\big\downarrow & & \big\downarrow{\scriptstyle B^A(u)} \\
\mathbb{C}(\mathsf{Y}(J){\times}A,B) & \xrightarrow{\ \cong\ } & \widehat{\mathbb{C}}(\mathsf{Y}(J),B^A) & \xrightarrow{\ \cong\ } & B^A(J)
\end{array}
$$

where $u : J \to I$ in $\mathbb{C}$. Yoneda tells us also how to construct the natural isomorphism

$$
\lambda : \widehat{\mathbb{C}}((-){\times}A,B) \xrightarrow{\ \cong\ } \widehat{\mathbb{C}}((-),B^A)
$$

as from commutation of the diagram

$$
\begin{array}{ccc}
\widehat{\mathbb{C}}(C{\times}A,B) & \xrightarrow{\ \lambda_C\ } & \widehat{\mathbb{C}}(C,B^A) \\
{\scriptstyle \widehat{\mathbb{C}}(c{\times}A,B)}\big\downarrow & & \big\downarrow{\scriptstyle \widehat{\mathbb{C}}(c,B^A)} \\
\widehat{\mathbb{C}}(\mathsf{Y}(I){\times}A,B) & \xRightarrow{\ \lambda_{\mathsf{Y}(I)}\ } & \widehat{\mathbb{C}}(\mathsf{Y}(I),B^A)
\end{array}
$$

— pretending that $\lambda_{\mathsf{Y}(I)}$ is identity — we get that

$$
(1) \qquad \big[\lambda_C(\tau)_I(c)\big]_J(u,a) = \tau_J(C(u)(c),a)
$$

for all $\tau : C{\times}A \to B$, $u : J \to I$ and $a \in A(J)$. As the representing element $\varepsilon : B^A{\times}A \to B$ has to satisfy $\lambda_{B^A}(\varepsilon) = \mathsf{id}_{B^A}$ we get from (1) that

$$
(2) \qquad \varepsilon_I(\varphi,a) = \varphi_I(\mathsf{id}_I,a)
$$

for $\varphi : \mathsf{Y}(I){\times}A \to B$ and $a \in A(I)$.

The following theorem says that the above considerations actually give rise to exponentials.

**Theorem 11.1** *Let $\mathbb{C}$ be a small category. Then the category of presheaves $\widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ over $\mathbb{C}$ is cartesian closed. For $A,B \in \widehat{\mathbb{C}}$ their exponential $B^A$ is given by*

$$
B^A(I) = \widehat{\mathbb{C}}(\mathsf{Y}(I){\times}A,B) \qquad and \qquad B^A(u) = \widehat{\mathbb{C}}(\mathsf{Y}(u){\times}A,B)
$$

65

*for $I \in \mathsf{Ob}(\mathbb{C})$ and $u : J \to I$ in $\mathbb{C}$. The counit $\varepsilon : B^A \times A \to B$ (also called evaluation map) is given by*

$$\varepsilon_I(\varphi, a) = \varphi_I(\mathsf{id}_I, a)$$

*for $\varphi \in B^A(I)$ and $a \in A(I)$. For $\tau : C \times A \to B$ its exponential transpose $\lambda(\tau) : C \to B^A$ is given by*

$$\left[\lambda(\tau)_I(c)\right]_J(u, x) = \tau_J(C(u)(c), x)$$

*for $c \in C(I)$, $u : J \to I$ in $\mathbb{C}$ and $x \in A(J)$.*

*Proof:* It is an easy exercise (left to the reader) to check naturality of $\varepsilon$ and $\lambda(\tau)$ as defined above. It remains to verify that

$(\beta) \quad \varepsilon \circ (\lambda(\tau) \times A) = \tau \qquad$ for all $\tau : C \times A \to B$ and

$(\eta) \quad \lambda(\varepsilon \circ (\sigma \times A)) = \sigma \qquad$ for all $\sigma : C \to B^A$

i.e. that $\varepsilon \circ ((-) \times A)$ is a natural bijection inverse to $\lambda$ as desired. Condition $(\beta)$ holds since

$$\left[\varepsilon \circ (\lambda(\tau) \times A)\right]_I(c, a) = \varepsilon_I(\lambda(\tau)_I(c), a)) =$$
$$= \lambda(\tau)_I(c)(\mathsf{id}_I, a) =$$
$$= \tau_I(C(\mathsf{id}_I)(c), a) =$$
$$= \tau_I(c, a)$$

and condition $(\eta)$ holds since we have

$$\lambda(\varepsilon \circ (\sigma \times A))_I(c)(u, a) = (\varepsilon \circ (\sigma \times A))_J(C(u)(c), a) =$$
$$= \varepsilon_J(\sigma_J(C(u)(c)), a) =$$
$$= \sigma_J(C(u)(c))(\mathsf{id}_J, a) = \qquad \text{(naturality of } \sigma)$$
$$= B^A(u)(\sigma_I(c))(\mathsf{id}_J, a) =$$
$$= \left[\sigma_I(c)\right]_J(\mathsf{Y}(u)(\mathsf{id}_J), a) =$$
$$= \sigma_I(c)(u, a)$$

where $u : J \to I$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Admittedly, the construction of exponentials in presheaf categories may appear as somewhat complicated at least at first sight. However, they tend to look much nicer in particular cases like the following ones.

(1) Suppose $\mathbb{C}$ is a group, i.e. a category $\mathbb{G}$ with a single object $*$ all whose morphisms are isomorphisms. Then for objects $A$ and $B$ in $\widehat{\mathbb{G}}$ according to the general construction above the underlying set $B^A(*)$ of the exponential $B^A$ consists of all natural transformations $\tau : \mathsf{Y}(*) \times A \to B$. Obviously, such a $\tau$ is uniquely determined by the map $t : A(*) \to B(*) : a \mapsto \tau(1, a)$ where we write $1$ for $\mathsf{id}_*$ because by naturality of $\tau$ we have

$$\tau(g, a) = \tau(g, A(g)(A(g^{-1})(a))) = B(g)(\tau(1, A(g^{-1})(a)) = B(g)(t(A(g^{-1})(a))$$

i.e. $\tau(g, -) = B(g) \circ t \circ A(g^{-1})$. Moreover, one easily checks that for arbitrary functions $t : A(*) \to B(*)$ there is a unique natural transformation $\tau$ with $\tau(1, a) = t(a)$, namely $\tau(g, a) = B(g)(t(A(g^{-1})(a)))$.[17] Thus, the exponential $B^A$ in $\widehat{\mathbb{G}}$ is isomorphic to the set $B(*)^{A(*)}$ on which $\mathbb{G}$ acts by sending $t \in B(*)^{A(*)}$ to $B^A(g)(t) = B(g) \circ t \circ A(g^{-1})$ for $g$ in $\mathbb{G}$.[18]
It is an instructive exercise(!) to explicitate how the global elements $1 \to B^A$ correspond to the morphisms from $A$ to $B$ in $\widehat{\mathbb{G}}$.

---

[17]We have to show that

$$B(h)(\tau(g, a)) = \tau(\mathsf{Y}(*)(h)(g), A(h)(a))$$

for which purpose we explicitate both sides of the equation. For the left hand side we have

$$B(h)(\tau(g, a)) = (B(h) \circ B(g) \circ t \circ A(g^{-1}))(a)$$

and for the right hand side we have

$$\tau(\mathsf{Y}(*)(h)(g), A(h)(a)) = \tau(gh, A(h)(a)) = (B(gh) \circ t \circ A((gh)^{-1}))(A(h)(a)) =$$
$$= B(gh) \circ t \circ A(h^{-1}g^{-1})(A(h)(a)) =$$
$$= B(gh) \circ t \circ A(g^{-1}) \circ A(h^{-1}) \circ A(h))(a) =$$
$$= (B(h) \circ B(g) \circ t \circ A(g^{-1}))(a)$$

and thus both sides of the equation are actually equal as desired.

[18]This action is actually *contravariant* as

$$B^A(gh)(t) = B(gh) \circ t \circ A(h^{-1}g^{-1}) = B(h) \circ B(g) \circ t \circ A(g^{-1}) \circ A(h^{-1}) = (B^A(h) \circ B^A(g))(t)$$

due to the contravariance of $A$ and $B$.

(2) Let **2** be the ordinal 2 considered as a category. We write $\alpha : 0 \to 1$ for the only nontrivial arrow of **2**. One readily checks (exercise!) that $\widehat{\mathbf{2}}$ is isomorphic to the comma category **Set↓Set** whose morphisms are commuting squares of the form

$$
\begin{array}{ccc}
X_1 & \xrightarrow{\ f_1\ } & Y_1 \\
\downarrow{\scriptstyle X_\alpha} & & \downarrow{\scriptstyle Y_\alpha} \\
X_0 & \xrightarrow[\ f_0\ ]{} & Y_0
\end{array}
$$

For objects $X = (X_1, X_\alpha, X_0)$ and $Y = (Y_1, Y_\alpha, Y_0)$ their exponential $Y^X$ can be constructed as follows

$(Y^X)_0 = Y_0^{X_0}$,

$(Y^X)_1 = \{(f_0, f_1) \in Y_0^{X_0} \times Y_1^{X_1} \mid Y_\alpha \circ f_1 = f_0 \circ X_\alpha\}$

$(Y^X)_\alpha(f_0, f_1) = f_0$.

The evaluation map $\varepsilon : Y^X \times X \to Y$ is given by

$$
\varepsilon_0(f, x) = f(x) \qquad \text{and} \qquad \varepsilon_1((f_0, f_1), x) = f_1(x) \ .
$$

We leave it as an exercise(!) to verify that these data do actually satisfy the required conditions.

## 11.2  Categorical semantics of typed $\lambda$-calculus in ccc's

We now define a formal language, the *typed $\lambda$-calculus*, and show how it serves as a so-called *internal language* for cartesian closed categories. Over the last 25 years the typed $\lambda$-calculus (originally invented in the 1930ies by A. Church) has experienced a *renaissance* in (theoretical) computer science as a foundation for *functional programming*. Objects of a ccc appear as interpretations of type expressions of the typed $\lambda$-calculus and its terms are interpreted as morphisms of the ccc under consideration. Contexts, i.e. variable declarations $\Gamma$ of the form $x_1:A_1, \ldots, x_n:A_n$, get interpreted as the cartesian product of (the interpretations of) the $A_i$ and variable $x_i$ in context $\Gamma$ gets interpreted as the projection on $A_i$.

But before describing the interpretation of typed $\lambda$-calculus in ccc's we carefully define its syntax.

The *types* (or *type expressions*) of simply typed $\lambda$–calculus are defined inductively from a collection of *base types* via the following rules

- every base type is a type

- if $A$ and $B$ are types then $A{\to}B$ is a type.

These are the type expressions of *traditional* typed $\lambda$-calculus. For the purposes of an interpretation in cartesian closed categories, i.e. from a semantical point of view, it turns out as convenient to close types under the following two additional rules

- 1 is a type (also called *unit type*)

- if $A$ and $B$ are types then $A{\times}B$ is a type.

*Contexts* are expressions of the form

$$x_1{:}A_1, \ldots, x_n{:}A_n$$

where $A_1, \ldots, A_n$ are type expressions and $x_1, \ldots, x_n$ is a list of pairwise distinct variables. The purpose of a context $x_1{:}A_1, \ldots, x_n{:}A_n$ is to *declare* the variables $x_1, \ldots, x_n$ together with their types $A_1, \ldots, A_n$. We use captial Greek letters $\Gamma, \Delta, \ldots$ as (meta)variables ranging over contexts.

Let $\mathbb{C}$ be a cartesian closed category. For interpreting the type expressions in $\mathbb{C}$ we first fix an assignment of $[\![B]\!] \in \mathsf{Ob}(\mathbb{C})$ for bases types $B$ which is extended to arbitrary type expressions via the following inductive clauses

- $[\![A{\to}B]\!] = [\![A]\!] \to [\![B]\!]$

- $[\![A{\times}B]\!] = [\![A]\!] \times [\![B]\!]$

- $[\![1]\!] = 1_{\mathbb{C}}$

where $X{\to}Y$ stands for the exponential $Y^X$ in $\mathbb{C}$ and $1_{\mathbb{C}}$ is a chosen terminal object in $\mathbb{C}$. Contexts are interpreted as cartesian products, i.e.

$$[\![x_1 : A_1, \ldots, x_n : A_n]\!] = [\![A_1]\!] \times \cdots \times [\![A_n]\!]$$

Next we will define the terms of typed $\lambda$-calculus. Terms not only have a type but are also defined *relative* to contexts in which their free variables are declared. We write

$$\Gamma \vdash t : A$$

for the judgement that *t is a term of type A in context* $\Gamma$. Traditionally, one would call already $t$ a term but, actually, one needs context $\Gamma$ to determine the meaning of $t$ (as we shall see the type $A$ of $t$ will be determined uniquely by $t$ and $\Gamma$). The interpretation of $\Gamma \vdash t : A$ will be a morphism from $[\![\Gamma]\!]$ to $[\![A]\!]$ in $\mathbb{C}$, i.e.

$$[\![\Gamma \vdash t : A]\!] : [\![\Gamma]\!] \to [\![A]\!]$$

and will be defined by structural recursion over terms. We will now introduce the various *term formation rules* together with the clauses fixing the interpretation of the constructed terms.

First of all variables declared in the context are terms, i.e.

$$\frac{}{x_1 : A_1, \ldots, x_n : A_n \vdash x_i : A_i} \; (\text{Var})$$

which is interpreted as

$$\pi_i : [\![x_1 : A_1, \ldots, x_n : A_n]\!] \longrightarrow [\![A_i]\!]$$

i.e. projection on the $i$-th component.

Explicit definition of functions is provided by the rule

$$\frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x{:}A.t : A{\to}B} \; (\lambda)$$

where

$$[\![\Gamma \vdash \lambda x{:}A.t : A{\to}B]\!] = \lambda([\![\Gamma, x : A \vdash t : B]\!])$$

i.e. $\varepsilon \circ ([\![\Gamma \vdash \lambda x{:}A.t : A{\to}B]\!] \times [\![A]\!]) = [\![\Gamma, x : A \vdash t : B]\!]$.

Application of functions to arguments is provided by the rule

$$\frac{\Gamma \vdash t : A{\to}B \qquad \Gamma \vdash s : A}{\Gamma \vdash t(s) : B} \; (\text{App})$$

where the resulting term is interpreted as

$$[\![\Gamma \vdash t(s)]\!] = \varepsilon \circ \langle [\![\Gamma \vdash t]\!], [\![\Gamma \vdash s]\!] \rangle$$

70

Notice that we have written $\Gamma \vdash t$ and $\Gamma \vdash s$ instead of $\Gamma \vdash t : A{\to}B$ and $\Gamma \vdash s$, respectively, which is justified by the fact that for derivable judgements $\Gamma \vdash s : A$ the type $A$ is determined uniquely simply because for each term former there is precisely one introduction rule.

This pleasant feature will be preserved by the remaining rules which we introduce next.

The following rule allows one to construct pairs

$$\frac{\Gamma \vdash t : A \quad \Gamma \vdash s : B}{\Gamma \vdash \langle t, s \rangle : A{\times}B} \; (\text{Pair})$$

where the constructed term will be interpreted as

$$[\![\Gamma \vdash \langle t, s \rangle : A{\times}B]\!] = \langle [\![\Gamma \vdash t : A]\!], [\![\Gamma \vdash s : B]\!] \rangle$$

The following two rules (for $i{=}1,2$)

$$\frac{\Gamma \vdash t : A_1{\times}A_2}{\Gamma \vdash \pi_i(t) : A_i} \; (\text{Proj}_i)$$

allow one to select components of pairs. Terms constructed by these rules will be interpreted as

$$[\![\Gamma \vdash \pi_i(t) : A_i]\!] = \pi_i \circ [\![\Gamma \vdash t : A_1{\times}A_2]\!]$$

with $\pi_i : [\![A_1]\!]{\times}[\![A_2]\!] \to [\![A_i]\!]$ the respective projection maps in $\mathbb{C}$.

The unique element of 1 may be constructed by the rule

$$\frac{}{\Gamma \vdash * : 1} \; (\text{Unit})$$

where $[\![\Gamma \vdash * : 1]\!] = !_{[\![\Gamma]\!]}$, the unique morphism from $[\![\Gamma]\!]$ to $1_{\mathbb{C}}$ in $\mathbb{C}$.

Up to now we have not considered equality between terms which we will do next. However, for this purpose we need a notion of *substitution* which on the semantical level will be interpreted as *composition* of morphisms in $\mathbb{C}$. Let $\Gamma \equiv x_1 : A_1, \ldots, x_n : A_n$ and $\Delta$ be contexts. A *substitution* $\sigma : \Delta \to \Gamma$ is an $n$-tuple $\sigma = \langle t_1, \ldots, t_n \rangle$ such that $\Delta \vdash t_i : A_i$ for $i{=}1, \ldots, n$ whose interpretation in $\mathbb{C}$ is given by

$$[\![\sigma]\!] = \langle [\![\Delta \vdash t_1]\!], \ldots, [\![\Delta \vdash t_n]\!] \rangle : [\![\Delta]\!] \to [\![\Gamma]\!] \; .$$

Now for $\Gamma \vdash t : A$ applying substitution $\sigma : \Delta \to \Gamma$ to it gives rise to a term $\Delta \vdash t[\sigma] : A$ whose interpretation is given by

$$\llbracket \Delta \vdash t[\sigma] : A \rrbracket = \llbracket \Gamma \vdash t : A \rrbracket \circ \llbracket \sigma : \Delta \to \Gamma \rrbracket$$

and which can be obtained syntactically by replacing (simultaneously) every free occurrence of $x_i$ by $t_i$, i.e.

$$t[\sigma] \equiv t[t_1, \ldots, t_n / x_1, \ldots, x_n]$$

for which we often employ the more readable notation $t[\vec{t}/\vec{x}]$, where, of course, one has to rename bound variables of $t$ in such a way that substituting $t_i$ for $x_i$ in $t$ does not bind free variables of $t_i$, i.e. we employ *capture-free substitution* as it is commonly called.

That capture-free substitution actually amounts to composition on the semantical level is the contents of the following lemma.

**Lemma 11.2** (Substitution Lemma)
*Whenever $\Gamma \vdash t : A$ and $\sigma : \Delta \to \Gamma$ then*

$$\llbracket \Delta \vdash t[\sigma] : A \rrbracket = \llbracket \Gamma \vdash t : A \rrbracket \circ \llbracket \sigma \rrbracket \quad .$$

*Proof:* We first show the claim for *variable substitutions* $v = \langle \vec{t} \rangle$ where all $t_i$ are variables declared in $\Delta$. Obviously, such variable substitutions are interpreted as target tuplings of projections.

We proceed by induction on derivations of $\Gamma \vdash t : A$. All cases are trivial with exception of the rule $(\lambda)$ for which we exhibit the argument. Suppose as induction hypothesis that for sequent $\Gamma, x{:}A \vdash t : B$ it holds that $\llbracket \Delta \vdash t[v] : B \rrbracket = \llbracket \Gamma, x{:}A \vdash t : A \rrbracket \circ \llbracket \sigma \rrbracket$ for all variable substitutions $v : \Delta \to \Gamma, x{:}A$. Let $v \equiv \langle v_1, \ldots, v_n \rangle : \Delta \to \Gamma$ be a variable substitution. Then we have

$$
\begin{aligned}
\llbracket \Gamma \vdash \lambda x{:}A.t \rrbracket \circ \llbracket v \rrbracket &= \lambda(\llbracket \Gamma, x{:}A \vdash t \rrbracket) \circ \llbracket v \rrbracket = \\
&= \lambda(\llbracket \Gamma, x{:}A \vdash t \rrbracket \circ (\llbracket v \rrbracket \times \mathsf{id}_{\llbracket A \rrbracket})) = \\
&= \lambda(\llbracket \Gamma, x{:}A \vdash t \rrbracket \circ \llbracket \langle v_1, \ldots, v_n, x \rangle \rrbracket) = \quad \text{(ind.hyp.)} \\
&= \lambda(\llbracket \Delta, x{:}A \vdash t[\vec{t}, x / \vec{x}, x] \rrbracket) = \\
&= \llbracket \Delta \vdash \lambda x{:}A.t[\vec{t}/\vec{x}] \rrbracket = \\
&= \llbracket \Delta \vdash \lambda x{:}A.t[v] \rrbracket
\end{aligned}
$$

where $\langle v_1, \ldots, v_n, x \rangle : \Delta, x : A \to \Gamma, x : A$ is again a variable substitution. Now we prove the claim for general substitutions $\sigma : \Delta \to \Gamma$. Again ee proceed by induction on derivations of $\Gamma \vdash t : A$. Again all cases are trivial with exception of the rule $(\lambda)$ for which we exhibit the argument. Suppose as induction hypothesis that for sequent $\Gamma, x{:}A \vdash t : B$ it holds that $[\![\Delta \vdash t[\sigma] : B]\!] = [\![\Gamma, x{:}A \vdash t : A]\!] \circ [\![\sigma]\!]$ for all substitutions $\sigma : \Delta \to \Gamma, x{:}A$. Now suppose $\tau \equiv \langle t_1, \ldots, t_n \rangle : \Delta \to \Gamma$. Then we have

$$[\![\Gamma \vdash \lambda x{:}A.t]\!] \circ [\![\tau]\!] = \lambda([\![\Gamma, x{:}A \vdash t]\!]) \circ [\![\tau]\!] =$$
$$= \lambda([\![\Gamma, x{:}A \vdash t]\!] \circ ([\![\tau]\!] \times \mathsf{id}_{[\![A]\!]})) = \quad (*)$$
$$= \lambda([\![\Gamma, x{:}A \vdash t]\!] \circ [\![\langle t_1, \ldots, t_n, x \rangle]\!]) = \quad \text{(ind.hyp.)}$$
$$= \lambda([\![\Delta, x{:}A \vdash t[\vec{t}, x/\vec{x}, x]]\!]) =$$
$$= [\![\Delta \vdash \lambda x{:}A.t[\vec{t}/\vec{x}]]\!] =$$
$$= [\![\Delta \vdash \lambda x{:}A.t[\tau]]\!]$$

where $\langle t_1, \ldots, t_n, x \rangle : \Delta, x : A \to \Gamma, x : A$. Notice that in $(*)$ we have used that $[\![\Delta, x : A \vdash t_i]\!] = [\![\Delta \vdash t_i]\!] \circ \pi$ where $\pi : [\![\Delta]\!] \times [\![A]\!] \to [\![\Delta]\!]$ as ensured by the already established claim for variable substitutions. $\qquad\square$

Next we will show that the $(\beta)$ and $(\eta)$ rule (as defined below in Theorem 11.2) of typed $\lambda$-calculus are valid under interpretations in arbitrary cartesian closed categories. For this purpose consider an arbitrary, but fixed ccc $\mathbb{C}$ together with an arbitrary, but fixed assignment of objects of $\mathbb{C}$ to base types.

**Theorem 11.2** (Soundness of typed $\lambda$-calculus)

(1) *If $\Gamma, x{:}A \vdash t : B$ and $\Gamma \vdash s : A$ then $[\![\Gamma \vdash (\lambda x{:}A.t)(s)]\!] = [\![\Gamma \vdash t[s/x]]\!]$, i.e. the rule*

$$(\beta) \quad \Gamma \vdash (\lambda x{:}A.t)(s) = t[s/x] : B$$

*holds.*

(2) *If $\Gamma \vdash t : A{\to}B$ then $[\![\Gamma \vdash \lambda x{:}A.t(x)]\!] = [\![\Gamma \vdash t]\!]$ for fresh $x$, i.e. the rule*

$$(\eta) \quad \Gamma \vdash \lambda x{:}A.t(x) = t : A{\to}B$$

*holds.*

*Proof:*

*ad* (1) : Let $\Gamma \equiv x_1{:}A_1, \ldots, x_n{:}A_n$ and

$$f = [\![\Gamma, x{:}A \vdash t : B]\!] \qquad \text{and} \qquad g = [\![\Gamma \vdash s : A]\!]$$

then

$$[\![\Gamma \vdash (\lambda x{:}A.t)(s)]\!] = \varepsilon \circ \langle \lambda(f), g \rangle = \varepsilon \circ (\lambda(f) \times \mathsf{id}_{[\![A]\!]}) \circ \langle \mathsf{id}_{[\![\Gamma]\!]}, g \rangle =$$
$$= f \circ \langle \mathsf{id}_{[\![\Gamma]\!]}, g \rangle = [\![\Gamma, x{:}A \vdash t]\!] \circ [\![\langle \vec{x}, s \rangle]\!] =$$
$$= [\![\Gamma \vdash t[s/x]]\!]$$

where the last equality follows from the Substitution Lemma 11.2.

*ad* (2) : Let $\Gamma \equiv x_1{:}A_1, \ldots, x_n{:}A_n$ then

$$[\![\Gamma \vdash \lambda x{:}A.t(x)]\!] = \lambda([\![\Gamma, x{:}A \vdash t(x)]\!]) =$$
$$= \lambda(\varepsilon \circ ([\![\Gamma \vdash t]\!] \times \mathsf{id}_{[\![A]\!]})) =$$
$$= [\![\Gamma \vdash t]\!]$$

where the second equality follows from $[\![\Gamma, x{:}A \vdash t]\!] = [\![\Gamma \vdash t]\!] \circ \pi_1$ (due to the Substitution Lemma 11.2) and $[\![\Gamma, x{:}A \vdash x]\!] = \pi_2$. $\qquad\square$

Traditionally, the (typed) $\lambda$-calculus is often considered without the rule $(\eta)$ (because $\lambda x{:}A.t(x)$ is considered anyway as more explicit than $t$). On the semantical level this can be simulated by considering the following weakening of the notion of cartesian closedness. One just requires for all objects $A$ and $B$ a (weak) exponential object $B^A$ together with an evaluation morphism $\varepsilon : B^A \times A \to B$ and for every $f : C \times A \to B$ a chosen morphism $\lambda(f) : C \to B^A$ satisfying $f = \varepsilon \circ (\lambda(f) \times \mathsf{id}_A)$ (without claiming that there is no other morphism $g$ with $f = \varepsilon \circ (g \times \mathsf{id}_A)$ !) where these choices are required to satisfy the constraint

$$\lambda(f) \circ h = \lambda(f \circ (h \times \mathsf{id}_A))$$

for all $h : D \to C$ in order to guarantee the validity of the Substitution Lemma. Categorically, this means that there is a (natural) section

$$\mathbb{C}((-){\times}A, B) \xrightarrow{\quad \lambda \quad} \mathbb{C}((-), B^A)$$

of the natural transformation $\varepsilon \circ ((-) \times \mathsf{id}_A) : \mathbb{C}((-), B^A) \to \mathbb{C}((-){\times}A, B)$.

74

**Completeness of Typed $\lambda$-Calculus w.r.t. CCCs**

For typed $\lambda$-calculus one may construct a *term model*, the so-called *classifying model*, where two terms (in context) receive the same interpretation if and only if they are provably equal. Moreover, every morphism in the classifying model arises as interpretation of some term, i.e. the model is *universal*[19] w.r.t. syntax.

The objects of the classifying model $\mathcal{C}$ are the contexts of typed $\lambda$-calculus and the morphisms from $\Delta$ to $\Gamma \equiv x_1{:}A_1, \ldots, x_n{:}A_n$ are tuples $\tau = \langle t_1, \ldots, t_n \rangle$ with $\Delta \vdash t_i : A_i$ for $i=1,...,n$ *modulo* provable equality, i.e. $\langle t_1, \ldots, t_n \rangle$ and $\langle t'_1, \ldots, t'_n \rangle$ get identified iff $\Delta \vdash t_i = t'_i : A_i$ is provable for $i=1,...,n$. Composition in $\mathcal{C}$ is given by (syntactical) substitution, i.e. for $\sigma : \Theta \to \Delta$ the composite $\tau \circ \sigma : \Theta \to \Gamma$ is given by $\langle t_1[\sigma], \ldots, t_n[\sigma] \rangle$. Binary products in $\mathcal{C}$ are given by juxtaposition, i.e. $\Gamma \times \Delta \equiv \Gamma, \Delta$ (where, of course, the variables declared in $\Gamma$ and $\Delta$ have to be made distinct by appropriate[20] renaming) and a terminal object in $\mathcal{C}$ is provided by the empty context. For contexts $\Gamma \equiv x_1{:}A_1, \ldots, x_n{:}A_n$ and $\Delta \equiv y_1{:}B_1, \ldots, y_m{:}B_m$ their exponential $\Delta^\Gamma$ is given by the context $f_1{:}B_1{}^\Gamma, \ldots, f_m{:}B_m{}^\Gamma$ where $B_j{}^\Gamma \equiv A_1 \to A_2 \to \ldots \to A_n \to B_j$. The evaluation map

$$\Delta^\Gamma \times \Gamma \xrightarrow{\varepsilon} \Delta$$

is given by $\varepsilon \equiv \langle f_1(x_1) \ldots (x_n), \ldots, f_m(x_1) \ldots (x_n) \rangle$.

Obviously, the interpretation of typed $\lambda$-calculus in $\mathcal{C}$ which interprets base types $B$ as $x{:}B$ has the property that types and contexts get interpreted as themselves and that $\Gamma \vdash t : A$ gets interpreted as the equivalence class of $\langle t \rangle$ modulo provable equivalence. Accordingly, the interpretation of $\Gamma \vdash t : A$ and $\Gamma \vdash t' : A$ in $\mathcal{C}$ are equal if and only if $\Gamma \vdash t = t' : A$ is provable. The model $\mathcal{C}$ is also universal because a morphism from $\Delta$ to $\Gamma$ as given by the equivalence class of a substitution $\tau \equiv \langle t_1, \ldots, t_n \rangle$ arises as the interpretation of this substitution $\tau$.

From these considerations the following theorem is immediate.

---

[19]This terminology comes from semantics of programming languages where a model $\mathcal{M}$ is called universal for a language $\mathcal{L}$ iff every object in $\mathcal{M}$ arises as interpretation of some program (i.e. closed term) of $\mathcal{L}$. Compare this with the notion of *Turing universal* meaning that for every computable function there is a program denoting it. From the latter point of view universality is a property of a language w.r.t. a given model whereas the former use of *universal* rather expresses a property of a model w.r.t. a given language!

[20]actually, we consider contexts up to renaming of the variables bound in them.

**Theorem 11.3** (Completeness of typed $\lambda$-calculus) *If $\Gamma \vdash t = s : A$ holds in all cartesian closed models then $\Gamma \vdash t = s : A$ is provable.*

Notice that our proof of completeness would not work if we required that all ccc's are *wellpointed* in the sense that morphisms $f, g : A \to B$ are already equal if $f \circ a = g \circ a$ for all $a : 1 \to A$ (which is the case for all set theoretic models).[21]

---

[21]Nevertheless, the typed $\lambda$-calculus is actually complete w.r.t. well-pointed models. The proof of this fact, however, requires a more intelligent technique and can be found in H. Friedman *Equality between functionals* pp.22–37, Lecture Notes in Math., Vol. 453, Springer, Berlin, 1975.

# 12  Elementary Toposes

The aim of this section is to introduce a class of categories, called *elementary toposes*, which are "well-behaved" to such an extent that they allow one to interpret (almost) all constructions which can be found in "real" mathematics traditionally based on set theory (see [LR]). As we have seen in the previous section in cartesian closed categories one may construct arbitrary function spaces. However, e.g. for the development of analysis, one also needs *power objects* $\mathcal{P}(A)$ which "internalize" the collection of subobjects of $A$ analogously to the way how the exponential $B^A$ "internalizes" the collection of all morphisms from $A$ to $B$. In particular, this implies that the object $\Omega = \mathcal{P}(1)$ of so-called *truth values* is available in an elementary topos. From this one will be able to construct the usual logical operations on $\Omega$ like $\wedge, \Rightarrow, \vee, \forall, \exists$ and $\neg$.

Elementary toposes were originally introduced by F. W. Lawvere and M. Tierney during a collaboration in winter term 1969/1970 at Dalhousie University in Canada.

**Definition 12.1** (elementary topos)
*An* (elementary) topos *is a category* $\mathbb{E}$ *such that*

(1) $\mathbb{E}$ *has finite limits*

(2) $\mathbb{E}$ *has all exponentials*

(3) $\mathbb{E}$ *has a* subobject classifier, *i.e. a mono* $\mathtt{t} : 1 \rightarrowtail \Omega$ *such that for every subobject* $m : P \rightarrowtail A$ *there exists a unique morphism* $\chi_m : A \to \Omega$ *(called the* classifying morphism for $m$*) such that*

$$
\begin{array}{ccc}
P & \longrightarrow & 1 \\
{\scriptstyle m}\downarrow & \lrcorner & \downarrow{\scriptstyle \mathtt{t}} \\
A & \longrightarrow & \Omega \\
& {\scriptstyle \chi_m} &
\end{array}
$$

*is a pullback square.*                                                    ◊

Notice that elementary toposes might be equivalently defined as cartesian closed categories $\mathbb{E}$ for which there exists a monomorphism $\mathtt{t} : 1 \to \Omega$ for which the following conditions hold

i) pullbacks of $t$ along arbitrary morphisms (with codomain $\Omega$) in $\mathbb{C}$ exist and

ii) for every mono $m : P \rightarrowtail A$ there exists a unique (classifying) map $\chi : A \to \Omega$ with

$$
\begin{array}{ccc}
P & \longrightarrow & 1 \\
\scriptstyle m \downarrow & \lrcorner & \downarrow \scriptstyle t \\
A & \longrightarrow & \Omega \\
& \scriptstyle \chi &
\end{array}
$$

i.e. $m \cong \chi^* t$.

At first sight this definition appears as somewhat weaker because it does not (explicitly) require the existence of pullbacks. The existence of the latter, however, follows from the fact that $\mathbb{E}$ automatically has all equalisers which can be seen as follows. Suppose $f, g : B \to A$ in $\mathbb{E}$. Let $\delta_A = \langle \mathsf{id}_A, \mathsf{id}_A \rangle : A \rightarrowtail A \times A$ be the diagonal on $A$ and $eq_A : A \times A \to \Omega$ be the classifying map for $\delta_A$. Consider the diagram

$$
\begin{array}{ccccc}
E & \longrightarrow & A & \longrightarrow & 1 \\
\scriptstyle e \downarrow & & \downarrow \scriptstyle \delta_A & & \downarrow \scriptstyle t \\
B & \xrightarrow[\langle f, g \rangle]{} & A \times A & \xrightarrow[eq_A]{} & \Omega
\end{array}
$$

where the right square and the outer rectangle are pullbacks. Then by Lemma 7.3 the left inner square is a pullback, too, from which it follows that $e$ is an equaliser for $f$ and $g$.

The paradigmatic[22] example of an elementary topos is **Set** where $\Omega = \{\mathsf{t}, \mathsf{f}\}$ and $\mathsf{t} : 1 \to \Omega : * \mapsto \mathsf{t}$. For a subobject $m : P \rightarrowtail A$ its classifying morphism is the function $\chi_m : A \to \Omega$ where $\chi_m(x) = \mathsf{t}$ iff $x \in m[P]$, i.e. $x$ is in the image of $P$ under $m$. In **Set** for every $\chi : A \to \Omega$ one may always choose the canonical pullback

$$
\begin{array}{ccc}
\chi^{-1}(\{\mathsf{t}\}) & \longrightarrow & 1 \\
\scriptstyle m \uparrow & \lrcorner & \downarrow \scriptstyle t \\
A & \longrightarrow & \Omega \\
& \scriptstyle \chi_m &
\end{array}
$$

---

[22]We will soon see that there is a lot of other elementary toposes, in particular presheaf categories $\widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ for small categories $\mathbb{C}$ though these are definitely not the only ones!

where $m$ is an inclusion, i.e. $m(x) = x$ for all arguments $x$ of $m$.

Such a notion of inclusion, however, is not available in arbitrary elementary toposes because, firstly, morphisms in a topos in general are not just set-theoretic functions and, secondly, there is *no universal* notion of equality for elements of different objects (if $m : P \rightarrowtail A$ and $a : 1 \to P$ then it does not make sense to ask whether $a : 1 \to P$ and $m \circ a : 1 \to A$ are equal because they are different morphisms in the topos!). The latter observation is the reason why subobjects of $A$ are conceptualized as equivalence classes of monos into $A$ as in the following

**Definition 12.2** (subobjects)

*Let $\mathbb{C}$ be a category. A* subobject *of $A \in \mathsf{Ob}(\mathbb{C})$ is a monomorphism into $A$. For subobjects $m : P \rightarrowtail A$ and $n : Q \rightarrowtail A$ we say that $m$ is contained in $n$ (notation $m \leq n$) iff there exists a morphism $f : P \to Q$ with $n \circ f = m$ as illustrated in the diagram[23]*

$$P \xrightarrow{\quad f \quad} Q$$

with $m$ and $n$ arrows to $A$ below.

*and we say that $m$ and $n$ are equal as subobjects (notation $m \sim n$) iff $m \leq n$ and $n \leq m$. We write $\mathsf{Sub}_{\mathbb{C}}(A)$ for the collection of subobjects of $A$ modulo the relation $\sim$. If $\mathbb{C}$ has pullbacks (of monomorphisms along arbitrary morphisms in $\mathbb{C}$) then $\mathsf{Sub}_{\mathbb{C}}$ extends to a functor $\mathsf{Sub}_{\mathbb{C}} : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ by putting $\mathsf{Sub}_{\mathbb{C}}(f)([m]_{\sim}) = [f^* m]_{\sim}$ where $f^* m$ is the pullback of $m$ along $f$.* $\diamond$

Notice that $\mathsf{Sub}_{\mathbb{C}}$ is a contravariant functor to **Set** only if $\mathbb{C}$ is well-powered, i.e. if $\mathsf{Sub}_{\mathbb{C}}(A)$ is small for all objects $A$ in $\mathbb{C}$, which, however, can always be achieved by assuming sufficiently many Grothendieck universes. Moreover, all conditions on $\mathsf{Sub}_{\mathbb{C}}$ that we consider in the context of toposes can be reformulated in a purely elementary way avoiding any reference to collections of the form $\mathsf{Sub}_{\mathbb{C}}(A)$ (see e.g. the presentation in [McL]). However, assuming the existence of $\mathsf{Sub}_{\mathbb{C}}$ as a presheaf allows one to express the existence of a subobject classifier in terms of representability of $\mathsf{Sub}_{\mathbb{C}}$ from which it follows immediately that subobject classifiers are unique up to isomorphism.

---

[23]necessarily $f$ is unique with this property and itself a monomorphism!

**Lemma 12.1** *Let $\mathbb{C}$ be a category with pullbacks (of monos along arbitrary morphisms). Then $\mathbb{C}$ has a subobject classifier iff $\mathsf{Sub}_{\mathbb{C}} : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}$ is a representable presheaf.*

*Proof:* If $\mathsf{t} : 1 \to \Omega$ is a subobject classifier for $\mathbb{C}$ then the natural transformation $\iota : \mathsf{Y}_{\mathbb{C}}(\Omega) \to \mathsf{Sub}_{\mathbb{C}}$ with $\iota_{\Omega}(\mathsf{id}_{\Omega}) = \mathsf{t}$ is an isomorphism in $\widehat{\mathbb{C}}$.

For the reverse direction suppose that $\iota : \mathsf{Y}_{\mathbb{C}}(\Omega) \to \mathsf{Sub}_{\mathbb{C}}$ is a natural isomorphism. Let $\top : U \rightarrowtail \Omega$ be an element of the equivalence class $\iota_{\Omega}(\mathsf{id}_{\Omega})$. Then for every $m : P \rightarrowtail A$ there exists a unique morphism $\chi : A \to \Omega$ with $m \sim \chi^* \top$, namely $\chi = \iota_A^{-1}([m])$. It remains to show that $U$ is terminal in $\mathbb{C}$. For every object $A$ in $\mathbb{C}$ we have

$$
\begin{array}{ccc}
A & \longrightarrow & U \\
\| & \lrcorner & \big\downarrow \top \\
A & \longrightarrow & \Omega
\end{array}
$$

giving rise to a map from $A$ to $U$ in $\mathbb{C}$. For unicity assume that $f_1, f_2 : A \to U$ in $\mathbb{C}$. Then we have

$$
\begin{array}{ccccc}
A & \xrightarrow{f_i} & U & =\!=\!= & U \\
\| & \lrcorner & \| & \lrcorner & \big\downarrow \top \\
A & \xrightarrow{f_i} & U & \underset{\top}{\rightarrowtail} & \Omega
\end{array}
$$

for $i=1,2$ from which it follows by uniqueness of classifying maps that $\top \cdot f_1 = \top \cdot f_2$. Thus, as $\top$ is monic we get $f_1 = f_2$ as desired. $\qquad\square$

From inspection of the proof it follows that a mono $\top : U \rightarrowtail \Omega$ is a subobject classifier iff $\top = \iota_{\Omega}(\mathsf{id}_{\Omega})$ for a natural isomorphism $\iota : \mathsf{Y}_{\mathbb{C}}(\Omega) \to \mathsf{Sub}_{\mathbb{C}}$. As representing objects for $\mathsf{Sub}_{\mathbb{C}}$ are unique up to isomorphism it follows that subobject classifiers are unique up to isomorphism.

As not only the existence of a subobject classifier, but also the existence of finite limits and exponentials can be formulated in terms of representability of certain presheaves being an elementary topos is a property of a category rather than (the requirement of) additional structure.

But notice that all these representability requirements can be expressed without any reference to $\mathbf{Set}$ (as e.g. in C. McLarty's book [McL]), i.e. in a purely elementary way which can be formalized in first order predicate logic. Thus,

the notion of *elementary topos* is *independent* from set theory as formalized by the first order theory ZFC. As we shall see soon toposes are structures which are strong enough to perform within them those usual set-theoretic constructions (cartesian products, function spaces, power sets etc.) that are needed in the actual practice of modern mathematics (analysis, algebra, geometry etc.). Thus, elementary toposes can be considered as a *foundation for mathematics* alternative to (axiomatic) set theory. This view of foundations has been forcefully propagated and developed by F. W. Lawvere since middle of 1960ies already. For a most readable textbook account of his view see his book *Sets for Mathematics* (CUP 2003) [LR] together with R. Rosebrugh.

The main difference between topos theory and axiomatic set theory lies in the different choice of basic concepts. In the topos case the basic notions are types (objects of the category) and functions between them (morphisms of the category) which is in accordance with the practice of modern mathematics. The basic assumption of ZFC is an *untyped* universe of all sets (which is a bit disputable from an ontological point of view!) and a binary relation $\in$ on this universe of sets telling which sets are elements of which sets. In ZFC functions show up as a derived notion, namely as particular sets of pairs where (following a suggestion of Kuratowski) a pair $\langle x, y \rangle$ is "implemented" as the set $\{\{x\}, \{x, y\}\}$. Natural numbers are also "implemented" (following a suggestion of J. von Neumann) as $0 = \emptyset$ and $n+1 = n \cup \{n\}$. All this kind of coding necessitated by ZFC's choice of basic notions makes it appear as somewhat artificial in the sense that its basic concepts do not match the basic concepts of actual mathematical practice.[24]

A further most useful aspect of topos theory is that it allows for a great variety of different models in contrast to ZFC for which it is fairly hard to construct different models (as e.g. P. Cohen's construction[25] of forcing models for ZFC where the Continuum Hypothesis fails).

The following theorem in one step provides us already with a great variety of different models for the axioms for an elementary topos.

---

[24]Notice, however that set theory is much stronger than (the logic of) toposes even if they are boolean, i.e. validate classical logic. The reason is the absence of ZFC's *replacement* axiom which allows one to transfinitely iterate $\mathcal{P}$. In most of modern mathematics this incredible strength of ZFC is not needed at all! However, there are exceptions like D. Martin's proof of *Borel Determinacy*!

[25]which, by the way, can be given a most understandable explanation in topos theoretic terms (see [MM] but known already in the early 1970ies!)

**Theorem 12.1** *For every small category $\mathbb{C}$ the category $\widehat{\mathbb{C}} = \mathbf{Set}^{\mathbb{C}^{op}}$ of presheaves over $\mathbb{C}$ is an elementary[26] topos.*

*Proof:* We have already seen in the previous section that $\widehat{\mathbb{C}}$ is cartesian closed. Using Yoneda (again) we can easily find a candidate for the subobject classifier $\Omega$ as it has to satisfy

$$(1) \qquad \Omega(I) \cong \widehat{\mathbb{C}}(\mathsf{Y}_{\mathbb{C}}(I), \Omega) \cong \mathsf{Sub}_{\widehat{\mathbb{C}}}(\mathsf{Y}_{\mathbb{C}}(I))$$

for all objects $I$ in $\mathbb{C}$. This suggests to define $\Omega$ as

$$(2) \qquad \Omega(-) = \mathsf{Sub}_{\widehat{\mathbb{C}}}(\mathsf{Y}_{\mathbb{C}}(-))$$

i.e. the morphism part of $\Omega$ operates by taking pullbacks

$$
\begin{array}{ccc}
\mathsf{Y}_{\mathbb{C}}(u)^{-1}S & \longrightarrow & S \\
{\scriptstyle \mathsf{Y}_{\mathbb{C}}(u)^* i} \downarrow \quad \lrcorner & & \downarrow {\scriptstyle i} \\
\mathsf{Y}_{\mathbb{C}}(J) & \xrightarrow[\mathsf{Y}_{\mathbb{C}}(u)]{} & \mathsf{Y}_{\mathbb{C}}(I)
\end{array}
$$

employing the fact that for subobjects $m : P \rightarrowtail A$ in $\widehat{\mathbb{C}}$ there exists a canonical[27] inclusion $i_m : S_m \hookrightarrow A$ where $S_m$ is the image of $P$ under $m$ in $A$. Canonical subobjects of $\mathsf{Y}_{\mathbb{C}}(I)$ are called *sieves of $I$* and can be identified with those sets $S$ of morphisms in $\mathbb{C}$ such that for all $u \in S$ it holds that $\mathsf{cod}(u) = I$ and $uv \in S$ for all $v \in \mathsf{Mor}(\mathbb{C})$ with $\mathsf{cod}(v) = \mathsf{dom}(u)$. Evidently, for a sieve $S$ of $I$ and morphism $u : J \to I$ in $\mathbb{C}$ the sieve $\mathsf{Y}_{\mathbb{C}}(u)^{-1}S$ consists of all morphisms $v : K \to J$ in $\mathbb{C}$ with $uv \in S$.

Condition (1) also suggest us how to construct the classifying morphism $\chi : A \to \Omega$ for $P \hookrightarrow A$. By naturality of the isomorphism $\widehat{\mathbb{C}}(\mathsf{Y}_{\mathbb{C}}(I), \Omega) \cong \mathsf{Sub}_{\widehat{\mathbb{C}}}(\mathsf{Y}_{\mathbb{C}}(I))$ we have

$$
\begin{array}{ccc}
\widehat{\mathbb{C}}(A, \Omega) & \xrightarrow{\;\cong\;} & \mathsf{Sub}_{\widehat{\mathbb{C}}}(A) \\
{\scriptstyle \widehat{\mathbb{C}}(a, \Omega)} \downarrow & & \downarrow {\scriptstyle \mathsf{Sub}_{\widehat{\mathbb{C}}}(a)} \\
\widehat{\mathbb{C}}(\mathsf{Y}_{\mathbb{C}}(I), \Omega) & \xrightarrow{\;\cong\;} & \mathsf{Sub}_{\widehat{\mathbb{C}}}(\mathsf{Y}_{\mathbb{C}}(I))
\end{array}
$$

---

[26] Actually, $\widehat{\mathbb{C}}$ is a *Grothendieck topos*, i.e. besides being an elementary topos it has small limits and colimits and a small generating family, namely $(\mathsf{Y}_{\mathbb{C}}(I))_{I \in \mathsf{Ob}(\mathbb{C})}$.

[27] Notice that this is possible only because $\mathbf{Set}^{\mathbb{C}^{op}}$ is so close to $\mathbf{Set}$ and, therefore, inherits enough structure from it!

for all generalised elements $a : \mathsf{Y}_\mathbb{C}(I) \to A$ from which it follows that

$$(3) \qquad \chi(a) = a^*P = \{u : J {\to} I \mid A(u)(a) \in P\}$$

according to our identification of $\Omega(I)$ with $\mathsf{Sub}_{\widehat{\mathbb{C}}}(\mathsf{Y}_\mathbb{C}(I))$. Condition (3) also tells us how to construct the subobject classifier $\mathsf{t} : 1 \to \Omega$. As $\mathsf{t}$ is the classifying map for the subobject $\mathsf{id}_1 : 1 \to 1$ it follows from (3) that

$$(4) \qquad \mathsf{t}_I(*) = \top_I = \{u \in \mathsf{Mor}(\mathbb{C}) \mid \mathsf{cod}(u) = I\}$$

i.e. $\mathsf{t}$ picks the maximal sieve of $I$ for every $I \in \mathsf{Ob}(\mathbb{C})$.

We leave it as a straightforward exercise(!) to the inclined reader to verify that $\mathsf{t} : 1 \to \Omega$ as defined by (2) and (4) is actually a subobject classifier where the construction of classifying maps is given by (3). □

Now let us look at $\Omega$ in some particular presheaf toposes.[28]

**Example 12.1** *If $I$ is a set then $\Omega$ in $\mathbf{Set}^I$ is given by the constant family $(\{0,1\})_{i \in I}$ where $\mathsf{t}_i(*) = 0$ (assuming that 0 codes "true"). Notice that this $\Omega$ has $2^{|I|}$ many global elements!*

**Example 12.2** *Let $\mathbb{C}$ be the category $0 \xrightarrow{\alpha} 1$. In $\widehat{\mathbb{C}}$ the truth value object $\Omega$ up to isomorphism looks as follows: $\Omega(1) = \{\mathsf{f}, \mathsf{u}, \mathsf{t}\}$, $\Omega(0) = \{\mathsf{t}, \mathsf{f}\}$ and $\Omega(\alpha)$ is the identity on $\{\mathsf{t}, \mathsf{f}\}$ and sends $\mathsf{u}$ to $\mathsf{t}$ (the "undecided" truth value $\mathsf{u}$ corresponds to the sieve on 1 consisting of just $\alpha$).*

**Example 12.3** *If $\mathbb{C}$ is a monoid $M$ (considered as a category) then $\Omega$ consists of all right ideals in $M$, i.e. subsets $I$ of $M$ such that $x \in I$ and $y \in M$ implies $xy \in I$, and $\Omega(x)(I) = \{y \in M \mid xy \in I\}$.*

**Example 12.4** *If $\mathbb{C}$ is a group $G$ then according to Example 12.3 the truth value object $\Omega$ of the topos $\widehat{\mathbb{C}}$ of $G$-actions has $\{G, \emptyset\}$ as underlying set because $G$ and $\emptyset$ are the only right ideals in $G$ which, moreover, are left invariant by all actions of group elements.*

The above examples are already sufficient for exhibiting the relations between the notions introduced in the next definition.

---

[28]We leave it as an exercise(!) to explicitate the construction of exponentials in these examples!

**Definition 12.3** (2-valued, boolean, wellpointed)
*Let $\mathbb{C}$ be a nonempty small category and $\mathbb{E} = \widehat{\mathbb{C}}$ the corresponding presheaf topos. Let $\mathsf{f} : 1 \to \Omega$ be the classifying map for the subobject $0 \rightarrowtail 1$ in $\mathbb{E}$.*

- *$\mathbb{E}$ is 2-valued iff $\Omega_\mathbb{E}$ has precisely 2 global elements, i.e. $|\mathbb{E}(1, \Omega)| = 2$.*

- *$\mathbb{E}$ is boolean iff $[\mathsf{t}, \mathsf{f}] : 1{+}1 \to \Omega$ is an isomorphism.*

- *$\mathbb{E}$ is wellpointed iff morphisms $f, g : A \to B$ in $\mathbb{E}$ are already equal whenever $fa = ga$ for all global elements $a$ of $A$, i.e. for all $a : 1 \to A$ in $\mathbb{E}$.* $\diamondsuit$

Obviously, for all (nonempty) sets $I$ the topos $\mathbf{Set}^I$ is boolean but not 2-valued if $I$ contains more than one element. On the other hand for every monoid $M$ the truth value object $\Omega$ of the topos $\widehat{M}$ of right $M$-actions (see Example 12.3) contains just 2 global elements, namely $M$ and $\emptyset$, (why?) and, accordingly, is 2-valued whereas $[\mathsf{t}, \mathsf{f}] : 1{+}1 \to \Omega$ is not an isomorphism if $M$ contains non-trivial right ideals (as e.g. the monoid $(\mathbb{N}, +, 0)$), i.e. $\widehat{M}$ is not boolean in such cases. Thus, the properties "2-valued" and "boolean" are independent in contrast to common superstition!
For every group $G$ the topos $\widehat{G}$ of $G$-actions is both boolean and 2-valued (see Example 12.4). But if $G$ is nontrivial then $\widehat{G}$ is not wellpointed as the representable object $G = \mathsf{Y}_G(*)$ has then no global elements but admits more than one endomap (namely, by Yoneda, as many as there are elements in the group $G$).
However, if a (presheaf) topos is wellpointed then it is 2-valued and boolean as we shall see next.

**Theorem 12.2** *Let $\mathbb{C}$ be a nonempty small category and $\mathbb{E} = \widehat{\mathbb{C}}$ the corresponding presheaf topos. Then $\mathbb{E}$ is 2-valued and boolean whenever $\mathbb{E}$ is wellpointed.*

*Proof:* Assume that $\mathbb{E} = \widehat{\mathbb{C}}$ is wellpointed.
Suppose $\mathbb{E}$ is not 2-valued, i.e. there exists a further global element $\mathsf{u} : 1 \to \Omega$ besides $\mathsf{t}$ and $\mathsf{f}$. Let $i : U \rightarrowtail 1$ be the subobject classified by $\mathsf{u}$. The maps $\mathsf{t} \circ i$ and $\mathsf{f} \circ i$ are different as they classify different subobjects of $U$, namely $U$ and $0$, respectively. Thus, by wellpointedness there exists a global element $u : 1 \to U$ separating $\mathsf{t} \circ i$ and $\mathsf{f} \circ i$. Obviously, the map $u$ is an isomorphism

from which it follows that $U \cong 1$ in contradiction to the assumption $\mathsf{t} \neq \mathsf{u}$. Thus $\mathbb{E}$ is 2-valued.

Let $\chi : \Omega \to \Omega$ be the classifying map for the monomorphism $[\mathsf{t}, \mathsf{f}] : 1{+}1 \rightarrowtail \Omega$. As $\mathsf{t}, \mathsf{f} \leq_\Omega [\mathsf{t}, \mathsf{f}]$ it follows that $\mathsf{t}^*[\mathsf{t}, \mathsf{f}] \cong \mathsf{id}_1 \cong \mathsf{f}^*[\mathsf{t}, \mathsf{f}]$. Thus $\chi \circ \mathsf{t} = \mathsf{t} = \chi \circ \mathsf{f}$ from which it follows by wellpointedness that $\chi = \mathsf{t} \circ !_\Omega$. Thus, we have $[\mathsf{t}, \mathsf{f}] \cong \mathsf{id}_\Omega$ and, accordingly, the map $[\mathsf{t}, \mathsf{f}]$ is an isomorphism, i.e. the topos $\mathbb{E}$ is boolean. $\square$

Notice that Definition 12.3 makes sense already for arbitrary elementary toposes because one can show (see e.g. [MM]) that they have all finite colimits, that $0 \rightarrowtail 1$ and that $[\mathsf{t}, \mathsf{f}] : 1{+}1 \rightarrowtail \Omega$.

Toposes $\mathbb{E}$ are called *localic* if the subobjects of $1_{\mathbb{E}}$ constitute a generating family for $\mathbb{E}$. Typical examples are presheaf toposes of the form $\widehat{P}$ where $P$ is a posetal category as e.g. the one considered in Example 12.2. One readily checks (reusing the second part of the proof of Theorem 12.2) that if a localic topos is 2-valued then it is also boolean.

Finally, we leave it as an exercise(!) to show that for small categories $\mathbb{C}$ the presheaf topos $\widehat{\mathbb{C}}$ is boolean iff $\mathbb{C}$ is a *groupoid*, i.e. all morphisms in $\mathbb{C}$ are isos. Thus $\widehat{\mathbb{C}}$ is boolean and 2-valued iff $\mathbb{C}$ is a connected groupoid ("connected" meaning that all homsets of $\mathbb{C}$ are inhabited). As $\mathbf{Set}^I$ is wellpointed iff $|I| = 1$ (why?) it follows that $\widehat{\mathbb{C}}$ is wellpointed if and only if $\mathbb{C}$ is equivalent to $\mathbf{1}$, the trivial nonempty category.

# 13  Logic of Toposes

We will show how to interpret *higher order constructive (intuitionistic) logic* (HOL) in an arbitrary elementary topos $\mathbb{E}$ which remains fixed throughout the whole section.

**Definition 13.1** *For $A \in \mathsf{Ob}(\mathbb{E})$ let $\leq_A$ be the partial order on $\mathsf{Sub}_{\mathbb{E}}(A)$ where $[m_1] \leq_A [m_2]$ iff there exists a (necessarily unique) monomorphism $m$ with $m_2 m = m_1$. Via the canonical isomorphism $\mathbb{E}(A, \Omega) \cong \mathsf{Sub}_{\mathbb{E}}(A)$ we consider $\leq_A$ also as a partial order on $\mathbb{E}(A, \Omega)$, the set of predicates on $A$, putting $\chi_1 \leq_A \chi_2$ iff $[\chi_1^* \mathsf{t}] \leq_A [\chi_2^* \mathsf{t}]$.* $\diamondsuit$

Pullbacks along arbitrary morphisms of $\mathbb{E}$ preserve this order.

**Lemma 13.1** *Let $f : B \to A$ be a morphism in $\mathbb{E}$. Then $\mathsf{Sub}_{\mathbb{E}}(f) = f^* : \mathsf{Sub}_{\mathbb{E}}(A) \to \mathsf{Sub}_{\mathbb{E}}(B)$ and, therefore, also $\mathbb{E}(f, \Omega) : \mathbb{E}(A, \Omega) \to \mathbb{E}(B, \Omega)$ are monotonic w.r.t. the order defined in Definition 13.1.*

*Proof:* Consider the diagram

$$
\begin{array}{ccc}
f^* A_1 & \longrightarrow & A_1 \\
{\scriptstyle f^* m} \downarrow & \lrcorner & \downarrow {\scriptstyle m} \\
f^* A_2 & \longrightarrow & A_2 \\
{\scriptstyle f^* m_2} \downarrow & \lrcorner & \downarrow {\scriptstyle m_2} \\
B & \xrightarrow{\ f\ } & A
\end{array}
$$

where $m_1 = m_2 m$. The claim for $\mathsf{Sub}_{\mathbb{E}}(f)$ follows from the fact that $f^* m_1 \cong f^* m_2 \circ f^* m$.

Then monotonicity of $\mathbb{E}(f, \Omega)$ follows from the fact that $\chi_i \circ f$ classifies $f^* m_i$ whenever $\chi_i$ classifies $m_i$. $\qquad\square$

In the following we write $\mathsf{t}_A$ as an abbreviation for $\mathsf{t} \circ !_A$.

**Lemma 13.2** *Let $\chi : A \to \Omega$ and $a : I \to A$. Then $\chi \circ a = \mathsf{t}_I$ iff $a$ factors through the mono $m : P \rightarrowtail A$ classified by $\chi$.*

*Proof:* If $\chi \circ a = \mathsf{t}_I = \mathsf{t}\circ!_I$ then there exists a (necessarily unique) morphism $f : I \to P$ with $m \circ f = a$ (and $!_P \circ f =!_I$). On the other hand if $m \circ f = a$ then $\chi \circ a = \chi \circ m \circ f = \mathsf{t} \circ !_P \circ f = \mathsf{t} \circ !_I = \mathsf{t}_I$. $\square$

Next we show that $\mathsf{Sub}_{\mathbb{E}}(A)$ has binary meets w.r.t. $\leq$ and that these are preserved by pullbacks along arbitrary morphisms $f : B \to A$.

**Lemma 13.3** *Let* $\wedge : \Omega \times \Omega \to \Omega$ *be the classifying map for the monomorphism* $\langle \mathsf{t}, \mathsf{t} \rangle : 1 \rightarrowtail \Omega \times \Omega$, *i.e.*

$$
\begin{array}{ccc}
1 & \longrightarrow & 1 \\
{\scriptstyle \langle \mathsf{t},\mathsf{t}\rangle}\big\downarrow & \lrcorner & \big\downarrow {\scriptstyle \mathsf{t}} \\
\Omega \times \Omega & \xrightarrow{\ \wedge\ } & \Omega
\end{array}
$$

*Then for* $\chi_1, \chi_2 : A \to \Omega$ *it holds that* $\wedge \circ \langle \chi_1, \chi_2 \rangle = \mathsf{t}_A$ *iff* $\chi_1 = \mathsf{t}_A = \chi_2$.

*Proof:* Obviously, we have $\chi_1 = \mathsf{t}_A = \chi_2$ iff $\langle \chi_1, \chi_2 \rangle$ factors through $\langle \mathsf{t}, \mathsf{t} \rangle$ which in turn by Lemma 13.2 is equivalent to $\wedge \circ \langle \chi_1, \chi_2 \rangle = \mathsf{t}_A$. $\square$

Next we introduce an equality predicate for arbitrary objects of $\mathbb{E}$.

**Definition 13.2** (equality predicate)
*For* $A \in \mathsf{Ob}(\mathbb{E})$ *let* $eq_A : A \times A \to \Omega$ *be the classifying map for the diagonal* $\delta_A = \langle \mathsf{id}_A, \mathsf{id}_A \rangle : A \to A \times A$. *For* $eq_\Omega : \Omega \times \Omega \to \Omega$ *we often write* $\Leftrightarrow$. $\diamond$

**Lemma 13.4** *Let* $a_1, a_2 : I \to A$. *Then* $eq_A \circ \langle a_1, a_2 \rangle = \mathsf{t}_I$ *iff* $a_1 = a_2$.

*Proof:* By Lemma 13.2 we have $eq_A \circ \langle a_1, a_2 \rangle = \mathsf{t}_I$ iff $\langle a_1, a_2 \rangle$ factors through $\delta_A$, i.e. $a_1 = a_2$. $\square$

Thus, in particular, for $\chi_1, \chi_2 : A \to \Omega$ we have $\Leftrightarrow \circ \langle \chi_1, \chi_2 \rangle = \mathsf{t}_A$ iff $\chi_1 = \chi_2$.

**Lemma 13.5** *Let* $\chi_1, \chi_2 : A \to \Omega$ *and* $a : I \to A$. *Then* $\chi_1 \circ a = \mathsf{t}_I = \chi_2 \circ a$ *if and only if* $\wedge \circ \langle \chi_1, \chi_2 \rangle \circ a = \mathsf{t}_I$.

*Proof:* By Lemma 13.3 we have $\chi_1 \circ a = \mathsf{t}_I = \chi_2 \circ a$ iff $\wedge \circ \langle \chi_1 \circ a, \chi_2 \circ a \rangle = \mathsf{t}_I$. The claim follows as $\langle \chi_1 \circ a, \chi_2 \circ a \rangle = \langle \chi_1, \chi_2 \rangle \circ a$. $\square$

**Theorem 13.1** (conjunction)
*For every $A \in \mathsf{Ob}(\mathbb{E})$ the poset $\mathsf{Sub}_{\mathbb{E}}(A)$ is a $\wedge$-semilattice w.r.t. $\leq$, i.e. has finite infima, that are preserved by $\mathsf{Sub}_{\mathbb{E}}(f) : \mathsf{Sub}_{\mathbb{E}}(A) \to \mathsf{Sub}_{\mathbb{E}}(B)$ for arbitrary morphisms $f : B \to A$ in $\mathbb{E}$. Moreover, if $\chi_1, \chi_2 : A \to \Omega$ classify subobjects $m_1 : P_1 \rightarrowtail A$ and $m_2 : P_2 \rightarrowtail A$, respectively, then $\wedge \circ \langle \chi_1, \chi_2 \rangle$ classifies the infimum of $m_1$ and $m_2$.*

*Proof:* Let $\chi_1$ and $\chi_2$ be the classifying maps for $m_1$ and $m_2$, respectively. Let $m : P \rightarrowtail A$ be the subobject classified by $\wedge \circ \langle \chi_1, \chi_2 \rangle$. We show that $m$ is the infimum of $m_1$ and $m_2$. By Lemma 13.2 and Lemma 13.5 for all $n : Q \rightarrowtail A$ we have $n \leq_A m$ iff $\wedge \circ \langle \chi_1, \chi_2 \rangle \circ n = \mathsf{t}_Q$ iff $\chi_1 \circ n = \mathsf{t}_Q$ and $\chi_2 \circ n = \mathsf{t}_Q$ iff $n \leq_A m_1$ and $n \leq_A m_2$. Thus $m = m_1 \wedge m_2$ as desired.
Now suppose $f : B \to A$. Then $\wedge \circ \langle \chi_1, \chi_2 \rangle \circ f = \wedge \circ \langle \chi_1 \circ f, \chi_2 \circ f \rangle$ classifies both $f^*m$ and the infimum of $f^*m_1$ and $f^*m_2$. Thus $f^*$ preserves meets of subobjects. $\qquad\square$

Instead of $\wedge \circ \langle \chi_1, \chi_2 \rangle$ we usually write $\chi_1 \wedge \chi_2$ in accordance with the fact that it is the infimum of $\chi_1$ and $\chi_2$ in $\mathbb{E}(A, \Omega)$.

Next we introduce implication as a binary operation on $\Omega$.

**Definition 13.3** (implication)
*Let $\Rightarrow : \Omega \times \Omega \to \Omega$ be defined as $\Leftrightarrow \circ \langle \wedge, \pi_1 \rangle$ where $\pi_1 : \Omega \times \Omega \to \Omega$ is the projection on the first component.* $\qquad\diamond$

Again, for reasons of readability, we often write $\chi_1 \Rightarrow \chi_2$ as an abbreviation for $\Rightarrow \circ \langle \chi_1, \chi_2 \rangle$.

**Lemma 13.6** *Let $\chi_1, \chi_2 : A \to \Omega$ then $\chi_1 \leq_A \chi_2$ iff $\Rightarrow \circ \langle \chi_1, \chi_2 \rangle = \mathsf{t}_A$.*

*Proof:* For $\chi_1, \chi_2 : A \to \Omega$ we have

$$\Rightarrow \circ \langle \chi_1, \chi_2 \rangle = \mathsf{t}_A \qquad \text{iff} \qquad \text{(Lemma 13.5)}$$
$$\wedge \circ \langle \chi_1, \chi_2 \rangle = \chi_1 \qquad \text{iff}$$
$$\chi_1 \leq_A \chi_2$$

where the second equality follows from the fact that in arbitrary posets $x \leq y$ iff $x$ is the infimum of $x$ and $y$. $\qquad\square$

**Lemma 13.7** *For maps $\chi_1$, $\chi_2$ and $\chi$ from $A$ to $\Omega$ we have*

$$\chi \wedge \chi_1 \leq_A \chi_2 \qquad \textit{iff} \qquad \chi \leq_A \chi_1 {\Rightarrow} \chi_2 \quad .$$

*Proof:* Let $m : P \rightarrowtail A$, $m_1 : P_1 \rightarrowtail A$ and $m_2 : P_2 \rightarrowtail A$ be the subobjects classified by $\chi$, $\chi_1$ and $\chi_2$, respectively. Then we have

| | | |
|---|---|---|
| $\chi \leq_A \chi_1 {\Rightarrow} \chi_2$ | iff | (Lemma 13.2) |
| $(\chi_1 {\Rightarrow} \chi_2) \circ m = \mathsf{t}_P$ | iff | (Lemma 13.6) |
| $\chi_1 \circ m \leq_P \chi_2 \circ m$ | iff | ($\chi_i \circ m$ classifies $m^* m_i$) |
| $m^* m_1 \leq m^* m_2$ | iff | ($m \circ (-)$ order reflecting) |
| $m \circ m^* m_1 \leq m \circ m^* m_2$ | iff | ($m \circ m^* m_i \cong m \wedge m_i$) |
| $m \wedge m_1 \leq_A m \wedge m_2$ | iff | (as $m \wedge m_1 \leq_A m$ holds anyway) |
| $m \wedge m_1 \leq_A m_2$ | iff | (as $m \wedge m_1$ is classified by $\chi \wedge \chi_1$) |
| $\chi \wedge \chi_1 \leq_A \chi_2$ | | |

i.e. the poset $\mathsf{Sub}_{\mathbb{E}}(A)$ has Heyting implication[29]. $\qquad\square$

Notice that pullback functors $f^*$ preserve Heyting implication because we have ${\Rightarrow} \circ \langle \chi_1, \chi_2 \rangle \circ a = {\Rightarrow} \circ \langle \chi_1 \circ a, \chi_2 \circ a \rangle$ for arbitrary $a : I \to A$.

Summarizing we have

**Theorem 13.2** *In $\mathbb{E}$ every $\mathsf{Sub}_{\mathbb{E}}(A)$ is a poset with finite infima and Heyting implication which structure is preserved by $f^{-1} : \mathsf{Sub}_{\mathbb{E}}(A) \to \mathsf{Sub}_{\mathbb{E}}(B)$ for all $f : B \to A$ in $\mathbb{E}$.*

*Proof:* For conjunctions and implication the claim is immediate from the previous lemmas. Every $\mathsf{Sub}_{\mathbb{E}}(A)$ has a greatest element $\mathsf{id}_A$ classified by $\mathsf{t}_A$. Preservation of greatest elements is immediate from $\mathsf{t}_A \circ f = \mathsf{t}_B$ for all $f : B \to A$. $\qquad\square$

Next we will introduce universal quantification over objects in $\mathbb{E}$.

---

[29] A poset $H$ with binary meets admits *Heyting implication* iff for all $a, b \in H$ there exists a necessarily unique $a {\Rightarrow} b \in H$ such that

$$x \leq a {\Rightarrow} b \quad \text{iff} \quad x \wedge a \leq b$$

for all $x \in H$.

For that purpose and also later on the following notational convention turns out as necessary and useful. If $f : A \to B$ is a map in $\mathbb{E}$ then we write $\ulcorner f \urcorner : 1 \to B^A$ as an abbreviation for

$$\lambda(1{\times}A \xrightarrow{\pi_2} A \xrightarrow{f} B)$$

called *name of* $f$ because it exhibits $f : A \to B$ as a global element of the exponential $B^A$.

**Definition 13.4** (universal quantification)
*We define $\forall_A : \Omega^A \to \Omega$ as the classifying map for $\ulcorner t_A \urcorner$, i.e.*



*for all objects $A$ of $\mathbb{E}$.* $\diamond$

**Lemma 13.8** *For $p : I \to \Omega^A$ we have $\forall_A \circ p = t_I$ iff $\varepsilon \circ (p{\times}\mathsf{id}_A) = t_{I{\times}A}$.*

*Proof:* Obviously $\forall_A \circ p = t_I$ iff the map $p$ factors through $\ulcorner t_A \urcorner$ (via $!_I$) which in turn is equivalent to $\varepsilon \circ ((\ulcorner t_A \urcorner \circ !_I){\times}\mathsf{id}_A) = \varepsilon \circ (p{\times}\mathsf{id}_A)$, i.e. $t_{I{\times}A} = \varepsilon \circ (p{\times}\mathsf{id}_A)$ because $\varepsilon \circ ((\ulcorner t_A \urcorner \circ !_I){\times}\mathsf{id}_A) = t_A \circ \pi_2 \circ (!_I{\times}\mathsf{id}_A) = t_{I{\times}A}$. $\square$

**Theorem 13.3** *For all $p : I \to \Omega^A$ and $\chi : I \to \Omega$ we have*

$$\chi \circ \pi \ \leq_{I{\times}A} \ \varepsilon \circ (p{\times}\mathsf{id}_A) \qquad iff \qquad \chi \ \leq_I \ \forall_A \circ p$$

*where $\pi : I{\times}A \to I$ is the first projection.*

*Proof:* Let $m : P \rightarrowtail I$ be the subobject classified by $\chi$. Then we have

| | | |
|---|---|---|
| $\chi \ \leq_I \ \forall_A \circ p$ | iff | (Lemma 13.2) |
| $\forall_A \circ p \circ m = t_P$ | iff | (Lemma 13.8) |
| $\varepsilon \circ ((p \circ m){\times}\mathsf{id}_A) = t_{P{\times}A}$ | iff | |
| $\varepsilon \circ (p{\times}\mathsf{id}_A) \circ (m{\times}\mathsf{id}_A) = t_{P{\times}A}$ | iff | |
| $\chi \circ \pi \ \leq_{I{\times}A} \ \varepsilon \circ (p{\times}\mathsf{id}_A)$ | | |

where the last equivalence follows from Lemma 13.2) and the fact that $\chi \circ \pi$ classifies the subobject $m \times \mathsf{id}_A$ (as $(\chi \circ \pi)^* \mathsf{t} \cong \pi^* \chi^* \mathsf{t} = \pi^* m = m \times \mathsf{id}_A$). $\qquad \square$

The following equivalent variant of Theorem 13.3 will turn out as more useful when we consider proof rules for the internal logic.

**Theorem 13.4** *Let $\varphi : I \times A \to \Omega$ and $\psi : I \to \Omega$ be maps in $\mathbb{E}$. Then for $\forall_A(\varphi) := \forall_A \circ \lambda(\varphi)$ it holds that*

$$\psi \circ \pi \leq_{I \times A} \varphi \qquad \textit{iff} \qquad \psi \leq_I \forall_A(\varphi)$$

*where $\pi : I \times A \to I$ is the first projection.*

*Proof:* Immediate from Theorem 13.3 instantiating $p$ by $\lambda(\varphi)$. $\qquad \square$

### The Internal Language

We have seen that in an elementary topos $\mathbb{E}$ one can interpret the logical operations $\top$ (truth), $\wedge$ (conjunction), $\Rightarrow$ (implication) and $\forall_A$ (universal quantification over $A$). We next will introduce a formal language, called *internal language of $\mathbb{E}$* [30], which is based on typed $\lambda$-calculus and contains constants for denoting the objects and morphisms of $\mathbb{E}$ including the above mentioned logical operations $\top$, $\wedge$, $\Rightarrow$ and $\forall_A$.

More precisely, the internal language of $\mathbb{E}$ is the typed $\lambda$-calculus over the collection $\mathsf{Ob}(\mathbb{E})$ of base types augmented by a constant of type $B^A$ for every morphism $f : A \to B$. As $\mathbb{E}$ is cartesian closed it is obvious how to interpret the internal language of $\mathbb{E}$ in the cartesian closed category $\mathbb{E}$ (a constant for object $A$ will be interpreted as $A$ and a constant for morphism $f : A \to B$ will be interpreted as $\ulcorner f \urcorner : 1 \to B^A$). The terms $\Gamma \vdash \varphi : \Omega$ of type $\Omega$ in context $\Gamma$ are called *propositions in context $\Gamma$* and will be interpreted as morphisms $[\![\Gamma]\!] \to \Omega_\mathbb{E}$, i.e. as subobjects of $[\![\Gamma]\!]$ via the isomorphism $\mathsf{Sub}_\mathbb{E}([\![\Gamma]\!]) \cong \mathbb{E}([\![\Gamma]\!], \Omega_\mathbb{E})$. If $\Gamma$ is a context and $\Gamma \vdash \varphi_i : \Omega$ ($i = 1, \ldots, n$) and $\Gamma \vdash \varphi : \Omega$ are propositions in context $\Gamma$ then we write

$$\Gamma \mid \varphi_1, \ldots, \varphi_n \vdash \varphi$$

---

[30] The internal language of $\mathbb{E}$ is also often called the Mitchell-Bénabou language because the internal language was introduced beginning of the 1970ies originally by Jean Bénabou and later (independently) by W. Mitchell.

for the *judgement*

$$[\![\Gamma \vdash \varphi_1]\!] \wedge \ldots \wedge [\![\Gamma \vdash \varphi_n]\!] \leq_{[\![\Gamma]\!]} [\![\Gamma \vdash \varphi]\!]$$

where the left hand side is the (finite) meet of the $[\![\Gamma \vdash \varphi_i]\!]$ in $\mathbb{E}([\![\Gamma]\!], \Omega)$. Often we write $\Phi$ as an abbreviation for the list $\varphi_1, \ldots, \varphi_n$ under which convention the general form of judgements is a *sequent*

$$\Gamma \mid \Phi \vdash \varphi$$

where $\Gamma \equiv x_1 : A_1, \ldots, x_n : A_n$ is called *variable context* and $\Phi \equiv \varphi_1, \ldots, \varphi_n$ is called *propositional context* or *context of assumptions*.

Next we formulate a few *proof rules* stating closure properties of sequents valid in arbitrary elementary toposes $\mathbb{E}$. As we shall see later these proof rules are *complete* in the sense that they are strong enough to *characterise validity in all elementary toposes*.

We begin with a few *structural* rules of fairly "bureaucratic" nature, (almost) too trivial to be found worth mentioning in informal reasoning but, nevertheless, absolutely necessary for formalizing even the most basic derivations.


**Structural Rules**

$$\frac{\sigma : \Delta \to \Gamma \qquad \Gamma \mid \Phi \vdash \varphi}{\Delta \mid \Phi[\sigma] \vdash \varphi[\sigma]} \ (\text{Subst}) \qquad\qquad \frac{\Gamma \mid \Phi \vdash \psi}{\Gamma \mid \Phi, \varphi \vdash \psi} \ (\text{Weak})$$

$$\frac{\Gamma \mid \Phi, \varphi, \varphi \vdash \psi}{\Gamma \mid \Phi, \varphi \vdash \psi} \ (\text{Contr}) \qquad\qquad \frac{\Gamma \mid \Phi_1, \varphi_1, \varphi_2, \Phi_2 \vdash \psi}{\Gamma \mid \Phi_1, \varphi_2, \varphi_1, \Phi_2 \vdash \psi} \ (\text{Perm})$$

$$\frac{\Gamma \vdash \varphi : \Omega}{\Gamma \mid \varphi \vdash \varphi} \ (\text{Ax}) \qquad\qquad \frac{\Gamma \mid \Phi \vdash \varphi \quad \Gamma \mid \Phi, \varphi \vdash \psi}{\Gamma \mid \Phi \vdash \psi} \ (\text{Cut})$$


Next we turn to the proper *logical* rules describing how to use the connectives $\top, \wedge, \Rightarrow$ and the quantifiers $\forall_A$.

Notice that for sake of readability we employ the notation $\forall x{:}A.\varphi$ instead of the more clumsy (but "official") notation $\forall_A(\lambda x{:}A.\varphi)$.

**Logical Rules**

$$\frac{}{\Gamma \mid \Phi \vdash \top} \ (\top)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi_1 \quad \Gamma \mid \Phi \vdash \varphi_2}{\Gamma \mid \Phi \vdash \varphi_1 \wedge \varphi_2} \ (\mathrm{I}\wedge)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi_1 \wedge \varphi_2}{\Gamma \mid \Phi \vdash \varphi_1} \ (\mathrm{E}\wedge_1)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi_1 \wedge \varphi_2}{\Gamma \mid \Phi \vdash \varphi_2} \ (\mathrm{E}\wedge_2)$$

$$\frac{\Gamma \mid \Phi, \varphi \vdash \psi}{\Gamma \mid \Phi \vdash \varphi \Rightarrow \psi} \ (\mathrm{I}\Rightarrow)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi \Rightarrow \psi \quad \Gamma \mid \Phi \vdash \varphi}{\Gamma \mid \Phi \vdash \psi} \ (\mathrm{E}\Rightarrow)$$

$$\frac{\Gamma, x{:}A \mid \Phi \vdash \varphi(x)}{\Gamma \mid \Phi \vdash \forall x{:}A.\varphi(x)} \ (\mathrm{I}\forall)$$

$$\frac{\Gamma \mid \Phi \vdash \forall x{:}A.\varphi(x)}{\Gamma, x{:}A \mid \Phi \vdash \varphi(x)} \ (\mathrm{E}\forall)$$

Notice that in the last two rules for $\forall$ we (implicitly) assume that for the $\varphi_i$ in $\Phi$ it holds that $\Gamma \vdash \varphi_i : \Omega$.

The attentive reader may have noticed the absence of the logical connectives $\bot$ (falsity), $\vee$ (disjunction) and $\exists$ (existential quantification). But due to a nice trick going back to B. Russell and D. Prawitz we may define these further logical operation as follows

($\bot$)    $\bot \equiv \forall p{:}\Omega.p$

($\vee$)    $\varphi \vee \psi \equiv \forall p{:}\Omega.\, (\varphi \Rightarrow p) \wedge (\psi \Rightarrow p) \Rightarrow p$

($\exists$)    $\exists x{:}A.\varphi(x) \equiv \forall p{:}\Omega.\, (\forall x{:}A.\varphi(x) \Rightarrow p) \Rightarrow p$

which is possible only because in contrast to first order logic we have quantification over $\Omega$ available. It is a straightforward, but most instructive exercise(!) to verify the following derived rules

$$\frac{}{\Gamma \mid \Phi, \bot \vdash \varphi} \ (\bot)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi_1}{\Gamma \mid \Phi \vdash \varphi_1 \vee \varphi_2} \ (\mathrm{I}\vee_1) \qquad \frac{\Gamma \mid \Phi \vdash \varphi_2}{\Gamma \mid \Phi \vdash \varphi_1 \vee \varphi_2} \ (\mathrm{I}\vee_2)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi_1 \vee \varphi_2 \quad \Gamma \mid \Phi, \varphi_1 \vdash \psi \quad \Gamma \mid \Phi, \varphi_2 \vdash \psi}{\Gamma \mid \Phi \vdash \psi} \ (\mathrm{E}\vee)$$

$$\frac{\Gamma \mid \Phi \vdash \varphi(t)}{\Gamma \mid \Phi \vdash \exists x{:}A.\varphi(x)} \ (\text{I}\exists) \qquad \frac{\Gamma \mid \Phi \vdash \exists x{:}A.\varphi(x) \quad \Gamma, x{:}A \mid \Phi, \varphi(x) \vdash \psi}{\Gamma \mid \Phi \vdash \psi} \ (\text{E}\exists)$$

where in the last two rules for $\exists$ we assume (implicitly) that $\Gamma, x{:}A \vdash \varphi(x) : \Omega$, $\Gamma \vdash \psi : \Omega$ and $\Gamma \vdash t : A$.

We leave it as an exercise(!) to show that in $\mathbb{E}$ we have

(I) $\qquad \Gamma, x{:}A \mid \Phi, \varphi(x) \vdash \psi \qquad$ iff $\qquad \Gamma \mid \Phi, \exists x{:}A.\varphi(x) \vdash \psi$

(II) $\qquad \Gamma, x{:}A \mid \Phi, \psi \vdash \varphi(x) \qquad$ iff $\qquad \Gamma \mid \Phi, \psi \vdash \forall x{:}A.\varphi(x)$

for $\Gamma \vdash \psi : \Omega$ and $\Gamma, x{:}A \vdash \varphi : \Omega$.

Reinterpreting the rules for $\bot$, $\vee$ and $\exists$ back to the (arbitrary elementary) topos $\mathbb{E}$ we observe that

**Theorem 13.5** *For all elementary toposes $\mathbb{E}$*

(1) *the posets $\mathsf{Sub}_{\mathbb{E}}(A)$ contain least elements $\bot_A$ that are preserved by $f^{-1}$ for arbitrary maps $f : B \to A$ in $\mathbb{E}$*

(2) *the posets $\mathsf{Sub}_{\mathbb{E}}(A)$ have binary joins that are preserved by $f^{-1}$ for arbitrary maps $f : B \to A$ in $\mathbb{E}$*

(3) *for all subobjects $r : R \rightarrowtail C{\times}A$ there exists a subobject $\exists_A(r) : I \rightarrowtail C$ such that for all subobjects $m : P \rightarrowtail C$ it holds that*

$$\exists_A(r) \leq_C m \qquad iff \qquad r \leq_{C \times A} \pi^* m$$

*where $\pi : C{\times}A \to C$ is the first projection and, moreover, for all maps $f : D \to C$ in $\mathbb{E}$ it holds that $f^* \exists_A(r) \cong \exists_A((f{\times}\mathsf{id}_A)^* r)$.*

*Proof:* From validity of rule $(\bot)$ it follows that $\mathsf{Sub}_{\mathbb{E}}(A)$ contains a least element $\bot_A$ classified by $[\![x{:}A \vdash \bot]\!]$. Validity of the rule (Subst) guarantees that $f^* \bot_A \cong \bot_B$ for all morphisms $f : B \to A$ in $\mathbb{E}$.

For existence of binary joins in $\mathsf{Sub}_{\mathbb{E}}(A)$ suppose that $m$ and $n$ are subobjects of $A$ classified by $\varphi : A \to \Omega$ and $\psi : A \to \Omega$, respectively. One easily checks that

$$x{:}A \mid \varphi \vee \psi \vdash \theta \qquad \text{iff} \qquad x{:}A \mid \varphi \vdash \theta \ \text{ and } \ x{:}A \mid \psi \vdash \theta$$

for all $x{:}A \vdash \theta : \Omega$. From this observation it follows that $m \vee n$ is classified by $x{:}A \vdash \varphi \vee \psi$. Validity of the rule (Subst) entails that $f^*(m \vee n) \cong f^* m \vee f^* n$ for all morphisms $f : B \to A$ in $\mathbb{E}$.

Suppose the subobject $r : R \rightarrowtail C{\times}A$ is classified by $\rho : C{\times}A \to \Omega$ in $\mathbb{E}$. Let $\exists_A(r)$ be defined as $\llbracket z{:}C \vdash \exists x{:}A.\rho(z,x) \rrbracket$. From the equivalence (I) above and the Substitution Lemma 11.2 it follows that for every subobject $m : P \rightarrowtail C$ it holds that $\exists_A(r) \leq_C m$ if and only if $r \leq_{C{\times}A} \pi^*m$. That $f^*\exists_A(r) \cong \exists_A((f{\times}\mathsf{id}_A)^*r)$ is immediate from the validity of the rule (Subst) in $\mathbb{E}$. $\qquad\square$

Notice that in the proof of Theorem 13.5 we have made essential use of reasoning in the internal logic of $\mathbb{E}$ in order to establish some purely categorical properties of $\mathbb{E}$. It is a typical phenomenon of working in topos theory that one jumps back and forth between categorical/diagramatic reasoning and logical reasoning valid in the internal logic. Of course, one usually chooses the point of view allowing the more transparent argument.

To illustrate the usefulness of the internal point of view we further analyse what Theorem 13.5(3) says in purely categorical terms. Instantiating the subobject $r$ by $\langle f, \mathsf{id}_A\rangle : A \rightarrowtail C{\times}A$, the *graph* of $f : A \to C$, condition (3) of Theorem 13.5 says that $\exists_A(\langle f, \mathsf{id}_A\rangle)$ is the least subobject $i : I \rightarrowtail C$ with



i.e. the *image* of $f : A \to C$. From leastness of $i$ it follows (exercise!) that $e$ is a *strong epi*, i.e. $e$ is an epi such that $m$ is an isomorphism whenever $e = mg$ for some monomorphism $m$. One easily checks that $e : A \to I$ is a strong epi in this sense iff $\exists_A(\langle e, \mathsf{id}_A\rangle) \cong \mathsf{id}_I$ (exercise!). Thus, condition (3) of Theorem 13.5 tell us that in $\mathbb{E}$ every map $f$ factors as a strong epi followed by a monomorphism and that such factorisations are stable under pullbacks along arbitrary morphisms in $\mathbb{E}$. Categories with finite limits and such pullback stable (strong epi/mono) factorisations are traditionally called *regular*. On the other hand in every regular category $\mathbb{C}$ we have existential quantification as given by the (strong epi/mono) factorisation of $\pi \circ r$. One easily checks (exercise!) that in a topos $\mathbb{E}$ the image of $f : A \to C$ is classified by (the interpretation of) $z : C \vdash \exists x{:}A.z{=}f(x)$. This allows us also to show that in an elementary topos every epimorphism is strong: suppose $e : A \twoheadrightarrow C$ is epic then by validity of the rule (Subst) we have

$$\llbracket z : C \vdash \exists x{:}A.z{=}e(x) \rrbracket \circ e = \llbracket x' : A \vdash \exists x{:}A.e(x'){=}e(x) \rrbracket = \mathsf{t}_A = \mathsf{t}_C \circ e$$

from which it follows that

$$[\![z : C \vdash \exists x{:}A.z{=}e(x)]\!] = \mathsf{t}_C$$

i.e. that $e$ is a strong epimorphism. Thus, every topos is *balanced* in the sense that a morphism is an isomorphism whenever it is epic and monic (as a monic strong epimorphism necessarily is an isomorphism).

We suggest it as a further exercise(!) to verify that in a topos $\mathbb{E}$ a subobject $r : R \rightarrowtail C{\times}A$ is (isomorphic to) the graph of a map from $A$ to $C$ iff for the classifying map $\rho : C{\times}A \to \Omega$ the judgement

$$x : A \mid \vdash \exists! z{:}C.\, \rho(z, x)$$

holds in $\mathbb{E}$ where, as usual, $\exists! z{:}C.\rho(z, x)$ stands as an abbreviation for

$$\exists z{:}C.\, (\rho(z, x) \wedge \forall z'{:}C.(\rho(z', x) \Rightarrow z{=}z'))$$

i.e. *unique existence*. Thus, in an elementary topos $\mathbb{E}$ the following *Axiom of Unique Choice*

(AC!) $\qquad \forall R : \mathcal{P}(B{\times}A).(\forall x{:}A.\exists! y{:}B.R(y, x)) \Rightarrow \exists f{:}B^A.\forall x{:}A.R(f(x), x)$

holds for all objects $A$ and $B$ in $\mathbb{E}$, saying that functional relations from $A$ to $B$ coincide with functions from $A$ to $B$.

Adding the following two valid axiom schemes

(Ext$_{\mathrm{fun}}$) $\qquad \forall f, g : B^A.\, (\forall x{:}A.f(x){=}g(x)) \Rightarrow f{=}g$

(Ext$_{\mathrm{prop}}$) $\qquad \forall u, v : \Omega.\, (u{\Leftrightarrow}v) \Rightarrow u{=}v$

makes our version of constructive (intuitionistic) higher order logic complete w.r.t. interpretation in elementary toposes.

Above we have tacitly assumed the validity of the usual equality axioms for $=$ whose verification we leave to the inclined reader as an exercise(!). We just notice that in presence of quantification over powerobjects one could introduce equality on $A$ as a derived notion defining

$$x =_A y \;\equiv\; \forall P{:}\mathcal{P}(A).\, (P(x){\Rightarrow}P(y))$$

which idea goes back to the baroque philosopher G.W.Leibniz whose point of view was that objects are equal iff they share the same properties.

As a further illustration of the power of the internal logic of toposes we show how it can be used to prove the existence of finite colimits in toposes. First of all the initial object of $\mathbb{E}$ is (up to isomorphism) given by the subobject $0 \rightarrowtail 1$ classified by $u{:}1 \vdash \bot : \Omega$, i.e. the least subobject of 1. If $A$ and $B$ are objects of the topos then their categorical sum $A{+}B$ appears as the subobject of $\mathcal{P}(A){\times}\mathcal{P}(B)$ classified by the predicate

$$P{:}\mathcal{P}(A), Q{:}\mathcal{P}(B) \vdash \big(\exists! x{:}A.P(x) \vee \exists! y{:}B.Q(y)\big) \wedge \big(\forall x{:}A, y{:}B.\neg(P(x) \wedge Q(y))\big)$$

A coequalizer of $f, g : A \to B$ can be constructed as follows: first define (via universal quantification over $\mathcal{P}(B{\times}B)$) the least equivalence relation $R$ on $B$ with $R(f(a), g(a))$ for all $a \in A$ (all understood in the internal sense) and then take as coequaliser of $f$ and $g$ the (strong) epi $q : B \twoheadrightarrow Q$ appearing in the (strong epi/mono) factorisation

$$B \xrightarrow{\;q\;} Q$$

with maps to $\mathcal{P}(B)$ labelled $\hat{1}(\rho)$ and $m$

where $\rho : B{\times}B \to \Omega$ is the classifying map for the subobject $R \rightarrowtail B{\times}B$. As a further application of the internal logic of toposes consider the following *extensivity property* of (binary sums).

**Theorem 13.6** (extensivity of sums)
*Suppose the squares*

$$\begin{array}{ccc} B_1 & \xrightarrow{g_1} & A_1 \\ {\scriptstyle b_1}\downarrow & & \downarrow{\scriptstyle a_1} \\ J & \xrightarrow{f} & I \end{array} \qquad \begin{array}{ccc} B_2 & \xrightarrow{g_2} & A_2 \\ {\scriptstyle b_2}\downarrow & & \downarrow{\scriptstyle a_2} \\ J & \xrightarrow{f} & I \end{array}$$

*commute in an elementary topos $\mathbb{E}$. Then both squares are pullbacks if and only if the square*

$$\begin{array}{ccc} B_1{+}B_2 & \xrightarrow{g_1+g_2} & A_1{+}A_2 \\ {\scriptstyle [b_1, b_2]}\downarrow & & \downarrow{\scriptstyle [a_1, a_2]} \\ J & \xrightarrow{f} & I \end{array}$$

*is a pullback.*

*Proof:* Employ the following construction of pullbacks using the internal logic of $\mathbb{E}$ : given maps $f : A \to C$ and $g : B \to C$ in $\mathbb{E}$ then their pullback is given by the subobject $P \rightarrowtail A \times B$ classified by $x{:}A, y{:}B \vdash f(x) =_C g(y)$. Details are left to the reader. $\square$

**Kripke-Joyal Semantics**

We have seen already how to interpret higher order logic in a topos where a formula $\varphi$ in context $x_1 : A_1, \ldots, x_n : A_n$ gets interpreted as a subobject of $A_1 \times \cdots \times A_n$.[31] This has the disadvantage that the most pleasant illusion of "elements" is totally lost. Well, for non-wellpointed toposes $\mathbb{E}$ one certainly doesn't know a predicate $\varphi : A \to \Omega$ when knowing the collection of global points $a : 1 \to A$ of $A$ with $\varphi \circ a = \mathsf{t}$. But, obviously, the predicate $\varphi : A \to \Omega$ is fully determined by the collection of all *generalised elements* $a : I \to A$ *satisfying* $\varphi$ (notation $a \Vdash \varphi$), i.e. $\varphi \circ a = \mathsf{t}_I$. Now Kripke-Joyal semantics exploits this observation by specifying (by structural recursion over $\varphi$) the collection of all generalised elements $a$ with $a \Vdash \varphi$ thus reestablishing the *form* of the Tarskian definition of truth.

Before giving the clauses of the Kripke-Joyal semantics we state the following two principles holding for $\Vdash$ (often called *forcing relation*)

**Monotonicity**    If $a \Vdash \varphi$ for $a : I \to A$ and $f : J \to I$ then $af \Vdash \varphi$

**Local Character**    If $a : I \to A$ and $e : J \twoheadrightarrow I$ epic with $ae \Vdash \varphi$ then $a \Vdash \varphi$.

whose straightforward verification we leave as an exercise(!) for the reader.

**Theorem 13.7** (Truth Conditions à la Kripke-Joyal)
*In an elementary topos $\mathbb{E}$ for predicates $\varphi, \psi : A \to \Omega$ and generalized elements $a : I \to A$ we have that*

($\top$) $a \Vdash \mathsf{t}_A$ *always holds*

($\wedge$) $a \Vdash \varphi \wedge \psi$ *iff* $a \Vdash \varphi$ *and* $a \Vdash \psi$

---

[31]Notice that at least from now on we often drop the semantic brackets $[\![ \cdot ]\!]$ when they are clear from the context. Moreover, in the end the distinction between object and metalanguage is not the greatest insight of logic after all!

$(\Rightarrow)$ $a \Vdash \varphi{\Rightarrow}\psi$ *iff for all* $f : J \to I$, $af \Vdash \varphi$ *implies* $af \Vdash \psi$

$(\bot)$ $a \Vdash \mathsf{f}_A$ *iff* $I \cong 0$

$(\vee)$ $a \Vdash \varphi \vee \psi$ *iff there exist jointly epic maps* $f : J \to I$ *and* $g : K \to I$, *i.e.* $[f, g] : J{+}K \twoheadrightarrow I$ *epic, such that* $af \Vdash \varphi$ *and* $ag \Vdash \psi$

*and for predicates* $\rho : C{\times}A \to \Omega$ *and generalised elements* $c : I \to C$ *we have*

$(\forall)$ $c \Vdash \forall_A(\rho)$ *iff* $c{\times}\mathsf{id}_A \Vdash \rho$

$(\exists)$ $c \Vdash \exists_A(\rho)$ *iff* $\langle ce, a \rangle \Vdash \rho$ *for some epi* $e : J \twoheadrightarrow I$ *and* $a : J \to A$.

*Proof:* Clause $(\top)$ is trivial and clause $(\wedge)$ follows from Lemma 13.3. Clause $(\bot)$ is left as an easy exercise!

$(\Rightarrow)$ : Suppose $a \Vdash \varphi{\Rightarrow}\psi$. Then by monotonicity for all $f : J \to I$ we have $af \Vdash \varphi{\Rightarrow}\psi$ from which it follows by Lemma 13.7 that $\varphi af \leq_J \psi af$. Thus, if $af \Vdash \varphi$, i.e. $\varphi af = \mathsf{t}_J$, then $\psi af = \mathsf{t}_J$, i.e. $af \Vdash \psi$. For the reverse direction suppose that for all $f : J \to I$, $af \Vdash \varphi$ implies $af \Vdash \psi$. By Lemma 13.6 for $a \Vdash \varphi{\Rightarrow}\psi$ it suffices to show that $\varphi a \leq_I \psi a$. Let $m : P \rightarrowtail I$ be the subobject classified by $\varphi a$. Then we have $\varphi am = \mathsf{t}_P$. Thus, by assumption it follows that $\psi am = \mathsf{t}_P$ and, therefore, for the subobject $n : Q \rightarrowtail I$ classified by $\psi a$ we have that $m \leq_I n$, i.e. $\varphi a \leq_I \psi a$ as desired.

$(\vee)$ : First observe that for subobjects $m : P \rightarrowtail X$ and $n : Q \rightarrowtail X$ of an object $X$ in $\mathbb{E}$ with $m \vee n \cong \mathsf{id}_X$ then $m$ and $n$ are jointly epic, i.e. $[m, n] : P{+}Q \twoheadrightarrow X$.

For $\varphi, \psi : A \to \Omega$ let $m$ and $n$ be the corresponding subobjects of $A$ and $m'$ and $n'$ their pullbacks along $a : I \to A$ as depicted in

$$
\begin{array}{ccc}
P' \longrightarrow P & \qquad & Q' \longrightarrow Q \\
m' \downarrow \quad \lrcorner \quad \downarrow m & \qquad & n' \downarrow \quad \lrcorner \quad \downarrow n \\
I \xrightarrow{\ a\ } A & \qquad & I \xrightarrow{\ a\ } A
\end{array}
$$

Now suppose that $a \Vdash \varphi{\vee}\psi$, i.e. $\varphi a \vee \psi a = (\varphi{\vee}\psi)a = \mathsf{t}_I$. Then $m' {\vee} n' \cong I$ as $m'$ and $n'$ are classified by $\varphi a$ and $\psi a$, respectively. Thus, by the above observation $m'$ and $n'$ are jointly epic. Moreover, we have $\varphi am' = \mathsf{t}_{P'}$ and $\psi an' = \mathsf{t}_{Q'}$. Thus, the desired claim holds putting $f = m'$ and $g = n'$.

For the reverse direction suppose that $f : J \to I$ and $g : K \to I$ are jointly epic with $af \Vdash \varphi$ and $ag \Vdash \psi$. Then $af$ and $ag$ factor through $m$ and $n$,

respectively. Accordingly, we have that both $af$ and $ag$ factor through $m{\vee}n$. Thus, their source tupling $a \circ [f,g] = [af, ag]$ factors through $m{\vee}n$ as well from which it follows that $\varphi{\vee}\psi \circ a \circ [f,g] = \mathsf{t}_{J+K}$. Thus, finally $\varphi{\vee}\psi \circ a = \mathsf{t}_I$ because $[f,g]$ is epic.

($\forall$) : First notice that $\forall_A(\rho) \circ c = \forall_A(\rho \circ (c{\times}\mathsf{id}_A))$. Thus, by Theorem 13.4 $c \Vdash \forall_A(\rho)$ is equivalent to $\rho \circ (c{\times}\mathsf{id}_A)) = \mathsf{t}_{I \times A}$, i.e. $c{\times}\mathsf{id}_A \Vdash \rho$ as desired.

($\exists$) : Let $r : R \rightarrowtail C{\times}A$ be the subobject classified by $\rho$ and let

$$
\begin{array}{ccc}
R & \xrightarrow{\ e'\ } & \mathsf{im}(\pi r) \\
& {}_{\pi r}\searrow & \big\downarrow{}^{m'} \\
& & C
\end{array}
$$

be the (strong epi/mono) factorisation of $\pi r$ where by Theorem 13.5 the subobject $m'$ is classified by $\exists_A(\rho)$.

Suppose $c \Vdash \exists_A(\rho)$. Then $c$ factors through $m'$ via some map $f : I \to \mathsf{im}(\pi r)$, i.e. $c = m'f$. Now consider the diagram

$$
\begin{array}{ccccc}
J & \xrightarrow{\ g\ } & R & \xrightarrow{\ r\ } & C{\times}A \\
{}_{e}\big\downarrow & \lrcorner & {}_{e'}\big\downarrow & & \big\downarrow{}^{\pi} \\
I & \xrightarrow{\ f\ } & \mathsf{im}(\pi r) & \xrightarrow{\ m'\ } & C
\end{array}
$$

and let $a : J \to A$ be the map with $rg = \langle ce, a \rangle$. Thus, the map $\langle ce, a \rangle$ factors (via $g$) through the subobject $r$ classified by $\rho$, i.e. $\langle ce, a \rangle \Vdash \rho$ as desired.

For the reverse direction suppose $e : J \twoheadrightarrow I$ is an epi and $a : J \to A$ with $\langle ce, a \rangle \Vdash \rho$. Then $\langle ce, a \rangle$ factors through $r$ via some map $g : J \to R$, i.e. $\langle ce, a \rangle = rg$. Then $ce = \pi rg = m'e'g$ from which it follows that $ce \Vdash \exists_A(\rho)$ because $\exists_A(\rho)$ classifies $m'$. Thus, due to the local character of $\Vdash$ from $ce \Vdash \exists_A(\rho)$ it follows that $c \Vdash \exists_A(\rho)$ because $e$ is epic. $\qquad\square$

In case of presheaf toposes Kripke-Joyal semantics takes the following even simpler form based on the observation that a predicate $\varphi : A \to \Omega$ is determined already by the set of generalized elements $a : \mathsf{Y}(I) \to A$ with $\varphi a = \mathsf{t}_{\mathsf{Y}(I)}$.

**Theorem 13.8** (Kripke-Joyal semantics for presheaf toposes)
*Let $\mathbb{E}$ be a presheaf topos $\widehat{\mathbb{C}}$ for some small category $\mathbb{C}$. If $A$ is an object of $\mathbb{E}$,*

$a \in A(I)$ and $f : J \to I$ is a morphism in $\mathbb{C}$ then we write $a{\restriction}f$ for $A(f)(a)$ to make notation more readable. For predicates $\varphi : A \to \Omega$ and $a \in A(I)$ we write $I \Vdash \varphi(a)$ iff $a \in P(I)$ for the subpresheaf $P \hookrightarrow A$ classified by $\varphi$. Then for predicates $\varphi, \psi : A \to \Omega$ and $a \in A(I)$ we have that

($\top$) $I \Vdash \mathsf{t}_A(a)$ always holds

($\wedge$) $I \Vdash (\varphi \wedge \psi)(a)$ iff $I \Vdash \varphi(a)$ and $I \Vdash \psi(a)$

($\Rightarrow$) $I \Vdash (\varphi \Rightarrow \psi)(a)$ iff for all $f : J \to I$, $J \Vdash \varphi(a{\restriction}f)$ implies $J \Vdash \psi(a{\restriction}f)$

($\bot$) $I \Vdash \mathsf{f}_A(a)$ never holds

($\vee$) $I \Vdash (\varphi \vee \psi)(a)$ iff $I \Vdash \varphi(a)$ or $I \Vdash \psi(a)$

and for predicates $\rho : C \times A \to \Omega$ and $c \in C(I)$ we have

($\forall$) $I \Vdash \forall_A(\rho)(c)$ iff $J \Vdash \rho(c{\restriction}f, a)$ for all $f : J \to I$ in $\mathbb{C}$ and $a \in A(J)$

($\exists$) $I \Vdash \exists_A(\rho)(c)$ iff $I \Vdash \rho(c, a)$ for some $a \in A(I)$.

*Proof:* Exploiting the 1-1-correspondence between $A(I)$ and morphism from $\mathsf{Y}_\mathbb{C}(I)$ to $A$ in $\widehat{\mathbb{C}}$ as ensured by Yoneda most of the claims are immediate from Theorem 13.7. For ($\bot$), ($\vee$) and ($\exists$) one has to exploit the specific nature of presheaf toposes where $0(I) = \emptyset$, $(P \vee Q)(I) = P(I) \cup Q(I)$ and where a map $e : A \to B$ is epic iff all $e_I : A(I) \to B(I)$ are surjective functions. Details are left to the reader as a straightforward exercise! $\square$

## Axiom of Choice in Toposes

We have seen already that the axiom of unique choice holds in every topos. But what about

(AC) $\quad \forall R : \mathcal{P}(B \times A).(\forall x{:}A.\exists y{:}B.R(y,x)) \Rightarrow \exists f{:}B^A.\forall x{:}A.R(f(x),x)$

the general axiom of choice in higher order logic?

**Theorem 13.9** (Diaconescu)
*Every elementary topos validating* (AC) *is already boolean.*

*Proof:* Let 0 and 1 be the two global elements of $2 = 1+1$ as given by left and right injection of 1 into 2. For $\varphi \in \Omega$ we define

$$V_i = \{x \in 2 \mid x{=}i \vee \varphi\} \in \mathcal{P}(2)$$

for $i{=}0, 1$. Thus it holds that $\forall V{\in}\{V_0, V_1\}.\exists x{\in}2.\, x \in V$. From this instantiating $A$ by $\{V_0, V_1\}$ and $B$ by 2 in (AC) it follows that

$$\exists f \in 2^{\{V_0, V_1\}}.\forall V{\in}\{V_0, V_1\}.f(V) \in V$$

Thus, it holds that $(f(V_0) \in V_0) \wedge (f(V_1) \in V_1)$, i.e. by definition of the $V_i$ that $(f(V_0){=}0 \vee \varphi) \wedge (f(V_1){=}1 \vee \varphi)$, from which it follows by distributivity that

$$(f(V_0){=}0 \wedge f(V_1){=}1) \vee \varphi$$

But the latter entails $\varphi \vee \neg\varphi$ which can be seen as follows. Of course, $\varphi$ entails $\varphi \vee \neg\varphi$. On the other hand $f(V_0){=}0 \wedge f(V_1){=}1$ implies $\neg\varphi$ because if $\varphi$ then $V_0 = V_1$ from which it follows that $0 = 1$, a contradiction. $\qquad\square$

Thus, unrestricted choice entails classical logic but not[32] *vice versa.* See e.g. Chapter VI of [MM] for boolean toposes not validating AC.
Validity of the logical scheme (AC) in a topos $\mathbb{E}$ is equivalent to the following principle called *Internal Axiom of Choice*

(IAC)    If $e$ is epic then $e^A$ is epic, too, for all objects $A$ of $\mathbb{E}$.

as shown on pp. 312-315 of [MM]. A stronger principle is the following principle called *External Axiom of Choice*

(EAC)    Every epimorphism $e$ splits, i.e. there is a section $s$ with $e \circ s = \mathsf{id}$.

Obviously, (EAC) entails (IAC) because every functor preserves split epimorphism and thus, in particular, the functors $(-)^A$ do. Notice that for nontrivial groups $\mathbb{G}$ in the presheaf topos $\widehat{\mathbb{G}}$ the unique map from the representable presheaf to 1 is a non-split epi because the representable presheaf does not have any global element. We leave it as an exercise(!) (see e.g. [FS]) that a topos satisfies (EAC) if and only if it satisfies (IAC) and 1 is *projective*, i.e. every $A$ with $A \twoheadrightarrow 1$ has a global element.[33] We leave it as an exercise(!) to show that a well-pointed topos satisfying (IAC) satisfies (EAC) as well.

---

[32]Notice that even the stronger theory ZF does not prove AC.

[33]Generally an object $C$ in $\mathbb{E}$ is called *projective* (or a *choice object*) iff $\mathbb{E}(C, -) : \mathbb{E} \to \mathbf{Set}$ preserves epimorphisms, i.e. for every epimorphism $e : A \twoheadrightarrow B$ and every $g : C \to B$ there exists an $f : C \to A$ with $e \circ f = g$.

**Natural Numbers in Toposes**

The category $\mathbf{Set}_f$ of finite sets and maps between them is an example of a well-pointed topos satisfying (EAC). Thus, in order to guarantee the existence of some infinite type as e.g. the type of natural numbers we need a further axiom.[34]

The following definition of *natural numbers object* (NNO) like many other notions in this field goes back to F.W.Lawvere.

**Definition 13.5** (Natural Numbers Object)
*A natural numbers object (NNO)* in a category $\mathbb{C}$ with terminal object $1$ is *an object $N$ in $\mathbb{C}$ together with morphisms $z : 1 \to N$ (zero) and $s : N \to N$* (successor) *such that for every object $A$ in $\mathbb{C}$ and morphisms $a : 1 \to A$ and* $t : A \to A$ *there exists a* unique *morphism $h : N \to A$ making the diagram*

$$
\begin{array}{ccc}
A & \xrightarrow{\ t\ } & A \\
{\scriptstyle h}\big\uparrow & & \big\uparrow{\scriptstyle h} \\
N & \xrightarrow[\ s\ ]{} & N
\end{array}
\qquad {\scriptstyle a}\nearrow \quad 1 \xrightarrow[z]{} N
$$

*commute.*                                                                 $\Diamond$

The intuition behind the definition of NNO is that for every $a \in A$ and $t : A \to A$ there exists a unique sequence $h : N \to A$ such that

$$ h(z) = a \qquad \text{and} \qquad h(s(n)) = t(h(n)) $$

i.e. the definition of NNO captures the idea of *iteration*. We leave it as an exercise(!) to the reader to show that in presence of exponentials (i.e. "higher types") from iteration one can derive the principle of *primitive recursion* guaranteeing for all maps $g : A \to B$ and $h : N{\times}A{\times}B \to B$ the existence of a unique map $R(g,h) = f : N{\times}A \to B$ such that

$$ f(z,a) = g(a) \qquad \text{and} \qquad f(s(n),a) = h(n,a,f(n,a)) $$

for all $n \in N$ and $a \in A$.

The definition of NNO not only guarantees the existence of sufficiently many algorithmic functions but also the usual *induction principle*

---

[34]Just as in ZFC which is modelled by the hereditary finite sets as long as one does not postulate the *infinity axiom*.

**Theorem 13.10** (induction principle for NNOs)
*Let $\mathbb{E}$ be an elementary topos with NNO $N$. If $m : P \rightarrowtail N$ is a subobject of $N$ such that*

(1) $z : 1 \to N$ *factors through $m$ and*

(2) $s \circ m$ *factors through $m$*

*then $m$ is an isomorphism.*
*Moreover, the topos $\mathbb{E}$ validates the induction principle*

$$(\mathrm{Ind}_N) \quad \forall P{:}\mathcal{P}(N).\ P(z) \Rightarrow (\forall n{:}N.P(n) \Rightarrow P(s(n))) \Rightarrow \forall n{:}N.P(n) \ .$$

*Proof:* Let $a : 1 \to P$ with $ma = z$ and $t : P \to P$ with $sm = mt$. Then there exists a (unique) map $h : N \to P$ such that $hz = a$ and $hs = th$. Thus we have



from which it follows that $mh = \mathsf{id}_N$. But then also $mhm = m$ from which it follows that $hm = \mathsf{id}_P$ because $m$ is monic. Thus $m$ is an isomorphism with $h$ as its inverse.
The verification of $(\mathrm{Ind}_N)$ is left as an exercise(!) to the reader. $\qquad\square$

It is routine to show that the image of the NNO in **Set** under the functor $\Delta : \mathbf{Set} \to \widehat{\mathbb{C}}$ is again a NNO in $\widehat{\mathbb{C}}$.

## Axiomatizing the Category $\mathcal{S}$ of Constant Sets

The foundational claim of topos theory as forcefully propagated by Lawvere (see [LR]) is that (most) mathematics can be performed in elementary toposes with NNO. As we have seen there are plenty of different such toposes.
In order to get more "classical" one might also postulate booleanness or even the axiom of choice. Lawvere has suggested (see e.g. [LR]) to axiomatize

the classical category $\mathcal{S}$ of *constant sets* as an elementary wellpointed topos admitting a NNO and satisfying the axiom of choice.

Notice, however, that not every such topos $\mathcal{S}$ will be isomorphic to **Set** as e.g. $\mathcal{S}$ may be countable due to the Theorem of Löwenheim and Skolem (guaranteeing the existence of countable models for consistent theories formulated in a countable language). But, actually, the category **Set** itself is also a relative notion, namely e.g. some model of ZFC which again may be countable (from the external point of view). Thus, in the end nothing (essential) is lost when replacing **Set** by an arbitrary topos $\mathcal{S}$ of constant sets satisfying Lawvere's axioms.

# 14 Some Exercises in Presheaf Toposes

An excellent reference for examples of and computations in presheaf toposes is [PRZ] which we strongly recommend for this purpose.

A topos is called *localic* iff subobjects of 1 (called *subterminals*) generate in which case we say that the topos is *localic*.

**Lemma 14.1** *Let $\mathbb{C}$ be a small category. Then $\widehat{\mathbb{C}}$ is localic iff $\mathbb{C}$ is posetal.*

*Proof:* Suppose $\mathbb{C}$ is posetal. Then all representable objects of $\widehat{\mathbb{C}}$ are subterminal. Since representable objects generate it follows that $\widehat{\mathbb{C}}$ is localic.
For the reverse direction suppose $\widehat{\mathbb{C}}$ is localic. We show that every representable object $\mathsf{Y}(I)$ is subterminal. For that purpose consider the subobject $S \subseteq \mathsf{Y}(I)$ consisting of all $u : J \to I$ with $uv_1 = uv_2$ for all $v_1, v_2 : K \to J$. Let $\chi : \mathsf{Y}(I) \to \Omega$ be the classifying map for $S$. Suppose $\chi \neq \top_{\mathsf{Y}(I)}$. Since by assumption $\widehat{\mathbb{C}}$ is localic there exists a $\tau : U \to \mathsf{Y}(I)$ with $U$ subterminal and $\chi \circ \tau \neq \top_U$. Thus there exists a map $v : J \to I$ with $\chi_J(v) \neq \top_J$ and $vw_1 = vw_2$ for all $w_1, w_2 : K \to J$, i.e. $v \in S(J)$ but $\chi_J(v) \neq \top_J$ which clearly is impossible since $\chi$ classifies $S \subseteq \mathsf{Y}(I)$. Thus we have shown that $\chi = \top_{\mathsf{Y}(I)}$ and, accordingly, we have $S = \mathsf{Y}(I)$ from which it follows that $\mathsf{id}_I \in S$ and therefore $v_1 = v_2$ whenever $v_1, v_2 : J \to I$. Since this holds for all objects $I$ of $\mathbb{C}$ it follows that $\mathbb{C}$ is indeed posetal. $\square$

**Corollary 14.1** *If for a small category $\mathbb{C}$ the topos $\widehat{\mathbb{C}}$ is localic and 2-valued then $\mathbb{C} \simeq \mathbf{1}$ and thus $\widehat{\mathbb{C}} \simeq \mathbf{Set}$.*

*Proof:* If $\widehat{\mathbb{C}}$ is localic then by the previous theorem $\mathbb{C}$ is posetal. If $\mathbb{C}$ were not equivalent to the terminal category $\mathbf{1}$ then in $\widehat{\mathbb{C}}$ there would exist nontrivial subterminals corresponding to nontrivial global elements of $\Omega$ which is impossible due to the assumption that $\widehat{\mathbb{C}}$ is 2-valued. $\square$

An alternative argument is the following one: if $\widehat{\mathbb{C}}$ is 2-valued then 0 and 1 are the only subterminals from which it follows by the assumption of $\widehat{\mathbb{C}}$ being localic that $\mathbb{C}$ is wellpointed which we have already seen to entail that $\mathbb{C}$ is equivalent to the terminal category.

**Theorem 14.1** *For a small category $\mathbb{C}$ the global sections functor $\Gamma : \widehat{\mathbb{C}} \to \mathbf{Set}$ has left adjoint $\Delta$ which in turn has a further left adjoint $\pi$.*

*Proof:* The left adjoint $\Delta$ to $\Gamma$ sends a set $S$ to the constant presheaf with value $S$ and $f : S \to T$ to the natural transformation $\Delta(f)$ with $\Delta(f)_I = f$ for all $I \in \mathbb{C}$. The unit of $\Delta \dashv \Gamma$ at $S$ sends $x \in S$ to the global element $\eta_S(x) : 1 \to \Delta(S)$ with $(\eta_S(x))_I(*) = x$. For $f : S \to \Gamma(A)$ the unique $\phi : \Delta(S) \to A$ with $\Gamma(\phi) \circ \eta_S = f$ is given by $\phi_I(x) = f(x)_I(*) \in A(I)$.

The left adjoint $\pi$ to $\Delta$ sends a presheaf $A$ to the set $\pi(A)$ of connected components of $\mathbf{Elts}(A)$. The unit $\eta_A : A \to \Delta(\pi(A))$ sends $a \in A$ to the connected components inhabitated by $a$. Obviously, a natural transformation $\phi : A \to \Delta(S)$ is constant on connected components and, therefore, the unique $f : \pi(A) \to S$ with $\Delta(f) \circ \eta_S = \phi$ sends a connected component of $A$ to the constant value of $\phi$ on it. $\qquad\square$

**Theorem 14.2** *Let $\mathbb{C}$ be a small category. Then $\Gamma : \widehat{\mathbb{C}} \to \mathbf{Set}$ has a right adjoint $\nabla$ iff the terminal object in $\widehat{\mathbb{C}}$ appears as retract of a representable object.*

*Proof:* If $\Gamma$ has a right adjoint then $\Gamma$ preserves colimits. Since 1 is a colimit of representable objects and $\Gamma$ preserves colimits it follows that $\Gamma(\mathsf{Y}(I)) \neq \emptyset$ for some object $I$ in $\mathbb{C}$. Thus there exists $e : 1 \to \mathsf{Y}(I)$ exhibiting 1 as retract of the representable object $\mathsf{Y}(I)$.

Suppose 1 appears as retract of some $\mathsf{Y}(I)$. Then there exists $e : 1 \to \mathsf{Y}(I)$. For every object $J$ of $\mathbb{C}$ we have $e_J : J \to I$ and $e_J \circ u = e_K$ for all $u : K \to J$ in $\mathbb{C}$. In particular, for $e_I$ we have $e_I e_I = e_I$, i.e. $e_I$ is a retract. Next observe that $\Gamma(A) \cong \{a \in A(I) \mid A(e_I)(a) = a\}$ by sending $\alpha : 1 \to A$ to $\alpha_I(*)$. Conversely, given $a \in A(I)$ with $A(e_I)(a) = a$ this induces an $\alpha : 1 \to A$ with $\alpha_J(*) = A(e_J)(a)$ for objects $J$ in $\mathbb{C}$. Moreover, for $f : A \to B$ and $a \in A(I)$ with $A(e_I)(a) = a$ we have $B(e)(f_I(a)) = f_I(A(e_I)(a)) = f_I(a)$. Thus, up to isomorphism $\Gamma$ looks as follows: $\Gamma(A) = \{a \in A(I) \mid A(e_I)(a) = a\}$ and $\Gamma(f)(a) = f_I(a)$. From now on we work with this isomorphic copy of $\Gamma$. Define $G : \mathbb{C} \to \mathbf{Set}$ as $G = \Gamma \circ \mathsf{Y}$. More explicitly, we have $G(J) = \{u : I \to J \mid u e_I = u\}$ and $G(v)(u) = vu$ for $v : J \to K$ in $\mathbb{C}$. Then the right adjoint to $\Gamma$ is given by $\nabla(S) = S^{G(-)}$. Its counit at $S$ is given by $\varepsilon_S : \Gamma(\nabla(S)) \to S : \gamma \mapsto \gamma(e_I)$. Suppose $f : \Gamma(A) \to S$. We have to show that there exists a unique $\phi : A \to \nabla(S)$ with $\varepsilon_S \circ \Gamma(\phi) = f$, i.e.

$\phi_I(a)(e_I) = f(a)$ for $a \in \Gamma(A)$. By naturality of $\phi$ we have for $u \in G(J)$ that

$$
\begin{array}{ccc}
A(J) & \xrightarrow{\phi_J} & S^{G(J)} \\
{\scriptstyle A(u)} \downarrow & & \downarrow {\scriptstyle S^{G(u)}} \\
A(I) & \xrightarrow[\phi_I]{} & S^{G(I)}
\end{array}
$$

and thus for $a \in A(J)$ we have (since $A(e_I)(A(u)(a)) = A(ue_I)(a) = A(u)(a)$ and thus $A(u)(a) \in \Gamma(A)$) that

$$
\begin{aligned}
\phi_J(a)(u) &= \phi_J(G(u)(e_I)) = S^{G(u)}(\phi_J(a))(e_I) = \\
&= (S^{G(u)} \circ \phi_J)(a)(e_I) = (\phi_I \circ A(u))(a)(e_I) = \\
&= \phi_I(A(u)(a))(e_I) = \\
&= f(A(u)(a))
\end{aligned}
$$

from which it follows that $\phi$ is determined uniquely $f$. Moreover, this observation tells us how $\phi$ has to look like provided it exists, namely

$$
\phi_J(a)(u) = f(A(u)(a))
$$

for $a \in A(J)$ and $u \in G(J)$. It is a straightforward exercise to show that the so defined $\phi$ is actually natural. Moreover, for $a \in \Gamma(A)$ we have

$$
(\varepsilon_S \circ \Gamma(\phi))(a) = \phi_I(a)(e_I) = f(A(e_I)(a)) = f(a)
$$

and thus we have $\varepsilon_S \circ \Gamma(\phi) = f$ as desired. $\qquad\square$

For the reverse implication of Theorem 14.2 we have the following *alternative Proof:* Suppose $e_0 : 1 \to \mathsf{Y}(I_0)$. Let $G = \Gamma \circ \mathsf{Y} : \mathbb{C} \to \mathbf{Set}$. The right adjoint to $\Gamma$ is given by $\nabla(S) = S^{G(-)}$. The unit $\eta_A : A \to \nabla(\Gamma(A))$ of the adjunction $\Gamma \dashv \nabla$ at $A$ is given by

$$
(\eta_A)_I(a)(i) = a \circ i
$$

for $a : I \to A$ and $i : 1 \to I$. It is a straightforward exercise to show that the so defined $\eta$ is actually natural. For showing that $\eta$ is the unit of $\Gamma \vdash \nabla$ we have to show that for $\phi : A \to \nabla(S)$ there exists a unique $f : \Gamma(A) \to S$ with $\nabla(f) \circ \eta_A = \phi$, i.e.

$$
\phi_I(a)(i) = \big[\nabla(f) \circ \eta_A\big]_I(a)(i) = f(a \circ i)
$$

108

for $a : I \to A$ and $i : 1 \to I$. For $\alpha : 1 \to A$ we may put $a := \alpha \circ !_{I_0} : I_0 \to A$ and then have

$$f(\alpha) = f(a \circ e_0) = \phi_{I_0}(a)(e_0) = \phi_{I_0}(\alpha \circ !_{I_0})(e_0)$$

from which it follows that $f$ is uniquely determined by $\phi$. Now for $a : I \to A$ and $i : 1 \to I$ we have

$$\big[\nabla(f) \circ \eta_A\big]_I(a)(i) = f(a \circ i) = \phi_{I_0}(a \circ i \circ !_{I_0})(e_0) = \phi_I(a)(i \circ !_{I_0} \circ e_0) = \phi_I(a)(i)$$

and thus $\nabla(f) \circ \eta_A = \phi$ as desired. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Notice that in case $\Gamma : \widehat{\mathbb{C}} \to \mathbf{Set}$ has a right adjoint $\nabla$ we have $\Gamma\Delta \cong \mathsf{Id}_{\mathbf{Set}}$. From Lemma 8.1 it follows that the counit $\eta : \mathsf{Id}_{\mathbf{Set}} \to \Gamma\Delta$ is a natural isomorphism and thus by Theorem 8.5 the functor $\Delta$ is full and faithful. By uniqueness of adjoints from $\Gamma\Delta \cong \mathsf{Id}_{\mathbf{Set}}$ it follows that $\Gamma\nabla \cong \mathsf{Id}_{\mathbf{Set}}$ and thus by the dual versions of Lemma 8.1 and Theorem 8.5 that $\nabla$ is also full and faithful and that $\varepsilon_S : \Gamma(\nabla(S)) \to S$ is an isomorphism for all $S \in \mathbf{Set}$. That's what Lawvere calls a UIAO situation (for *Unity and Identity of Adjoint Opposites*), i.e. $L \dashv F \dashv R$ with $F \circ L = \mathsf{Id} = F \circ R$ where as above from Lemma 8.1 it follows that both $L$ and $R$ are full and faithful.

For the Sierpiński topos $\widehat{\mathbf{2}}$ the global sections functor is given by $\Gamma(A) = A(1)$. Its right adjoint $\nabla$ is given by $\nabla(S)(1) = S$ and $\nabla(S)(0) = \{*\}$. The left adjoint of $\Gamma$ is given by $\Delta(S)(0{\to}1) = \mathsf{id}_S$. The left adjoint $\pi$ of $\Delta$ is given by $\pi(A) = A(0)$. The functor $\pi$ has a further left adjoint given by $L(S)(1) = \emptyset$ and $L(S)(0) = S$.

Let $\mathbb{M}_n$ be the monoid whose elements besides the unit 1 are $e_0, \ldots, e_{n-1}$ with $e_i e_j = e_i$ for $1 \leq i, j < n$.

The presheaf topos $\widehat{\mathbb{M}_1}$ consists of retractions and maps between them since for an object $X$ of $\widehat{\mathbb{M}_1}$ the map $r_X(x) = x \cdot e_0$ is a retraction (because $r_X(r_X(x)) = x \cdot e_0 \cdot e_0 = x \cdot e_0 e_0 = x \cdot e_0 = r_X(x)$) and for a morphism $f : X \to Y$ we have $f(r_X(x)) = f(x \cdot e_0) = f(x) \cdot e_0 = r_Y(f(x))$. For $X$ in $\widehat{\mathbb{M}_1}$ we have $\Gamma(X) = \{x \in X \mid r_X(x) = x\}$. For a set $S$ let $\Delta(S)$ be the set $S$ with right action of $\mathbb{M}_1$ given by $x \cdot e_0 = x$ for $x \in S$. One easily checks that $\Delta \vdash \Gamma \vdash \Delta$, i.e. that $\Delta$ and $\nabla$ coincide.[35] Thus, also $\pi$ coincides with $\Gamma$.

The presheaf topos $\widehat{\mathbb{M}_2}$ may be considered as the topos of *reflexive graphs* where for an object $X$ of $\widehat{\mathbb{M}_2}$ the underlying set of $X$ is thought of as the set

---

[35]The transpose of $f : S \to \Gamma(X)$ is the morphism $f : \Delta(S) \to X$ and the transpose of $g : \Gamma(X) \to S$ is the morphism $\check{g} : X \to \Delta(S)$ sending $x \in X$ to $\check{g}(x) = g(x \cdot e_0)$

of *edges* and $\Gamma(X) = \{x \in X \mid x \cdot e_0 = x = x \cdot e_1\}$ is thought of as the set of *loops* (which are identified with the *nodes* of the graph). Since for $x \in X$ we have $x \cdot e_i \cdot e_j = x \cdot e_i e_j = x \cdot e_i$ it follows that $x \cdot e_0$ and $x \cdot e_1$ are loops corresponding to the source and target node of edge $x$, respectively. Notice that $\Gamma(X)$ is the set of global elements of $X$ and for $h : X \to Y$ the map $\Gamma(h)$ is the restriction of $h$ to $\Gamma(X)$ (which factors through $\Gamma(Y)$). The left adjoint $\Delta$ of the global sections functor $\Gamma$ sends a set $S$ to the set $S$ on which $\mathbb{M}_2$ acts as $x \cdot e_i = x$ for $i = 0, 1$. The right adjoint $\nabla$ to $\Gamma$ sends a set $S$ to the $\mathbb{M}_2$-action $\nabla(S)$ whose underlying set is $S \times S$ on which $\mathbb{M}_2$ acts as $(x_0, x_1) \cdot e_i = (x_i, x_i)$ for $i = 0, 1$. Notice that $\nabla(S)$ is the "chaotic" graph where for all $x, y \in S$ there exists precisely one edge from $x$ to $y$ whereas $\Delta(S)$ is the "discrete" graph with as few edges as possible. The left adjoint $\pi$ to $\Delta$ sends $X$ to the connected components of the graph $X$, i.e. $\Gamma(X)/\sim_X$ where $\sim_X$ is the least equivalence relation on $\Gamma(X)$ containing all pairs of the form $(x \cdot e_0, x \cdot e_1)$. For $f : X \to Y$ the map $f$ respects $\sim_X$ to $\sim_Y$ and, therefore, the map $\pi(f) : \pi(X) \to \pi(Y) : [x]_{\sim_X} \mapsto [f(x)]_{\sim_Y}$ is well defined and provides the morphism part of the functor $\pi$.

Notice that for the topos of graphs $\widehat{\mathbb{G}}$ where $\mathbb{G}$ is the category

$$V \underset{d_1}{\overset{d_0}{\rightrightarrows}} E$$

the situation is different because $\Gamma : \widehat{\mathbb{G}} \to \mathbf{Set}$ does *not* have a right adjoint $\nabla$ since none of the representable objects of $\widehat{\mathbb{G}}$ has a global element. As emphasized by F. W. Lawvere this may be seen as a qualitative distinction between $\widehat{\mathbb{M}_2}$ and $\widehat{\mathbb{G}}$.

# 15   Sheaves

Sheaves (germ. *Garben*) are discussed in detail in [MM]. Here we just give a short introduction to this vast field, mostly without proofs.

**Definition 15.1** *Let $X$ be a topological space and $\mathcal{O}(X)$ the lattice of open subsets of $X$. A* sheaf *over $X$ is a presheaf $A : \mathcal{O}(X)^{\mathsf{op}} \to \mathbf{Set}$ such that for every $U \in \mathcal{O}(X)$ and sieve $\mathcal{U}$ on $U$ with $\bigcup \mathcal{U} = U$ every $f : \mathcal{U} \to A$ has a unique extension $\bar{f} : \mathcal{O}(X)/U \to A$.*
*We write $\mathsf{Sh}(X)$ for the full subcategory of $\mathbf{Set}^{\mathcal{O}(X)^{\mathsf{op}}}$ on sheaves over $X$.* $\diamond$

Notice that a natural transformation $f : \mathcal{U} \to A$ amounts to a choice of an $f(V) \in A(V)$ for all $V \in \mathcal{U}$ which is *compatible* in the sense that

$$A(V \cap W \hookrightarrow V)(f(V)) = f(V \cap W) = A(V \cap W \hookrightarrow W)(f(W))$$

for all $V, W \in \mathcal{U}$. The requirement that $f$ has a unique extension $\bar{f} : \mathcal{O}(X)/U \to A$ along the inclusion of $\mathcal{U}$ into $\mathcal{O}(X)/U$ amounts to the requirement that there exists a *unique* $a = \bar{f}(U) \in A(U)$ such that $f(V) = A(V \hookrightarrow U)(a)$ for all $V \in \mathcal{U}$.
A typical example of a sheaf over $X$ is $R^d$ where $R^d(U)$ is the set of all continuous functions from $U$ to $\mathbb{R}$ and $R^d(V \hookrightarrow U)(f) = f{\restriction}V$, the restriction of $f : U \to \mathbb{R}$ to the open set $V \subseteq U$. This makes sense also when replacing $\mathbb{R}$ by an arbitrary topological space. In case of a discrete space $I$ we write $\Delta(I)$ for the ensuing sheaf. Obviously $\Delta(I)(U)$ consists of all locally constant maps from $U$ to $I$.

**Theorem 15.1** *For a topological space $X$ the category $\mathsf{Sh}(X)$ is closed under limits taken in $\mathbf{Set}^{\mathcal{O}(X)^{\mathsf{op}}}$, is an exponential ideal in $\mathbf{Set}^{\mathcal{O}(X)^{\mathsf{op}}}$, i.e. $B^A$ is a sheaf whenever $A \in \mathbf{Set}^{\mathcal{O}(X)^{\mathsf{op}}}$ and $B \in \mathsf{Sh}(X)$, and $\mathsf{Sh}(X)$ is a topos where $\Omega(U) = \{V \in \mathcal{O}(X) \mid V \subseteq U\}$, $\Omega(V \hookrightarrow U)(W) = V \cap W$ and $\top : 1 \to \Omega$ is given by $\top_U = U$.*

*Proof:* It is straightforward but tedious to verify the first two claims (see [MM] for details).
It is easy to see that $\Omega$ is a sheaf. Suppose that $A \in \mathsf{Sh}(X)$ and $P$ is a subsheaf of $A$. Then $P \hookrightarrow A$ is classified by the map $\chi : A \to \Omega$ sending $a \in A(V)$ to the greatest open subset $V$ of $U$ such that $A(V \hookrightarrow U)(a) \in P(V)$. Notice

that for verifying the existence of a greatest such $V$ one needs that $P$ itself is a sheaf. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

A. Grothendieck – motivated by some questions in algebraic geometry – has generalised the notion of "topology" to arbitrary small categories $\mathbb{C}$. We will present the notion of *Grothendieck topology* a bit later on. First we present the more accessible notion of *coverage* as used in [Joh].

**Definition 15.2** *A* coverage Cov *on a small category* $\mathbb{C}$ *assigns to every object $I$ of $\mathbb{C}$ a set* $\mathsf{Cov}(I)$ *of sieves on $I$ such that for every $S \in \mathsf{Cov}(I)$ and $u : J \to I$ in $\mathbb{C}$ there exists $R \in \mathsf{Cov}(J)$ with $R \subseteq u^*S$.*
*A* sheaf *w.r.t.* Cov *is a presheaf $A$ over $\mathbb{C}$ such that for every $S \in \mathsf{Cov}(I)$ and $f : S \to A$ there exists a unique $\bar{f} : \mathsf{Y}(I) \to A$ making the diagram*

$$
\begin{array}{ccc}
S & \lhook\joinrel\longrightarrow & \mathsf{Y}(I) \\
{\scriptstyle f}\Big\downarrow & \diagdown\diagup\diagup\; \bar{f} & \\
A &
\end{array}
$$

*commute. We write* $\mathsf{Sh_{Cov}}(\mathbb{C})$ *for the full subcategory of* $\mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ *on sheaves w.r.t.* Cov. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Diamond$

In the following we say that a "presheaf $A$ has the sheaf property w.r.t. a sieve $S$ on $I$" iff every $f : S \to A$ has a unique extension to a $\bar{f} : \mathsf{Y}(I) \to A$. Recall also the fact that $\bar{f} : \mathsf{Y}(I) \to A$ extends $f : S \to A$ iff $\bar{f} \circ \mathsf{Y}(u) = f(u)$ for all $u \in S$ (identifying $A(I)$ with $\widehat{\mathbb{C}}(\mathsf{Y}(I), A)$ by Yoneda).
Next we prove to lemmas allowing one to augment a covering without changing the sheaves.

**Lemma 15.1** *Let* Cov *be a coverage on $\mathbb{C}$ and $A$ a* Cov-*sheaf. If $R$ is a sieve on $I$ and $S \in \mathsf{Cov}(I)$ with $S \subseteq R$ then $A$ has the sheaf property w.r.t. $R$.*

*Proof:* Suppose $f : R \to A$. Since $A$ is a sheaf there exists a unique $\bar{f} : \mathsf{Y}(I) \to A$ coinciding with $f$ on $S$. It remains to show that $\bar{f}$ coincides with $f$ also on $R$.
Let $u : J \to I$ in $R$. Then for all $v : K \to J$ in $u^*S$ we have

$$
\bar{f} \circ \mathsf{Y}(u) \circ \mathsf{Y}(v) = \bar{f} \circ \mathsf{Y}(uv) = f(uv) = f(u) \circ \mathsf{Y}(v)
$$

where the second equality holds because $uv \in S$. Since $u^*S$ contains a cover (in the sense of Cov) it follows that $\bar{f} \circ \mathsf{Y}(u) = f(u)$. Thus $\bar{f}$ extends $f$. $\quad\square$

**Lemma 15.2** *Let* Cov *be a coverage on* $\mathbb{C}$ *and* $A$ *be a* Cov*-sheaf. If* $R$ *is a sieve on* $I$ *and* $S \in \mathsf{Cov}(I)$ *such that* $u^*R$ *contains a cover in the sense of* Cov *for all* $u : J \to I$ *in* $S$ *then* $A$ *has the sheaf property w.r.t.* $R$.

*Proof:* Suppose $f : R \to A$. For $u : J \to I$ in $S$ let $\varphi_u : u^*R \to R$ be defined as $\varphi_u(v) = uv$ and $f_u = f \circ \varphi_u : u^*R \to A$. Then there exists a unique morphism $\widetilde{f_u} : \mathsf{Y}(J) \to A$ with $\widetilde{f_u} \circ m = f_u$ where $m : u^*R \hookrightarrow \mathsf{Y}(Y)$. Let $a_u = \widetilde{f_u}(\mathsf{id}_J)$. We next show that $u \mapsto a_u$ gives rise to a natural transformation $\widetilde{f} : S \to A$. Suppose $u : J \to I$ in $S$ and $v : K \to J$. Let $\psi : (uv)^*S \to u^*S$ with $\psi(w) = vw$ and $n : (uv)^*S \hookrightarrow \mathsf{Y}(K)$. Obviously, we have $f_u\psi = f_{uv}$ and $m \circ \psi = \mathsf{Y}(v) \circ n$ as indicated in the following diagram

$$
\begin{array}{ccccc}
\mathsf{Y}(J) & \xleftarrow{\;m\;} & u^*S & \xrightarrow{\;f_u\;} & A \\
\uparrow{\scriptstyle \mathsf{Y}(v)} & & \uparrow{\scriptstyle \psi} & \nearrow{\scriptstyle f_{uv}} & \\
\mathsf{Y}(K) & \xleftarrow{\;n\;} & (uv)^*S & &
\end{array}
$$

Thus we have $\widetilde{f_u} \circ \mathsf{Y}(v) \circ n = \widetilde{f_u} \circ m \circ \psi = f_u \circ \psi = f_{uv} = \widetilde{f_{uv}} \circ n$ from which it follows that $\widetilde{f_u} \circ \mathsf{Y}(v) = \widetilde{f_{uv}}$ (since $A$ has the sheaf property w.r.t. $(uv)^*S$) from which it follows that $\widetilde{f}$ is actually a natural transformation from $S$ to $A$. Since $A$ is a Cov-sheaf and $S \in \mathsf{Cov}(I)$ there exists a unique morphism $\bar{f} : \mathsf{Y}(I) \to A$ extending $\widetilde{f}$ along the inclusion $S \hookrightarrow \mathsf{Y}(I)$. We will show that $\bar{f}$ extends $f$ and is unique with this property.
Let $u : J \to I$ in $R$. Then for $v : K \to J$ in $u^*S$ we have $uv \in R \cap S$ and $\widetilde{f_{uv}} = f(uv) = f(u) \circ \mathsf{Y}(v)$ (notice that $(uv)^*S = \mathsf{Y}(K)$ since $uv \in S$) and thus

$$\bar{f} \circ \mathsf{Y}u \circ \mathsf{Y}v = \bar{f} \circ \mathsf{Y}(uv) = \widetilde{f_{uv}} = f(u) \circ \mathsf{Y}(v)$$

from which it follows that $\bar{f} \circ \mathsf{Y}(u) = f(u)$ since $u^*S$ contains a cover in the sense of Cov. Thus, we have shown that for all $u \in R$ we have $\bar{f} \circ \mathsf{Y}(u) = f(u)$, i.e. that $\bar{f}$ extends $f$ as desired.
Suppose $g : \mathsf{Y}(I) \to A$ with $g \circ \mathsf{Y}(u) = f(u)$ for all $u \in R$. Let $u : J \to I$ in $S$. Then for $v \in u^*R$ we have $uv \in R \cap S$ and $\widetilde{f_{uv}} = f(uv) = f(u) \circ \mathsf{Y}(v)$ and thus

$$g \circ \mathsf{Y}(u) \circ \mathsf{Y}(v) = g \circ \mathsf{Y}(uv) = f(uv) = f(u) \circ \mathsf{Y}(v) = \widetilde{f_{uv}} = \widetilde{f_u} \circ \mathsf{Y}(v)$$

from which it follows that $g \circ \mathsf{Y}(u) = \widetilde{f_u}$ since by assumption $u^*R$ contains a cover in the sense of $\mathsf{Cov}$. Thus, we have shown that for all $u \in S$ we have $g \circ \mathsf{Y}(u) = \widetilde{f_u} = \bar{f}(u)$ from which it follows that $g = \bar{f}$ since $S \in \mathsf{Cov}(I)$ and $A$ is a $\mathsf{Cov}$-sheaf. Thus we have shown uniqueness of $\bar{f}$. $\qquad\square$

We have the following analogue of Theorem 15.1

**Theorem 15.2** *Let* $\mathsf{Cov}$ *be a coverage on a small category* $\mathbb{C}$. *Then* $\mathsf{Sh}_{\mathsf{Cov}}(\mathbb{C})$ *is closed under limits taken in* $\mathbf{Set}^{\mathbb{C}^{op}}$, *is an exponential ideal in* $\mathbf{Set}^{\mathbb{C}^{op}}$, *i.e.* $B^A$ *is a sheaf whenever* $A \in \mathbf{Set}^{\mathbb{C}^{op}}$ *and* $B \in \mathsf{Sh}_{\mathsf{Cov}}(\mathbb{C})$, *and* $\mathsf{Sh}_{\mathsf{Cov}}(\mathbb{C})$ *is a topos.*
*The subobject classifier* $\Omega$ *for* $\mathsf{Sh}_{\mathsf{Cov}}(\mathbb{C})$ *can be described as follows. For* $I \in$ $\mathbb{C}$, $\Omega(I)$ *consists of all* $\mathsf{Cov}$-*closed sieves on* $I$, *i.e. sieves* $S$ *on* $I$ *such that* $u : J \to I$ *is in* $S$ *whenever* $u^*S \in \mathsf{Cov}(J)$, $\Omega(u)(S) = u^*S$ *and* $\top_I = \top_I$.

*Proof:* It is straightforward but tedious to verify the first two claims (see [MM] for details).
It is easy to see that $\Omega$ is a sheaf. Suppose that $A \in \mathsf{Sh}_{\mathsf{Cov}}(\mathbb{C})$ and $P$ is a sheaf of $A$. Then $P \hookrightarrow A$ is classified by the map $\chi : A \to \Omega$ sending $a \in A(V)$ to the sieve $\chi_I(a) = \{u : J \to I \mid A(u)(a) \in P(J)\}$. Notice that for verifying that $\chi_I(a)$ is closed one needs that $P$ itself is a sheaf. $\qquad\square$

A coverage $\mathsf{Cov}$ on $\mathbb{C}$ can be saturated to a so-called Grothendieck topology on $\mathbb{C}$ which notion we define next.

**Definition 15.3** *A* Grothendieck topology *on a small category* $\mathbb{C}$ *is a coverage* $\mathcal{J}$ *with* $\top_I \in \mathcal{J}(I)$ *for all* $I \in \mathbb{C}$ *and satisfying the following* locality *property*

(L) *if* $R$ *is a sieve on* $I$ *and* $S \in \mathcal{J}(I)$ *with* $u^*R \in \mathcal{J}(J)$ *for all* $u : J \to I$ *in* $S$ *then* $R \in \mathcal{J}(I)$. $\qquad\diamond$

Notice that the locality property entails that $\mathcal{J}(I)$ is upward closed within sieves on $I$.
Obviously, Grothendieck topologies on $\mathbb{C}$ are closed under (componentwise) intersections. Thus, for every coverage $\mathsf{Cov}$ there exists a least Grothendieck topology $\mathcal{J}$ on $\mathbb{C}$ with $\mathsf{Cov} \subseteq \mathcal{J}$. Notice that this way $\mathcal{J}$ is obtained from $\mathsf{Cov}$ by an inductive definition which in general stabilizes at a very transfinite ordinal.
The next lemma says that sheaves w.r.t. $\mathsf{Cov}$ and sheaves w.r.t. the induced $\mathcal{J}$ coincide.

**Lemma 15.3** *Let $\mathcal{J}$ be the least Grothendieck topology containing* Cov. *Then a presheaf $A \in \mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ is a* Cov-*sheaf iff it is a $\mathcal{J}$-sheaf.*

*Proof:* Obviously, since $\mathsf{Cov} \subseteq \mathcal{J}$ every $\mathcal{J}$-sheaf is also a Cov-sheaf.
For the reverse inclusion it suffices to show that the collection $\mathcal{J}_A$ of all sieves $S$ on some $I$ such that for all $u : J \to I$, $A$ satisfies the sheaf condition w.r.t. $u^*S$ forms a Grothendieck topology. Obviously $\mathcal{J}_A$ is a coverage containing all maximal sieves $\top_I$. So it remains to show that $\mathcal{J}_A$ satisfies the locality property (L). Suppose $R$ is a sieve on $I$ and $S \in \mathcal{J}_A(I)$ with $u^*R \in \mathcal{J}_A(J)$ for all $u : J \to I$ in $S$. We have to show that $R \in \mathcal{J}_A(I)$, i.e. that $A$ has the sheaf property w.r.t. to all reindexings of $R$. Let $u : J \to I$. Then $u^*S \in \mathcal{J}_A(J)$ and for all $v : K \to J$ in $u^*S$ we have $v^*u^*R = (uv)^*R \in \mathcal{J}_A(K)$ since $uv \in S$. Thus by Lemma 15.2 the presheaf $A$ satisfies the sheaf condition w.r.t. $u^*R$ as desired. $\qquad\square$

One can show that a subobject $\mathcal{J}$ of $\Omega$ in $\widehat{\mathbb{C}}$ is a Grothendieck topology if and only if the classifying map $j : \Omega \to \Omega$ for $\mathcal{J} \hookrightarrow \Omega$ satisfies the conditions

(j1) $j \circ \top = \top$

(j2) $j \circ \wedge = \wedge \circ (j \times j)$

(j3) $j \circ j = j$.

In an arbitrary topos $\mathbb{E}$ it makes sense to consider maps $j : \Omega \to \Omega$ satisfying the conditions (j1)-(j3). Such maps are called *Lawvere-Tierney toplogies* on topos $\mathbb{E}$. A mono $m : P \rightarrowtail X$ in $\mathbb{E}$ is called *j-dense* iff its characteristic map $\chi : A \to \Omega$ satisfies $j \circ \chi = \top_A$. One then defines an object $A$ of $\mathbb{E}$ to be a *j-sheaf* iff $\mathbb{E}(m, A) : \mathbb{E}(X, A) \to \mathbb{E}(P, A)$ is a bijection for all *j*-dense monos $m$ in $\mathbb{E}$. One can show that the full subcategory $\mathbb{E}_j$ of $\mathbb{E}$ on *j*-sheaves is a topos and that the inclusion $i : \mathbb{E}_j \hookrightarrow \mathbb{E}$ has a finite limit preserving left adjoint $a$ called *sheafification*. Moreover, if $j$ is the characteristic map for a Grothendieck topology $\mathcal{J}$ on $\mathbb{C}$ then $\widehat{\mathbb{C}}_j$ coincides with $\mathsf{Sh}_{\mathcal{J}}(\mathbb{C})$. Accordingly, the inclusion of $\mathsf{Sh}_{\mathcal{J}}(\mathbb{C})$ into $\widehat{\mathbb{C}}$ has a finite limit preserving left adjoint (also called sheafification.
By definition a *Grothendieck topos* is a topos equivalent to one of the form $\mathsf{Sh}_{\mathcal{J}}(\mathbb{C})$ where $\mathcal{J}$ is a Grothendieck topology on a small category $\mathbb{C}$.
Grothendieck toposes can be characterized more abstractly as follows.

**Theorem 15.3** $\mathbb{E}$ *is a Grothendieck topos iff $\mathbb{E}$ is a locally small elementary topos with small sums and a small generating family.*

*Proof:* See e.g. [MM]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

For Grothendieck toposes $\mathsf{Sh}_{\mathcal{J}}(\mathbb{C})$ one can give a Kripke-Joyal semantics which differs from the one for $\mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ only for $\bot$, $\vee$ and $\exists$, namely as follows

($\bot$) $I \Vdash \bot$ iff $\emptyset \in \mathcal{J}(I)$

($\vee$) $I \Vdash (\phi \vee \psi)(a)$ iff there exists a $\mathcal{J}$-covering family $(u_j : I_j \to I)_{j \in J}$ such that for all $j \in J$, $I_j \Vdash \phi(au_j)$ or $I_j \Vdash \psi(au_j)$

($\exists$) $I \Vdash \exists x{:}A.\rho(x,c)$ iff there exists a $\mathcal{J}$-covering family $(u_j : I_j \to I)_{j \in J}$ and a family $\big(a_j \in A(I_j)\big)_{j \in J}$ such that $I_j \Vdash \rho(a_j, cu_j)$ for all $j \in J$

where $(u_j : I_j \to I)_{j \in J}$ is $\mathcal{J}$-covering iff the induced sieve on $I$ is in $\mathcal{J}(I)$. Thus, for disjunctions and existential statements it is easier to be valid in $\mathsf{Sh}_{\mathcal{J}}(\mathbb{C})$ than in $\mathbf{Set}^{\mathbb{C}^{\mathrm{op}}}$ because it suffices for them to hold only "locally" and they need not necessarily hold "globally".