Issues of logic, algebra and topology in ontology

Steven Vickers
School of Computer Science, University of Birmingham,
Birmingham, B15 2TT, UK.
s.j.vickers@cs.bham.ac.uk

January 31, 2008

Abstract

When one uses a particular logical formalism, one makes an *ontological commitment* to being able to interpret the symbols involved. We discuss this in a case study of geometric logic, being aided by a presentation of the logic as a sequent calculus. We also discuss the connections of geometric logic with topology and algebra.

Contents

1 Introduction

In my book "Topology via Logic" [Vic89] I motivated the application of topology to computer science by presenting topologies as *observational* accounts of (e.g.) computer programs, with open sets representing observable properties and the axioms of topology reflecting a logic of observations. This developed ideas of Smyth and Abramsky, and seemed to provide a useful explanation of how topologies were used in denotational semantics.

The logic involved (*geometric* logic) is well known in topos theory, and has a predicate form going somewhat beyond the propositional logic of my book. This chapter is presented as an ontological examination of the logic, developing the observational ideas in my book. (To some extent these were already sketched in [Vic92].)

The title mentions logic, algebra and topology, which together cover vast parts of mathematics, and indeed geometric logic does have deep and subtle connections with all those. It should, however, be clear that a short chapter such as this cannot give a comprehensive survey of the connections between those parts of mathematics and ontology. Instead, we shall focus on geometric logic as a case study for an ontological examination, and briefly mention its connections with those broader fields. The logic brings together topology and algebra in some rather remarkable ways, and has an inherent continuity.

We shall describe in fair detail the presentation of geometric logic as given in [Joh02b], and from this technical point of view there is little new. However, we shall also use that as the basis for a novel ontological discussion, with particular emphasis on the question *What is the ontological commitment of the logic?* The logic in itself avoids certain ontological problems with classical logic, and that is our primary reason for choosing it as a case study. However, we shall also see that the mode of presentation in [Joh02b], using sequents rather than sentences, in itself facilitates the analysis. This is because it makes a clear distinction between *formulae* and *axioms,* and that reflects an ontological distinction between *observations* and *hypotheses.*

Our naive view of ontological questions is that they concern the connection between symbols and the world. Russell [Rus45, Chap. XVIII, Knowledge and Perception in Plato] says that for a symbolic formula to *exist* we must demonstrate an instance of it in the world. If we say "lions exist, but unicorns don't", then to prove our point with regard to lions, we go to the zoo and identify something that we can agree is an instance of "lion". Of course, this presupposes an understanding of what would constitute such an instance, but we think the scenario is a plausible one. Another day we instead go looking for a unicorn and this time our outing is less successful. Despite our seeing a rhinoceros, a goat with a horn missing, a plastic toy unicorn and a royal coat of arms, none of these seems truly satisfactory and we return home still not knowing whether unicorns exist. Nonetheless, we can agree it was still worth trying.

This ontological connection between symbols and world is clearly not in itself part of formal logic. Nonetheless, we shall argue informally how formal features of the logic can make it easier to analyse the informal connection.

Our main thesis has two parts.

First, we shall be arguing that a formal logic (with connectives and rules of inference) carries a certain *ontological commitment* to how it could be interpreted “in the real world”. Classical first-order logic uses various symbols that on the face of it have a straightforward relationship with concepts of everyday life: $\wedge$ (*conjunction*) means “and”, $\vee$ (*disjunction*) means “or”, $\neg$ (*negation*) means “not”, $\rightarrow$ (*implication*) means “implies”, $\forall$ (*universal quantification*) means “for all” and $\exists$ (*existential quantification*) means “for some”.However, we shall argue that the way classical logic deals with these adds up to a very strong ontological commitment that could be problematic in reality. Specifically, negation ($\neg$), implication ($\rightarrow$) and universal quantification ($\forall$) cannot be expected to have a uniform interpretation on the same level as conjunction ($\wedge$) and disjunction ($\vee$). This suggests a need to consider other less standard logics to describe “the real world”.

In fact, even in formal mathematics this need can make itself felt. The strong ontological commitment of classical logic can be sustained in formal mathematics (in particular, in set theory), but only because that ontological commitment is already built in to the way set theory is formalized. In other settings it may cause problems. One example we shall briefly mention later (Section 6.1) is *sheaves over a topological space X*. The geometric logic that we shall describe was invented to be used with sheaves, and indeed for more general contexts known as *toposes*. That background is not needed here, but it means there is a well established mathematical setting in which geometric logic can be interpreted.

The second part of our thesis is that it is fruitful to examine the ontological commitment of geometric logic, and explore how it might be interpreted in “the real world”. We do not claim that it is *the* right logic to use, but it avoids the more immediate problems of classical logic.

2 Ingredients of logic

We first review the ingredients of logic. We shall adopt a sequent approach, specifically that set out in [Joh02b]. Technically, this is all well established (apart from some of the notation); what is new is the ontological discussion. We include enough in this Section to show how the sequent approach facilitates a more careful ontological analysis. In Section 3 we shall discuss in more detail geometric logic and an ontology for it.

A *many-sorted, first-order signature* has a set of sorts, a set of predicate symbols, and a set of function symbols. Each predicate or function symbol has an *arity* stipulating the number and sorts of its arguments, and (for a function) the sort of its result. A predicate symbol with no arguments is *propositional*, while a function with no arguments is a *constant*. We shall express the arities of predicates and functions thus:

$$
\begin{array}{ll}
P \subseteq A_1, ..., A_n & \text{(for a predicate)} \\
P \subseteq 1 & \text{(for a proposition)} \\
f : A_1, ..., A_n \to B & \text{(for a function)} \\
c : B & \text{(for a constant)}
\end{array}
$$

These symbols in the signature are *extra-logical* – outside the logic. They are meaningless until interpreted. Since the nature of the interpretation will be very important in our ontological discussion, we shall introduce a non-standard notation that makes the interpretation quite explicit. Suppose we have an interpretation that we call M. Mathematically, M must interpret each sort A as a *set*, the *carrier* for A, which we shall write as $\{M|A\}$.

Note that we do *not* presume that $\{M|A\}$ has any elements. The normal account of classical logic requires each carrier to be non-empty, but this is actually a big ontological commitment.

A function symbol $f : A_1, ..., A_n \to B$ is used to construct terms of sort B by applying f to n arguments of sorts $A_1, \ldots, A_n$. In M, therefore, the interpretation of f should tell us how, if we are given arguments in the form of values $a_i \in \{M|A_i\}$, there is then a corresponding result $f(a_1, \ldots, a_n) \in \{M|B\}$. Hence f is interpreted as a function from the cartesian product $\prod_{i=1}^{n}\{M|A_i\}$ to $\{M|B\}$. To simplify notation we shall write $\{M|A_1, \ldots, A_n\}$ for that cartesian product, and to simplify it further we shall often use vector notation $\{M|\vec{A}\}$. Then the vector $(a_1, \ldots, a_n) = \vec{a} \in \{M|\vec{A}\}$ and the interpretation of f is as a function

$$\{M|f\} : \{M|\vec{A}\} \to \{M|B\}.$$

A constant $c : B$ is a function in the special case of having no arguments and so is interpreted as an element

$$\{M|c\} \in \{M|B\}.$$

For a predicate $P \subseteq A_1, ..., A_n$, the interpretation needs to say for which argument tuples $\vec{a} \in \vec{A}$ the predicate $P(a_1, \ldots, a_n)$ is true. Hence it is equivalent to specifying a subset of the set of all tuples (of the right sorts):

$$\{M|P\} \subseteq \{M|\vec{A}\}.$$

Just as with constants, a proposition $P \subseteq 1$ is a predicate in the special case of having no arguments. This is, as one would expect, interpreted as a truth value. However, we can also see this as a special case of predicates with arguments. If the vector $\vec{A}$ is empty – its length is zero – then for $\{M|\vec{A}\}$ we are looking for the "cartesian product of no sets". The most natural interpretation of this is the 1-element set whose only element is the empty (zero length) vector ε (say). A subset $\{M|P\} \subseteq \{\varepsilon\}$ is determined solely by the truth value of $\varepsilon \in \{M|P\}$. If the truth value is **true** then $\{M|P\} = \{\varepsilon\}$, while if the truth value is **false** then $\{M|P\} = \emptyset$. Any one-element set will do for this purpose, which is why we write $P \subseteq 1$ to say that P is a propositional symbol.

Once the signature is given, *terms* can be built up in the usual way. A term will usually contain variables, and if $\vec{x}$ is a list of distinct variables x_i, each with a stipulated sort $\sigma(x_i)$, then we say that a term t is *in context* $\vec{x}$ if all its variables are amongst the x_is. We also say that $(\vec{x}.t)$ is a term in context.

- Each variable x is a term of sort $\sigma(x)$.
- Suppose $f : A_1, ..., A_n \to B$ is a function symbol in the signature, and for each i $(1 \leq i \leq n)$, t_i is a term in context $\vec{x}$ of sort A_i. Then $f(t_1, \ldots t_n)$(or $f(\vec{t})$) is a term in context $\vec{x}$ of sort B.

If an interpretation M is given for the signature, then it extends to all terms. Consider a term in context $(\vec{x}.t)$ of sort $\sigma(t)$. If values $\vec{a} \in \{M|\sigma(\vec{x})\}$ are given, then they can be substituted for the variables $\vec{x}$ and then the whole expression can be evaluated in an obvious way to get an element of $\{M|\sigma(t)\}$. Thus the term in context is interpreted as a function

$$\{M|\vec{x}.t\} : \{M|\sigma(\vec{x})\} \to \{M|\sigma(t)\}.$$

More systematically, we can say how to evaluate this using the structure of t.

The simplest case is when t is just one of the variables, say x_i. Then the function is the projection function

$$\{M|\vec{x}.x_i\}(\vec{a}) = a_i.$$

Note something important here. The context has variables that are not used in the term, but they still influence the way the term is interpreted. We cannot define the interpretation of the term x_i without also knowing what context $\vec{x}$ it is taken to be in.

Now suppose we have a term of the form $f(\vec{t})$ where each t_i is a term in context $\vec{x}$. Once we have calculated $\{M|\vec{x}.t_i\}(\vec{a})$ for these subterms, then we can say

$$\{M|\vec{x}.f(\vec{t})\}(\vec{a}) = \{M|f\}(\{M|\vec{x}.t_1\}(\vec{a}), \ldots, \{M|\vec{x}.t_n\}(\vec{a})).$$

In the special case where the context $\vec{x}$ is empty (so the term is *closed* – it has no variables), we use exactly the same procedure but bearing in mind that the vector $\vec{a}$ is empty (ε). This gives us a function from $\{\varepsilon\}$ to $\{M|\sigma(t)\}$, which is equivalent to picking out an element $\{M|\varepsilon.f(\vec{t})\} \in \{M|\sigma(f(\vec{t}))\}$.

Next we look at formulae in context. We start by describing all the ways that formulae can be constructed in classical first-order predicate logic. However, we shall later retreat from this in the face of problems that are essentially *ontological* in nature, problems of what kind of interpretations are envisaged. The standard account does not meet these problems. The reason is that it takes its interpretation in the formal set theory of mathematics, and that formal theory already presupposes the ontological commitment of classical logic.

- $\top$ (true) and $\bot$ (false) are formulae in any context.

- Suppose $P \subseteq A_1, ..., A_n$ is a predicate in the signature, and for each i ($1 \leq i \leq n$), t_i is a term in context $\vec{x}$ of sort A_i. Then $P(t_1, \ldots t_n)$ is a formula in context $\vec{x}$.

- If s and t are terms in context $\vec{x}$, and their sorts $\sigma(s)$ and $\sigma(t)$ are the same, then $s = t$ is a formula in context $\vec{x}$. (We make equality an explicit part of the logic, rather than relying on its introduction as a predicate in the signature.)

- If ϕ and ψ are formulae in context $\vec{x}$, then so too are $\phi \wedge \psi$ (ϕ *and* ψ), $\phi \vee \psi$ (*or*), $\phi \rightarrow \psi$ (*implies*) and others to taste.

- If ϕ is a formula in context $\vec{x}$, then so is $\neg\phi$.

- If ϕ is a formula in context $\vec{x}y$, then $(\exists y)\phi$ and $(\forall y)\phi$ are formulae in context $\vec{x}$.

Note that it is the *free* variables of a formula that appear in its context – the bound variables do not. However, it is possible for a context to include variables that are not used in the formula. For example, $\top$ and $\perp$ have no free variables but can be considered in any context.

Just as with terms, the interpretation $\{M|\vec{x}.\phi\}$ of a formula in context depends on the context, not just the formula. $\{M|\vec{x}.\phi\}$ will be a subset of $\{M|\sigma(\vec{x})\} = \prod_{i=1}^{n}\{M|\sigma(x_i)\}$. This allows us to discuss the interpretation of formulae in a more discerning way than if we just took the free variables of a formula to be its context (which is what the standard account in effect does).

The interpretation of formulae in context can now be defined from their structure. Here are the rules. We take $\vec{a} \in \{M|\sigma(\vec{x})\}$, in other words $\vec{a}$ is a list with each a_i in $\{M|\sigma(x_i)\}$.

$\{M|\vec{x}.\top\} = \{M|\sigma(\vec{x})\}$, $\{M|\vec{x}.\perp\} = \emptyset$. (Note how the interpretation depends on the context.)

$\vec{a} \in \{M|\vec{x}.s = t\}$ if $\{M|\vec{x}.s\}(\vec{a}) = \{M|\vec{x}.t\}(\vec{a})$ as elements of $\{M|\sigma(s)\}$.

$\vec{a} \in \{M|\vec{x}.\phi \wedge \psi\}$ if $\vec{a} \in \{M|\vec{x}.\phi\}$ and $\vec{a} \in \{M|\vec{x}.\psi\}$.

$\vec{a} \in \{M|\vec{x}.\phi \vee \psi\}$ if $\vec{a} \in \{M|\vec{x}.\phi\}$ or $\vec{a} \in \{M|\vec{x}.\psi\}$.

$\vec{a} \in \{M|\vec{x}.\neg\phi\}$ if $\vec{a} \notin \{M|\vec{x}.\phi\}$.

$\vec{a} \in \{M|\vec{x}.\phi \rightarrow \psi\}$ if $\vec{a} \notin \{M|\vec{x}.\phi\}$ or $\vec{a} \in \{M|\vec{x}.\psi\}$. (This conforms with the logical equivalence $(\phi \rightarrow \psi) \equiv (\neg\phi \vee \psi)$.)

$\vec{a} \in \{M|\vec{x}.(\exists y)\phi\}$ if there is some $b \in \{M|\sigma(y)\}$ such that $\vec{a}b \in \{M|\vec{x}y.\phi\}$.

$\vec{a} \in \{M|\vec{x}.(\forall y)\phi\}$ if for every $b \in \{M|\sigma(y)\}$ we have $\vec{a}b \in \{M|\vec{x}y.\phi\}$.

In essence, this is the Tarskian definition of semantics: $\wedge$ is "and", $\vee$ is "or", etc.

2.1 Interpretations and ontology

If ontology is the discussion of *being*, or *existence*, then our position is that interpretations are the basis of this discussion. It is the interpretation that provides the instances of formulae. As we have stated it, these instances are elements of *sets*, and at first we understand those as mathematical constructs in formal set theory. However, for any kind of philosophical or applicational discussion we shall want to be able to conceive of M as the "real-world interpretation" of the signature, with each $\{M|\vec{x}.\phi\}$ a collection of real-world things. Though this connection is informal, we shall later look at how this ambition might affect the formal logic.

The use of particular connectives represents an *ontological commitment* that those connectives should have meaning in the setting where we find our interpretations. To see how this could be a problem, let us examine negation. In the example of "there is a lion", we went to the zoo, saw a lion, and believed. But what about its negation, "there is no lion"? How do we ascertain the truth of that? Certainly it is not enough to visit the zoo and fail to see a lion. Maybe there are lions at the zoo, but they all happen to be asleep in a private part of the cage, or we looked in the sealion pool by mistake. Or maybe there are no lions at the zoo, but there are some on the African savannah. We know how to recognize lions, and we know how to ascertain their existence by seeing one. But that does not tell us at all how to ascertain their non-existence. In other words, there is no uniform ontological account of negation.

Implication is even worse than negation, since negation is a special case of it – $\neg\phi$ is equivalent to $\phi \to \bot$.

Similarly, there is no uniform ontological account of universal quantification. We might know how to recognize brownness in lions, but that would not tell us how to ascertain the truth of "all lions are brown".

We shall admit only those formulae that use the connectives to which we are prepared to make the ontological commitment in the interpretations we are considering. For those connectives, we shall take it that the rules given above for determining $\{M|\vec{x}.\phi\}$ still make sense, so $\{M|\vec{x}.\phi\}$ is well defined as a "set" in what ever interpretational sense it is that we have in mind.

2.2 Theories and models

If Σ is a signature, it is usual to define a *theory* over Σ to be a set of *sentences* over Σ, where a sentence is a formula in the empty context. However, we have now envisaged making an ontological restriction to the admissible formulae, and that may rule out implication and negation. It is hardly possible to conduct logic without them, since they lie at the heart of the notion of logical deduction. We shall give a slightly different definition of "theory" that allows for this. This *sequent* form of logic is well known. We shall follow closely the presentation in [Joh02b].

Definition 1 *A* sequent *over Σ is an expression $\phi \vdash^{\vec{x}} \psi$ where ϕ and ψ are formulae (with whatever connectives we are using) in context $\vec{x}$.*

This can be read as meaning the sentence $(\forall x_1 \cdots \forall x_n)(\phi \to \psi)$, but in logics without $\to$ and $\forall$ this will not be a formula.

Definition 2 *A* theory *over* Σ *is a set* T *of sequents over* Σ*, called the* axioms *of* T*. An interpretation* M satisfies *the sequent* $\phi \vdash^{\vec{x}} \psi$ *if* $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\}$*, and it is a* model *of a theory* T *if it satisfies every axiom in* T*.*

As part of the logic, we shall need to say not only what are the admissible connectives but also what are the *rules of inference*. Each will be presented in the form of a schema

$$\frac{\alpha_1 \ \cdots \ \alpha_n}{\beta}$$

where each α_i (a *premiss*) and β (the *conclusion*) is a sequent. We shall not list rules yet, but typical would be the *cut* rule

$$\frac{\phi \vdash^{\vec{x}} \psi \ \psi \vdash^{\vec{x}} \chi}{\phi \vdash^{\vec{x}} \chi}$$

The *soundness* of a rule is then that if an interpretation satisfies all the premisses it must also satisfy the conclusion. This would normally have to be justified in terms of the ontological explanation of the connectives. For the cut rule it would usually be plain that if $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\}$ and $\{M|\vec{x}.\psi\} \subseteq \{M|\vec{x}.\chi\}$ then $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\chi\}$.

Using the rules of inference, one can infer, or *derive*, many more sequents from the axioms of a theory. If the rules are all sound, then a model of a theory also satisfies all the sequents derived from the axioms.

Note that the sequent formulation (and in particular the explicit context on the turnstile) makes it easier to deal correctly with empty carriers. As an example, consider the two valid entailments $(\forall y)\phi \vdash^{x} \phi[x/y]$ and $\phi[x/y] \vdash^{x} (\exists y)\phi$. Applying the cut rule to these we obtain $(\forall y)\phi \vdash^{x} (\exists y)\phi$. Even if $\sigma(x)$ has an empty carrier this is valid, since then $\{M|x\}$ is the empty set and so are both $\{M|x.(\forall y)\phi\}$ and $\{M|x.(\exists y)\phi\}$. However, the rules do not allow us to deduce $(\forall y)\phi \vdash (\exists y)\phi$ (with empty context), and it would not be valid with the empty carrier because we would have $\{M|\varepsilon.(\forall y)\phi\} = 1$ but $\{M|\varepsilon.(\exists y)\phi\} = \emptyset$.

Unlike the case with formulae, with axioms we make no ontological commitment to being able to ascertain that an interpretation satisfies even a single sequent, let alone a possibly infinite set of them in a theory. There is thus a definite ontological distinction between formulae and sequents. We should understand theories as being like *scientific hypotheses* or *background assumptions*. In fact there is a Popperian flavour to theories.

Suppose we have a theory T and an interpretation M. Suppose also we find some elements in M and ascertain for them some properties from the signature of T. This amounts to finding an element of $\{M|\vec{x}.\phi\}$ for some formula in context $(\vec{x}.\phi)$. The ontological commitment is that we know what is required for our claim to have found such elements. Now suppose also that from the axioms of T we can, using the inference rules, logically deduce the sequent $\phi \vdash^{\vec{x}} \bot$. It should follow that $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\bot\}$. But this is nonsense, since $\{M|\vec{x}.\bot\}$ is by definition empty but we have found an element of $\{M|\vec{x}.\phi\}$. Thus the interpretation M *cannot possibly be a model of* T. If M is a "real world" interpretation, then we cannot simply reject it. Possibly we made a mistake in the way we interpreted ϕ. But if not, then we were mistaken in thinking the axioms of T would apply in the real world. Our observations have led to a Popperian Big No to our theory T.

Note that this process can be carried through only if a sequent $\phi \vdash^{\vec{x}} \bot$ can be derived from the theory T, with ϕ not logically equivalent to $\bot$. In other words, T must be *falsifiable*. In the example of geometric logic, to which we turn next, this can happen only if T has explicit axioms of the form $\phi \vdash^{\vec{x}} \bot$.

3 Geometric logic

We now turn to *geometric logic*, a positive logic that rejects negation (and also implication and universal quantification) in its formulae. Its ontological commitment is to conjunction, disjunction, equality and existential quantification. Note that we are not claiming it as *the* absolute irreducible logic. We just say that, because of problems with negation and universal quantification, geometric logic is more likely to be applicable in "real world interpretations". It still carries ontological commitments of its own. For example, consider conjunction. We said that, if ϕ and ψ are formulae in context $\vec{x}$, then to ascertain that $\vec{a}$ is in $\phi \wedge \psi$ we have to ascertain that $\vec{a}$ is in ϕ and $\vec{a}$ is in ψ. This makes assumptions about our ability to form tuples of things. The logic also presupposes that the two tasks, ascertaining that $\vec{a}$ is in ϕ and ascertaining that $\vec{a}$ is in ψ, do not interfere with each other and can be done in either order. Also, our use of equality means that we expect to be able to ascertain equality between things, but not necessarily inequality. This says something about the kind of things we are prepared to talk about.

The connectives of geometric logic are $\top$, $\bot$, $\wedge$, $\vee$, $=$ and $\exists$. However, note one peculiarity: we allow *infinitary* disjunctions $\bigvee$. If S is a set of formulae, then $\bigvee S$ is also a formula, the disjunction of all the elements of S, and $\{M|\vec{x}.\bigvee S\}$ is defined in the obvious way. This does lead to subtle ontological questions of its own, since we should examine the nature of the set S and how its members are found. Is S also intended to be a "real-world" collection? If it is purely a mathematical construct, what mathematics are we using? Once we start thinking about different logics, it raises the question of what logic to use for mathematics itself. We shall largely ignore these questions here, except (Section 5) to say some brief hints about a fascinating connection between the infinitary disjunctions and algebra.

We shall first present the formal logic, still following [Joh02b], and then (Section 3.4) discuss its ontology.

3.1 Rules of inference

The rules of inference for geometric logic as given here are taken from [Joh02b]. The first group are *propositional*, in the sense that they have no essential interaction with the terms or variables. The propositional rules are *identity*

$$\phi \vdash^{\vec{x}} \phi,$$

cut

$$\frac{\phi \vdash^{\vec{x}} \psi \qquad \psi \vdash^{\vec{x}} \chi}{\phi \vdash^{\vec{x}} \chi},$$

the *conjunction* rules

$$\phi \vdash^{\vec{x}} \top, \quad \phi \wedge \psi \vdash^{\vec{x}} \phi, \quad \phi \wedge \psi \vdash^{\vec{x}} \psi, \quad \frac{\phi \vdash^{\vec{x}} \psi \quad \phi \vdash^{\vec{x}} \chi}{\phi \vdash^{\vec{x}} \psi \wedge \chi},$$

the *disjunction* rules

$$\phi \vdash^{\vec{x}} \bigvee S \quad (\phi \in S), \quad \frac{\phi \vdash^{\vec{x}} \psi \quad (\text{all } \phi \in S)}{\bigvee S \vdash^{\vec{x}} \psi}$$

and *frame distributivity*

$$\phi \wedge \bigvee S \vdash^{\vec{x}} \bigvee \{\phi \wedge \psi \mid \psi \in S\}.$$

Note that $\bigvee \emptyset$ plays the role of $\bot$ (false). To find an element of $\{M|\vec{x}.\bigvee \emptyset\}$ we must find a formula ϕ in $\emptyset$ and then find an element of $\{M|\vec{x}.\phi\}$. But clearly there can be no such ϕ, so $\{M|\vec{x}.\bigvee \emptyset\}$ is empty. From the general disjunction rules we can then derive the rule of *ex falso quodlibet,*

$$\bot \vdash^{\vec{x}} \psi$$

for any ψ in context $\vec{x}$.

Next come the rules specific to predicate logic. These involve terms and variables.

For the first rule, *substitution*, we use the following notation. Suppose ϕ is a formula in context $\vec{x}$, and $\vec{s}$ is a vector of terms in another context $\vec{y}$ such that the vector $\vec{s}$ has the same length and sorts as $\vec{x}$ – we can write $\sigma(\vec{s}) = \sigma(\vec{x})$. Then $\phi[\vec{s}/\vec{x}]$ is ϕ with $\vec{s}$ *substituted* for $\vec{x}$ – the variables in $\vec{x}$ are all replaced by the corresponding terms in $\vec{s}$. Some notes:

- Since the terms s_i may have their own free variables, taken from $\vec{y}$, $\phi[\vec{s}/\vec{x}]$ is in the context $\vec{y}$ instead of $\vec{x}$.
- There is no particular problem if $\vec{x}$ and $\vec{y}$ have variables in common. For example, suppose ϕ is the formula (in context x) $g(x) = a$ and s is the term $f(x)$ where $f : \sigma(x) \to \sigma(x)$. We can substitute $f(x)$ for x,

$$(g(x) = a)[f(x)/x] \equiv (g(f(x)) = a).$$

- There can be a problem of "capture of variables" if one of the context variables in $\vec{y}$ is also used as a bound (quantified) variable in ϕ. To avoid this, the bound variables should be renamed to be distinct from the context variables.

The *substitution* rule is

$$\frac{\phi \vdash^{\vec{x}} \psi}{\phi[\vec{s}/\vec{x}] \vdash^{\vec{y}} \psi[\vec{s}/\vec{x}]}$$

The next rules are: the *equality* rules

$$\top \vdash^{x} x = x, \qquad (\vec{x} = \vec{y}) \wedge \phi \vdash^{\vec{z}} \phi[\vec{y}/\vec{x}]$$

In the second $\vec{z}$ has to include all the variables in $\vec{x}$ and $\vec{y}$, as well as those free in ϕ, and the variables in $\vec{x}$ have to be distinct. Our substitution $\phi[\vec{y}/\vec{x}]$ is not quite in accordance with the definition, since $\vec{x}$ is not the whole of the context. However, we can easily replace it by a licit substitution $\phi[\vec{t}/\vec{z}]$ where $\vec{t}$ is defined as follows. If z_i is x_j for some j, then t_i is defined to be y_j. Otherwise, t_i is defined to be z_i.

The substitution rule justifies *context weakening*

$$\frac{\phi \vdash^{\vec{x}} \psi}{\phi \vdash^{\vec{x},y} \psi}.$$

In other words, a deduction in one context will still be valid if we add extra variables, though not if we remove unused variables (which is what would be done for a deduction of $(\forall x)\ \phi(x) \vdash (\exists x)\ \phi(x)$). Note that ϕ here (and ψ likewise) is in two separate contexts: $\vec{x}y$ and $\vec{x}$. We shall consider it given as in context $\vec{x}$. Then since $\vec{x}$ can be considered to be a vector of terms in context $\vec{x}y$, we can get ϕ in the extended context as $(\vec{x}y.\phi[\vec{x}/\vec{x}])$.

The *existential* rules are

$$\frac{\phi \vdash^{\vec{x},y} \psi}{(\exists y)\phi \vdash^{\vec{x}} \psi}, \qquad \frac{(\exists y)\phi \vdash^{\vec{x}} \psi}{\phi \vdash^{\vec{x},y} \psi}.$$

The *Frobenius* rule is

$$\phi \wedge (\exists y)\psi \vdash^{\vec{x}} (\exists y)(\phi \wedge \psi).$$

3.2 Soundness

In a mathematical semantics, the soundness of most of the rules can be readily justified from the semantics of connectives given above. For example, for the final conjunctive rule one has that if $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\}$ and $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\chi\}$ then

$$\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\} \cap \{M|\vec{x}.\chi\} = \{M|\vec{x}.\psi \wedge \chi\}$$

from the definition of $\{M|\vec{x}.\psi \wedge \chi\}$. (In more general semantics we shall see how the rules have subtle consequences for the ontological commitment.)

However, where substitution is involved we have to be more careful. The semantics of a formula is defined in terms of how that formula is structured using the connectives. When a formula is described using a substitution, as in

$\phi[\vec{s}/\vec{x}]$, that substitution is not part of the connective structure of the formula and so we do not have a direct *definition* of the "semantics of substitution". It is nevertheless possible to describe the semantic effect of substitution, but it has to be proved as a *Substitution Lemma*. The Substitution Lemma in effect analyses how substitution interacts with the different connectives.

Recall that each term in context $(\vec{y}.s_i)$ gets interpreted as a function $\{M|\vec{y}.s_i\} : \{M|\sigma(\vec{y})\} \to \{M|\sigma(s_i)\}$. Putting these together, we get

$$\{M|\vec{y}.\vec{s}\} : \{M|\sigma(\vec{y})\} \to \prod_i \{M|\sigma(s_i)\} = \{M|\sigma(\vec{s})\} = \{M|\sigma(\vec{x})\}$$

defined by $\{M|\vec{y}.\vec{s}\}(\vec{a}) = (\{M|\vec{y}.s_1\}(\vec{a}), \ldots, \{M|\vec{y}.s_n\}(\vec{a}))$.

Lemma 3 (Substitution Lemma) *Let $\vec{x}$ and $\vec{y}$ be contexts, and let $\vec{s}$ be a vector of terms in context $\vec{y}$ with $\sigma(\vec{s}) = \sigma(\vec{x})$.*

1. *If $\vec{t}$ is a vector of terms in context $\vec{x}$ then $\{M|\vec{y}.\vec{t}[\vec{s}/\vec{x}]\}$ is the composite*
$$\{M|\vec{x}.\vec{t}\} \circ \{M|\vec{y}.\vec{s}\} : \{M|\sigma(\vec{y})\} \to \{M|\sigma(\vec{s})\} = \{M|\sigma(\vec{x})\} \to \{M|\sigma(\vec{t})\}.$$
2. *If ϕ is a formula in context $\vec{x}$, then $\{M|\vec{y}.\phi[\vec{s}/\vec{x}]\}$ is the inverse image under $\{M|\vec{y}.\vec{s}\}$ of $\{M|\vec{x}.\phi\}$, in other words if $\vec{a} \in \{M|\sigma(\vec{y})\}$ then*
$$\vec{a} \in \{M|\vec{y}.\phi[\vec{s}/\vec{x}]\} \text{ iff } \{M|\vec{y}.\vec{s}\}(\vec{a}) \in \{M|\vec{y}.\phi\}.$$

Proof. Induction on the structure of t or ϕ. ■

This Lemma is needed for the soundness of the substitution and equality rules. As an illustration of how it is used, consider the substitution rule

$$\frac{\phi \vdash^{\vec{x}} \psi}{\phi[\vec{s}/\vec{x}] \vdash^{\vec{y}} \psi[\vec{s}/\vec{x}]}$$

If $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\}$ then

$$\begin{aligned}\vec{a} \in \{M|\vec{y}.\phi[\vec{s}/\vec{x}]\} &\Leftrightarrow \{M|\vec{y}.\vec{s}\}(\vec{a}) \in \{M|\vec{y}.\phi\} \\ &\Rightarrow \{M|\vec{y}.\vec{s}\}(\vec{a}) \in \{M|\vec{y}.\psi\} \Leftrightarrow \vec{a} \in \{M|\vec{y}.\phi[\vec{s}/\vec{x}]\}.\end{aligned}$$

One of the more interesting rules here is the second equality rule,

$$(\vec{x} = \vec{y}) \wedge \phi \vdash^{\vec{z}} \phi[\vec{y}/\vec{x}].$$

Recall that $\vec{x}$ here is a sequence of distinct variables from the context $\vec{z}$, and $\vec{y}$ is a sequence of variables from $\vec{z}$, not necessarily distinct, that is sort-compatible with $\vec{x}$. Actually, we might as well assume that $\vec{x}$ is the whole of $\vec{z}$, since by reflexivity we can add extra equations, for the variables of $\vec{z}$ that are not in $\vec{x}$, to say that they are equal to themselves. We are therefore justifying $(\vec{x} = \vec{y}) \wedge \phi \vdash^{\vec{x}} \phi[\vec{y}/\vec{x}]$ where each y_i is a variable $x_{\alpha(i)}$, say. Now an element $\vec{a} \in \{M|\vec{x}.(\vec{x} = \vec{y}) \wedge \phi\}$ is an element $\vec{a} \in \{M|\vec{x}.\phi\}$ such that for each possible index i of the sequence $\vec{x}$, we have $a_i = a_{\alpha(i)}$. Now consider $\{M|\vec{x}.\phi[\vec{y}/\vec{x}]\}$. Since each y_i is a term in context $\vec{x}$ we have a substitution function $\{M|\vec{x}.\vec{y}\} : \{M|\sigma(\vec{x})\} \to \{M|\sigma(\vec{x})\}$ mapping $\vec{b}$ to $\vec{c}$, defined by $c_i = b_{\alpha(i)}$. By the Substitution Lemma, we have $\vec{b} \in \{M|\vec{x}.\phi[\vec{y}/\vec{x}]\}$ if $\vec{c} \in \{M|\vec{x}.\phi\}$. Now given our $\vec{a}$ as above, we can take $\vec{b} = \vec{c} = \vec{a}$ and the required conditions are satisfied. Hence $\vec{a} \in \{M|\vec{x}.\phi[\vec{y}/\vec{x}]\}$.

3.3 Beyond rules of inference

Each inference rule operates within a single signature, and this imposes a limit on what can be expressed with them. There are more subtle intensions regarding the way different signatures relate to each other. Our main example of this for the moment is the property for mathematical sets that a function is equivalent to a total, single-valued relation – its graph. To express this in logical terms, suppose $(\vec{x}y.\Gamma)$ is a formula in context that is total and single-valued. In other words, it satisfies the properties

$$\top \vdash^{\vec{x}} (\exists y)\Gamma$$

$$\Gamma \wedge \Gamma[y'/y] \vdash^{\vec{x}yy'} y = y'.$$

Then in any model there is a unique function $f : \sigma(\vec{x}) \to \sigma(y)$ such that Γ holds iff $y = f(\vec{x})$. This principle is not a consequence of the rules of geometric logic. Indeed, there are mathematical systems (geometric categories [Joh02a] that happen not to be toposes) in which the rules are all sound, but the principle does not hold. Nonetheless, the principle does hold in those systems (toposes) in which geometric logic was first identified, and we take it to be an implicit part of geometric logic. In other words, geometric logic is not just logic (connectives and inference rules). We shall not try to give a complete account of these non-logical principles, though we shall meet some more later.

These principles carry their own ontological commitments. In the above example, the interpretation of a function symbol must be the same an that of a total, single-valued predicate.

3.4 Geometric ontology

We now examine, as carefully as we can, the ontological commitments implicit in geometric logic.

The ontological commitment of the connectives as such does not seem deep. Their interpretation as given above is more or less that of Tarski: $\wedge$ is "and", $\vee$ is "or", $\exists$ is "there exists", etc. But note that the logic does expect something of the "sets" used as carriers. Clearly we must know something about how to find elements of them – how to *apprehend* elements, to use the word of [Vic92]. To form cartesian products $\{M|\vec{A}\}$, we must also know how to form tuples of elements. This is perhaps not so obvious as it seems. How do you apprehend a tuple of lions? Is it just a bunch of lions? But that would not allow a tuple with the same lion in more than one component (e.g. ⟨Elsa, Lenny, Elsa, Parsley⟩), which is certainly allowed by the logic. (Otherwise the equality relation is empty.) So clearly the components of the tuple are more like pointers, "that lion over there, Elsa". And is it properly understood how the interpretation works with observations made at different times? Next, because $=$ (though not $\neq$) is built in to the logic, we must know something about how to ascertain equality between a pair of apprehended elements.

Let us suppose – in some interpretation M – we know how to apprehend elements and ascertain equality for each sort. (The discussion is not quite finished yet, because we need to examine what properties these ingredients have. We shall return to it later.) Let us suppose we also know how to form tuples. Equality between tuples will be ascertained componentwise. This will then tell us about the sets $\{M|\vec{A}\}$ for each sort tuple $\vec{A}$. For a predicate $P \subseteq \vec{A}$, the interpretation $\{M|P\} \subseteq \{M|\vec{A}\}$ must tell us what it takes to ascertain $P(\vec{a})$ for $\vec{a} \in \{M|\vec{A}\}$. This then lifts to formulae in context $(\vec{x}.\phi)$.

Note that there may be different ways of ascertaining $\phi(\vec{a})$ for the same $\vec{a}$, hence different manifestations of the same element of $\{M|\vec{x}.\phi\}$. What is important is that equality between them is determined by equality for the underlying $\vec{a}$. An illuminating example is when ϕ is of the form $(\exists y)\psi$. To ascertain that $\vec{a}$ is in $\{M|\vec{x}.\phi\}$, one must actually apprehend an element $\vec{a}b$ of $\{M|\vec{x}y.\psi\}$. Hence apprehending an element of $\{M|\vec{x}.\phi\}$ is exactly the same as apprehending an element of $\{M|\vec{x}y.\psi\}$. But ascertaining equality between them is different, since in the former case the y component is ignored.

Now there is a rather fundamental question about the meaning of a sequent $\phi \vdash^{\vec{x}} \psi$. We have already explained it as meaning $\{M|\vec{x}.\phi\} \subseteq \{M|\vec{x}.\psi\}$. But what does this mean in terms of apprehension? Suppose an element $\vec{a}$ is apprehended in $\{M|\vec{x}.\phi\}$. What does it mean to say it is also in $\{M|\vec{x}.\psi\}$? To put it another way, is it possible to apprehend some $\vec{b}$ in $\{M|\vec{x}.\psi\}$ such that $\vec{b}$ and $\vec{a}$ are equal as elements of $\{M|\sigma(\vec{x})\}$? Three possible interpretations spring to mind.

1. "Already done": Whatever it took to apprehend $\vec{a}$ as an element of $\{M|\vec{x}.\phi\}$, that is already enough to apprehend a suitable $\vec{b}$.
2. "Nearly done": A well defined program of extra work will yield a suitable $\vec{b}$ given $\vec{a}$.
3. "Can be done": There is some suitable $\vec{b}$, though we don't necessarily know how to find it.

The "already done" interpretation would be extremely strong, since it means that validity of sequents follows directly from knowing how formulae are interpreted. This is clearly incompatible with the idea mentioned above that theory axioms represent background assumptions, or scientific hypotheses.

The "nearly done" interpretation is less strong, since some ingenuity might be required to find the "well defined program of extra work". In fact, this interpretation is roughly speaking the standard one for intuitionistic logic. There one thinks of the elements of $\{M|\vec{x}.\phi\}$ as the *proofs* of ϕ. A proof of $(\forall\vec{x})(\phi \to \psi)$ (and so of the sequent $\phi \vdash^{\vec{x}} \psi$) is an algorithm that takes a tuple $\vec{a}$ and a proof of $\phi(\vec{a})$ (in other words, an element of $\{M|\vec{x}.\phi\}$ for some M) and returns a proof of $\psi(\vec{a})$. Nonetheless, it is hard to see this as compatible with the idea of axioms as scientific hypotheses.

We shall follow the "can be done" interpretation.

Note that this makes the cut rule,

$$\frac{\phi \vdash^{\vec{x}} \psi \qquad \psi \vdash^{\vec{x}} \chi}{\phi \vdash^{\vec{x}} \chi},$$

more subtle than it looks. Suppose we believe the sequents $\phi \vdash^{\vec{x}} \psi$ and $\psi \vdash^{\vec{x}} \chi$ for an interpretation M, and we want to justify $\phi \vdash^{\vec{x}} \chi$. Suppose we have $\vec{a}$ in $\{M|\vec{x}.\phi\}$. The first sequent tells us that there is, somewhere out there waiting to be found, a $\vec{b}$ in $\{M|\vec{x}.\psi\}$ equal to $\vec{a}$ as elements of $\{M|\vec{x}.\psi\}$. However, it does not tell us how to find it. The second sequent tells us that when we do find it, we can then believe there is a $\vec{c}$ in $\{M|\vec{x}.\chi\}$ equal to $\vec{b}$. The cut rule asserts that we do not have to go to the trouble of finding $\vec{b}$. Our belief that it is there, and one day might be found, is already enough to justify us in believing in $\vec{c}$. Hence we justify the sequent $\phi \vdash^{\vec{x}} \chi$.

We can put this another way. Our explanation of the "can be done" interpretation of a sequent $\phi \vdash^{\vec{x}} \psi$, was that if *we have* an element of $\{M|\vec{x}.\phi\}$, then *there is* (out there somewhere) an equal element of $\{M|\vec{x}.\psi\}$. The cut rule uses the idea that we can equivalently weaken on the left hand side, and start from *there is* an element of $\{M|\vec{x}.\phi\}$.

For geometric logic, "can be done" governs how we interpret function symbols. Recall that a function $f : \vec{A} \to B$ is expected to be logically equivalent to its graph, a predicate $\Gamma_f \subseteq \vec{A}, B$ (or, more generally, a formula in context) that is total and single-valued:

$$\top \vdash^{\vec{x}} (\exists y)\Gamma_f(\vec{x}, y)$$

$$\Gamma_f(\vec{x}, y) \wedge \Gamma_f(\vec{x}, y') \vdash^{\vec{x},y,y'} y = y'.$$

These sequents too are given a "can be done" interpretation. Think of the graph Γ_f as being a *specification* of the function. Given arguments $\vec{a}$ and a candidate result b, Γ_f provides a way for ascertaining whether b is indeed the result for $f(\vec{a})$, but it does not in any way tell us how to find b. (That would in fact be a "nearly done" interpretation.) The totality axiom tells us (or hypothesizes) that there is such a b waiting to be found, and single-valuedness says that any two such bs are equal. It follows that when we talk about a "function" between "real world sets", we must not in general expect this to be a method or algorithm.

This style of interpretation can actually be internalized in the logic by eliminating function symbols in favour of predicates for their graphs (together with axioms for totality and single-valuedness). Suppose we have a graph predicate Γ_f for each function symbol f, characterized by

$$\Gamma_f(\vec{x}, y) \dashv\vdash^{\vec{x}y} y = f(\vec{x}).$$

Then we can define a graph formula in context $(\vec{x}\vec{y}.\Gamma_{\vec{t}})$ for each term tuple in context $(\vec{x}.\vec{t})$, where $\sigma(\vec{y}) = \sigma(\vec{t})$. For a single term t, if t is a variable x_i then Γ_t is just the formula $y = x_i$. If t is $f(\vec{s})$, suppose we have defined $\Gamma_{\vec{s}}$ in context $\vec{x}\vec{z}$. Then we can define $\Gamma_{f(\vec{s})}$ in context $\vec{x}y$ as $(\exists\vec{z})(\Gamma_{\vec{s}} \wedge \Gamma_f(\vec{z}, y))$. Once that is done, formulae can be replaced by alternatives without function symbols. For example, $P(\vec{t})$ can be replaced by $(\exists\vec{y})(\Gamma_{\vec{t}} \wedge P(\vec{y}))$.

If we look at the inference rules, we find that some of them are obvious, but some have hidden subtleties. We have already mention the cut rule.

The next interesting ones are the conjunction rules. Examining conjunction itself in more detail, to apprehend an element of $\{M|\vec{x}.\phi \wedge \psi\}$, we must apprehend elements $\vec{a}$ and $\vec{b}$ of $\{M|\vec{x}.\phi\}$ and $\{M|\vec{x}.\psi\}$ and then ascertain that they are equal as elements of $\{M|\sigma(\vec{x})\}$. Now consider the rule

$$\frac{\phi \vdash^{\vec{x}} \psi \quad \phi \vdash^{\vec{x}} \chi}{\phi \vdash^{\vec{x}} \psi \wedge \chi}.$$

For the conclusion, suppose we have apprehended $\vec{a}$ in $\{M|\vec{x}.\phi\}$. The premiss sequents tell us that there are $\vec{b}$ and $\vec{c}$ in $\{M|\vec{x}.\psi\}$ and $\{M|\vec{x}.\chi\}$ such that $\vec{b}$ and $\vec{a}$ are equal in $\{M|\sigma(\vec{x})\}$, and so are $\vec{c}$ and $\vec{a}$. To deduce that there is an element of $\{M|\vec{x}.\psi\wedge\chi\}$, clearly we need to make assumptions about "ascertaining equality" – it needs to be symmetric and transitive. In fact the equality rule $\top \vdash^{x} x = x$ will need it to be reflexive too, so it must be an equivalence relation.

For the rules involving substitution, we need to consider the Substitution Lemma. This is most conveniently understood in terms of the logical style explained above, in which function symbols are replaced by predicates for their graphs. The Substitution Lemma (or at least, part (ii) of it) then says the following. Suppose ϕ is in context $\vec{x}$, and $\vec{s}$ in context $\vec{y}$ is sort compatible with $\vec{x}$. Then finding an element $\vec{a} \in \{M|\vec{y}.\phi[\vec{s}/\vec{x}]\}$ is equivalent (in the "can be done" sense) to finding elements $\vec{a}\vec{b} \in \{M|\vec{y}\vec{x}.\Gamma_{\vec{s}}\}$ and $\vec{c} \in \{M|\vec{x}.\phi\}$ such that $\vec{b}$ is equal to $\vec{c}$.

4 Topology

The links between geometric logic and topology arise from a very direct correspondence: the disjunctions and finite conjunctions in the logic correspond to the unions and finite intersection that characterize the behaviour of open sets. There is a then a rough correspondence between propositional geometric theories and topological spaces: the space is the space of models for the theory, topologized using the logical formulae.

Using the theories instead of topological spaces is generally known as "point-free topology", and has been found useful in various fields, especially in constructive mathematics (e.g. as "locales" [Joh82], in topos theory, and as "formal topologies" [Sam87] in predicative type theory). The applications in computer science, based on ideas of observational theory, could even be read as suggesting that topology in some sense arises from an ontological shift in the understanding of propositions.

A major idea in topos theory is to generalize this correspondence to predicate theories, leading to Grothendieck's new notion of *topos* as "generalized topological space". The theory then corresponds to its "classifying topos", representing (in an indirect way) the "space of models". These ideas are implicit in the standard texts on toposes, such as [MLM92], [Joh02a], [Joh02b], though often hidden. [Vic07] attempts to bring them out more explicitly.

In the space available here it has only been possible to hint at the deep connections between geometric logic and topology, but the curious reader is encouraged to explore the references suggested.

5 Algebra

We now turn to a feature of geometric logic that makes essential use of the infinitary disjunctions, and sets it quite apart from finitary logics. The effect is that geometric logic can be considered to embrace a variety of set-theoretic constructions on sorts, and we shall examine the ontological aspects of this.

5.1 Lists and finite sets

In any geometric theory T, suppose A is a sort. Consider now an extended theory that also has a sort B, together with function symbols

$$\begin{aligned} &\varepsilon : B \\ &\gamma : A \times B \to B. \end{aligned}$$

We shall in fact use infix notation for γ, writing $x \circ y$ for $\gamma(x, y)$. We also add axioms

$$\begin{aligned} x \circ y = \varepsilon &\vdash^{xy} \bot \\ x \circ y = x' \circ y' &\vdash^{xyx'y'} x = x' \wedge y = y' \\ \top &\vdash^{y} \bigvee_{n \in \mathbb{N}} (\exists x_1) \cdots (\exists x_n)(y = x_1 \circ \ldots \circ x_n \circ \varepsilon). \end{aligned}$$

Here $\mathbb{N}$ denotes the set $\{0, 1, 2, 3, \ldots\}$ of natural numbers, so the right hand side of the last axiom is

$$y = \varepsilon \vee (\exists x_1) y = x_1 \circ \varepsilon \vee (\exists x_1)(\exists x_2) y = x_1 \circ (x_2 \circ \varepsilon) \vee \cdots .$$

In any model M of this extended theory, each list $(a_1, \ldots, a_n)$ of elements of $\{M|A\}$ gives an element $a_1 \circ \cdots \circ a_n \circ \varepsilon$ of $\{M|B\}$. The third axiom says that *any* element of $\{M|B\}$ can be got this way, and the first two axioms say that the list is unique – if

$$a_1 \circ \cdots \circ a_m \circ \varepsilon = a'_1 \circ \cdots \circ a'_n \circ \varepsilon$$

then $m = n$ and each $a_i = a'_i$. It follows that $\{M|B\}$ is isomorphic with the set of finite lists of elements of $\{M|A\}$, which we write $\{M|A\}^*$.

This ability to characterize list sets (up to isomorphism) by logic relies essentially on the infinitary disjunctions in geometric logic. It cannot be done in finitary logic. It means that in effect geometric logic embraces sort constructors. Instead of adding all the axioms explicitly, we could allow ourselves to write a derived sort A^*, with the interpretation $\{M|A^*\} = \{M|A\}^*$.

Moreover, this fits with our previous ontology. To apprehend an element of $\{M|A^*\}$, we should apprehend a tuple $\vec{a}$ of elements of $\{M|A\}$. The tuple can have any finite length. To ascertain that $\vec{a}$ and $\vec{a}'$ are equal, we should find that they have the same length and then that each component of $\vec{a}$ is equal to the corresponding component of $\vec{a}'$.

In a similar way, we can use geometric logic to characterize the *finite power set*, $\mathcal{F}A$. We use the same symbols ε and $\circ$, but now ε is to mean the empty set $\emptyset$ and $a \circ b$ means $\{a\} \cup b$. Hence $a_1 \circ \cdots \circ a_n \circ \varepsilon$ means $\{a_1, \ldots a_n\}$. We keep the third axiom, but we replace the first two so as to give a different definition of equality. For this we take axioms

$$x_1 \circ \cdots \circ x_m \circ \varepsilon = x'_1 \circ \cdots \circ x'_n \circ \varepsilon \dashv\vdash^{\vec{x}\vec{x}'} \bigwedge_{i=1}^{m} \bigvee_{j=1}^{n} x_i = x'_j \wedge \bigwedge_{j=1}^{n} \bigvee_{i=1}^{m} x'_j = x_i$$

for all possible m, n. This in effect says $\{a_1, \ldots a_m\} = \{a'_1, \ldots a'_n\}$ iff $\{a_1, \ldots a_m\} \subseteq \{a'_1, \ldots a'_n\}$ (i.e. every a_i is equal to at least one of the a'_js) and $\{a'_1, \ldots a'_n\} \subseteq \{a_1, \ldots a_m\}$.

Again, this fits our ontology. To apprehend an element of $\{M|\mathcal{F}A\}$, we should apprehend a tuple $\vec{a}$ of elements of $\{M|A\}$, just as we did for $\{M|A^*\}$. However, this time the equality is different.

5.2 Free algebras

The list sets and finite power sets are both examples of a much more general construction, of *free algebras*. These arise from a particular kind of geometric theory, namely *algebraic* theories. An algebraic theory is defined by *operators* and *equational laws*, and in terms of geometric theories as defined above this means there are no predicates, and the axioms are all of the form

$$\top \vdash^{\vec{x}} s = t.$$

The models are then often called *algebras*.

Many examples are widely known, for example the theories of groups, rings, vector spaces and Boolean algebras.

The fact that algebraic theories are geometric is interesting, but not very deep. A much more significant fact about geometric theories emerges when one considers *free algebras*, and this is something that relies on very specific properties of geometric logic, and in particular its use of infinitary disjunctions.

Let T be an algebraic theory, with only one sort, A. (Similar results hold for theories with more than one sort, but they are more complicated to state.) A *free* algebra, on a set X, is constructed in two stages. First, we consider all the terms that can be formed, in the empty context, using the operators of T, and also using the elements of X as constants. Next, we define two terms s and t to be *congruent* if the sequent $\top \vdash s = t$ can be inferred (using the inference rules of geometric logic) from the axioms of T. The set of congruence classes is an algebra for T, and is called the *free* T-algebra on X, denoted $T\langle X\rangle$. It can be proved to have a characteristic property that is actually rather fundamental: given any T-algebra A, then any function $f : X \to A$ extends uniquely to a T-homomorphism from $T\langle X\rangle$, got by evaluating the terms (representing elements of $T\langle X\rangle$) in A.

List sets and finite powersets are both examples of free algebras.

The *logical* significance of these constructions is that in geometric theories, geometric structure and axioms can be used to constrain the carrier for one sort to be isomorphic to a free algebra over the carrier of another. (This is not possible in *finitary* first order predicate logic.) Hence geometric logic may be understood as including an inherent *type theory*, with constructions that can be applied to sorts. The list sets and finite powersets were the first examples. This again is something that goes beyond the strict logic, analogous to issues discussed in Section 3.3.

On the other hand, there is also an *ontological* significance. Construction of terms can be understood as a process of apprehending elements (by gathering together other elements in a structured way), and then finding a proof of congruence is ascertaining equality between elements. Thus we may see a "real-world" significance in the free algebra constructions, as typified by list sets and finite powersets.

6 Categories

We have described a connection between, on the one hand, the formal structure of formulae, constructed using formal symbols such as $\wedge$ and $\vee$, and, on the other, the informal ideas of how we might interpret those formulae in the real world. At first sight the interpretation is straightforward, once we have assigned meaning to the primitive symbols of the signature. After that one might think it is just a matter of interpreting $\wedge$ as 'and', $\vee$ as 'or', and so on. However, we saw that particular connectives could easily be problematic. Having particular connectives and particular logical rules about their use imposes an ontological commitment on our interpretations.

This comparison between the logic and the real world may seem unavoidably vague, because of the transformation from formal to informal. However, it actually has two separate transformations bundled up together: one is from formal to informal, but the other is from a logical formalism of terms and formulae to an explanation that is more about collections and functions. There is a way to separate these out using *category theory*. A category is a mathematical structure whose 'objects' and 'morphisms' may embody intuitive ideas of collections and functions between them. In the formal world, we can transform from the logical style to the collections style, by interpreting the logic inside a category. This is known as *categorical logic*. The ontological commitment can now be discussed in a precise mathematical way in terms of the properties of the category. The transformation from formal to informal still has to be made, but the formal category structure may now be closer in kind to the informal structure of the real world that we want to capture.

For example, one of the assumptions we made, in Section 3.4, of real world objects was that it is possible to yoke them together in pairs or longer tuples. This corresponds directly to the categorical idea of *product*. If X and Y are two collections, then there should be another collection $X \times Y$ whose elements are pairs of elements, one from X and one from Y. More generally, if we have two functions $f : Z \to X$ and $g : Z \to Y$, then we should be able to pair their results to get a function $\langle f, g\rangle : Z \to X \times Y$. Category theory uses this idea to characterize $X \times Y$ as the "product" of X and Y, so the informal idea of pairing elements corresponds naturally to the formal idea that the category has products.

Other chapters in this book describe categories in more technical detail. For a fuller description of how geometric logic corresponds to categorical structure, and for pointers to further reading, see [Vic07].

6.1 Sheaves

Sheaves provide a fundamental example of a formal setting where geometric logic can be interpreted, but other parts of ordinary logic – including negation – go wrong. They cannot support the ontological commitment of full classical logic.

We shall not define sheaves in detail here. A good intuition is that if we have a topological space X, then a sheaf over it is a set "continuously parametrized" by a point of x. For each x there is a set A_x (the *stalk* of the sheaf at x), and as x varies, the stalk, the set A_x, varies with it in a continuous way – no sudden jumps. If $a \in A_x$ then there is a neighbourhood U of x such that for each $y \in U$, the stalk A_y has an element corresponding to a. Also, if there are two such ways of choosing "elements corresponding to a", then there is some neighbourhood of x where the two choices agree. That is very vague, but it can be made precise and defines the notion of "local homeomorphism" (see, e.g., [Vic07] again).

Without saying any more about the general notion, we can describe a very simple example where the problems with negation are easy to see. *Sierpinski space* has two points, $\perp$ and $\top$. The topology can be described using the idea of *neighbourhoods*, referred to above. $\{\top\}$ is a neighbourhood of $\top$, but the only neighbourhood of $\perp$ is $\{\perp, \top\}$. When one works out what a sheaf is, it turns out to be a pair of sets $A_\perp$ and $A_\top$ (the stalks), together with a function $f : A_\perp \to A_\top$. The function is needed because, for each $a \in A_\perp$, the definition of sheaf requires a neighbourhood U of $\perp$ (and in this case U can only be $\{\perp, \top\}$), and an element $a_\top \in A_\top$ corresponding to a. The function f shows how to pick $a_\top$ for each a.

Subsheaves are analogous to subsets. A subsheaf of the sheaf A (given by $f : A_\perp \to A_\top$) is a pair of subsets $B_\perp \subseteq A_\perp$ and $B_\top \subseteq A_\top$ such that when f is restricted to $B_\perp$, it maps into $B_\top$:

$$\begin{array}{ccc} B_\perp & \subseteq & A_\perp \\ \downarrow & & \downarrow f \\ B_\top & \subseteq & A_\top \end{array}$$

Now suppose we have another subsheaf, C. We can try to define more subsheaves $B \cup C$ and $B \cap C$ "stalkwise" by

$$(B \cup C)_x = B_x \cup C_x,$$
$$(B \cap C)_x = B_x \cap C_x.$$

for $x = \perp, \top$. But are these subsheaves? The question is whether f restricts properly. In fact, for $\cup$ and $\cap$ it works. This shows that the geometric connectives $\vee$ and $\wedge$ can be interpreted in the expected way.

Now let us look at $\neg$. We try to define a subsheaf $\neg B$ by $(\neg B)_x = A_x - B_x = \{a \in A_x \mid a \notin B_x\}$. Our question about f now amounts to the following. We know that if $a \in B_\perp$ then $f(a) \in B_\top$. Can we deduce that if $a \notin B_\perp$ then $f(a) \notin B_\top$? No, in general. For a simple example, take $B_\perp = \emptyset$, $B_\top = A_\top$.

For an intuitive idea of what is happening here, think of $A_\top$ as "the reality of A", and $A_\perp$ as "what we have seen of it". f translates our observations into real things. However, (i) we may not have seen everything – there may be elements of $A_\top$ that are not f of anything; and (ii) we may have observed two things that are in reality one and the same. Now $\vee$ and $\wedge$ work just as well for our observations as for reality, but $\neg$ doesn't. Failure to observe (calculating $A_\perp - B_\perp$) does not map to non-existence ($A_\top - B_\top$).

References

[Joh82] P.T. Johnstone, *Stone spaces*, Cambridge Studies in Advanced Mathematics, no. 3, Cambridge University Press, 1982.

[Joh02a] ———, *Sketches of an elephant: A topos theory compendium, vol. 1*, Oxford Logic Guides, no. 44, Oxford University Press, 2002.

[Joh02b] ———, *Sketches of an elephant: A topos theory compendium, vol. 2*, Oxford Logic Guides, no. 44, Oxford University Press, 2002.

[MLM92] S. Mac Lane and I. Moerdijk, *Sheaves in geometry and logic*, Springer-Verlag, 1992.

[Rus45] Bertrand Russell, *History of western philosophy*, Simon and Schuster, New York, 1945.

[Sam87] G. Sambin, *Intuitionistic formal spaces – a first communication*, Mathematical Logic and its Applications (Dimiter G. Skordev, ed.), Plenum, 1987, pp. 187–204.

[Vic89] Steven Vickers, *Topology via logic*, Cambridge University Press, 1989.

[Vic92] S.J. Vickers, *Geometric theories and databases*, Applications of Categories in Computer Science (M.P. Fourman, P.T. Johnstone, and A.M. Pitts, eds.), LMS Lecture Note Series, no. 177, Cambridge University Press, 1992, pp. 288–314.

[Vic07] Steven Vickers, *Locales and toposes as spaces*, Handbook of Spatial Logics (Marco Aiello, Ian E. Pratt-Hartmann, and Johan F.A.K. van Benthem, eds.), Springer, 2007, pp. 429–496.