THE MITCHELL-BÉNABOU LANGUAGE SEMINAR: *TOPOI, LOGIC, AND FORCING*

AREEB S.M.

Abstract. The axioms of set theory are stated in first order logic. The language of set theory contains, along with standard first order logical symbols the membership symbol "$\in$". In a topos, not only can we canonically interpret the symbols of first order logic, we can also interpret "membership" in this sense. Doing so yields a versatile tool for mimicking set theoretic constructions inside a topos.

Contents

1. Formal languages

1.1. **Motivation.** One uses language to perform mathematics. In this effort, one describes mathematical structures by making statements, sequences of words that can be parsed . Formal languages are introduced as a method to turn language into a "meta-mathematical" object, to attempt precision in describing what can be described in a rigorous manner.

We start with a *vocabulary*, which is a collection of symbols, called *words*. One then considers *sentences*, finite (possibly empty) sequences of words. A *formal language* is a collection of sentences, which are called *axioms*.

Typically a formal language is generated by a *grammar*, a collection of string manipulations called *formation rules* which are intuitively recipes to construct new well formed sentences from others.

Remark 1.1. Many authors (for instance, many in computability theory) replace the word-sentence metaphor with an alphabet-word one, and speak of *accepted words* instead of well formed sentences.

Date: 21.06.2022.

First order languages are a class of formal languages for describing structures. By "first order" we mean that the existential and universal quantification takes place over "elements" or "inhabitants" of the structure being described[1].

Example 1.2. One could attempt to write a language for monoids by having

- A "constant" word/symbol "1".
- A "function" symbol "$\cdot$".
- "Equality" and "universal quantification" symbols "=" and "$\forall$".
- "Variable" symbols "g_1, g_2, g_3".
- Parentheses and the full-stop symbol.

so that we can write the sentences $\forall g_1.(g_1 \cdot 1 = g_1 = 1 \cdot g_1)$ and $\forall g_1.\forall g_2.\forall g_3.(g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3)$.

Remark 1.3. We will ultimately only use a sentence of the form $\forall x.\phi$ as an abuse of notation. We will really mean a sentence $\forall (x \in X).\phi$, or $\forall (x : X).\phi$. Here the symbol X is the *sort* of x, intuitively the *type* of object that x is. We will always quantify over the inhabitants of a fixed sort.

For instance, in writing a language for 1-categories, we use separate sorts for objects and arrows, as we want to make sense of source and target symbols which intuitively take an entity of sort "arrow" to entities of sort "object", and similarly for the identity map assignment. We also need to quantify over objects and arrows separately in the statements encoding the axioms. In the case of one object, the language for monoids works as we need only quantify over arrows, so we survive with a single implicit sort.

Notation 1.4. For convenience, we will write a string $\forall (x_1 : X_1, x_2 : X_2, \ldots, x_n : X_n)$. when we mean the string $\forall (x_1 : X_1).\forall (x_2 : X_2).\ldots\forall (x_n : X_n)$. (also to aid readability).

1.2. **First order logic.** The vocabulary over which we consider first order languages includes standard symbols in every first order language (such as equality), as well as a few symbols that are additional data. These additional symbols belong to certain distinguished classes, and the notion of a "signature" organizes this data.

Remark 1.5. The main source for this section is [Car14, Section 4.1]. Also mentioned, are various intermediate fragments of first order logic that have been studied in the literature.

Definition 1.6. The data of a *first order signature* Σ consists of

- A set $\Sigma-\mathsf{Sort}$ of *sorts* or *types*.
- A set $\Sigma-\mathsf{Fun}$ of *function symbols*, as well as for each function symbol $f \in \Sigma-\mathsf{Fun}$ a finite **nonempty** sequence of sorts $A_1, A_2, \ldots, A_n, B$ called the sort of f. The natural number n here is called the *arity* of f.
- A set $\Sigma-\mathsf{Rel}$ of *relation symbols*, as well as for each relation symbol $R \in \Sigma-\mathsf{Rel}$ a finite list of sorts $A_1, A_2, A_3, \ldots, A_n$ called the sort of R. Again, the number n here is called the *arity* of R.
- For each sort $X \in \Sigma-\mathsf{Sort}$ a collection of mutually distinct symbols called *variables*[2].

We also call function symbols of arity 0 *constants*.

Notations 1.7. If f is a function symbol of sort $A_1, A_2, \ldots, A_n, B$, we will write

$$f\colon A_1 \times A_2 \times \cdots \times A_n \to B$$

[1]This is in contrast with say, second order logic where we have a notion of quantifying over "collections of inhabitants".

[2]In particular, each variable has a uniquely determined sort.

for ease of readability.

Similarly, if R is a relation symbol of sort $A_1, A_2, \dots, A_n$, we denote it as

$$R \rightarrowtail A_1 \times A_2 \times \dots A_n$$

Remark 1.8. For us, it will be convenient to assume a countably infinite collection of variables for each sort.

Construction 1.9. Consider a first order signature Σ. Then we define inductively a class $\Sigma-\mathsf{Term}$ of *terms* over Σ and a type/sort *judgment* $t : A$ with $A \in \Sigma-\mathsf{Sort}$ for each $t \in \Sigma-\mathsf{Term}$, as well as a (finite) collection of *free* variables $\mathsf{FV}(t)$.

One starts by considering every variable x of sort X as a term, with the judgment $x : X$ and $\mathsf{FV}(x) := \{x\}$. One then continues by introducing for each function symbol $f \colon A_1 \times A_2 \times \cdots \times A_n \to B$ and terms $t_1 : A_1, t_2 : A_2, \dots, t_n : A_n$ the term $f(t_1, t_2, \dots, t_n)$. This is given the typing judgment $f(t_1, t_2, \dots, t_n) : B$ and one sets $\mathsf{FV}(f(t_1, t_2, \dots, t_n)) := \bigcup_{i=1}^n \mathsf{FV}(t_i)$

Intuitively, terms are determinations of inhabitants of a structure. We will introduce "formulae", which are intuitively well formed strings that are meant to have a "truth value". We will inductively construct a class of "first order formulae" over a signature through a list of string manipulations called "formation rules".

Definitions 1.10. A formation rule is a string manipulation that takes a family of formulae $(\phi_i)_{i \in I}$ and a free variable assignment $\mathsf{FV}(\phi_i)$ for each i and produces a string ψ and a collection of *free* variables $\mathsf{FV}(\psi)$.

The *first order formation rules* are

- **Relation**: For a relation symbol $R \rightarrowtail A_1 \times A_2 \times \cdots \times A_n$ and Σ-terms $t_1 : A_1, t_2 : A_2, \dots, t_n : A_n$ we have a formula $R(t_1, t_2, \dots, t_n)$, where we set $\mathsf{FV}(R(t_1, t_2, \dots, t_n))$ to be the collection $\bigcup_{i=1}^n \mathsf{FV}(t_i)$ of all variables occurring in one of the terms t_i.
- **Equality**: For terms t_1, t_2 of the same sort we have a formula $t_1 = t_2$ with $\mathsf{FV}(t_1 = t_2) := \mathsf{FV}(t_1) \cup \mathsf{FV}(t_2)$.
- **Truth**: There is a formula $\top$ with no free variables ($\mathsf{FV}(\top) := \emptyset$).
- **Binary conjunction**: For Σ-formulae ϕ, ψ there is a formula $\phi \wedge \psi$ where we set $\mathsf{FV}(\phi \wedge \psi) := \mathsf{FV}(\phi) \cup \mathsf{FV}(\psi)$.
- **Existential quantification**: For a formula ϕ and $x : A$, we have a formula $\exists(x : A).\phi$ where $\mathsf{FV}(\exists(x : A)\phi) := \mathsf{FV}(\phi) \setminus \{x\}$.
- **Falsehood**: There is a formula $\bot$ with no free variables.
- **Binary disjunction**: For formulae ϕ, ψ we have a formula $\phi \vee \psi$ where we set $\mathsf{FV}(\phi \vee \psi) := \mathsf{FV}(\phi) \cup \mathsf{FV}(\psi)$.
- **Implication**: For formulae ϕ, ψ we have a formula $\phi \Rightarrow \psi$ where we define $\mathsf{FV}(\phi \Rightarrow \psi) := \mathsf{FV}(\phi) \cup \mathsf{FV}(\psi)$.
- **Negation**: For a formula ϕ, there is a formula $\neg\phi$ with the same free variables.
- **Universal Quantification**: For a formula ϕ and a variable $x : A$, there is a formula $\forall(x : A).\phi$ with $\mathsf{FV}(\forall(x : A).\phi) := \mathsf{FV}(\phi) \setminus \{x\}$

The class $\Sigma-\mathsf{Form}$ of *first order formulae* over Σ is the smallest class of strings, whose symbols are either in Σ or are the auxiliary ones $\exists, \forall, \dots,$ that is closed under the first order formation rules.

A first order *statement* is a first order formula ϕ such that $\mathsf{FV}(\phi) = \emptyset$.

A *first order theory* over a signature Σ is a collection of first order statements over Σ.

Remarks 1.11. Our motivating examples are all descriptions by first order theories. In fact, the standard algebraic examples of groups, rings, etc. can all be modeled by first order languages over a single sort.

As a degenerate case, a signature with no sorts simply reduces to propositional logic (intuitively, there are no variables to quantify over).

2. Internal language

2.1. **Motivation.** A topos behaves like a "universe of sets". In particular, the objects of a topos should behave like sets. A fundamental operation one can perform on sets is that, given a set A and a first order formula ϕ in a single free variable x one can form a subset $\{x \in A \mid \phi(x)\}$.

Remark 2.1. This is in fact one of the two *axiom schema*[3] of ZFC, the axiom schema of specification[4].

One needs to restrict to forming subsets of a given set, in order to avoid allowing for "too large sets" such as the one constructed in Russel's Paradox (cf. [Rus96])

The corresponding notion for a subset in a topos is that of a subobject. Thus, we might hope to be able to determine subobjects of an object X in a topos by means of formulae ϕ with unique free variables $x : X$. In particular, we wish to determine subobjects of X in a manner analogous to $\{x \in X \mid \phi(x)\}$.

Remark 2.2. The main sources for this section are [MacMoe94, §VI.5] and [Car14, Section 5]

2.2. **The internal language of a Heyting Category.** The Mitchell-Bénabou language of a topos is a first order language in the sense of the previous section. However, it also comes with an "interpretation", a way of making sense of the notion of a topos "satisfying" certain formulae.

Definition 2.3. If $\mathcal{C}$ is a category with finite products, the *canonical signature*[5] is the first order signature $\Sigma_{\mathcal{C}}$ which has

- a sort A for every object A of $\mathcal{C}$.
- a function symbol $\mathsf{f}\colon \mathsf{A}_1 \times \mathsf{A}_2 \times \cdots \times \mathsf{A}_n \to \mathsf{B}$ for every arrow $A_1 \times A_2 \times \cdots \times A_n \xrightarrow{f} B$ in $\mathcal{C}$.
- a relation symbol $\mathsf{R} \rightarrowtail \mathsf{A}_1 \times \mathsf{A}_2 \times \cdots \times \mathsf{A}_n$ for every subobject R of $A_1 \times A_2 \times \cdots \times A_n$ in $\mathcal{C}$.
- Variables for each sort.

There is a general technique for "interpreting" first order terms over a signature Σ in a category $\mathcal{C}$ with finite products. The necessary data is that of a Σ-structure[6].

Definition 2.4. For a first order signature Σ and a category $\mathcal{C}$ with finite products, the data of a Σ-*structure* M *in* $\mathcal{C}$ is that of

- an object MA of $\mathcal{C}$ for every sort $A \in \Sigma{-}\mathsf{Sort}$.
- an arrow $Mf\colon MA_1 \times MA_2 \times \cdots \times MA_n \to MB$ in $\mathcal{C}$ for every function symbol $f\colon A_1 \times A_2 \times \cdots \times A_n \to B \in \Sigma{-}\mathsf{Fun}$.
- a subobject $MR \rightarrowtail M_1 \times MA_2 \times \cdots \times MA_n$ in $\mathcal{C}$ for every relation symbol $R \rightarrowtail A_1 \times A_2 \times \cdots \times A_n \in \Sigma{-}\mathsf{Rel}$.

Remark 2.5. The Σ-structures in a category $\mathcal{C}$ with finite products organise into a category (cf. [Car14]).

---

[3] A family of axioms parameterized by first order formulae

[4] Also called the axiom schema of separation or restricted comprehension

[5] Also called the *internal language*.

[6] An extension of the Tarskian notion of structure one encounters in classical model theory.

Observation 2.6. *The canonical signature of a category $\mathcal{C}$ with finite products has a natural structure in $\mathcal{C}$, called the tautological structure (denoted $\mathcal{S}_{\mathcal{C}}$). It simply runs the definition of the canonical signature in reverse, assigning to each sort, function symbol and relation symbol the corresponding object, arrow and subobject respectively.*

The upshot now is that given a Σ-structure in a $\mathcal{C}$ with finite products, we can make sense of the first order terms in $\Sigma-\mathsf{Term}$. It will be convenient to work with a slightly more general notion, that of a term (resp. formula) "in context"

Definitions 2.7. Given a signature, a *context* is a finite list of distinct variables.

A term (resp. formula) is said to be *interpretable* in a context $\bar{x}$ if its free variables are all contained in $\bar{x}$.

Any term (resp. formula) has a *canonical* context that it is interpretable in, given by its free variables (in order of occurrence, say).

A term (resp. formula) *in context* is the data of a context $\bar{x}$ and a term (resp. formula) ϕ that it is interpretable in. We will denote it $\phi[\bar{x}]$ (or $\{\bar{x}.\phi\}$).

Construction 2.8. Consider a Σ-structure in a category $\mathcal{C}$ with finite products. Then we define for each Σ-term $t : B$ in a context $\bar{x} = x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n$, an arrow

$$Mt[\bar{x}] := Mt\colon MA_1 \times MA_2 \times \cdots \times MA_n \to MB$$

If t is a variable $x : A$, x is necessarily x_i for some i and we set $Mt := \pi_i$, the projection onto the i'th component. We then define recursively

$$M(f(t_1, t_2, \ldots, t_m)[\bar{x}]) := MA_1 \times MA_2 \times \cdots \times MA_n \xrightarrow{(Mt_j[\bar{x}])} MC_1 \times MC_2 \times \cdots \times MC_m \xrightarrow{Mf} B$$

for a term made by substituting in a function symbol $f\colon C_1 \times C_2 \times \cdots \times C_m \to B$ terms $t_1 : A_1, t_2 : A_2, \ldots, t_m : A_m$ in context $\bar{x}$.

Notation 2.9. From this point on, if we have a first order signature Σ, a Σ-structure M in a category $\mathcal{C}$ with finite products, and a context $\bar{x} = x_1 : A_1, x_2 : A_2, \ldots, x_n : A_n$ we will set

$$M\bar{x} := A_1 \times A_2 \times \cdots \times A_n$$

We have so far been able to interpret terms assuming only that the category $\mathcal{C}$ has finite products. To start interpreting formulae, we will require more categorical structure.

Definition 2.10. A *Heyting category* is a category $\mathcal{C}$ with finite limits, such that

- Every morphism has an image (epi-mono factorization) and these are stable under pullback.
- The poset of subobjects[7] $\mathsf{Sub}_{\mathcal{C}}(X)$ of each object X has finite joins (least upper bounds/ finite coproducts).
- The pullback functor $f^{\bullet}\colon \mathsf{Sub}_{\mathcal{C}}(b) \to \mathsf{Sub}_{\mathcal{C}}(a)$ associated to any arrow $f\colon a \to b$ has a right adjoint $\forall_f$.

Construction 2.11. The pullback functor $f^{\bullet}\colon \mathsf{Sub}_{\mathcal{C}}(b) \to \mathsf{Sub}_{\mathcal{C}}(a)$ associated to any arrow $f\colon a \to b$ in a Heyting category also has a left adjoint $\exists_f$. This is constructed by simply sending a subobject represented by an arrow a to the image of $f \circ a$.

Observation 2.12. *If $\mathcal{C}$ is a Heyting category, the meet operation on $\mathsf{Sub}_{\mathcal{C}}(E)$ given by $A \wedge B := A \times_E B$.*

When $\mathcal{C}$ has coproducts, the join operation is such that $A \vee B$ is the image of the map $A \coprod B \to E$.

[7]We will assume that the subobjects form, at most some kind of "large" set. That is, we will assume our foundations to be such that this is a genuine poset, in the sense of being a set with a partial order.

As the joins and meets are respectively, coproducts and products in the subobject poset, the pullback maps preserve both as they admit both left and right adjoints.

Example 2.13. We have already seen in the course of the seminar so far that topoi satisfy the heyting category axioms.

For an example of a Heyting category that isn't a topos[8], consider the full subcategory of the category of sets comprising the at most countable sets. This is a Heyting category as images and posets of subobjects can be computed as in the category of sets. However, it cannot be cartesian closed as a consequence of Cantor's diagonal argument.

Proposition 2.14. *(cf.* [Joh02, A1.4.13]*) If $\mathcal{C}$ is a Heyting category, there is a binary operation $\Rightarrow$ on each $\mathsf{Sub}_{\mathcal{C}}(X)$ such that for subobjects A and B, $A \Rightarrow B$ is the largest subobject C such that $C \wedge A \leq B$. This operation further commutes with pullbacks . Thus, each $\mathsf{Sub}_{\mathcal{C}}(E)$ is a Heyting algebra[9], and furthermore the pullback maps are algebra homomorphisms.*

Remark 2.15. In [Car14], one finds also a discussion of "fragments" of first order logic, which produce subsets of formulae that can be interpreted in more general classes of categories. However, we will concern ourselves with full finitary first order logic alone, so that Heyting categories will suffice.

Construction 2.16. Given a Σ-structure M in a Heyting category $\mathcal{C}$, we will construct for each formula ϕ in context $\bar{x} = x_1 : A_1, x_2 : A_2, \dots, x_n : A_n$ a subobject $M\phi[\bar{x}] \rightarrowtail M\bar{x}$ by structural induction.

- If $\phi = R(t_1, t_2, \dots, t_m)$ is obtained by the **relation** formation rule applied to a relation symbol $R \rightarrowtail B_1 \times B_2 \times \cdots \times B_m$ and terms $t_1 : B_1, t_2 : B_2, \dots, t_m : B_m$, one defines the subobject $M\phi[\bar{x}]$ as the pullback

$$\begin{tikzcd} M\phi(\bar{x}) \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & MR \arrow[d, tail] \\ M\bar{x} \arrow[r, "(Mt_i(\bar{x}))"] & MB_1 \times MB_2 \times \cdots \times MB_m \end{tikzcd}$$

- If ϕ is the formula $s = t$ obtained by the **equality** formation rule we define $M(s = t)[\bar{x}]$ to be the equalizer of the arrows $Ms[\bar{x}], Mt[\bar{x}] \colon M\bar{x} \to MB$, where B is the sort of s (and t).
- $M\top[\bar{x}]$ is the maximal subobject $1_{M\bar{x}}$. $M\bot[\bar{x}]$ is similarly the minimal subobject $\bot \in \mathsf{Sub}_{\mathcal{C}}(M\bar{x})$.
- For formulae obtained by applying the various propositional connectives $\vee, \wedge, \neg, \Rightarrow$ to formulae in context, say $\phi[\bar{x}]$ and $\psi[\bar{x}]$, we simply apply the Heyting algebra operations of $\mathsf{Sub}_{\mathcal{C}}(\bar{x})$ to $M\phi[\bar{x}]$ and $M\psi[\bar{x}]$.
- For a formula ϕ in context $\bar{x}, y : B$, denote by $\pi \colon M\bar{x} \times MB \to M\bar{x}$ the projection. One then sets $M\exists(y : B)\phi[\bar{x}] := \exists_\pi(M\phi[\bar{x}, y : B])$ and similarly $M\forall(y : B)\phi[\bar{x}] := \forall_\pi(M\phi[\bar{x}, y : B])$.

Definition 2.17. We will say that a formula ϕ in context $\bar{x}$ is *satisfied* by a Heyting category $\mathcal{C}$ if $M\phi[\bar{x}]$ is the top element $1_{M\bar{x}}$. In this case we write $\mathcal{C} \vDash \phi[\bar{x}]$.

Observation 2.18. *For a formula ϕ in context $\bar{x}, x : A$, we have $\mathcal{C} \vDash \phi[\bar{x}, x : A]$ if and only if $\mathcal{C} \vDash \forall(x : A).\phi[\bar{x}]$.*

---

[8] The author learned of this example from Andreas Blass.

[9] One sets $\neg A := A \Rightarrow \bot$ as usual

Proof. Assume that $\mathcal{C} \vDash \phi[\bar{x}, x : A]$, i.e. $M\phi[\bar{x}, x : A] = \top$. Then, as the interpretation of $\forall(x : A)\phi[\bar{x}]$ is obtained by applying the right adjoint $\forall_\pi$, this preserves the terminal object $\top$ and hence $\mathcal{C} \vDash \forall(x : A)\phi[\bar{x}]$.

For the converse, assume that $\mathcal{C} \vDash \forall(x : A)\phi[\bar{x}]$, so that $\forall_\pi M\phi[\bar{x}, x : A] = \top$. Then on applying the right adjoint $\pi^\bullet$ of $\exists_\pi$ we see $\pi^\bullet\forall_\pi M\phi[\bar{x}, x : A] = \top$. But in light of the counit map $\pi^\bullet\forall_\pi M\phi[\bar{x}, x : A] \xrightarrow{\epsilon_{M\phi[\bar{x}, x:A]}} M\phi[\bar{x}, x : A]$, we have

$$\top = \pi^\bullet\forall_\pi M\phi[\bar{x}, x : A] \leq M\phi[\bar{x}, x : A]$$

hence $M\phi[\bar{x}, x : A] = \top$, proving the converse. □

In particular, to see satisfiability of first order formulae it suffices to consider the case of first order formulae with no free variables.

3. The Mitchell-Bénabou Language of a topos

3.1. **Motivation.** We have seen that in a Heyting category, one can interpret full first order logic (over say, the canonical signature). A topos has two notable categorical properties that we have not referenced so far – exponential objects[10], and the existence of a subobject classifier.

Remark 3.1. Intuitively, Heyting categories have intrinsically the "first order aspects of first order logic". For instance, exponentials and subobject classifiers are intuitively "second order" objects, as they allow for a notion of parameterizing over families.

In topoi we can in fact make sense of an extension of the internal language, called the *Mitchell-Bénabou* language. It will have, in addition to the terms that one forms with the grammar of the internal language, terms of the form "$f(t) : B$" (where t is a term of type A and f is a term of type B^A), "$\lambda(x : X).\phi$" and new formulae of the form "$x \in E$".

Remark 3.2. The term $f(t)$ intuitively corresponds to the evaluation of such an f at t. Complementing it is the term $\lambda(x : X).\phi$. Here we have a term $\phi : B$ interpretable in context $\bar{x}, x : X$ and produce one interpretable in the context $\bar{x}$. Intuitively this is the assignment "$x \mapsto \phi(x)$".

The formula, $x \in E$, as the notation suggests corresponds to a sort of membership.

3.2. **The grammar of the Mitchell-Bénabou language.** We will provide a formal description of the Mitchell-Bénabou language of topos $\mathcal{E}$. We will restrict attention to the canonical signature of $\mathcal{E}$, and also its tautological structure.

Definition 3.3. The term formation rules of the Mitchell-Bénabou language consist of the first order formation ones, and the following:

- **Application**: For terms $\sigma : B^A$ and $a : A$, we have a term $\sigma(t) : B$, with $\mathsf{FV}(\sigma(t)) := \mathsf{FV}(\sigma) \cup \mathsf{FV}(t)$.
- **Abstraction**: For a term $\phi : B$ and a variable $x : A$, we have a "lambda-term" $\lambda(x : A).\phi : B^A$ with $\mathsf{FV}(\lambda(x : A).\phi) := \mathsf{FV}(\phi) \setminus \{x\}$.
- **Pairing**: For terms $a : A$ and $b : B$ we have a term $(a, b) : A \times B$ with $\mathsf{FV}\,((a, b)) := \mathsf{FV}(a) \cup \mathsf{FV}(b)$.

Warning 3.4. We unfortunately have two ways of generating terms of the form $f(t)$, one from terms $f : B^A$ and one from function symbols $f \colon B \to A$. We will attempt to be explicit about the types of f and t in all instances, and distinguish the two by context.

Definition 3.5. The formula formation rules of the Mitchell-Bénabou language consist of the first order formation rules, and

[10] For instance, the Heyting category of at most countable sets does not even have all exponentials.

- **Membership**: For terms $a : A$ and $E : \mathsf{P}\,A$, we have a formula $a \in E$, with $\mathsf{FV}(a \in E) := \mathsf{FV}(a) \cup \mathsf{FV}(E)$.

Construction 3.6. We extend the canonical interpretation M of the internal language of a topos in itself to the Mitchell-Bénabou language by extending the inductive construction of the previous section. On terms:

- Consider a term $\sigma(t)$ in context $\bar{x}$ obtained by applying the **Application** term formation rule to terms $\sigma[\bar{x}]\colon B^A$ and $t[\bar{x}]\colon A$. Then we define $M\sigma(t)[\bar{x}]$ to be the composite $M\bar{x} \xrightarrow{(M\sigma, Mt)} B^A \times A \xrightarrow{\mathsf{ev}} B$ where ev is the counit of the adjunction $(-) \times A \dashv (-)^A$.
- Consider a term in context $\lambda(x : A).\phi[\bar{x}] : B^A$ obtained by the **Abstraction** term formation rule from a term $\phi[\bar{x}, x : A] : B$, for a variable $x : A$. Then we define the arrow $M\lambda(x : A).\phi[\bar{x}]\colon M\bar{x} \to B^A$ to be the exponential transpose of $M\phi[\bar{x}, x : A]\colon M\bar{x} \times A \to B$.
- Consider a term in context $(a, b)[\bar{x}] : A \times B$ obtained from terms $a : A$ and $b : B$ in context $\bar{x}$ by the **Pairing** rule. We then interpret to be the universal map $M(a, b)[\bar{x}] := (Ma[\bar{x}], Mb[\bar{x}])\colon M\bar{x} \to A \times B$

Remark 3.7. Note that to interpret all the terms of the Mitchell-Bénabou language, it suffices to have finite products and exponentials, that is to work in a cartesian closed category. Cartesian closed categories are in some sense the categories that one can interpret simply typed λ-calculus (indeed, related to the suggestive λ-abstraction rule).

In light of this one can say that the Mitchell-Benabou language of a topos combines both the expressive power of first order logic (which required the structure of a heyting category) and that of simply typed λ-calculus (as topoi are also cartesian closed).

And on formulae, if $(a \in E)[\bar{x}]$ is a formula in context obtained from terms $a[\bar{x}] : A$ and $E[\bar{x}] : \mathsf{P}\,A$ from the **Membership** formula formation rule, we define $M(a \in E)[\bar{x}] \rightarrowtail M\bar{x}$ by the pullback

$$\begin{tikzcd}
M(a \in E)[\bar{x}] \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \text{``}\{(x, T) \mid x \in T\}\text{''} \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & * \arrow[d, tail, "\top"] \\
M\bar{x} \arrow[r, "(Ma[\bar{x}], ME[\bar{x}])"] & A \times \mathsf{P}\,A \arrow[r, "\in_A"] & \Omega
\end{tikzcd}$$

3.3. **An alternate Mitchell-Bénabou language.** We have interpreted a formula ϕ in context $\bar{x}$ as a subobject of $M\bar{x}$. In a topos, we can use the subobject classifier to equivalently interpret $\phi[\bar{x}]$ as an arrow $M\bar{x} \to \Omega$. This is the point of view Mitchell takes in [Mit72].

However, here one still distinguishes formulae and terms, even though a formula and a term of type Ω in context $\bar{x}$ are both interpreted as an arrow $M\bar{x} \to \Omega$. For instance, one cannot construct a formula of the form $(a \in E) = (b \in F)$, as this would require the **Equality** formation rule to be applied to two formulae obtained from the **Membership** formation rule. This is in contrast to the approach of Maclane and Moerdijk in [MacMoe94], where one translates the formula formation rules into term formation rules, and simply defines formulae as those terms of type Ω (with the same free variables).

Precisely, we have

- The **Relation** formation rule now forms out of a relation symbol $R \rightarrowtail B_1 \times B_2 \times \cdots \times B_m$ and terms $t_1 : B_1, t_2 : B_2, \ldots, t_m : B_m$ in context $\bar{x}$ a term $R(t_1, t_2, \ldots, t_m) : \Omega$, interpreted as the composite of $M\bar{x} \xrightarrow{(Mt_i[\bar{x}])} B_1 \times B_2 \times \cdots \times B_m$ with the map $B_1 \times B_2 \times \cdots \times B_m \to \Omega$ the subobject R classifies.

- Instead of an **Equality** formula formation rule, we have for terms s, t of the same sort E a term $s = t$ of type Ω. Given a context $\bar{x}$ that $s = t$ is interpretable in, we set the interpretation of $(s = t)[\bar{x}]$ to be the composite of $M\bar{x} \xrightarrow{(Ms[\bar{x}],Mt[\bar{x}])} E \times E$ with the map $\chi_{\delta_E} \colon E \times E \to \Omega$ classifying the diagonal $\delta_E \colon E \to E \times E$.
- **Truth** and **Falsehood** simply specify terms $\top, \bot$ of type Ω and are interpreted as the corresponding truth values $* \to \Omega$.
- The propositional connectives $\vee, \wedge, \Rightarrow, \neg$ now take terms of type Ω to terms of type Ω and can in fact be seen as being induced by the maps $\neg \colon \Omega \to \Omega$ and $\vee, \wedge, \Rightarrow \colon \Omega \times \Omega \to \Omega$ in the topos.
- The descriptions of Existential and universal quantification are essentially the same, except that one works with classifying maps instead of subobjects. However, one can show that the adjoint functors on the subobject posets are really a consequence of the two "internal adjoints" $\exists_X, \forall_X \colon \Omega^X \to \Omega$.
- The **Membership** rule now produces a term $(a \in E) : \Omega$ from $a : A$ and $E : \Omega^A$, and in a context $\bar{x}$ it is just interpreted as the composite

$$M\bar{x} \xrightarrow{(ME[\bar{x}],Ma[\bar{x}])} \Omega^A \times A \to \Omega$$

Examples 3.8. To end this section, we list some examples of familiar notions being described by the Mitchell-Bénabou language

- An arrow $f \colon X \to Y$ in a topos $\mathcal{E}$ is a monomorphism if and only if

$$\mathcal{E} \vDash \forall(x, x' : X).(f(x) = f(x') \Rightarrow x = x')$$

- A topos $\mathcal{E}$ is Boolean if and only if $\mathcal{E} \vDash \forall(p : \Omega).(p \vee \neg p)$
- A topos $\mathcal{E}$ satisfies the internal axiom of choice of the previous talk if and only if for every pair of objects X, Y of $\mathcal{E}$

$$\mathcal{E} \vDash \forall(f : Y^X).((\forall(y : Y).\exists(x : X).(f(x) = y)) \Rightarrow (\exists(g : X^Y).\forall(y : Y).(f(g(y)) = y)))$$

- One encodes the axiom of dependent choice as

$$\mathcal{E} \vDash \forall(x : X)\exists!(y : Y).\phi(x, y) \Rightarrow \exists(f : Y^X).\forall(x : X).\phi(x, f(x))$$

  for every formula ϕ interpretable in the context $x : X, y : Y$.
- The object of epimorphisms, $\mathsf{Epi}(X, Y) \rightarrowtail Y^X$ that we encountered in previous talks can be described as the subobject that is that interpretation of the formula

$$\forall(y : Y).\exists(x : X).(f(x) = y)$$

  with free variable $f : Y^X$.

Remark 3.9. To see that the above examples do work as we have claimed, one can unwind all the definitions and check directly. However, the technique of working with a semantics (as we will introduce in the next section) provides a versatile tool for showing these kinds of equivalences.

4. Kripke-Joyal semantics

4.1. **Motivation.** Semantics is intuitively the process of turning a formula in some formal language (the "internal" system) into a statement in the "ordinary/naive" language that we work in (the "external" system). For instance, a statement of the form "$x \in \{a \mid \phi(x) \vee \psi(x)\}$ if and only if $\phi(a)$ or $\psi(a)$" would be an application of a semantics (for set theory, say).

In practice, when one proves a statement of the form $\forall(x \in X).\phi$ one usually proves $\phi(a)$ for a "general element" a of X. Classical (Tarskian) Kripke semantics is an extension of this, giving a notion of satisfaction (called forcing, as Kripke was originally inspired by Cohen's forcing) of a formula by a tuple of elements.

In a topos, one must replace elements $\alpha \in X$ by "generalized elements", that is arrows $\alpha\colon U \to X$. We will define for a formula ϕ in context $\bar{x}$, and a generalized element $U \xrightarrow{\alpha} X$ a notion of "α forcing $\phi[\bar{x}]$", denoted $U \Vdash \phi[\alpha/\bar{x}]$ in terms of the external categorical properties of the topos.

As the Yoneda lemma determines an object by its generalized elements, Kripke-Joyal semantics are an excellent tool in analyzing structures defined internal to a topos in terms of the Mitchell-Bénabou language.

4.2. **A semantics for the Mitchell-Bénabou language.** We will work with a topos $\mathcal{E}$ throughout, and our formulas will be in its Mitchell-Bénabou language[11].

Definition 4.1. For a formula ϕ in context $\bar{x}$ and a generalized element $\alpha\colon U \to M\bar{x}$, one says $U \Vdash \phi[\alpha/\bar{x}]$ if $\mathsf{Im}\alpha \le M\phi[\bar{x}]$, that is we have a diagram

$$\begin{tikzcd} & M\phi[\bar{x}] \arrow[r] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & * \arrow[d, tail, "\top"] \\ U \arrow[ur, dashed] \arrow[r, "\alpha"] & M\bar{x} \arrow[r, "{\phi[\bar{x}]}"'] & \Omega \end{tikzcd}$$

Example 4.2. For instance, if ϕ is a formula with no free variables (so that the canonical context is empty) and $* \Vdash \phi$ if and only if ϕ is $\top$.

Observation 4.3. *(Monoticity) Consider a formula in context $\phi[\bar{x}]$ and a pair of composable arrows $U' \xrightarrow{\beta} U \xrightarrow{\alpha} M\bar{x}$ in $\mathcal{E}$. If $U \Vdash \phi[\alpha/\bar{x}]$, then $U' \Vdash \phi[(\alpha \circ \beta)/\bar{x}]$.*

Remark 4.4. Classically, one considers the notion of a poset of states, that may or may not force a formula. One has monoticity within the poset, this is intuitively gaining more information upon increasing in the poset[12].

In a topos, one can thus think of restricting along β as being more specific about the generalized elements we consider.

Observation 4.5. *(Local character) Consider a formula in context $\phi[\bar{x}]$, as well as in $\mathcal{E}$ a generalized element $U \xrightarrow{\alpha} M\bar{x}$ and an epimorphism $U' \twoheadrightarrow U$. If $U' \Vdash \phi[(\alpha \circ \beta)/\bar{x}]$, then $U \Vdash \phi[\alpha/\bar{x}]$.*

Proof. The hypothesis is the data of a diagram

$$\begin{tikzcd} & & M\phi[\bar{x}] \arrow[r, "p_{M\phi[\bar{x}]}"] \arrow[d, tail, "\iota"'] \arrow[dr, phantom, "\lrcorner", very near start] & * \arrow[d, tail, "\top"] \\ U' \arrow[urr, dashed, "u"] \arrow[r, two heads, "\beta"'] & U \arrow[r, "\alpha"'] & M\bar{x} \arrow[r, "{\phi[\bar{x}]}"'] & \Omega \end{tikzcd}$$

Recall that the monomorphism ι is the equalizer of the map $\phi[\bar{x}]\colon M\bar{x} \to \Omega$ and the map $M\bar{x} \xrightarrow{p_{M\bar{x}}} * \xrightarrow{\top} \Omega$. Now

$$\phi[\bar{x}] \circ \alpha \circ \beta = \top \circ p_{M\phi[\bar{x}]} \circ u = \top \circ p_{M\bar{x}} \circ \iota \circ u = \top \circ p_{M\bar{x}} \circ \alpha \circ \beta$$

[11]We will therefore use then the extension of the canonical signature and tautological structure described in the previous section .

[12]Indeed, in temporal logic, one is concerned with formalizing statements that are "time dependent". There are Kripke-semantics for temporal logics, where ascending in the poset will directly correspond to a progression of time.

Consequently as β is epic,

$$\phi[\bar{x}] \circ \alpha = \top \circ p_{M\bar{x}} \circ \alpha$$

and therefore α factors into the equalizer $M\phi[\bar{x}]$, that is $U \Vdash \phi[\alpha/\bar{x}]$. □

Observation 4.6. *In a topos $\mathcal{E}$, and a formula ϕ in context $\bar{x}$ in the Mitchell-Bénabou language, we have $\mathcal{E} \vDash \phi[\bar{x}]$ if and only if for every generalized element $\alpha \colon U \to M\bar{x}$, $U \Vdash \phi[\alpha/\bar{x}]$.*

Remark 4.7. In practice, we will mostly be interested in concluding that a topos satisfies a formula in context by checking that generalized elements force it. In light of image factorization, it will suffice to check that subobjects force the proposition. In other words the conditions of the observation are equivalent to the fact that for every subobject $\alpha \colon U \rightarrowtail M\bar{x}$, $U \Vdash \phi[\alpha/\bar{x}]$.

Classically, the notion of forcing is defined by a structural induction on the formation rules of formulae. For instance it can be shown that for terms $t, s : A$ in context $\bar{x}$, a generalized element $U \xrightarrow{\alpha} M\bar{x}$ forces $(t = s)[\bar{x}]$ if and only the arrow α equalizes the arrows $Mt, Ms \colon M\bar{x} \to MA = A$.

It is in light of this equivalent definition[13] that the above observation gains significance. It gives us a way to demonstrate the satisfaction of certain properties by a topos, which is of use in practice.

References

[Car14] Caramello, Olivia. *Topos-theoretic background.* Available online at `https://www.oliviacaramello.com/Unification/ToposTheoreticPreliminariesOliviaCaramello.pdf`.

[Joh02] Johnstone, Peter. *Sketches of an Elephant: a topos theory compendium*, Volumes 1 and 2. Oxford University Press (2002).

[Joh14] Johnstone, Peter. *Topos Theory*, Dover Publications (2014).

[MacMoe94] Maclane, Saunders. Moerdijk, Ieke. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory.* DOI: `https://doi.org/10.1007/978-1-4612-0927-0`

[Mit72] Mitchell, William. *Boolean topoi and the theory of sets.* DOI: `https://doi.org/10.1016/0022-4049(72)90006-0`

[Rus96] Russell, Bertrand. *The Principles of Mathematics. 2d. ed.* Reprint, New York: W. W. Norton & Company, 1996. (First published in 1903.)

[13] This will be the subject of subsequent talks.