Topos Theory

The interpretation of logic in categories

Olivia Caramello

Interpreting first-order logic in categories

- In Logic, first-order languages are a wide class of formal languages used for talking about mathematical structures of any kind (where the restriction 'first-order' means that quantification is allowed only over individuals rather than over collections of individuals or higher-order constructions on them).
- A first-order language contains sorts, which are meant to represent different *kinds* of individuals, terms, which denote individuals, and formulae, which make assertions about the individuals. Compound terms and formulae are formed by using various logical operators.
- It is well-known that first-order languages can always be interpreted in the context of (a given model of) set theory. In this lecture, we will show that these languages can also be meaningfully interpreted in a category, provided that the latter possesses enough categorical structure to allow the interpretation of the given fragment of logic. In fact, sorts will be interpreted as objects, terms as arrows and formulae as subobjects, in a way that respects the logical structure of compound expressions.

Signatures

Definition

A first-order signature Σ consists of the following data.

a) A set Σ-Sort of sorts.

b) A set Σ-Fun of function symbols, together with a map assigning to each $f \in \Sigma$-Fun its *type*, which consists of a finite non-empty list of sorts: we write

$$f : A_1 \cdots A_n \to B$$

to indicate that f has type $A_1, \ldots, A_n, B$ (if $n = 0$, f is called a constant of sort B).

c) A set Σ-Rel of relation symbols, together with a map assigning to each Σ-Rel its *type*, which consists of a finite list of sorts: we write

$$R \rightarrowtail A_1 \cdots A_n$$

to indicate that R has type $A_1, \ldots A_n$.

Terms

For each sort A of a signature Σ we assume given a supply of variables of sort A, used to denote individuals of kind A. Starting from variables, terms are built-up by repeated 'applications' of function symbols to them, as follows.

Definition

Let Σ be a signature. The collection of terms over Σ is defined recursively by the clauses below; simultaneously, we define the sort of each term and write $t : A$ to denote that t is a term of sort A.

a) $x : A$, if x is a variable of sort A.

b) $f(t_1, \ldots, t_n) : B$ if $f : A_1 \cdots A_n \to B$ is a function symbol and $t_1 : A_1, \ldots, t_n : A_n$.

Formation rules for formulae I

Consider the following formation rules for recursively building classes of formulae F over Σ, together with, for each formula ϕ, the (finite) set $FV(\phi)$ of free variables of ϕ.

(i) Relations: $R(t_1,\dots,t_n)$ is in F, if $R \rightarrowtail A_1 \cdots A_n$ is a relation symbol and $t_1 : A_1,\dots,t_n : A_n$ are terms; the free variables of this formula are all the variables occurring in some t_i.

(ii) Equality: $(s = t)$ is in F if s and t are terms of the same sort; $FV(s = t)$ is the set of variables occurring in s or t (or both).

(iii) Truth: $\top$ is in F; $FV(\top) = \emptyset$.

(iv) Binary conjunction: $(\phi \wedge \psi)$ is in F, if ϕ and ψ are in F; $FV(\phi \wedge \psi) = FV(\phi) \cup FV(\psi)$.

(v) Falsity: $\bot$ is in F; $FV(\bot) = \emptyset$.

(vi) Binary disjunction: $(\phi \vee \psi)$ is in F, if ϕ and ψ are in F; $FV(\phi \vee \psi) = FV(\phi) \cup FV(\psi)$.

(vii) Implication: $(\phi \Rightarrow \psi)$ is in F, if ϕ and ψ are in F; $FV(\phi \Rightarrow \psi) = FV(\phi) \cup FV(\psi)$.

(viii) Negation: $\neg\phi$ is in F, if ϕ is in F; $FV(\neg\phi) = FV(\phi)$.

Formation rules for formulae II

(ix) Existential quantification: $(\exists x)\phi$ is in F, if ϕ is in F and x is a variable; $FV((\exists x)\phi) = FV(\phi) \setminus \{x\}$.

(x) Universal quantification: $(\forall x)\phi$ is in F, if ϕ is in F and x is a variable; $FV((\forall x)\phi) = FV(\phi) \setminus \{x\}$.

(xi) Infinitary disjunction: $\bigvee_{i \in I} \phi_i$ is in F, if I is a set, ϕ_i is in F for each $i \in I$ and $FV(\bigvee_{i \in I} \phi_i) := \bigcup_{i \in I} FV(\phi_i)$ is finite.

(xii) Infinitary conjunction: $\bigwedge_{i \in I} \phi_i$ is in F, if I is a set, ϕ_i is in F for each $i \in I$ and $FV(\bigwedge_{i \in I} \phi_i) := \bigcup_{i \in I} FV(\phi_i)$ is finite.

A context is a finite list $\vec{x} = x_1, \ldots, x_n$ of distinct variables (the empty context, for $n = 0$ is allowed and indicated by []).

Notation: We will often consider formulae-in-context, that is formulae ϕ equipped with a context $\vec{x}$ such that all the free variables of ϕ occur among $\vec{x}$; we will write either $\phi(\vec{x})$ or $\{\vec{x} \,.\, \phi\}$.

Classes of formulae

Definition

In relation to the above-mentioned forming rules:

- The set of atomic formulae over Σ is the smallest set closed under *Relations* and *Equality*).
- The set of Horn formulae over Σ is the smallest set containing the class of atomic formulae and closed under *Truth* and *Binary conjunction*.
- The set of regular formulae over Σ is the smallest set containing the class of atomic formulae and closed under *Truth*, *Binary conjunction* and *Existential quantification*.
- The set of coherent formulae over Σ is the smallest set containing the set of regular formulae and closed under *False* and *Binary disjunction*.
- The set of first-order formulae over Σ is the smallest set closed under all the forming rules except for the infinitary ones.
- The *class* of geometric formulae over Σ is the smallest class containing the class of coherent formulae and closed under *Infinitary disjunction*.
- The *class* of infinitary first-order formulae over Σ is the smallest class closed under all the above-mentioned forming rules.

Sequents

Definition

- By a sequent over a signature Σ we mean a formal expression of the form $(\phi \vdash_{\vec{x}} \psi)$, where ϕ and ψ are formulae over Σ and $\vec{x}$ is a context suitable for both of them. The intended interpretation of this expression is that ψ is a logical consequence of ϕ in the context $\vec{x}$, i.e. that any assignment of individual values to the variables in $\vec{x}$ which makes ϕ true will also make ψ true.
- We say a sequent $(\phi \vdash_{\vec{x}} \psi)$ is Horn (resp. regular, coherent, ...) if both ϕ and ψ are Horn (resp. regular, coherent, ...) formulae.

Notice that, in full first-order logic, the general notion of sequent is not really needed, since the sequent $(\phi \vdash_{\vec{x}} \psi)$ expresses the same idea as $(\top \vdash_{[]} (\forall \vec{x})(\phi \Rightarrow \psi))$.

First-order theories

Definition

- By a theory over a signature Σ, we mean a set $\mathbb{T}$ of sequents over Σ, whose elements are called the (non-logical) axioms of $\mathbb{T}$.
- We say that $\mathbb{T}$ is an algebraic theory if its signature Σ has a single sort and no relation symbols (apart from equality) and its axioms are all of the form $\top \vdash_{\vec{x}} \phi$ where ϕ is an atomic formula $(s = t)$ and $\vec{x}$ its canonical context.
- We say $\mathbb{T}$ is a Horn (resp. regular, coherent, ...) theory if all the sequents in $\mathbb{T}$ are Horn (resp. regular, coherent, ...).

Deduction systems for first-order logic I

- To each of the fragments of first-order logic introduced above, we can naturally associate a deduction system, in the same spirit as in classical first-order logic. Such systems will be formulated as *sequent-calculi*, that is they will consist of inference rules enabling us to derive a sequent from a collection of others; we will write

$$\frac{\Gamma}{\sigma}$$

  to mean that the sequent σ can be inferred by a collection of sequents Γ. A double line instead of the single line will mean that each of the sequents can be inferred from the other.
- Given the axioms and inference rules below, the notion of proof is the usual one, and allowing the axioms of theory $\mathbb{T}$ to be taken as premises yields the notion of proof relative to a theory $\mathbb{T}$.

Consider the following rules.

Deduction systems for first-order logic II

- The rules for finite conjunction are the axioms

$$(\phi \vdash_{\vec{x}} \top) \quad ((\phi \wedge \psi) \vdash_{\vec{x}} \phi) \quad ((\phi \wedge \psi) \vdash_{\vec{x}} \psi)$$

  and the rule

$$\frac{(\phi \vdash_{\vec{x}} \psi)(\phi \vdash_{\vec{x}} \chi)}{(\phi \vdash_{\vec{x}} (\psi \wedge \chi))}$$

- The rules for finite disjunction are the axioms

$$(\bot \vdash_{\vec{x}} \phi) \quad (\phi \vdash_{\vec{x}} (\phi \vee \psi)) \quad (\psi \vdash_{\vec{x}} \phi \vee \psi)$$

  and the rule

$$\frac{(\phi \vdash_{\vec{x}} \chi)(\psi \vdash_{\vec{x}} \chi)}{((\phi \vee \psi) \vdash_{\vec{x}} \chi)}$$

- The rules for infinitary conjunction (resp. disjunction) are the infinitary analogues of the rules for finite conjunction (resp. disjunction).

Deduction systems for first-order logic III

- The rules for implication consist of the double rule

$$\frac{(\phi \wedge \psi \vdash_{\vec{x}} \chi)}{\overline{(\psi \vdash_{\vec{x}} (\phi \Rightarrow \chi))}}$$

- The rules for existential quantification consist of the double rule

$$\frac{(\phi \vdash_{\vec{x},y} \psi)}{\overline{((\exists y)\phi \vdash_{\vec{x}} \psi)}}$$

provided that y is not free in ψ.

- The rules for universal quantification consist of the double rule

$$\frac{(\phi \vdash_{\vec{x},y} \psi)}{\overline{(\phi \vdash_{\vec{x}} (\forall y)\psi)}}$$

- The distributive axiom is

$$((\phi \wedge (\psi \vee \chi)) \vdash_{\vec{x}} ((\phi \wedge \psi) \vee (\phi \wedge \chi)))$$

Deduction systems for first-order logic IV

- The Frobenius axiom is

$$((\phi \wedge (\exists y)\psi) \vdash_{\vec{x}} (\exists y)(\phi \wedge \psi))$$

  where y is a variable not in the context $\vec{x}$.

- The Law of excluded middle is

$$(\top \vdash_{\vec{x}} \phi \vee \neg\phi)$$

Fragments of first-order logic

Definition

In addition to the usual structural rules of *sequent-calculi* (Identity axiom, Equality rules, Substitution rule, and Cut rule), our deduction systems consist of the following rules:

Horn logic	finite conjunction
Regular logic	finite conjunction, existential quantification and Frobenius axiom
Coherent logic	finite conjunction, finite disjunction, existential quantification, distributive axiom and Frobenius axiom
Geometric logic	finite conjunction, infinitary disjunction, existential quantification, 'infinitary' distributive axiom, Frobenius axiom
Intuitionistic first-order logic	all the finitary rules except for the law of excluded middle
Classical first-order logic	all the finitary rules

Provability in fragments of first-order logic

Definition

We say a sequent σ is provable in an algebraic (regular, coherent, ...) theory $\mathbb{T}$ if there exists a derivation of σ relative to $\mathbb{T}$, in the appropriate fragment of first-order logic.

In geometric logic, intuitionistic and classical provability of geometric sequents coincide.

Theorem

If a geometric sequent σ is derivable from the axioms of a geometric theory $\mathbb{T}$ using 'classical geometric logic' (i.e. the rules of geometric logic plus the Law of Excluded Middle), then there is also a constructive derivation of σ, not using the Law of Excluded Middle.

Categorical semantics

- Generalizing the classical Tarskian definition of satisfaction of first-order formulae in ordinary set-valued structures, one can obtain, given a signature Σ, a notion of Σ-structure in a category with finite products, and define, according to the categorical structure present on the category, a notion of interpretation of an appropriate fragment of first-order logic in it.
- Specifically, we will introduce various classes of 'logical' categories, each of them providing a semantics for a corresponding fragment of first-order logic:

Cartesian categories	Horn logic
Regular categories	Regular logic
Coherent categories	Coherent logic
Geometric categories	Geometric logic
Heyting categories	First-order logic

Structures in categories

Definition

Let $\mathscr{C}$ be a category with finite products and Σ be a signature. A Σ-structure M in $\mathscr{C}$ is specified by the following data:

(i) A function assigning to each sort A in Σ-Sort, an object MA of $\mathscr{C}$. For finite strings of sorts, we define $M(A_1, \ldots, A_n) = MA_1 \times \cdots \times MA_n$ and set $M([])$ equal to the terminal object 1 of $\mathscr{C}$.

(ii) A function assigning to each function symbol $f : A_1 \cdots A_n \to B$ in Σ-Fun an arrow $Mf : M(A_1, \ldots, A_n) \to MB$ in $\mathscr{C}$.

(iii) A function assigning to each relation symbol $R \rightarrowtail A_1 \cdots A_n$ in Σ-Rel a subobject $MR \rightarrowtail M(A_1, \ldots, A_n)$ in $\mathscr{C}$.

Homomorphisms of structures

Definition

A Σ-structure homomorphism $h : M \to N$ between two Σ-structures M and N in $\mathscr{C}$ is a collection of arrows $h_A : MA \to NA$ in $\mathscr{C}$ indexed by the sorts of Σ and satisfying the following two conditions:

(i) For each function symbol $f : A_1 \cdots A_n \to B$ in Σ-Fun, the diagram

$$\begin{array}{ccc} M(A_1, \ldots, A_n) & \xrightarrow{Mf} & MB \\ \downarrow{\scriptstyle h_{A_1} \times \cdots \times h_{A_n}} & & \downarrow{\scriptstyle h_B} \\ N(A_1, \ldots, A_n) & \xrightarrow[Nf]{} & NB \end{array}$$

commutes.

(ii) For each relation symbol $R \rightarrowtail A_1 \cdots A_n$ in Σ-Rel, there is a commutative diagram in $\mathscr{C}$ of the form

$$\begin{array}{ccc} MR & \longrightarrow & M(A_1, \ldots, A_n) \\ \downarrow & & \downarrow{\scriptstyle h_{A_1} \times \cdots \times h_{A_n}} \\ NR & \longrightarrow & M(A_1, \ldots, A_n) \end{array}$$

The category of Σ-structures

Definition

Given a category $\mathscr{C}$ with finite products, Σ-structures in $\mathscr{C}$ and Σ-homomorphisms between them form a category, denoted by $\Sigma\text{-str}(\mathscr{C})$. Identities and composition in $\Sigma\text{-str}(\mathscr{C})$ are defined componentwise from those in $\mathscr{C}$.

Remark

If $\mathscr{C}$ and $\mathscr{D}$ are two categories with finite products, then any functor $T : \mathscr{C} \to \mathscr{D}$ which preserves finite products and monomorphisms induces a functor $\Sigma\text{-str}(T) : \Sigma\text{-str}(\mathscr{C}) \to \Sigma\text{-str}(\mathscr{D})$ in the obvious way.

The interpretation of terms

Definition

Let M be a Σ-structure in a category $\mathscr{C}$ with finite products. If $\{\vec{x}\ .\ t\}$ is a term-in-context over Σ (with $\vec{x} = x_1, \ldots, x_n$, $x_i : A_i$ ($i = 1, ..., n$) and $t : B$, say), then an arrow

$$[[\vec{x}\ .\ t]]_M : M(A_1, \ldots, A_n) \to MB$$

in $\mathscr{C}$ is defined recursively by the following clauses:

a) If t is a variable, it is necessarily x_i for some unique $i \leq n$, and then $[[\vec{x}\ .\ t]]_M = \pi_i$, the ith product projection.

b) If t is $f(t_1, \ldots, t_m)$ (where $t_i : C_i$, say), then $[[\vec{x}\ .\ t]]_M$ is the composite

$$M(A_1, \ldots, A_n) \xrightarrow{([[\vec{x}.t_1]]_M, \ldots, [[\vec{x}.t_m]]_M)} M(C_1, \ldots, C_m) \xrightarrow{Mf} MB$$

Interpreting formulae in categories

- In order to interpret formulae in categories, we need to have a certain amount of categorical structure present on the category in order to give a meaning to the logical connectives which appear in the formulae.
- In fact, the larger is the fragment of logic, the larger is the amount of categorical structure required to interpret it. For example, to interpret finitary conjunctions, we need to form pullbacks, to interpret disjunctions we need to form unions of subobjects, etc.
- Formulae will be interpreted as subobjects in our category; specifically, given a category $\mathscr{C}$ and a Σ-structure M in it, a formula $\phi(\vec{x})$ over Σ where $\vec{x} = (x_1^{A_1}, \ldots, x_n^{A_n})$, will be interpreted as a subobject

$$[[\vec{x} \, . \, \phi]]_M \rightarrowtail M(A_1, \ldots, A_n)$$

defined recursively on the structure of ϕ.

Cartesian categories

Recall that by a finite limit in a category $\mathscr{C}$ we mean a limit of a functor $F : \mathscr{J} \to \mathscr{C}$ where $\mathscr{J}$ is a finite category (i.e. a category with only a finite number of objects and arrows).
In any category $\mathscr{C}$ with pullbacks, pullbacks of monomorphisms are again monomorphisms; thus, for any arrow $f : a \to b$ in $\mathscr{C}$, we have a pullback functor

$$f^* : \mathrm{Sub}_{\mathscr{C}}(b) \to \mathrm{Sub}_{\mathscr{C}}(a) .$$

Definition

A cartesian category is any category with finite limits.

As we shall see below, in cartesian categories we can interpret atomic formulae as well as finite conjunctions of them; in fact, conjunctions will be interpreted as pullbacks (i.e. intersections) of subobjects.

Regular categories

Definition

- Given two subobjects $m_1 : a_1 \rightarrowtail c$ and $m_2 : a_2 \rightarrowtail c$ of an object c in a category $\mathscr{C}$, we say that m_1 factors through m_2 if there is a (necessarily unique) arrow $r : a_1 \to a_2$ in $\mathscr{C}$ such that $m_2 \circ r = m_1$. (Note that this defines a preorder relation $\leq$ on the collection $\mathrm{Sub}_{\mathscr{C}}(c)$ of subobjects of a given object c.)
- We say that a cartesian category $\mathscr{C}$ has images if we are given an operation assigning to each morphism of $\mathscr{C}$ a subobject $Im(f)$ of its codomain, which is the least (in the sense of the preorder $\leq$) subobject of $cod(f)$ through which f factors.
- A regular category is a cartesian category $\mathscr{C}$ such that $\mathscr{C}$ has images and they are stable under pullback.

Fact

Given an arrow $f : a \to b$ in a regular category $\mathscr{C}$, the pullback functor $f^* : \mathrm{Sub}_{\mathscr{C}}(b) \to \mathrm{Sub}_{\mathscr{C}}(a)$ has a left adjoint $\exists_f : \mathrm{Sub}_{\mathscr{C}}(a) \to \mathrm{Sub}_{\mathscr{C}}(b)$, which assigns to a subobject $m : c \rightarrowtail a$ the image of the composite $f \circ m$.

As we shall see below, in regular categories we can interpret formulae built-up from atomic formulae by using finite conjunctions and existential quantifications; in fact, the existential quantifiers will be interpreted as images of certain arrows.

Coherent categories

Definition

A coherent category is a regular category $\mathscr{C}$ in which each $Sub_{\mathscr{C}}(c)$ has finite unions and each $f^* : \mathrm{Sub}_{\mathscr{C}}(b) \to \mathrm{Sub}_{\mathscr{C}}(a)$ preserves them.

As we shall see below, in coherent categories we can interpret formulae built-up from atomic formulae by using finite conjunctions, existential quantifications, and finite disjunctions; in fact, finite disjunctions will be interpreted as finite unions of subobjects.

Note in passing that, if coproducts exist, a union of subobjects of an object c may be constructed as the image of the induced arrow from the coproduct to c.

Geometric categories

Definition

- A (large) category $\mathscr{C}$ is said to be well-powered if each of the preorders $Sub_{\mathscr{C}}(a)$, $a \in Ob(\mathscr{C})$, is equivalent to a small category.
- A geometric category is a well-powered regular category whose subobject lattices have arbitrary unions which are stable under pullback.

As we shall see below, in coherent categories we can interpret formulae built-up from atomic formulae by using finite conjunctions, existential quantifications, and infinitary disjunctions; in fact, disjunctions will be interpreted as unions of subobjects.

Quantifiers as adjoints

Let X and Y be two sets. For any given subset $S \subseteq X \times Y$, we can consider the sets

$$\forall_p S := \{y \in Y \mid \text{for all } x \in X, (x, y) \in S\} \text{ and}$$

$$\exists_p S := \{y \in Y \mid \text{there exists } x \in X, (x, y) \in S\} .$$

The projection map $p : X \times Y \to Y$ induces a map (taking inverse images) at the level of powersets $p^* : \mathscr{P}(Y) \to \mathscr{P}(X \times Y)$. If we regard these powersets as poset categories (where the order-relation is given by the inclusion relation) then this map becomes a functor; also, the assignments $S \to \forall_p S$ and $S \to \exists_p S$ yield functors $\forall_p, \exists_p : \mathscr{P}(X \times Y) \to \mathscr{P}(Y)$.

Theorem

The functors $\exists_p$ and $\forall_p$ are respectively left and right adjoints to the functor $p^* : \mathscr{P}(Y) \to \mathscr{P}(X \times Y)$ which sends each subset $T \subseteq Y$ to its inverse image $p^* T$ under p.

The theorem generalizes to the case of an arbitrary function in place of the projection p.

Heyting categories I

Definition

A Heyting category is a coherent category $\mathscr{C}$ such that for any arrow $f : a \to b$ in $\mathscr{C}$ the pullback functor $f^* : \mathrm{Sub}_{\mathscr{C}}(b) \to \mathrm{Sub}_{\mathscr{C}}(a)$ has a right adjoint $\forall_f : \mathrm{Sub}_{\mathscr{C}}(a) \to \mathrm{Sub}_{\mathscr{C}}(b)$ (as well as its left adjoint $\exists_f : \mathrm{Sub}_{\mathscr{C}}(a) \to \mathrm{Sub}_{\mathscr{C}}(b)$).

Theorem

Let $a_1 \rightarrowtail a$ and $a_2 \rightarrowtail a$ be subobjects in a Heyting category. Then there exists a largest subobject $(a_1 \Rightarrow a_2) \rightarrowtail a$ such that $(a_1 \Rightarrow a_2) \cap a_1 \leq a_2$. Moreover, the binary operation on subobjects thus defined is stable under pullback.

In particular, all the subobject lattices in a Heyting category are Heyting algebras.

Thus, in a Heyting category we may interpret full finitary first-order logic.

Fact

Any geometric category is a Heyting category.

Heyting categories II

Theorem

Any elementary topos is a Heyting category.

Sketch of proof.

Let $\mathscr{E}$ be an elementary topos. The existence of a left adjoint to the pullback functor follows from the existence of images, while the existence of the right adjoint follows from the cartesian closed structure.

The object Ω has the structure of an internal Heyting algebra in $\mathscr{E}$; in fact, the Heyting algebra structure of the subobject lattices in $\mathscr{E}$ is induced by this internal structure via the Yoneda Lemma. □

The interpretation of first-order formulae I

Let M be a Σ-structure in a category $\mathscr{C}$ with finite limits. A formula-in-context $\{\vec{x} \,.\, \phi\}$ over Σ (where $\vec{x} = x_1, \ldots, x_n$ and $x_i : A_i$, say) will be interpreted as a subobject $[[\vec{x} \,.\, \phi]]_M \rightarrowtail M(A_1, \ldots, A_n)$ according to the following recursive clauses:

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

$$\begin{array}{ccc} [[\vec{x} \,.\, \phi]]_M & \longrightarrow & MR \\ \downarrow & & \downarrow \\ M(A_1, \ldots, A_n) & \xrightarrow{([[\vec{x}.t_1]]_M, \ldots, [[\vec{x}.t_m]]_M)} & M(B_1, \ldots, B_m) \end{array}$$

- If $\phi(\vec{x})$ is $(s = t)$, where s and t are terms of sort B, then $[[\vec{x} \,.\, \phi]]_M$ is the equalizer of $[[\vec{x} \,.\, s]]_M, [[\vec{x} \,.\, t]]_M : M(A_1, \ldots, A_n) \to MB$.
- If $\phi(\vec{x})$ is $\top$ then $[[\vec{x} \,.\, \phi]]_M$ is the top element of $\mathrm{Sub}_{\mathscr{C}}(M(A_1, \ldots, A_n))$.

The interpretation of first-order formulae II

- If ϕ is $\psi \wedge \chi$ then $[[\vec{x} \,.\, \phi]]_M$ is the intersection (= pullback)

$$\begin{array}{ccc} [[\vec{x} \,.\, \phi]]_M & \longrightarrow & [[\vec{x} \,.\, \chi]]_M \\ \downarrow & & \downarrow \\ [[\vec{x} \,.\, \psi]]_M & \longrightarrow & M(A_1, \ldots, A_n) \end{array}$$

- If $\phi(\vec{x})$ is $\bot$ and $\mathscr{C}$ is a coherent category then $[[\vec{x} \,.\, \phi]]_M$ is the bottom element of $\mathrm{Sub}_{\mathscr{C}}(M(A_1, \ldots, A_n))$.
- If ϕ is $\psi \vee \chi$ and $\mathscr{C}$ is a coherent category then $[[\vec{x} \,.\, \phi]]_M$ is the union of the subobjects $[[\vec{x} \,.\, \psi]]_M$ and $[[\vec{x} \,.\, \chi]]_M$.
- If ϕ is $\psi \Rightarrow \chi$ and $\mathscr{C}$ is a Heyting category, $[[\vec{x} \,.\, \phi]]_M$ is the implication $[[\vec{x} \,.\, \psi]]_M \Rightarrow [[\vec{x} \,.\, \chi]]_M$ in the Heyting algebra $\mathrm{Sub}_{\mathscr{C}}(M(A_1, \ldots, A_n))$ (similarly, the negation $\neg\psi$ is interpreted as the pseudocomplement of $[[\vec{x} \,.\, \psi]]_M$).

The interpretation of first-order formulae III

- If ϕ is $(\exists y)\psi$ where y is of sort B, and $\mathscr{C}$ is a regular category, then $[[\vec{x} \,.\, \phi]]_M$ is the image of the composite

$$[[\vec{x}, y \,.\, \psi]]_M \longrightarrow M(A_1, \ldots, A_n, B) \xrightarrow{\pi} M(A_1, \ldots, A_n)$$

  where π is the product projection on the first n factors.
- If ϕ is $(\forall y)\psi$ where y is of sort B, and $\mathscr{C}$ is a Heyting category, then $[[\vec{x} \,.\, \phi]]_M$ is $\forall_\pi([[\vec{x}, y \,.\, \psi]]_M)$, where π is the same projection as above.
- If ϕ is $\bigvee_{i \in I} \phi_i$ and $\mathscr{C}$ is a geometric category then $[[\vec{x} \,.\, \phi]]_M$ is the union of the subobjects $[[\vec{x} \,.\, \phi_i]]_M$.
- If ϕ is $\bigwedge_{i \in I} \phi_i$ and $\mathscr{C}$ has arbitrary intersections of subobjects then $[[\vec{x} \,.\, \phi]]_M$ is the intersection of the subobjects $[[\vec{x} \,.\, \phi_i]]_M$.

Models of first-order theories in categories

Definition

Let M be a Σ-structure in a category $\mathscr{C}$.

a) If $\sigma = \phi \vdash_{\vec{x}} \psi$ is a sequent over Σ interpretable in $\mathscr{C}$, we say that σ is satisfied in M if $[[\vec{x} \, . \, \phi]]_M \leq [[\vec{x} \, . \, \psi]]_M$ in $\mathrm{Sub}_{\mathscr{C}}(M(A_1, \ldots, A_n))$.
b) If $\mathbb{T}$ is a theory over Σ interpretable in $\mathscr{C}$, we say M is a model of $\mathbb{T}$ if all the axioms of $\mathbb{T}$ are satisfied in M.
c) We write $\mathbb{T}$-mod($\mathscr{C}$) for the full subcategory of Σ-str($\mathscr{C}$) whose objects are models of T.

We say that a functor $F : \mathscr{C} \to \mathscr{D}$ between two cartesian (resp. regular, coherent, geometric, Heyting) categories is cartesian (resp. regular, coherent, geometric, Heyting) if it preserves finite limits (resp. finite limits and images, finite limits and images and finite unions of subobjects, finite limits and images and arbitrary unions of subobjects, finite limits and images and Heyting implications between subobjects).

Theorem

If $\mathbb{T}$ is a regular (resp. coherent, ...) theory over Σ, then for any regular (resp. coherent, ...) functor $T : \mathscr{C} \to \mathscr{D}$ the functor Σ-str$(T) : \Sigma$-str$(\mathscr{C}) \to \Sigma$-str$(\mathscr{D})$ defined above restricts to a functor $\mathbb{T}$-mod$(T) : \mathbb{T}$-mod$(\mathscr{C}) \to \mathbb{T}$-mod$(\mathscr{D})$. If T is moreover conservative (that is, reflects isomorphisms) then the functor Σ-str(T) reflects the property of being a $\mathbb{T}$-model.

Examples

- A topological group can be seen as a model of the theory of groups in the category of topological spaces.
- Similarly, an algebraic (resp. Lie) group is a model of the theory of groups in the category of algebraic varieties (resp. the category of smooth manifolds).
- A sheaf of rings (more generally, a sheaf of models of a Horn theory $\mathbb{T}$) on a topological space X can be seen as a model of the theory of rings (resp. of the theory $\mathbb{T}$) in the topos $\mathbf{Sh}(X)$ of sheaves on X.
- A sheaf of models of a geometric theory $\mathbb{T}$ over a signature Σ in a topos $\mathbf{Sh}(X)$ of sheaves on a topological space X is a Σ-structure in $\mathbf{Sh}(X)$ whose stalks are models of $\mathbb{T}$.
- A bunch of set-based models of a theory $\mathbb{T}$ indexed over a set I can be seen as a model of $\mathbb{T}$ in the functor category $[I, \mathbf{Set}]$. More generally, we have that $\mathbb{T}\text{-mod}([\mathscr{C}, \mathbf{Set}]) \simeq [\mathscr{C}, \mathbb{T}\text{-mod}(\mathbf{Set})]$.

Soundness and completeness

Theorem (Soundness)

Let $\mathbb{T}$ be a Horn (resp. regular, coherent, first-order, geometric) theory over a signature $\mathbb{T}$, and let M be a model of $\mathbb{T}$ in a cartesian (resp. regular, coherent, Heyting, geometric) category $\mathscr{C}$. If σ is a sequent (in the appropriate fragment of first-order logic over Σ) which is provable in $\mathbb{T}$, then σ is satisfied in M.

Theorem (Completeness)

Let $\mathbb{T}$ be a Horn (resp. regular, coherent, first-order, geometric) theory. If a Horn (resp. regular, coherent, Heyting, geometric) sequent σ is satisfied in all models of $\mathbb{T}$ in cartesian (resp. regular, coherent, Heyting, geometric) categories, then it is provable in $\mathbb{T}$.

Toposes as mathematical universes

We say that a first-order formula $\phi(\vec{x})$ over a signature Σ is valid in an elementary topos $\mathscr{E}$ if for every Σ-structure M in $\mathscr{E}$ the sequent $\top \vdash_{\vec{x}} \phi$ is satisfied in M.

Theorem

Let Σ be a signature and $\phi(\vec{x})$ a first-order formula over Σ. Then $\phi(\vec{x})$ is provable in intuitionistic (finitary) first-order logic if and only if it is valid in every elementary topos.

Sketch of proof.

The soundness result is part of a theorem mentioned above. The completeness part follows from the existence of canonical Kripke models and the fact that, given a poset P and a Kripke model $\mathscr{U}$ on P there is a model $\mathscr{U}^*$ in the topos $[P, \mathbf{Set}]$ such that the first-order sequents valid in $\mathscr{U}$ are exactly those valid in $\mathscr{U}^*$. □

Hence an elementary topos can be considered as a mathematical universe in which one can do mathematics similarly to how one does it in the classical context of sets (with the only exception that one must in general argue constructively).

The internal language of a topos I

Given a category $\mathscr{C}$ with finite products, in particular an elementary topos, one can define a first-order signature $\Sigma_{\mathscr{C}}$, called the internal language of $\mathscr{C}$, for reasoning about $\mathscr{C}$ in a set-theoretic fashion, that is by using 'elements'.

Definition

The signature $\Sigma_{\mathscr{C}}$ has one sort $\ulcorner A \urcorner$ for each object A of $\mathscr{C}$, one function symbol $\ulcorner f \urcorner : \ulcorner A_1 \urcorner, \cdots, \ulcorner A_n \urcorner \to \ulcorner B \urcorner$ for each arrow $f : A_1 \times \cdots \times A_n \to B$ in $\mathscr{C}$, and one relation symbol $\ulcorner R \urcorner \rightarrowtail \ulcorner A_1 \urcorner \cdots \ulcorner A_n \urcorner$ for each subobject $R \rightarrowtail A_1 \times \cdots \times A_n$.

Note that there is a canonical $\Sigma_{\mathscr{C}}$-structure in $\mathscr{C}$, which assigns A to $\ulcorner A \urcorner$, f to $\ulcorner f \urcorner$ and R to $\ulcorner R \urcorner$.

The usefulness of this definition lies in the fact that properties of $\mathscr{C}$ or constructions in it can often be formulated in terms of satisfaction of certain formulae over $\Sigma_{\mathscr{C}}$ in the canonical structure; the internal language can thus be used for proving things about $\mathscr{C}$.

The internal language of a topos II

If $\mathscr{C}$ is an elementary topos, we can extend the internal language by allowing the formation of formulae of the kind $\tau \in \Gamma$, where τ is a term of sort A and Γ is a term of sort Ω^A. Indeed, we may interpret this formula as the subobject whose classifying arrow is the composite

$$W \xrightarrow{\langle \tau, \Gamma \rangle} A \times \Omega^A \xrightarrow{\in_A} \Omega$$

where W denotes the product of (the objects representing the) sorts of the variables occurring either in τ or in Γ (considered without repetitions) and $\langle \tau, \Gamma \rangle$ denotes the induced map to the product.
Note that an object A of $\mathscr{C}$ gives rise to a constant term of type Ω^A.
Thus in a topos we can also interpret all the common formulas that we use in Set Theory.

Kripke-Joyal semantics I

Kripke-Joyal semantics represents the analogue for toposes of the usual Tarskian semantics for classical first-order logic.

In the context of toposes, it makes no sense to speak of elements of a structure in a topos, but we can replace the classical notion of element of a set with that of generalized element of an object: a generalized element of an object c of a topos $\mathscr{E}$ is simply an arrow $\alpha : u \to c$ with codomain c.

Definition

Let $\mathscr{E}$ be a topos and M be a Σ-structure in $\mathscr{E}$. Given a first-order formula $\phi(x)$ over Σ in a variable x of sort A and a generalized element $\alpha : U \to MA$ of MA, we define

$$U \models_M \phi(\alpha) \quad \text{iff} \quad \alpha \text{ factors through } [[x . \phi]]_M \rightarrowtail MA$$

Of course, the definition can be extended to formulae with an arbitrary (finite) number of free variables.

In the following proposition, the notation $+$ denotes binary coproduct.

Kripke-Joyal semantics II

Proposition

If $\alpha : U \to MA$ is a generalized element of MA while $\phi(x)$ and $\psi(x)$ are formulas with a free variable x of sort A, then

- $U \models (\phi \wedge \psi)(\alpha)$ *if and only if* $U \models \phi(\alpha)$ *and* $U \models \psi(\alpha)$.
- $U \models (\phi \vee \psi)(\alpha)$ *if and only if there are arrows* $p : V \to U$ *and* $q : W \to U$ *such that* $p + q : V + W \to U$ *is epic, while both* $V \models \phi(\alpha \circ p)$ *and* $W \models \psi(\alpha \circ q)$.
- $U \models (\phi \Rightarrow \psi)(\alpha)$ *if and only if for any arrow* $p : V \to U$ *such that* $V \models \phi(\alpha \circ p)$, *then* $V \models \psi(\alpha \circ p)$.
- $U \models (\neg\phi)(\alpha)$ *if and only if whenever* $p : V \to U$ *is such that* $V \models \phi(\alpha \circ p)$, *then* $V \cong 0_{\mathscr{E}}$.

If $\phi(x, y)$ has an additional free variable y of sort B then

- $U \models (\exists y)\phi(\alpha, y)$ *if and only if there exist an epi* $p : V \to U$ *and a generalized element* $\beta : V \to B$ *such that* $V \models \phi(\alpha \circ p, \beta)$.
- $U \models (\forall y)\phi(\alpha, y)$ *if and only if for every object* V, *for every arrow* $p : V \to U$ *and every generalized element* $c : V \to B$ *one has* $V \models \phi(\alpha \circ p, \beta)$.

For further reading

- R. I. Goldblatt.
  Topoi. The categorial analysis of logic, vol. 98 of *Studies in Logic and the Foundations of Math.*
  North-Holland, 1979 (revised second edition, 1984).
- P. T. Johnstone.
  Sketches of an Elephant: a topos theory compendium. Vols. 1-2, vols. 43-44 of *Oxford Logic Guides*
  Oxford University Press, 2002.
- S. Mac Lane and I. Moerdijk.
  Sheaves in geometry and logic: a first introduction to topos theory
  Springer-Verlag, 1992.