Introduction to Categorical Logic

Finn Klein

October 10, 2024

Please note: This is a **draft version!** It will be updated throughout the week.

1 What is a logics and how does it work

Suppose (or recall) that you are a curious human being who likes to learn about the world around you. Through observation, you realise that the following seems to be true.

When it rains, the street gets wet.

To verify your intuition, you could keep a record of all the days on which it rains and all days on which the street is wet. With the exception of some few outliers, you might then be able to verify

$$\{\text{Rainy days}\} \subseteq \{\text{Days with wet street}\}.$$

Unfurling the definition of *rain*s, using analytical thinking and knowledge about the world, you might even figure out the mechanism behind this correlation: If it rains, water falls from the sky onto the street, making it wet.

Let's now consider an analogous, but more stripped-down situation: Through observation, you realise that the following seems to be true

Natural numbers ≥ 6 **are also** ≥ 5**.**

In set notation, this amounts to the following inclusion:

$$\{n \in \mathbb{N} \mid n \geq 6\} \subseteq \{n \in \mathbb{N} \mid n \geq 5\}.$$

The joys of mathematics allow us to explain this observation using the following simple* steps.

1. Define the (purely formal) **theory** of a partial order. This is a formal language in which we can construct formulae. Using a fixed deduction system, derive the fact that for symbols a, b, c satisfying $a \geq b$ and $b \geq c$, one has $a \geq c$ from the axioms of the theory. (Call a logical statement like this a **sequent** and denote this specific sequent by $a \geq b \wedge b \geq c \vdash_{\{a,b,c\}} a \geq c$.)

2. Define a **model** of the theory in the following way: Choose a universe U of objects which we can quantify over. To each formula ϕ in the theory of partial orders, assign a collection $\{x : U \mid \phi(x)\}$ of those objects of U which satisfy ϕ. (That is, from ϕ assign a truth value to every object of U and keep only those with truth value "true".) In our case, we choose the universe $\mathbb{N}$ and interpret $\leq$ as the classical total order on $\mathbb{N}$.

3. Conclude the "proof" by realising that the purely formal consideration from step 1 now implies the required containment of subsets of $\mathbb{N}$: That is, the sequent $x \geq 6 \vdash_x x \geq 5$ "realises" to the containment

$$\{n \in \mathbb{N} \mid n \geq 6\} \subseteq \{n \in \mathbb{N} \mid n \geq 5\}.$$

Step one happens purely on the level of **Syntax**: It is pure formality without content. The construction of a model is on the level of **Semantics**: It's pure observation of content without structure. The real magic of logic happens in bringing those two worlds together. Two desirable properties of this step are outlined in the following non-definition:

Definition 1.0.1. An assignment of model to theory is **sound** if every formally provable sequent $\phi_1 \vdash \phi_2$ realises to a containment

$$\{x \in U \mid \phi_1(x)\} \subseteq \{x \in U \mid \phi_2(x)\}.$$

It is **complete** if the reverse implication holds.

Designing deduction systems which satisfy both soundness and completeness is one of the big goals in logic: Not only can we prove every observably true statement, we can also not prove anything that is observably false. One might thus ask whether such deduction systems always exist for, say, first-order logic. Unfortunately, the answer is no: Gödel's incompleteness theorem shows that logical theories which are able to express the theory of the natural numbers cannot both be sound and complete. Despite this, many logical theories and fragments of first-order logic do satisfy soundness and completeness properties.

Categorical logic is the study of logic, specifically of the interplay of syntax and semantics, using the tools of category theory.

2 Categorical basics

2.1 Subobjects

Definition 2.1.1. Let $\mathbf{C}$ be a locally small category, and let $X, Y \in \mathbf{C}$ be objects. A morphism $f : X \to Y$ is a **monomorphism** if for all objects $W \in \mathbf{C}$ and maps $g, h : W \to X$, the equality $f \circ g = f \circ h$ implies that $g = h$.

Remark. Monomorphisms in the category of sets are exactly injective maps. To see this, let $x, y \in X$ and define two maps $g, h : \{*\} \to X$, where $g(*) = x$ and $h(*) = y$. The monomorphism property amounts to saying that $f(x) = f(y)$ implies $x = y$, which is exactly injectivity.

Definition 2.1.2. For any object $X \in \mathbf{C}$, let $\mathrm{Mon}(X)$ denote the class of monomorphisms with codomain X. Define a preorder $\leq$ on $\mathrm{Mon}(X)$ as follows: If $f : Y_1 \to X$ and $g : Y_2 \to X$ are monomorphisms, one has $f \leq g$ iff f factors as $g \circ h$ for a (mono)morphism[1] $Y_1 \to Y_2$. The preorder $\leq$ induces an equivalence relation $\sim$ by defining $f \sim g$ iff $f \leq g$ and $g \leq f$. The set of subobjects $\mathrm{Sub}(X)$ of X is the set of $\sim$-equivalence classes in $\mathrm{Mon}(X)$. (One can show that this is infact a set. TODO: Subject to smallness conditions??)

Example 2.1.3. Let X be a finite set. $\mathrm{Sub}(X)$ is the set obtained from the powerset $\mathbb{2}^X$ by quotienting out bijections. In other words, if $|X| = n$, then $\mathrm{Sub}(X)$ is isomorphic to the set $\{0, 1, \ldots, n\}$.

TODO: In finite limit category, the poset of subobjects forms a meet-semilattice.

3 Logic

3.1 Signatures and theories

We will define certain language formation rules which will allow us to write down sentences which resemble formalised mathematics. None of the formulae in this chapter will have any mathematical content, they're just symbols pasted together in a specific way.

Definition 3.1.1. A (first-order) **signature** is a tuple $\Sigma = (S, F, R)$ where

1. S is a set of *sorts*,
2. F is a set of *function symbols* of the form $f : A_1, \ldots, A_n \to B$, where the A_i and B are sorts,
3. R is a set of *relation symbols* $R \hookrightarrow A_1, \ldots A_n$, where the A_i are sorts.

Remark. For any sort A, we consider a class $V(A)$ of variables of type A. For any $x \in V(A)$, we say that x is a variable of sort (or type) A.

A first-order signature provides the building blocks from which we can build terms.

Definition 3.1.2. Let Σ be a first-order signature. A **term** of Σ is one of the following:

1. If A is a sort of Σ, then any variable $a \in V(A)$ is a term (of sort A).
2. If $A_1, \ldots, A_n, B$ are sorts and $a_1 : A_1, \ldots, a_n : A_n$ are terms of their respective sorts, then $f(a_1, \ldots, a_n)$ is a term (of sort B).

Starting from terms, we can build **formulae** over Σ inductively. Every formula ϕ has an associated set of free variables $FV(\phi)$, which we will also define inductively.

[1] Any morphism that satisfies $f = g \circ h$ is automatically a monomorphism. (Exercise)

1. If A is a sort and $a_1, a_2 : A$ are terms of sort A, then $a_1 = a_2$ is a formula. The set of free variables $FV(a_1 = a_2)$ is the set of variables occurring in either a_1 or a_2.

2. If $R \hookrightarrow A_1, \ldots, A_n$ is a relation symbol and $a_1 : A_1, \ldots, a_n : A_n$, then $R(a_1, \ldots, a_n)$ is a formula. $FV(R(a_1, \ldots, a_n))$ is the set of variables occurring in one of the a_i.

3. True ($\top$) and False ($\bot$) are formulae. $FV(\top) = FV(\bot) = \emptyset$.

4. If ϕ is a formula, then its negation $\neg\phi$ is a formula. One has $FV(\neg\phi) = FV(\phi)$.

5. If ϕ_1, ϕ_2 are formulae, then we can form their binary disjunction $\phi_1 \vee \phi_2$ and binary conjunction $\phi_1 \wedge \phi_2$. One has $FV(\phi_1 \vee \phi_2) = FV(\phi_1 \wedge \phi_2) = FV(\phi_1) \cup FV(\phi_2)$.

6. If ϕ_1, ϕ_2 are formulae, then we can form the implication $\phi_1 \Rightarrow \phi_2$. One has $FV(\phi_1 \Rightarrow \phi_2) = FV(\phi_1) \cup FV(\phi_2)$.

7. If $(\phi_i)_{i \in I}$ is a family of formulae (possibly infinite) such that the union $U := \bigcup_{i \in I} FV(\phi_i)$ is a finite set, we may form their infinitary disjunction $\bigvee_{i \in I} \phi_i$ and infinitary conjunction $\bigwedge_{i \in I} \phi_i$. Both these formulae have free variables equal to U.

8. If ϕ is a formula and $x \in FV(\phi)$, then $(\exists x)\phi$ and $(\forall x)\phi$ are formulae. One has $FV((\exists x)\phi) = FV((\forall x)\phi) = FV(\phi) \setminus \{x\}$.

Remark. Intuitively speaking, formulae over a signature Σ are things about whose truthfulness we can make inquiries.

Example 3.1.3. Let $\Sigma = (\{\mathbb{N}\}, \emptyset, \{<\})$. Consider the following formulae and their associated free variables.

1. $\phi_1 = x < y$, where x, y are variables of type $\mathbb{N}$. One has $FV(\phi_1) = \{x, y\}$.

2. $\phi_2 = (\exists y) x < y$. One has $FV(\phi_2) = \{x\}$.

3. $\phi_3 = (\forall x)(\exists y) x < y$. One has $FV(\phi_3) = \emptyset$.

Definition 3.1.4. (Context and Sequents).

1. Let ϕ be a formula. A **context** for ϕ is a finite set $\vec{x}$ of variables such that $FV(\phi) \subseteq \vec{x}$.

2. A **sequent** is a statement of the form $\phi_1 \vdash_{\vec{x}} \phi_2$, where ϕ_1, ϕ_2 are formulae and let $\vec{x}$ be a common context for both.

We are now ready to define what a logical theory is.

Definition 3.1.5. A **logical theory** $\mathbb{T}$ over a first-order signature Σ is a set of sequents over Σ, the so-called axioms. We call $\mathbb{T}$ an atomic (Horn, regular, coherent, geometric, (infinitary) first-order) theory if it only uses sequents built from a limited subset of logical constructions, as outlined in Figure 1.

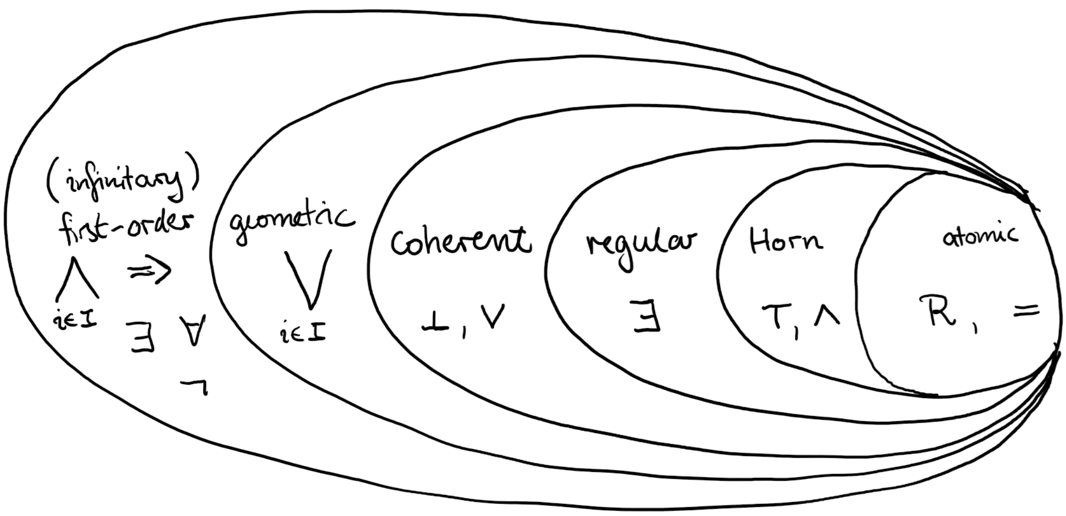


Figure 1: The onion of first-order theories.

Example 3.1.6. Every category with finite products $\mathbf{C}$ comes with a built-in first-order signature $\Sigma_{\mathbf{C}}$, the so-called **internal language** of $\mathbf{C}$, which is constructed as follows:

1. The sorts of $\Sigma_{\mathbf{C}}$ are the objects of $\mathbf{C}$. If X is an object of $\mathbf{C}$, we denote the corresponding sort by $[X]$ the corresponding sort.

2. For every morphism $f : X_1 \times \ldots \times X_n \to Y$ in $\mathbf{C}$ we have a function symbol

$$[f] : [X_1] \ldots [X_n] \to [Y].$$

3. For every subobject $R \in \mathrm{Sub}(X_1 \times \ldots \times X_n)$ we have a relation symbol

$$[R] \hookrightarrow [X_1] \ldots [X_n].$$

3.2 Deduction systems

Given now a completely formal language in which we can formulate our sequents, we would like to have a framework to decide whether one sequent can be derived from the other. This is achieved using another (completely formal) system called **sequent calculus.** There are many flavours of this calculus available and we will present a tiered collection of them here, as we have with the first-order theories before.

Definition 3.2.1. An **inference rule** is a pair (H, C) of finite sets of sequents. The set H is called the set of hypotheses of the deduction, and C is called the set of conclusions. If $H = \{h_1, \ldots, h_n\}$ and $C = \{c_1, \ldots, c_m\}$, we notate such an inference rule as follows:

$$\frac{h_1 \ldots h_n}{c_1 \ldots c_m}$$

Remark. Here's another gentle reminder that an inference rule is just a formal collection of symbols without mathematical meaning. We will later assign meaning to them when we start talking about models in categories. Morally, however, an inference rule (H, C) should be read as follows: If all hypotheses in H are provably true, then so is any conclusion in C.

We will be using subsets of the following set inference rules, depending on the fragment of logic we're in (here, Φ denotes an arbitrary set of hypotheses and *Gamma* denotes an arbitrary finite context):

1. Deduction rules regarding equality:

$$\frac{}{t = t} \qquad \frac{t_1 = t_2}{t_2 = t_1} \qquad \frac{t_1 = t_2 \quad t_2 = t_3}{t_1 = t_3}$$

2. Rules regarding truth, falsity and redundancy:

$$\frac{}{\phi_1 \dots \phi_m \vdash \phi_i} \qquad \frac{}{\Phi \vdash \top} \qquad \frac{}{\Phi \perp \vdash \phi}$$

Here, $i \in \{1, \dots, m\}$, and ϕ is arbitrary. The last rule is sometimes called *Ex falso quotlibet*, i.e. "From falsehood [follows] whatever you want."

3. Algebraic axioms of $\wedge$, $\vee$ and $\exists$:

$$\frac{}{\phi \wedge (\psi \vee \chi) \vdash \phi \wedge \psi \vee \phi \wedge \chi} \qquad \frac{}{\phi \wedge (\exists x : \psi) \vdash \exists x : (\phi \wedge \psi)}$$

These are called distributivity axiom and Frobenius axiom, respectively.

4. Introduction and elimination of finitary disjunction:

$$\frac{\Phi, \phi_1 \vdash \psi \qquad \Phi, \phi_2 \vdash \psi}{\Phi, \phi_1 \vee \phi_2 \vdash \psi}$$

5. Introduction and elimination of finitary conjunction:

$$\frac{\Phi \vdash \phi_1 \qquad \Phi \vdash \phi_2}{\Phi \vdash \phi_1 \wedge \phi_2}$$

6. Introduction and elimination of implication:

$$\frac{\Phi, \phi \vdash \psi}{\Phi \vdash \phi \Rightarrow \psi}$$

7. Substitution:

$$\frac{\Phi, x = y \vdash_{\Gamma, x, y} \phi}{\Phi \vdash_{x,y} \phi[x/y]}$$

where y is not a free variable of any formula in Φ and $\phi[x/y]$ denotes the formula obtained from ϕ by substituting any occurrence of x with y.

8. Introduction and elimination of existential quantification:

$$\frac{\Phi, \phi \vdash_{\Gamma,x} \psi}{\Phi, (\exists x : \phi) \vdash_{\Gamma} \psi}$$

where x is not a free variable of any formula in Φ, nor a free variable of ψ.

9. Introduction and elimination of universal quantification:

$$\frac{\Phi \vdash_{\Gamma,x} \psi}{\Phi \vdash_{\Gamma} (\forall x : \psi)}$$

where x is not a free variable of any formula in Φ, nor a free variable of ψ.

Remark. Introduction and elimination of disjunction and conjunction also have infinitary analogues.

4 Interpreting logical theories in categories

Given a logical theory, we want to be able to interpret it in a category which has enough structure (limits, colimits, exponentials and certain extra structure on the poset of subobjects) to express the logical connectives. "Weaker" logical theories like atomic or Horn theories can be interpreted in categories with a smaller amount of extra structure, while interpreting a full first-order theory requires a lot of extra structure.

4.1 Σ-structures

Let Σ be a first-order signature and let $\mathbf{C}$ be a category with finite products. We want to be able to *interpret* first-order theories over Σ inside of $\mathbf{C}$. [TODO: What does that actually mean?] First, we need to find an incarnation of the sorts, function symbols and relation symbols of Σ inside of $\mathbf{C}$.

Definition 4.1.1. A Σ-structure M in $\mathbf{C}$ is the following data:

1. For every sort A of Σ, an object $MA \in \mathbf{C}$.
2. For every function symbol $f : A_1 \dots A_n \to B$ of Σ, a morphism
$$Mf : MA_1 \times \dots \times MA_n \to MB.$$
3. For every relation symbol $R \hookrightarrow A_1 \dots A_n$, a subobject
$$MR \in \mathrm{Sub}(MA_1 \times \dots \times MA_n).$$

If M and N are Σ-structures in $\mathbf{C}$, a morphism of Σ-structures $\phi : M \to N$ consists of a collection of morphisms $\phi_A : MA \to NA$, where A ranges over all sorts of Σ, such that the following squares commute:

$$\begin{array}{ccc} MA_1 \times \dots \times MA_n & \xrightarrow{Mf} & MB \\ \downarrow{\scriptstyle \phi_{A_1} \times \dots \times \phi_{A_n}} & & \downarrow{\scriptstyle \phi_B} \\ NA_1 \times \dots \times NA_n & \xrightarrow{Nf} & NB \end{array}$$

$$\begin{array}{ccc} MR & \hookrightarrow & MA_1 \times \ldots \times MA_n \\ {\scriptstyle \phi_R}\downarrow & & \downarrow{\scriptstyle \phi_{A_1}\times\ldots\times\phi_{A_n}} \\ NR & \hookrightarrow & NA_1 \times \ldots \times NA_n \end{array}$$

where the top and bottom arrows of the bottom diagram are some representatives of the corresponding subobjects.

We denote by $\Sigma - \mathrm{str}(\mathbf{C})$ the category of Σ-structure in $\mathbf{C}$ and their homomorphisms.

4.2 Motivating example: Interpreting formulae in the category of sets

As a motivating example, consider the polynomial inequality $x \leq 1 - x^2$. To motivate this section, we want to express the solutions to this inequality in the most categorical way possible. In a way, we have one "sort", the real numbers $\mathbb{R}$, one function symbol $g : \mathbb{R} \to \mathbb{R}, x \mapsto 1 - x^2$ and one relation symbol, $\leq$. To be precise, $(\mathbb{R}, g, \leq)$ is one possible Σ-structure for the signature

$$\Sigma = (\{A\}, \{f : A \to A\}, \{R \hookrightarrow A^2\}).$$

Let X now be the set of solutions of the inequality. We can write

$$X = \{x \in \mathbb{R} : x \subseteq f(x)\} = \{x \in \mathbb{R} : (x, f(x)) \in \leq\},$$

and in that representation it is clear that X is given by the following pullback:

$$\begin{array}{ccc} X & \longrightarrow & \leq \\ \downarrow & \lrcorner & \downarrow \\ \mathbb{R} & \xrightarrow{\langle id, f\rangle} & \mathbb{R} \times \mathbb{R} \end{array}$$

Suppose additionally that we wanted to restrict to only those solutions x which themselves are greater than or equal to 0. (Formally, we have added a new constant symbol or 0-aric function to our signature Σ.) Let Z be the set of those solutions. It is clear that $Z = X \cap \{x \in \mathbb{R} : 0 \leq x\}$. The latter is easily seen to be the pullback of the diagram

$$\begin{array}{ccc} & & \leq \\ & & \downarrow \\ \mathbb{R} & \xrightarrow{\langle id, !_0\rangle} & \mathbb{R} \times \mathbb{R} \end{array}$$

where $!_0$ is the constant zero map, which factors through the terminal object as $\mathbb{R} \to \{*\} \to \mathbb{R}$. Let's call this pullback Y. Finally, the intersection Z can be realised as yet another pullback: $Z = X \times_{\mathbb{R}} Y$ with the obvious maps.

Remark. The point of this is not to make it easier to determine the set of solutions (because obviously it doesn't make it easier), but to demonstrate how one can interpret elements satisfying certain logical formulae using purely categorical constructions.

4.3 Horn logic and finite limit categories

Recall that Horn logic has formulae built out of equalities, relations, truth and finite conjunctions. It turns out that Horn logic can be interpreted in any category $\mathbf{C}$ which has finite limits (equivalently, finite products and equalisers.)

Definition 4.3.1 (Interpreting terms)**.** Let $\mathbf{C}$ have finite limits and let M be a Σ-structure in $\mathbf{C}$. Let $\vec{x}.t$ be a term with context. Assume $\vec{x} = \{x_1, \ldots, x_n\}$, where $x_i : A_i$, and that t is of type B. An interpretation of t is a morphism

$$[[\vec{x}.t]] : MA_1 \times \ldots \times MA_n \to MB.$$

Specifically, one constructs

1. If t is a variable x_i, then $[[\vec{x}.t]]$ is the projection map

$$MA_1 \times \ldots \times MA_n \to MA_i.$$

2. if $t = f(t_1, \ldots, t_m)$, where the t_i are of types B_i and f is a function symbol from $B_1 \ldots B_m$ to C, then $[[\vec{x}.t]]$ is the composite

$$MA_1 \times \ldots \times MA_n \xrightarrow{\langle [[\vec{x}.t_1]], \ldots, [[\vec{x}.t_m]] \rangle} MB_1 \times \ldots \times MB_m \xrightarrow{Mf} MC.$$

Definition 4.3.2 (Interpreting Horn formulae)**.** In the same setup as the previous definition, let $\vec{x}.\phi$ be a formula with context. Assume again that $\vec{x} = \{x_1, \ldots, x_n\}$, where $x_i : A_i$. An interpretation of $\vec{x}.\phi$ in $\mathbf{C}$ is giving a subobject

$$[[\vec{x}.\phi]] \hookrightarrow MA_1 \times \ldots MA_n,$$

according to the following case distinction:

1. If $\phi(\vec{x}) = R(t_1, \ldots, t_m)$, where R is a relation symbol of type $B_1 \ldots B_m$, then $[[\vec{x}.\phi]]$ is the pullback

$$\begin{array}{ccc} [[\vec{x}.\phi]] & \longrightarrow & MR \\ \downarrow & & \downarrow \\ MA_1 \times \ldots \times MA_n & \xrightarrow{\langle [[\vec{x}.t_1]], \ldots, [[\vec{x}.t_m]] \rangle} & MB_1 \times \ldots \times MB_m \end{array}$$

2. If $\phi(\vec{x}) = t_1 = t_2$, where $t_1, t_2 : B$, then $[[\vec{x}.\phi]]$ is the equaliser

$$[[\vec{x}.\phi]] \longrightarrow MA_1 \times \ldots \times MA_n \underset{[[\vec{x}.t_2]]}{\overset{[[\vec{x}.t_1]]}{\rightrightarrows}} MB.$$

3. If $\phi(\vec{x}) = \phi_1 \wedge \phi_2$, where ϕ_1, ϕ_2 are formulae in the same context, then $[[\vec{x}.\phi]]$ is the intersection (pullback) of the subobjects $[[\vec{x}.\phi_1]]$ and $[[\vec{x}.\phi_2]]$.

4. If $\phi(\vec{x}) = \top$, then $[[\vec{x}.\phi]]$ is the top subobject

$$\mathrm{id} : MA_1 \times \ldots \times MA_n \to MA_1 \times \ldots \times MA_n.$$

4.4 Regular logic and regular categories

We begin by an example in the category of Sets motivating the interpretation of existential quantification using adjoints.

Example 4.4.1 (Existential quantifiers are adjoints)**.** Let $\{x, y\}.\phi(x, y)$ be a formula in a two-sorted signature Σ. After choosing a Σ-structure in **Set**, we can identify the interpretation of ϕ with a subset $S = \{(x, y) \in X \times Y \mid \phi(x, y)\}$. The projection map $\pi : X \times Y \to Y$ induces a morphism of posets $\pi^* : 2^X \to 2^{X\times Y}$. By abstract nonsense, this morphism has a left adjoint g. We claim that, for all $U \subseteq X \times Y$, we have that $g(U)$ is given by

$$\exists U := \{y \in Y \mid \exists x \in X : (x, y) \in U\}.$$

To see this, we just need to verify that this set satisfies the adjunction property: i.e., for all $T \subseteq Y$, one has

$$\exists U \subseteq T \iff U \subseteq \pi^*(T) = X \times T.$$

Let's start with the implication $\Rightarrow$. Suppose that $\exists_x U \subseteq T$. Let $u \in U$. Obviously, $U \subseteq X \times \exists_x U$. But $X \times \exists_x U \subseteq X \times T$ by assumption. To see the inverse implication, suppose that $U \subseteq X \times T$. Let $y \in Y$ such that there exists an $x \in X$ satisfying $(x, y) \in U$. Then, one obviously has $(x, y) \in U \subseteq X \times T$, where the last inclusion is by assumption. It follows that $y \in T$ and we are done.

The advantage of this reformulation of existential quantification is that it works in any *regular* category.

Definition 4.4.2 (Image)**.** Let **C** be a category and let $f : X \to Y$ be a morphism. The **image** $\mathrm{im}(f)$ of f, if it exists, is the smallest subobject of Y through which f factors. **C** is called **regular** if it has finite limits and $\mathrm{im}(f)$ exists for every morphism f in **C**.

Example 4.4.3. The category of sets is regular.

Proposition 4.4.4. Let **C** be a regular category and let $f : X \to Y$ be a morphism in **C**. The pullback map $f^* : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ has a left adjoint

$$\exists_f : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$$
$$m \mapsto \mathrm{im}(f \circ m).$$

Example 4.4.5. Let's try to reconcile this proposition with Example 4.4.1. Recall that our map $\exists_x$ maps a subobject $U \hookrightarrow X \times Y$ to the subset

$$\exists_x U := \{y \in Y \mid \exists x \in X : (x, y) \in U\}$$

of Y. This set is easily seen to be the image of the composite $U \hookrightarrow X \times Y \to Y$.

With this in mind, the way in which formulae involving existential quantification are interpreted inside of regular categories should not come as a surprise.

Definition 4.4.6 (Interpreting existential quantification)**.** Let **C** be a regular category. In the same setup as in Definition 4.3.2, let $\vec{x}.\phi$ be a formula with context. Assume again that $\vec{x} = \{x_1, \ldots, x_n\}$, where $x_i : A_i$. If $\phi(\vec{x}) = (\exists y)\psi(\vec{x}, y)$, where $y : B$, define the interpretation $[[\vec{x}.\phi]]$ to be the image of the composite map

$$[[\vec{x}, y.\, \psi]] \to MA_1 \times \ldots \times MA_n \times MB \to MA_1 \times \ldots \times MA_n.$$

4.5 Coherent logic

Lemma 4.5.1. Let $\mathbf{C}$ be a regular category with finite coproducts. For every X, the poset of subobjects of X is a lattice.

Proof. We've already shown that $\mathrm{Sub}(X)$ forms a meet-semilattice for every finite limit category (TODO). It remains to show that $\mathrm{Sub}(X)$ admits finite joins and a minimal element.

1. The minimal element of $\mathrm{Sub}(X)$ is given by the unique map from the initial object (empty coproduct) $\perp$ into X. It is easy to check that this is a monomorphism, hence defines a subobject. It is minimal by initiality.

2. Let S_1, S_2 be subobjects of X. Consider the map $S_1 \sqcup S_2 \to X$ (not a monomorphism) induced by the universal property of the coproduct. Since $\mathbf{C}$ is regular, this map uniquely factors through its image, which is a monomorphism (and hence defines a subobject.) It is left as an exercise to check that this construction satisfies the axioms for a join operation.

□

Remark. The condition that all subobject posets form lattices is strictly weaker than the requirement that finite colimits exist. TODO: Example. We go with the latter one for the sake of convenience.

Using the lattice structure of the subobject posets, we can interpret finite conjunctions and falsehood:

Definition 4.5.2. Let $\mathbf{C}$ be a regular category admitting finite coproducts. In the same setup as in Definition 4.3.2, let $\vec{x}.\phi$ be a formula with context. Assume again that $\vec{x} = \{x_1, \ldots, x_n\}$, where $x_i : A_i$.

1. If $\phi = \phi_1 \vee \phi_2$, where ϕ_1 and ϕ_2 are formulae with the same context, then define $[[\vec{x}.\phi]]$ as the join of the subobjects $[[\vec{x}.\phi_1]]$ and $[[\vec{x}.\phi_2]]$ of $MA_1 \times \ldots \times MA_n$.

2. If $\phi = \perp$, then define $[[\vec{x}.\phi]]$ as the bottom element $\perp$ of $MA_1 \times \ldots \times MA_n$.

4.6 Geometric logic and geometric categories

Definition 4.6.1. Let $\mathbf{C}$ be a regular category admitting all images.

1. $\mathbf{C}$ is **well-powered** if for all $X \in \mathbf{C}$, the category of subobjects $\mathrm{Sub}(X)$ is equivalent to a small category.

2. $\mathbf{C}$ is **geometric** if it is well-powered and if for all $X \in \mathbf{C}$ and collections of subobjects $(X_i \hookrightarrow X)_{i \in I}$, the infinitary join $\bigvee_{i \in I} X_i$ (smallest common supremum to all the X_i) exists in $\mathrm{Sub}(X)$ and is stable under pullback: For all $f : Y \to X$, one has
$$\bigvee_{i \in I} f^* X_i = f^* \bigvee_{i \in I} X_i$$
in $\mathrm{Sub}(Y)$.

The structure of geometric categories allows us to interpret infinitary conjunctions (and hence all of geometric logic):

Definition 4.6.2. Let $\mathbf{C}$ be a regular category admitting finite coproducts. In the same setup as in Definition 4.5.2, let $\vec{x}.\phi$ be a formula with context. Assume again that $\vec{x} = \{x_1, \ldots, x_n\}$, where $x_i : A_i$, and that $\phi = \bigvee_{\imath \in I} \phi_i$, where ϕ_i are formulae in the same context. Define $[[\vec{x}.\phi]]$ as the (infinitary) join of the subobjects $[[\vec{x}.\phi_i]]$ in $\mathrm{Sub}(MA_1 \times \ldots \times MA_n)$, where i ranges over I.

4.7 $\forall, \Rightarrow, \neg$ and Heyting categories

TODO add blurb

Definition 4.7.1. A coherent category $\mathbf{C}$ is called a Heyting category if for all $f : X \to Y$ in $\mathbf{C}$, the pullback functor $f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$ has a right adjoint, which we will denote by $\forall_f : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$.

It turns out that this additional structure allows us to define the missing logical operations: Implication, negation and universal quantification.

Example 4.7.2 (Heyting implication)**.** Let $Y \hookrightarrow X$ be a subobject and consider the inclusion map $f : Y \to X$. Pulling back a subobject Y' of X by f amounts to computing the meet $Y \wedge Y'$. If $\mathbf{C}$ is a Heyting category, we obtain a right adjoint $\forall_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$. Let Y' be any subobject of X and define the **Heyting implication**

$$Y \to Y' := \forall_f(Y \wedge Y') \in \mathrm{Sub}(X).$$

Exercise 4.7.3. Verify that this defines the usual implication in the category of sets. More specifically, let X be a set and consider two formulae $x.\phi$ and $x.\psi$, where $x : X$, and their interpretations $[[x.\phi]], [[x.\psi]] \subseteq X$. Denote by f the inclusion of $[[x.\phi]]$ into X.

1. Verify that $\forall_f$ is given by

$$\begin{aligned} \mathrm{Sub}([[x.\phi]]) &\to \mathrm{Sub}(X) \\ W &\mapsto \{x \in X \mid \phi(x) \Rightarrow x \in W\}. \end{aligned}$$

2. Verify that

$$[[x.\phi]] \to [[x.\psi]] = \{x \in X \mid \phi(x) \Rightarrow \psi(x)\}.$$

Exercise 4.7.4. Define interpretations of formulae of the form $\vec{x}.(\phi \to \psi)$ in a general Heyting category $\mathbf{C}$.

Heyting categories get their name from the fact that the subobject lattices of any object form a Heyting algebra.

Definition 4.7.5. A (distributive, bounded) lattice X is called a **Heyting algebra** if for all $x \in X$ there exists an (order-theoretic/Galois) right adjoint $x \to (-)$ to the map

$$x \wedge (-) : X \to X.$$

Remark. I am not sure if the distributivity property is necessary or if it follows from properties of the Heyting implication. For a great reference on Heyting algebras, the reader is referred to [TODO: Johnstone, Stone spaces]

Definition 4.7.6. Let X be a Heyting algebra and let $x \in X$. Let $\perp$ denote the lowest element of X. The **pseudocomplement** $\neg X$ of X is defined to be the element $x \to \perp$ of X.

Exercise 4.7.7. Let X be a topological space.

1. Verify that the set of open subsets of X forms a Heyting algebra. (What is the implication?)

2. Compute the pseudocomplement $\neg Y$ for an open set $Y \subseteq X$.

Finally, we will show how to interpret universal quantification in arbitrary Heyting categories. The setup is similar to that of Example 4.4.1, but we are now considering the right adjoint.

Exercise 4.7.8. Let $\{x, y\}.\phi(x, y)$ be a formula in a two-sorted signature Σ. After choosing a Σ-structure in **Set**, we can identify the interpretation of ϕ with a subset $S = \{(x, y) \in X \times Y \mid \phi(x, y)\}$. The projection map $\pi : X \times Y \to Y$ induces a morphism of posets $\pi^* : 2^X \to 2^{X \times Y}$. In a Heyting category, this morphism has a right adjoint $\forall_\pi$.

1. Show that for all $U \subseteq X \times Y$, one has that

$$\forall_\pi(U) = \{y \in Y \mid \forall x \in X : (x, y) \in U\}.$$

2. Define interpretations of formulae of the form $\vec{x}.(\forall y)\phi(\vec{x}, y)$ in a general Heyting category **C**.

Remark (Truth values and interpretations)**.** Suppose ϕ is a formula with empty context. An interpretation of ϕ in a Σ-structure in some category **C** with finite products is a subobject of the empty product, or terminal object, of **C**. These subobjects are also called **truth values**. Compare with the situation in **Set** - the terminal object is the singleton set $\{*\}$, which has two subobjects: $\{*\}$ itself, and the empty set - these can be identified with truth and falsehood, respectively. Giving an interpretation of ϕ in **Set** is then nothing else but assigning a truth value to ϕ.

In the same spirit, if ψ is a formula with nonempty context $\vec{x} = MA_1 \times \ldots \times MA_n$, an interpretation of ψ in **Set** can be understood as the subset of all elements

$$\vec{x} \in MA_1 \times \ldots \times MA_n$$

which render $\psi(\vec{x})$ to be true.

4.8 Models of a theory

Definition 4.8.1 (Satisfaction of sequents)**.** Let $\sigma = (\phi \vdash_{\vec{x}} \psi)$ be a sequent over a signature Σ which is interpretable in a category $\mathbf{C}$ (i.e. ϕ and ψ are interpretable) and let M be a Σ-structure in $\mathbf{C}$. Assume again that $\vec{x} = \{x_1, \ldots, x_n\}$ and that $x_i : A_i$.

1. We say that σ is **satisfied** in M if $[[\vec{x}.\phi]] \leq [[\vec{x}.\psi]]$ as subobjects of

$$MA_1 \times \ldots \times MA_n.$$

   If σ is satisfied in M, we write $M \vDash \sigma$.

2. Let $\mathbb{T}$ be a logical theory and suppose that all axioms of $\mathbb{T}$ are interpretable in $\mathbf{C}$. Then we call M a **model** of $\mathbb{T}$ if all axioms of $\mathbb{T}$ are satisfied in M. Denote by $\mathbb{T} - \mathrm{mod}(\mathbf{C})$ the full subcategory of $\Sigma - \mathrm{str}(\mathbf{C})$ spanned by the models of $\mathbb{T}$ in $\mathbf{C}$.

Lemma 4.8.2. Let M and N be Σ-structures in $\mathbf{C}$ and let $\vec{x}.\phi$ be a geometric formula interpretable in $\mathbf{C}$. For any morphism $f : M \to N$ of Σ-structures, one obtains a commutative square

$$\begin{array}{ccc} [[\vec{x}.\phi]]_M & \longrightarrow & MA_1 \times \ldots \times MA_n \\ \downarrow & & \downarrow \\ [[\vec{x}.\phi]]_N & \longrightarrow & NA_1 \times \ldots \times NA_n \end{array}$$

Proof. The proof is by induction on the structure of the formula. We just give a few examples to get acquainted with the technicalities:

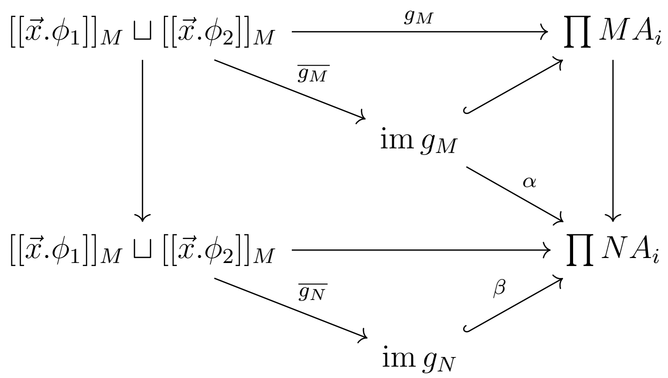
1. If ϕ is of the form $t_1 = t_2$, where the t_i are terms of type B in the right context, then the morphism between the interpretations is obtained from the universal property of the equaliser and the following diagram:

$$\begin{array}{ccccc} [[\vec{x}.t_1 = t_2]]_M & \longrightarrow & MA_1 \times \ldots \times MA_n & \rightrightarrows & MB \\ \downarrow & & \downarrow & & \downarrow \\ [[\vec{x}.t_1 = t_2]]_N & \longrightarrow & NA_1 \times \ldots \times NA_n & \rightrightarrows & NB \end{array}$$

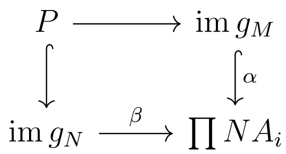
2. If ϕ is of the form $\phi_1 \wedge \phi_2$, then the same argument works by replacing "equaliser" with "pullback".

3. If ϕ is of the form $\phi_1 \vee \phi_2$, we have to argue using the defining property of the image as follows. Suppose we already have morphisms $[[\vec{x}.\phi_i]]_M \to [[\vec{x}.\phi_i]]_N$

for $i = 1, 2$. Consider the following diagram:



Here, $\alpha = \prod f_{A_i} \circ g_M$. (Also, recall from the definition that the $\operatorname{im} g_M$ are exactly the joins in the subobject lattice.) Form the pushout square



and observe that g_M factors through P by the universal property of the pullback. By minimality of $\operatorname{im} g_M$, it follows that $P \cong \operatorname{im} g_M$ and we have our desired morphism.

4. We can proceed similarly for existential quantifiers and (infinitary) joins.

□

Remark. It is important to note that interpretations of formulae which make use of the Heyting category structure (in particular, formulae which use infinitary meets, universal quantification and implication) are not preserved in this way. (TODO: Example and double check)

Definition 4.8.3. Let $p \in \{\text{cartesian}, \text{regular}, \text{coherent}, \text{geometric}, \text{first-order}\}$. Let $\mathbf{C}$ and $\mathbf{D}$ be two categories in which p logic can be interpreted. A functor $F : \mathbf{C} \to \mathbf{D}$ which preserves finite products and monomorphisms is a p functor if it preserves all of the categorical structure used in the construction of interpretations of p logic.

Lemma 4.8.4. In the same setting as above, let M be a model of a p theory $\mathbb{T}$ in $\mathbf{C}$ and let $F : \mathbf{C} \to \mathbf{D}$ be a p functor. Define $F(M)$ as the Σ-structure in $\mathbf{D}$ defined by applying F to all objects MA_i, morphisms Mf and subobjects MR. Then, $F(M)$ is a model of $\mathbb{T}$ in $\mathbf{D}$. In other words, F induces a functor $\mathbb{T}-\text{mod}(\mathbf{C}) \to \mathbb{T}-\text{mod}(\mathbf{D})$.

5 Theories are categories, models are functors

We have done a lot of work setting up a categorical framework for interpreting logical statements. The following theorem allows us to reap what we have sowed - we can now use the powerful tools from category theory to examine models of logical theories.

Theorem 5.1. Let $p \in \{\text{cartesian}, \text{regular}, \text{coherent}, \text{geometric}, \text{first-order}\}$ and denote by $\mathbf{Cat}_p$ the 2-category whose objects are those categories $\mathbf{C}$ having enough structure for p logic to be interpretable in $\mathbf{C}$, whose 1-morphisms are functors respecting that structure and whose 2-morphisms are natural transformations. Let $\mathbb{T}$ be a p theory. Then, the 2-functor

$$\begin{aligned} \mathbf{Cat}_p &\to \mathbf{Cat} \\ \mathbf{D} &\mapsto \mathbb{T} - \text{mod}(\mathbf{D}) \end{aligned}$$

is representable: There exists a category $\mathbf{C}_{\mathbb{T}}$, the so-called **syntactic category** of $\mathbb{T}$, which gives an equivalence of categories

$$\text{Fun}_p(\mathbf{C}_{\mathbb{T}}, \mathbf{D}) \cong \mathbb{T} - \text{mod}(\mathbf{D})$$

natural in $\mathbf{D}$.

Proof. TODO □

Corollary 5.0.1. There is a model $M_{\mathbb{T}} \in \mathbf{C}_{\mathbb{T}}$, called the **universal model** of $\mathbb{T}$ with the property that a p sequent $(\phi \vdash_{\vec{x}} \psi)$ is satisfied in $M_{\mathbb{T}}$ if and only if it is provable in $\mathbb{T}$.

TODO:

1. Entailment, provability

6 Example time!!!

6.1 The theory of graphs

The theory $\mathbb{T}_\Gamma$ of graphs has two sorts V, E, two function symbols $s, t : E \to V$, no relation symbols and no axioms. Denote by Σ_Γ the underlying signature.

Lemma 6.1.1. Let $\mathbf{C}$ be a category with equalisers. Every Σ_Γ-structure in $\mathbf{C}$ is a model for $\mathbb{T}_\Gamma$.

Corollary 6.1.2. Let Γ be the small category

$$\bullet \rightrightarrows \bullet$$

Then, one has an equivalence of categories

$$\mathbb{T}_\Gamma - \text{mod}(\mathbf{C}) \simeq \text{Fun}(\Gamma^{op}, \mathbf{C})$$

which is natural in $\mathbf{C}$.

Proof. One has a sequence of natural equivalences

$$\mathbb{T}_\Gamma - \text{mod}(\mathbf{C}) \cong \Sigma_\Gamma - \text{str}(\mathbf{C}) \cong \text{Fun}(\Gamma^{op}, \mathbf{C}).$$

□

Note that Γ is not the syntactic category of $\mathbb{T}_\Gamma$: Neither does it have all finite products nor does it have the appropriate universal property. With a bit more work (and the theory of so-called *flat functors*, one can show the following:

Corollary 6.1.3. The syntactic category for the theory of graphs $\mathbf{C}_{\mathbb{T}_\Gamma}$ is given by the category of presheaves on Γ.

One says that $\mathbb{T}_\Gamma$ is **of presheaf type**.

6.2 Algebraic theories

Algebraic theories are Horn theories which have one sort and no relation symbols. They can be interpreted in any finite limit category.

Example 6.2.1 (The theory of groups)**.** The theory of groups has one sort G, one constant symbol $1 : G$, two function symbols $\mu : G \times G \to G$ and $\iota : G \to G$, subject to the axioms

$$\begin{aligned}
&\top \vdash_{\{x,y,z\}} \mu(\mu(x,y),z) = \mu(x,\mu(y,z))\\
&\top \vdash_{\{x,y\}} \mu(1,x) = x\\
&\top \vdash_{\{x,y\}} \mu(x,1) = x\\
&\top \vdash_{\{x,y\}} \mu(x,\iota(x)) = 1\\
&\top \vdash_{\{x,y\}} \mu(\iota(x),x) = 1.
\end{aligned}$$

Example 6.2.2 (The theory of rings)**.** The theory of rings has one sort R, one constant symbol $0 : R$ and three function symbols: Multiplication $\mu : R \times R \to R$, Addition $\alpha : R \times R \to R$ and additive inverse $- : R \to R$, subject to the axioms

$$\begin{aligned}
&\top \vdash_{\{x,y,z\}} \alpha(\alpha(x,y),z) = \alpha(x,\alpha(y,z))\\
&\top \vdash_{\{x,y\}} \alpha(x,y) = \alpha(y,x)\\
&\top \vdash_{\{x\}} \alpha(0,x) = x\\
&\top \vdash_{\{x\}} \alpha(x,-(x)) = 0\\
&\top \vdash_{\{x,y,z\}} \mu(\mu(x,y),z) = \mu(x,\mu(y,z))\\
&\top \vdash_{\{x,y\}} \mu(x,y) = \mu(y,x)\\
&\top \vdash_{\{x,y,z\}} \mu(x,\alpha(y,z)) = \alpha(\mu(x,y),\mu(x,z)).
\end{aligned}$$

Example 6.2.3 (The theory of unital rings)**.** The theory of unital rings is obtained from the theory of rings by adding a constant symbol $1 : R$ and imposing the additional axiom

$$\top \vdash_{\{x\}} \mu(1,x) = x.$$

6.3 Non-algebraic Horn theories

Example 6.3.1 (The theory of posets)**.** The theory of posets has one sort X, no function symbols, one relation symbol $\leq$ and the following axioms:

$$\begin{aligned} \top &\vdash_{\{x\}} x \leq x \\ x \leq y \wedge y \leq x &\vdash_{\{x,y\}} x = y \\ x \leq y \wedge y \leq z &\vdash_{\{x,y,z\}} x \leq z. \end{aligned}$$

Example 6.3.2. The theories of bounded posets can be obtained from the theory of posets by adding constant symbols and adding the obvious axioms.

6.4 Regular theories

Example 6.4.1 (The theory of categories)**.** The theory of categories has two sorts, O (objects) and M (morphisms), three function symbols: $dom : M \to O$ (domain of a morphism), $cod : M \to O$ (codomain of a morphism) and $id : O \to M$ (assignment of identity morphism to each object), as well as a relation symbol $C \hookrightarrow M \times M \times M$ (composition) subject to the following axioms:

$$\begin{aligned} \top &\vdash_{\{X:O\}} dom(id(X)) = X \wedge cod(id(X)) = X \\ dom(g) = cod(f) &\vdash_{\{f,g,h:M\}} (\exists h).C(f,g,h) \\ \top &\vdash_{\{f:M\}} C(f, id(cod(f)), f) \wedge C(id(dom(f)), f, f) \\ C(f,g,h) = C(f,g,h') &\vdash_{\{f,g,h,h':M\}} h = h' \\ C(f,g,k) \wedge C(g,h,l) \wedge C(k,h,m) &\vdash_{\{f,g,h,k,l,m:M\}} C(f,l,m). \end{aligned}$$

6.5 Coherent theories

Example 6.5.1 (The theory of integral domains)**.** The theory of integral domains can be obtained from the theory of unital rings by imposing the additional axiom

$$\mu(x,y) = 0 \vdash_{\{x,y\}} x = 0 \vee y = 0.$$

Example 6.5.2 (The theory of graphs with diamond completion)**.** TODO

6.6 Geometric theories

Example 6.6.1 (The theory of finite sets)**.** The theory of finite sets has one sort X, no function symbols and one relation symbol $R_n \hookrightarrow X^n$ for each natural number n, subject to the following axioms:

1. For each natural number, one has one axiom

$$R_n(x_1, \ldots, x_n) \vdash_{\{x_1,\ldots,x_n,y\}} \bigvee_{i=1}^{n} y = x_i.$$

2. Every finite set is exhausted by some finite set of elements:

$$\top \vdash_{\emptyset} \bigvee_{n \in \mathbb{N}} (\exists x_1) \cdots (\exists x_n) R_n(x_1, \ldots, x_n).$$

3. For each surjection $f : \{1, \dots, m\} \to \{1, \dots, n\}$, one has one axiom

$$R_n(x_1, \dots, x_n) \vdash_{\{x_1,\dots,x_n\}} R_m(x_{f(1)}, \dots, x_{f(m)}).$$

4. For each $n \in \mathbb{N}$ and $i < j \in \{1, \dots, n\}$, one has one axiom

$$(R_n(x_1, \dots, x_n) \wedge x_i = x_j) \vdash_{\{x_1,\dots,x_n} R_{n-1}(x_1, \dots, \widehat{x_i}, \dots, x_n).$$

The relations $R_n(x_1, \dots, x_n)$ may be understood as the sentence *the collection of the* x_i *exhausts the finite set.*