

Functorial semantics and HSP type theorems

Michael Barr*

1999-06-11

Dedicated to the memory of Evelyn Nelson and Alan Day

Abstract

We show that if \mathcal{C} is the category of models for a theory in the sense of Linton over an arbitrary base category, then a full subcategory of \mathcal{C} is closed under Homomorphic images that split in the underlying category, under \mathcal{M} -Subobjects for a class \mathcal{M} of monomorphisms and under Products if and only if it is an intersection of a nest of subcategories, each determined from the preceding by a class of “Horns”, in which the crucial arrow lies in the class \mathcal{E} of epimorphisms orthogonal to \mathcal{M} .

1 Introduction

A celebrated theorem of Garrett Birkhoff states (in modern language) that a subcategory of the category of models of a (finitary) equational theory is defined by equations if and only if it is closed under Homomorphic images, Subobjects and Products (HSP) [Birkhoff, 1935]. The theorem survives unchanged when the finitary restriction is dropped. Recently Vaughan Pratt asked if there was such a theorem for theories based on the category of partially ordered sets (hereafter called posets). The answer in case of posets is developed in [Barr, to appear], which may be read as an introduction to this paper.

There are several differences between the theory developed here and Birkhoff’s. In the first place, the result depends on how you define HSP, in particular, what kind of subobjects are allowed. Second, the idea of satisfying an equation has to be generalized to that of satisfying a certain Horn clause, *using only the predicates available in the base category*. Third, a single set (or even class) of equations may not suffice, since an HSP subcategory may allow new equations that were not even be stateable in the original.

The second point above needs some amplification. The kinds of Horn clauses that are allowed have the form

$$\bigwedge \phi_i(\vec{x}) \Rightarrow \psi(\vec{\tau}(\vec{x}))$$

*This research has been supported by grants from the NSERC of Canada and the FCAR du Québec

where the ϕ_i and ψ are predicates stateable in the base category, \vec{x} is a string of variables of the appropriate type and $\vec{\tau}$ is a string of operations that can apply to a string of type \vec{x} . It is assumed that these implications are properly typed. The only predicates available in the category of sets are membership and equality. A Horn clause built using only those two predicates is the same thing as an equation. An equation specifies membership by its domain and equality of variables by using a diagonal. For example, the Horn clause $x = y \Rightarrow \tau(x, y, z) = z$ is normally rendered in an equational theory by the equation $\tau(x, x, z) = z$. For a theory defined over posets, this would allow the use of Horn clauses using both equalities and inequalities.

The notion of theory we use is essentially that of Linton [1969]: a category that has the same objects as the base and whose arrows include all the arrows of the base and more. The extra arrows correspond to the operations that a model must possess. We prefer to work in the dual of Linton’s theory category, but that is a matter of convenience only.

2 Preliminaries

2.1 Factorization systems. We will be dealing with a *factorization system* on a category. This consists of two classes of arrows, which we will denote \mathcal{E} and \mathcal{M} , each closed under composition and including all the isomorphisms. These have the property that each arrow of the category factors uniquely, up to isomorphism, as an arrow in \mathcal{E} , followed by an arrow of \mathcal{M} . We use a doubled-headed arrow \twoheadrightarrow to indicate that an arrow is in \mathcal{E} and a tailed arrow \rightarrowtail to indicate that an arrow is in \mathcal{M} . *In this paper, we will assume that all arrows in \mathcal{E} are epimorphisms and all arrows in \mathcal{M} are monomorphisms.* This is well known to imply that all regular epimorphisms are in \mathcal{E} and all regular monomorphisms are in \mathcal{M} . It also follows that in any square

$$\begin{array}{ccc} A & \xrightarrow{e} & B \\ f \downarrow & & \downarrow g \\ C & \xrightarrow{m} & D \end{array}$$

with $e \in \mathcal{E}$ and $m \in \mathcal{M}$, there is a unique arrow $h: B \rightarrow C$, called the “diagonal fill-in”, for which $h \circ e = f$ and $m \circ h = g$. One way of stating this condition, which we will use later in the paper is the following. Fix $e \in \mathcal{E}$ and $m \in \mathcal{M}$. Then the condition states that for any $f: A \rightarrow C$ and $g: B \rightarrow D$ such that $m \circ f = g \circ e$, then there is a unique $h: B \rightarrow C$ such that $h \circ e = f$ and $m \circ h = g$. This is precisely the statement

that

$$\begin{array}{ccc}
\mathrm{Hom}(B, C) & \xrightarrow{\mathrm{Hom}(e, C)} & \mathrm{Hom}(A, C) \\
\mathrm{Hom}(B, m) \downarrow & & \downarrow \mathrm{Hom}(A, m) \\
\mathrm{Hom}(B, D) & \xrightarrow{\mathrm{Hom}(e, D)} & \mathrm{Hom}(A, D)
\end{array}$$

is a pullback.

Another important property of factorization systems is that an arrow is in $\mathcal{E} \cap \mathcal{M}$ if and only if it is an isomorphism. One way is assumed and the other follows from the uniqueness, up to isomorphism, of the factorization.

We will denote this factorization system \mathcal{E}/\mathcal{M} .

2.2 Proposition. *Suppose \mathcal{E}/\mathcal{M} is a factorization system. Suppose A is an object and $\{A_i\}$ is a class of \mathcal{M} -subobjects of A whose intersection exists. Then that intersection is also an \mathcal{M} -subobject of A .*

This is relevant to the application of Proposition 5.10. Note that we do not assume that this class is a set, only that it have an intersection.

Proof. Suppose we factor the inclusion $f: \bigcap A_i \rightarrow A$ as $f = m \circ e$ with $m \in \mathcal{M}$ and $e \in \mathcal{E}$. Then we have, for any i , a diagram

$$\begin{array}{ccc}
\bigcap A_i & \xrightarrow{e} & A' \\
\downarrow & & \downarrow m \\
A_j & \xrightarrow{\quad} & A
\end{array}$$

The diagonal fill-in shows that $A' \subseteq A_j$ for each j and hence $A' \subseteq \bigcap A_j$. The opposite inclusion is given by e , which has to be monic, despite being in \mathcal{E} . \square

2.3 Proposition. *Let \mathcal{B} be a complete category. Suppose \mathcal{E}/\mathcal{M} is a factorization system in \mathcal{B} and assume that either \mathcal{B} is \mathcal{M} -well-powered or that an arbitrary class of \mathcal{M} -subobjects of an objects of \mathcal{B} has an intersection. Let \mathbf{T} be a triple on \mathcal{B} and suppose $\mathcal{C} = \mathcal{B}^{\mathbf{T}}$ is the category of Eilenberg-Moore algebras with $U: \mathcal{C} \rightarrow \mathcal{B}$ and $F: \mathcal{B} \rightarrow \mathcal{C}$ the free and underlying functors, respectively. Then there is a unique factorization system $\mathcal{E}^{\mathbf{T}}/\mathcal{M}^{\mathbf{T}}$ on \mathcal{C} for which $\mathcal{M}^{\mathbf{T}} = U^{-1}(\mathcal{M})$. Moreover, $F(\mathcal{E}) \subseteq \mathcal{E}^{\mathbf{T}}$.*

Proof. Let $\mathcal{M}^{\mathbf{T}} = U^{-1}(\mathcal{M})$. We say that the arrow $C_1 \rightarrow C_2$ is in $\mathcal{E}^{\mathbf{T}}$ if it cannot be factored as $C_1 \rightarrow C' \rightarrow C_2$ with $C' \rightarrow C_2$ a non-isomorphism in $U^{-1}(\mathcal{M})$. I claim that any $f: C_1 \rightarrow C_2$ in $\mathcal{E}^{\mathbf{T}}$ is an epimorphism. If not, there would be distinct arrows $g, h: C_2 \rightarrow C_3$ with $g \circ f = h \circ f$. But then the equalizer $k: C_4 \rightarrow C_2$ of g and h would be a proper subobject of C_2 through which f factors. But Uk is a regular

monic and therefore belongs to the \mathcal{M} part of every epi/mono factorization system so that $k \in \mathcal{M}^{\mathbf{T}}$. This contradicts the assumption that f factors through no proper $\mathcal{M}^{\mathbf{T}}$ subobject of C_2 .

Given an arbitrary morphism $f: C_1 \rightarrow C_2$, Let $C_3 \subseteq C_2$ be the intersection of all the $\mathcal{M}^{\mathbf{T}}$ -subalgebras of C_2 . These exist because U creates arbitrary limits and we are assuming that such intersections exist in \mathcal{B} . The preceding proposition implies that the inclusion $C_3 \rightarrow C_2$ is in $\mathcal{M}^{\mathbf{T}}$. It is also clear that $C_1 \rightarrow C_3$ is in $\mathcal{E}^{\mathbf{T}}$ from its definition. This gives the required factorization in \mathcal{C} .

Suppose that $f: B_1 \rightarrow B_2$ belongs to \mathcal{E} . If

$$\begin{array}{ccc} FB_1 & \xrightarrow{Ff} & FB_2 \\ \downarrow & & \downarrow \\ C_1 & \xrightarrow{g} & C_2 \end{array}$$

is a commutative square in \mathcal{C} with $g \in \mathcal{M}^{\mathbf{T}}$, it corresponds by adjointness to a square

$$\begin{array}{ccc} B_1 & \xrightarrow{f} & B_2 \\ \downarrow & & \downarrow \\ UC_1 & \xrightarrow{Ug} & UC_2 \end{array}$$

which gives a diagonal fill-in $B_2 \rightarrow UC_1$ that corresponds via adjointness to an arrow $FB_2 \rightarrow C_1$. This shows that $Ff \in \mathcal{E}^{\mathbf{T}}$. \square

3 \mathcal{B} -based sketches

3.1 Sketches. Let \mathcal{B} be a category. By a \mathcal{B} -based sketch we mean a diagram

$$\begin{array}{ccc} & \mathcal{B}_0 & \\ & \swarrow V & \searrow F \\ \mathcal{B} & & \mathcal{K} \end{array}$$

in which \mathcal{B}_0 is a graph and $V: \mathcal{B}_0 \rightarrow \mathcal{B}$ is the inclusion of a subgraph, \mathcal{K} is a graph with diagrams and $F: \mathcal{B}_0 \rightarrow \mathcal{K}$ is a graph morphism that is an isomorphism on nodes (objects).

One way to think of this is that \mathcal{K} is a graph with diagrams whose nodes are objects of \mathcal{B} and some of whose arrows are identified as arrows of \mathcal{B} . In accordance with this perception, we will often say, “Let \mathcal{K} be a \mathcal{B} -based sketch,” even though all the components are needed.

There are two kinds of sketches that are of special interest. In the first, \mathcal{B}_0 is a small graph, often finite. In the second kind, \mathcal{B}_0 is the graph underlying \mathcal{B} , the morphism V is the identity, \mathcal{K} is the graph with diagrams underlying a category and $F \circ V^{-1}$ is a functor. In that case $F \circ V^{-1}: \mathcal{B} \rightarrow \mathcal{K}$ is a theory in the sense defined below (Section 4) with $\mathcal{T} = \mathcal{K}^{\text{op}}$. Let us call such a sketch a *theory-type* sketch. Since V is the identity we will ignore it and say that $F: \mathcal{B} \rightarrow \mathcal{K}$ is a theory-type sketch.

3.2 Models. Given such a sketch, a **model** in the category of sets is a pair (B, M) where B is an object of \mathcal{B} and $M: \mathcal{K}^{\text{op}} \rightarrow \mathbf{Set}$ is a graph morphism that takes the diagrams of \mathcal{K} to commutative diagrams. Thus M is a model of the graph with diagrams \mathcal{K} . In addition B and M must fit together so that

$$MF = \mathcal{B}(V-, B)$$

A morphism $(f, \phi): (B, M) \rightarrow (B', M')$ consists of a morphism $f: B \rightarrow B'$ and a natural transformation $\phi: M \rightarrow M'$ such that

$$\begin{array}{ccc} MFB_0 & \xrightarrow{\phi FB_0} & M'FB_0 \\ \downarrow = & & \downarrow = \\ \mathcal{B}(VB_0, B) & \xrightarrow{\mathcal{B}(VB_0, f)} & \mathcal{B}(VB_0, B') \end{array}$$

commutes for every object B_0 of \mathcal{B}_0 .

The notation is intended to suggest that a model is an object of \mathcal{B} with structure and that M defines the structure. Let us see that this is the case. Let (B, M) be a model. Every object of \mathcal{K} has the form FB_0 for a unique object B_0 of \mathcal{B}_0 . Thus M is determined on objects by the formula $MFB_0 = \mathcal{B}(VB_0, B)$. The other thing that M provides is an arrow $Mf: \mathcal{B}(VB_0, B) \rightarrow \mathcal{B}(VB'_0, B)$ for each arrow $f: FB_0 \rightarrow FB'_0$ of \mathcal{K} , subject to the condition $Mfg = \mathcal{B}(Vg, B)$ for $g \in \mathcal{B}_0$ and subject also to the commutativity conditions determined by the diagrams of \mathcal{K} .

The idea of this is that a morphism from B_0 to B is a B_0 -type in B . Then a model tells how to give, for each $FB_0 \rightarrow FB'_0$ in \mathcal{K} , a map from B_0 -types in B to B'_0 -types in B and is thus a kind of structure on B .

3.3 Adjoints. Let $\text{Mod}(\mathcal{K})$ denote the category of models as just described. There is an obvious functor $U: \text{Mod}(\mathcal{K}) \rightarrow \mathcal{B}$ such that $U(B, M) = B$ and $U(f, \phi) =$

f . It is easy to see that this functor creates limits. We are interested in knowing conditions under which this functor has a left adjoint. Although it does not in general, the following three theorems give conditions under which the adjoint exists that cover most cases of interest.

3.4 Theorem. *If \mathcal{B} is a locally presentable category and the graph \mathcal{B}_0 is small, then U has a left adjoint.*

Proof. To give the object B the structure of a model requires giving an arrow $VB'_0 \rightarrow B$ for every arrow $FB'_0 \rightarrow FB_0$ in \mathcal{K} and every arrow $VB_0 \rightarrow B$ in \mathcal{B} , subject to certain conditions. Leaving the conditions aside for the moment, this requires an arrow $SB \rightarrow B$ where

$$SB = \sum \mathcal{K}(FB'_0, FB_0) \times \mathcal{B}(VB_0, B) \times VB'_0$$

the sum taken over all pairs (B_0, B'_0) of objects of \mathcal{K} . We use the notation $X \times B$ for a set X and object B of \mathcal{B} to denote the sum of an X -fold of copies of B . The category of objects B equipped with an arrow $SB \rightarrow B$ is denoted $(S : \mathcal{B})$.

At this point, we require,

3.5 Lemma. *The category $(S : \mathcal{B})$ is accessible and the underlying functor to \mathcal{B} is an accessible functor with a left adjoint.*

Proof. Since \mathcal{B}_0 is small, the functor $\widehat{S}(B_0) : \mathcal{B} \rightarrow \mathbf{Set}$ given by

$$\widehat{S}(B_0)B = \mathcal{B}(VB_0, B)$$

is accessible, essentially by definition. The products in the definition of S are symbolic and really represent sums. Since colimits commute with colimits, the functor $S : \mathcal{B} \rightarrow \mathcal{B}$ is also accessible. The category $(S : \mathcal{B})$ is an example of a weighted bilimit and thus the lemma follows from Theorem 5.1.6 of [Makkai & Paré, 1990]. \square

The category of models of \mathcal{K} is a full subcategory of $(S : \mathcal{B})$ determined by conditions of two kinds. The first is that for any object $f : SB \rightarrow B$, the two composites

$$\sum \mathcal{B}_0(B'_0, B_0) \times \mathcal{B}(VB_0, B) \times VB'_0 \xrightarrow{\sum(F \times \text{id} \times \text{id})}$$

$$\sum \mathcal{K}(FB'_0, FB_0) \times \mathcal{B}(VB_0, B) \times VB'_0 = SB \xrightarrow{f} B$$

and

$$\sum \mathcal{B}_0(B'_0, B_0) \times \mathcal{B}(VB_0, B) \times VB'_0 \xrightarrow{\sum(V \times \text{id} \times \text{id})}$$

$$\sum \mathcal{B}(VB'_0, VB_0) \times \mathcal{B}(VB_0, B) \times VB'_0 \xrightarrow{\sum(\text{compose} \times \text{id})}$$

$$\sum \mathcal{B}(VB'_0, B) \times VB'_0 \xrightarrow{\sum(\text{apply})} \sum B \xrightarrow{\nabla} B$$

be equal. Here ∇ is the codiagonal arrow, that is the morphism on the sum that is the identity on each component.

In order to state the second condition we require some notation. Let $h: SB \rightarrow B$ be an object of $(S: \mathcal{B})$, let (B'_0, B''_0) be a pair of objects of \mathcal{B}_0 and let $f: FB'_0 \rightarrow FB''_0$ be an arrow of \mathcal{K} . Then by restricting h to the summand of SB indexed by the pair (B'_0, B''_0) , we get an arrow

$$h(B'_0, B''_0): \mathcal{K}(FB'_0, FB''_0) \times \mathcal{B}(VB''_0, B) \times VB'_0 \rightarrow B$$

Applied to $f \in \mathcal{K}(FB'_0, B''_0)$, we get an arrow

$$h(B'_0, B''_0, f): \mathcal{B}(VB''_0, B) \times VB'_0 \rightarrow B$$

The second condition is that for any diagram

$$\begin{array}{ccccccccc} FB_0 & \xrightarrow{f_0} & FB_1 & \xrightarrow{f_1} & \cdots & \xrightarrow{f_{n-2}} & FB_{n-1} & \xrightarrow{f_{n-1}} & FB_n \\ \downarrow = & & & & & & & & \downarrow = \\ FB_0 & \xrightarrow{g_0} & FC_1 & \xrightarrow{g_1} & \cdots & \xrightarrow{g_{m-2}} & FC_{m-1} & \xrightarrow{g_{m-1}} & FB_n \end{array} \quad (*)$$

in \mathcal{K} , we have

$$h(B_0, B_n, f_{n-1} \circ \cdots \circ f_0) = h(B_0, B_n, g_{m-1} \circ \cdots \circ g_0): \mathcal{B}(VB_n, B) \times VB_0 \rightarrow B$$

Since \mathcal{B}_0 is small, we can define $W: (S: \mathcal{B}) \rightarrow \mathcal{B}$ by

$$W(SB \rightarrow B) = \sum \mathcal{B}_0(B'_0, B_0) \times \mathcal{B}(VB_0, B) \times VB'_0 + \sum \mathcal{B}(VB_n, B) \times VB_0$$

The first sum is taken over all pairs (B_0, B'_0) of objects of \mathcal{B}_0 and the second sum is taken over all the diagrams of the form $(*)$. Then what we have described above are two natural transformations $W \rightarrow U$ such that a T -algebra is a model of \mathcal{K} if and only if the two natural transformations agree on it. It is straightforward to show that W preserves limits and hence by Corollary 4.4 of [Barr, 1991] the full subcategory on which these two transformations agree is reflective in $(S: \mathcal{B})$. \square

3.6 Theorem. *If \mathcal{B} is locally presentable and $F: \mathcal{B} \rightarrow \mathcal{K}$ is a theory-type sketch in which \mathcal{K} is an accessible category and F an accessible functor, then U has a left adjoint.*

Proof. The argument is similar, for in this case even if \mathcal{K} is large, it might as well be small. In fact, choose a cardinal κ large enough that \mathcal{B} and \mathcal{K} are κ -accessible categories and F is a κ -accessible functor. Let \mathcal{B}_0 and \mathcal{K}_0 be the full subcategories

of \mathcal{B} and \mathcal{K} , respectively, consisting of the κ -accessible objects. Then any model of \mathcal{K}_0 has a unique extension to a model of \mathcal{K} . In fact, let (B, M) be such a model. We must provide, for any objects FB'_0 and FB_0 of \mathcal{K} a map

$$\mathcal{K}(FB'_0, FB_0) \times \mathcal{B}(VB_0, B) \rightarrow \mathcal{B}(VB'_0, B)$$

Suppose first that B'_0 is κ -accessible. Then any element $f \in \mathcal{K}(FB'_0, FB_0)$ factors through a κ -accessible object $B''_0 \rightarrow B_0$. We have

$$\mathcal{K}(FB'_0, FB''_0) \times \mathcal{B}(VB_0, B) \rightarrow \mathcal{K}(FB'_0, FB''_0) \times \mathcal{B}(VB''_0, B) \rightarrow \mathcal{B}(VB'_0, B)$$

since now both FB'_0 and FB''_0 belong to \mathcal{K}_0 . Write a general B'_0 as a κ -filtered colimit of κ -accessible objects and take limits on both sides. The uniqueness follows similarly from the fact that these categories are κ -accessible. \square

3.7 Theorem. *Suppose that $F: \mathcal{B} \rightarrow \mathcal{K}$ is a theory-type sketch and that F has a right adjoint. Then U has a left adjoint.*

Proof. Let R be right adjoint to F . Begin by defining an embedding $J: \mathcal{K} \rightarrow \text{Mod}(\mathcal{K})$. It is, up to isomorphism, the Yoneda embedding. For an object A of \mathcal{K} , we let $J(A)$ be the pair (RFA, M) where $M(FB) = \mathcal{B}(B, RFA) \cong \mathcal{K}(FB, FA)$. If $f: FA' \rightarrow FA$ is an arrow, then Jf is the composite

$$\mathcal{B}(B, RFA') \cong \mathcal{K}(FB, FA') \xrightarrow{\mathcal{K}(FB, f)} \mathcal{K}(FB, FA) \rightarrow \mathcal{B}(B, RFA)$$

Then we have

$$\begin{array}{ccc} \mathcal{K} & \xrightarrow{J} & \text{Mod}(\mathcal{K}) \\ & \searrow F & \swarrow U \\ & \mathcal{B} & \end{array}$$

R

with F left adjoint to R . I claim that JF is left adjoint to U . In fact,

$$\begin{aligned} \text{Mod}(\mathcal{K})(JFA, (M'P')) &\cong NT(\mathcal{K}(-, FA), M') \cong M'FA \\ &\cong \mathcal{B}(A, P') \cong \mathcal{B}(A, U(M', P')) \end{aligned}$$

This completes the proof. \square

3.8 The Kleisli category. Let $\mathbf{T} = (T, \eta, \mu)$ be a triple on \mathcal{B} . The **Kleisli category** [Kleisli, 1965] \mathcal{K} of the triple has as objects those of \mathcal{B} with homsets given by

$$\mathcal{K}(A, B) = \mathcal{B}(A, TB)$$

The multiplication and unit of the triple give the composition and identities of \mathcal{K} . We have functors $F: \mathcal{B} \rightarrow \mathcal{K}$ and right adjoint $R: \mathcal{K} \rightarrow \mathcal{B}$. The functor F is the identity function on objects, while on arrows it is given by $f: A \rightarrow B$ goes to $\eta B \circ f = Tf \circ \eta A$. On objects, $RB = TB$ while on an arrow $f: A \rightarrow B$ in \mathcal{K} , meaning $f: A \rightarrow TB$ in \mathcal{B} , we let $Uf = \mu B \circ Tf$. Details are found in [Barr & Wells, 1985].

The situation $F: \mathcal{B} \rightarrow \mathcal{K}$ is, with V the identity functor, an example of a \mathcal{B} -based sketch.

We also have the well-known category $\mathcal{B}^{\mathbf{T}}$ of Eilenberg-Moore algebras, together with the functors $U^{\mathbf{T}}: \mathcal{B}^{\mathbf{T}} \rightarrow \mathcal{B}$ right adjoint to $F^{\mathbf{T}}: \mathcal{B} \rightarrow \mathcal{B}^{\mathbf{T}}$. The relation between the Eilenberg-Moore category and the Kleisli category is given by the following.

3.9 Theorem. *The Kleisli category of a triple is isomorphic to the full image of the left adjoint $F^{\mathbf{T}}$. The Eilenberg-Moore category is isomorphic to the category of models of the Kleisli category.*

The first assertion is well-known and is found in many places, for example in [Barr & Wells, 1985]. For the second, see Theorem 4.1 below.

3.10 Horns. Let us now assume fixed an \mathcal{E}/\mathcal{M} factorization system in \mathcal{B} . By a **Horn** on \mathcal{K} (relative to the given factorization system), we mean a diagram in \mathcal{K} of the form

$$\begin{array}{ccc} FA' & & FA'' \\ & \swarrow Ff & \nearrow g \\ & FA & \end{array}$$

in which $f \in \mathcal{E}$.

A model (B, M) of \mathcal{K} **satisfies** the Horn if there is a function $h: MFA'' \rightarrow MFA'$ such that the diagram

$$\begin{array}{ccc} MFA' & \xleftarrow{h} & MFA'' \\ & \searrow MFf & \swarrow Mg \\ & MFA & \end{array}$$

commutes. This is equivalent to the commutation of

$$\begin{array}{ccc} \mathcal{B}(A', B) & \xleftarrow{h} & \mathcal{B}(A'', B) \\ & \searrow \mathcal{B}(f, B) & \swarrow Mg \\ & \mathcal{B}(A, B) & \end{array}$$

where $\mathcal{B}(f, B)$ is monic since f is epic. That implies that the factorization, if it exists, is unique. Thus there is a full subcategory of models consisting of those for which such a factorization exists.

The fact that the objects appearing in the Horn are free is what determines that the Horn clauses are restricted to predicates available in \mathcal{B} .

3.11 HSP subcategories. Again we assume fixed a factorization system on \mathcal{B} and the associated system in the category \mathcal{C} of models. By an HSP subcategory $\mathcal{C}_0 \subseteq \mathcal{C}$ (relative to the given factorization system), we mean a full subcategory that is closed under U -split epimorphisms, \mathcal{M} monics and products.

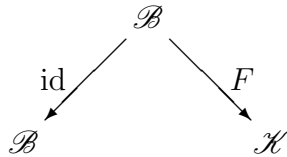
4 Functorial semantics

The primary reference for functorial semantics in a general setting (that is other than over the category of sets) is [Linton, 1969]. We give an outline of the main construction here.

Let \mathcal{B} be a category. A **theory** over \mathcal{B} is a category \mathcal{T} equipped with a functor $G: \mathcal{B}^{\text{op}} \rightarrow \mathcal{T}$ that is an isomorphism on objects. It is usual to suppose that G preserves limits, but in fact, we really require that it have a left adjoint. If we were to restrict to accessible categories and functors, these assumptions are equivalent.

Assuming that G has a left adjoint then, up to equivalence, $\mathcal{K} = \mathcal{T}^{\text{op}}$ is the Kleisli category for the triple associated to the adjunction. We find it more convenient to work with the Kleisli category directly. Hence for us the functorial semantics setting is a category \mathcal{B} a category \mathcal{K} and a functor $F: \mathcal{B} \rightarrow \mathcal{K}$ that has a right adjoint and is an isomorphism on objects.

A Kleisli category can always be viewed as the theory-type sketch



A model of the Kleisli category is exactly the same as for any other sketch. It consists of the assignment, corresponding to each morphism $\sigma: FB_1 \rightarrow FB_2$ of \mathcal{K} , of a morphism $\sigma B: \mathcal{B}(B_1, B) \rightarrow \mathcal{B}(B_2, B)$ subject to two restrictions:

1. If $\sigma = \tau \circ \rho$, then $\sigma B = \tau B \circ \rho B$.
2. If $\sigma = Ff$, then $\sigma B = \mathcal{B}(f, B)$.

We will freely use this description of a model. A morphism of models is as above in 3.2.

Let $\mathcal{C} = \text{Mod}(\mathcal{K})$ denote the category of models of \mathcal{K} (although they are really models of $F: \mathcal{B} \rightarrow \mathcal{K}$, but we will make the usual abuse of notation) and natural transformations between the functors. Since a natural transformation between representable functors is induced by a unique morphism between the representing objects a natural transformation between the functors also gives an arrow between the objects. Thus there is a functor $U: \mathcal{C} \rightarrow \mathcal{B}$ that takes the pair (B, M) to the object B . This functor has an adjoint that takes the object B to the pair $(\mathcal{K}(-, FB), TB)$.

The following is more-or-less proved (but not stated explicitly) in [Linton, 1969] (see Section 6, especially 6.2) and stated (with proof) in the unpublished [Linton, 1974].

4.1 Theorem. *Let \mathbf{T} be a triple on \mathcal{B} with Kleisli category \mathcal{K} and let \mathcal{C} be the category of models of the resultant theory-type sketch. Then the underlying functor $U: \mathcal{C} \rightarrow \mathcal{B}$ is tripleable and the triple is \mathbf{T} .*

Proof. We begin by showing that the underlying functor reflects isomorphisms. If $(f, \mu): (B, M) \rightarrow (B', M')$ is such that f is an isomorphism then for every $\sigma: FB_1 \rightarrow FB_2$ in \mathcal{K} , the horizontal arrows in the commutative diagram

$$\begin{array}{ccc} \mathcal{B}(B_1, B) & \xrightarrow{\mathcal{B}(B_1, f)} & \mathcal{B}(B_1, B') \\ \sigma B \downarrow & & \downarrow \sigma B' \\ \mathcal{B}(B_2, B) & \xrightarrow{\mathcal{B}(B_2, f)} & \mathcal{B}(B_2, B') \end{array}$$

are isomorphisms, their inverses given by $\mathcal{B}(B_1, f^{-1})$ and $\mathcal{B}(B_2, f^{-1})$, respectively. But if a square of this form commutes, it continues to commute with the inverses and hence f^{-1} is also a morphism of models, clearly inverse to the original one.

Finally, we verify Beck's criterion. Suppose

$$(B', M') \begin{array}{c} \xrightarrow{(d^0, \mu^0)} \\ \xrightarrow{(d^1, \mu^1)} \end{array} (B, M)$$

is a U -split coequalizer, which means there is a split coequalizer diagram in \mathcal{B}

$$B' \begin{array}{c} \xrightarrow{d^0} \\ \xrightarrow{d^1} \end{array} B \xrightarrow{d} B''$$

This means that for any $\sigma: FB_1 \rightarrow FB_2$ in \mathcal{K} , the top (and bottom) row of the following diagram is a coequalizer

$$\begin{array}{ccccc} \mathcal{B}(B_1, B') & \rightrightarrows & \mathcal{B}(B_1, B) & \longrightarrow & \mathcal{B}(B_1, B'') \\ \sigma B' \downarrow & & \downarrow \sigma B & & \\ \mathcal{B}(B_2, B') & \rightrightarrows & \mathcal{B}(B_2, B) & \longrightarrow & \mathcal{B}(B_2, B'') \end{array}$$

That implies there is a unique arrow $\sigma B'': \mathcal{B}(B_1, B'') \rightarrow \mathcal{B}(B_2, B'')$ such that the right hand square commutes. This defines the arrows $\sigma B''$ for all the operations of the theory. That the equations are satisfied is a similar computation that uses the fact that $\mathcal{B}(B_1, B) \twoheadrightarrow \mathcal{B}(B_1, B'')$ is surjective and can be canceled from the relation for B . The commutativity of the right hand square in the diagram above (when it is filled in) shows that the arrow $B \rightarrow B''$ is a morphism of models and the fact that it is the required coequalizer follows similarly from the cancellability of surjections. In fact, given $(f, \mu): (B, M) \rightarrow (B''', M''')$ with $f \circ d^0 = f \circ d^1$, we first use the fact that B'' is the coequalizer of d^0 and d^1 to get the required arrow $B'' \rightarrow B'''$. The only thing left is to show that it is a morphism of models. For any $\sigma: FB_1 \rightarrow FB_2$, the outer square of

$$\begin{array}{ccccc} \mathcal{B}(B_1, B) & \longrightarrow & \mathcal{B}(B_1, B'') & \longrightarrow & \mathcal{B}(B_1, B''') \\ \sigma B \downarrow & & \downarrow \sigma B'' & & \downarrow \sigma B''' \\ \mathcal{B}(B_2, B) & \longrightarrow & \mathcal{B}(B_2, B'') & \longrightarrow & \mathcal{B}(B_2, B''') \end{array}$$

commutes and the surjection can be canceled to show that the right hand square also commutes, which completes the proof. \square

The next theorem implies that an HSP subcategory of a category tripleable over \mathcal{B} is also tripleable over \mathcal{B} .

4.2 Theorem. *Let $U: \mathcal{C} \rightarrow \mathcal{B}$ be tripleable and suppose $I: \mathcal{D} \rightarrow \mathcal{C}$ is the inclusion of a reflective subcategory that is closed under U -split epimorphic images. Then UI is also tripleable.*

Proof. The functor UI has a left adjoint because both U and I do. Similarly, it reflects isomorphisms. Suppose that

$$D_1 \begin{array}{c} \xrightarrow{d^0} \\ \rightrightarrows \\ \xrightarrow{d^1} \end{array} D_0$$

is a UI -split coequalizer pair. Then there is an object C of \mathcal{C} and an arrow $d: D_0 \rightarrow C$ such that

$$D_1 \begin{array}{c} \xrightarrow{d^0} \\ \rightrightarrows \\ \xrightarrow{d^1} \end{array} D_0 \xrightarrow{d} C$$

is a coequalizer that is preserved by U . But then d is a U -split epi whose domain is in \mathcal{D} and hence so is C . \square

5 The main HSP theorem

5.1 Theorem. *Let \mathcal{B} be a complete base category with a factorization system \mathcal{E}/\mathcal{M} . Let $F: \mathcal{B} \rightarrow \mathcal{K}$ the free functor for a Kleisli category and \mathcal{H} a class of Horns with respect to the factorization system. Then the full subcategory of $\mathcal{C} = \text{Mod}(\mathcal{K})$ consisting of all models that satisfy all the Horns of \mathcal{H} is an HSP subcategory of \mathcal{C} .*

Proof. Let \mathcal{C}_0 denote the full subcategory. It is immediate that this subcategory is closed under limits. Let (B, M) be an object of \mathcal{C}_0 and $(B_0, M_0) \twoheadrightarrow (B, M)$ be an arrow of \mathcal{M} . Let

$$\begin{array}{ccc} FA' & & FA'' \\ & \swarrow Ff & \nearrow g \\ & FA & \end{array}$$

be a Horn in \mathcal{H} . We must show we can fill in the top rear arrow in the diagram

$$\begin{array}{ccccc} \mathcal{B}(A', B_0) & & & & \mathcal{B}(A'', B_0) \\ & \searrow & & \swarrow & \\ & & \mathcal{B}(A, B_0) & & \\ & \downarrow & & \downarrow & \\ \mathcal{B}(A', B) & \longleftarrow & & \longrightarrow & \mathcal{B}(A'', B) \\ & \searrow & & \swarrow & \\ & & \mathcal{B}(A, B) & & \end{array}$$

To do this it is sufficient to show that the left front square is a pullback. But this follows from the properties of factorization systems as described in 2.1, applied to the square

$$\begin{array}{ccc} \mathcal{B}(A', B_0) & \longrightarrow & \mathcal{B}(A, B_0) \\ \downarrow & & \downarrow \\ \mathcal{B}(A', B) & \longrightarrow & \mathcal{B}(A, B) \end{array}$$

Finally, let $(B, M) \rightarrow (B_0, M_0)$ be such that $B \rightarrow B_0$ is a split epi and suppose that (B, M) is in \mathcal{C}_0 . We must be able to fill in the lower rear arrow in the diagram

$$\begin{array}{ccccc}
 & & \mathcal{B}(A', B) & \longleftarrow & \mathcal{B}(A'', B) \\
 & & \searrow & & \swarrow \\
 & & \mathcal{B}(A, B) & & \\
 & \swarrow & \downarrow & & \searrow \\
 \mathcal{B}(A', B_0) & & & & \mathcal{B}(A'', B_0) \\
 & \searrow & \downarrow & & \swarrow \\
 & & \mathcal{B}(A, B_0) & &
 \end{array}$$

This comes from the diagonal fill-in in the square

$$\begin{array}{ccc}
 \mathcal{B}(A'', B) & \longrightarrow & \mathcal{B}(A'', B_0) \\
 \downarrow & & \downarrow \\
 \mathcal{B}(A'', B_0) & & \\
 \downarrow & & \downarrow \\
 \mathcal{B}(A', B) & \longleftarrow & \mathcal{B}(A, B_0)
 \end{array}$$

□

The converse is not true as it stands. The reason is that when Horns are imposed, the possibility of new Horns emerges. Thus the process may be iterated. That is, we may now look at Horns on the Kleisli category of \mathcal{C}_0 and get a new category \mathcal{C}_1 and so on. And it can happen that these Horns couldn't have been imposed on \mathcal{C} since the necessary elements weren't there. It is time for an example.

5.2 Example. Let \mathcal{B} be the category of posets and let $\mathbf{2}, \mathbf{3}, \dots$ denote the chains with two elements, three elements, \dots . Let $1, 1 + 1, 1 + 1 + 1, \dots$ denote the discrete posets with one element, two elements, three elements, \dots . Begin with a theory that has, in addition to the operations provided by \mathcal{B} one $\mathbf{2}$ -ary operation we denote σ and one equation $\sigma(x, x) = x$. More precisely, begin with the category \mathcal{C} for which an object is a poset B together with an operation that assigns to each chain $x \leq y$ of B an element $\sigma(x, y) \in B$, subject to the condition that $\sigma(x, x) = x$. It is immediate that the underlying functor $U: \mathcal{C} \rightarrow \mathcal{B}$ has a left adjoint and is tripleable. The Kleisli category in question is that of this triple.

The free algebra on $\mathbf{2}$ is actually $1 + \mathbf{2}$, consisting of $x \leq y$ and $\sigma(x, y)$. The lack of any comparison between x or y and $\sigma(x, y)$, combined with the equation we assume implies that there are no other elements in this free algebra. Now we impose

the following Horn condition:

$$\begin{array}{ccc}
 \mathbf{F2} & & \mathbf{F2} \\
 & \swarrow Ff & \nearrow \langle y, \sigma \rangle \\
 & F(1+1) &
 \end{array}$$

Here we continue to denote the two elements of $\mathbf{2}$ by $x \leq y$. The function $f: 1+1 \rightarrow \mathbf{2}$ is the inclusion function and is immediately seen to be epic. Thus this is a Horn condition with respect to the epi/regular-mono factorization system in \mathcal{B} . The effect of this condition on an algebra is to force the condition $y \leq \sigma(x, y)$. The result of this condition is that the free algebra on $\mathbf{2}$ now contains not only the three elements it used to, but also elements such as $\sigma(x, \sigma(x, y))$, $\sigma(y, \sigma(x, y))$ and much, much more. More to the point, there is also the possibility of new conditions that couldn't even be stated in the previous theory. As an example, the equation $\sigma(x, \sigma(x, y)) = \sigma(x, y)$ is an equation that could not be stated in the previous theory, because one of the terms wasn't there. Instead of an equality, we could have introduced an inequality. Thus the possibility of iterating the Horn condition construction must be allowed in any converse to the HSP theorem. First we state the obvious corollary to the theorem.

5.3 Corollary. *Suppose we have an ordinal chain of full subcategories*

$$\dots \subseteq \mathcal{C}_{\kappa+1} \subseteq \mathcal{C}_{\kappa} \subseteq \dots \subseteq \mathcal{C}_1 \subseteq \mathcal{C}_0 = \mathcal{C}$$

such that each is determined from the next by a class of Horn conditions and at limit ordinals κ , $\mathcal{C}_{\kappa} = \bigcap_{\lambda < \kappa} \mathcal{C}_{\lambda}$. Then $\bigcap_{\kappa} \mathcal{C}_{\kappa}$ is an HSP subcategory of \mathcal{C} .

Proof. This is an easy consequence of the facts that being HSP is transitive and closed under intersection. \square

5.4 Theorem. *Let \mathcal{C} be \mathcal{E} -well-co-powered (for example, suppose that \mathcal{C} is accessible). Suppose that $F: \mathcal{B} \rightarrow \mathcal{K}$ is the free functor into a Kleisli category for which \mathcal{C} is the category of algebras. Suppose that \mathcal{D} is an HSP subcategory of \mathcal{C} . Then there is an ordinal chain*

$$\dots \subseteq \mathcal{C}_{\kappa+1} \subseteq \mathcal{C}_{\kappa} \subseteq \dots \subseteq \mathcal{C}_1 \subseteq \mathcal{C}_0 = \mathcal{C}$$

of subcategories defined by Horns and by intersections at the limit ordinals such that $\bigcap_{\kappa} \mathcal{C}_{\kappa} = \mathcal{D}$.

We begin with:

5.5 Proposition. *Let FB be a free object of \mathcal{C} that is not in \mathcal{D} . Then there is a Horn that is satisfied by every object of \mathcal{D} , but not by FB .*

Proof. Let $I: \mathcal{C} \rightarrow \mathcal{D}$ be the inclusion. If C is an object of \mathcal{C} and D an object of \mathcal{D} any arrow $f: C \rightarrow D$ has an \mathcal{E}/\mathcal{M} factorization $C \twoheadrightarrow D_0 \twoheadrightarrow D$ and D_0 is in \mathcal{D} .

Since \mathcal{C} is well-co-powered, it follows that the set of such D_0 is a solution set for the adjoint. Hence \mathcal{D} is a reflective subcategory. It is also immediate that if $L: \mathcal{C} \rightarrow \mathcal{D}$ is the reflector with adjunction natural transformations $\alpha: \text{id} \rightarrow IL$ and $\beta: LI \rightarrow \text{id}$ adjunction natural transformations, then $\alpha C \in \mathcal{E}^{\mathbf{T}}$ for every object C of \mathcal{C} . Let $\epsilon: FU \rightarrow \text{id}$ and $\eta: \text{id} \rightarrow UF$ be the natural transformations corresponding to the adjunction $U \dashv F$.

Since αFB is in $\mathcal{E}^{\mathbf{T}}$ and not an isomorphism, it is not in $\mathcal{M}^{\mathbf{T}} = U^{-1}(\mathcal{M})$ and hence $U\alpha FB$ is not in \mathcal{M} . Then we can factor $U\alpha FB = g \circ f$ with $f: UFB \rightarrow B'$ in \mathcal{E} and $g \in \mathcal{M}$ and f is not an isomorphism. Now we consider the Horn

$$\begin{array}{ccc} FB' & & FB \\ & \swarrow Ff & \nearrow \epsilon FB \\ & FUFB & \end{array}$$

I claim that every object of \mathcal{D} satisfies this Horn. We must show that there is an arrow $h: \mathcal{C}(FB, D) \rightarrow \mathcal{C}(FB', D)$ such that the diagram

$$\begin{array}{ccc} \mathcal{C}(FB', D) & \xleftarrow{h} & \mathcal{C}(FB, D) \\ & \swarrow \mathcal{C}(Ff, D) & \searrow \mathcal{C}(\epsilon FB, D) \\ & \mathcal{C}(FUFB, D) & \end{array}$$

commutes. So suppose $k: FB \rightarrow D$ is an element of $\mathcal{C}(FB, D)$. We must show that there is an arrow $l: FB' \rightarrow D$ such that

$$\begin{array}{ccc} FUFB & \xrightarrow{\epsilon UB} & FB \\ Ff \downarrow & & \downarrow k \\ FB' & \xrightarrow{l} & D \end{array}$$

commutes. The universal mapping property of α implies that there is an arrow $m: ILFB \rightarrow D$ such that $k = m \circ \alpha FB$. Then the diagram

$$\begin{array}{ccccc}
 FUFB & \xrightarrow{Ff} & FB' & \xrightarrow{Fg} & FUILFB \\
 \downarrow \epsilon FB & & & & \downarrow \epsilon ILFB \\
 FB & \xrightarrow{\alpha FB} & & & ILFB \\
 & \searrow k & & \swarrow m & \\
 & & D & &
 \end{array}$$

commutes. The top square commutes by the naturality of ϵ since $Fg \circ Ff = FU\alpha FB$. Then $l = m \circ \epsilon ILFB \circ Fg$ is the required map.

Now we must show that FB does not satisfy the Horn. If FB satisfies the Horn, there is an $h: \mathcal{C}(FB, FB) \rightarrow \mathcal{C}(FB', FB)$ such that the triangle

$$\begin{array}{ccc}
 \mathcal{C}(FB', FB) & \xleftarrow{h} & \mathcal{C}(FB, FB) \\
 \mathcal{C}(Ff, FB) & & \mathcal{C}(\epsilon FB, FB) \\
 & \searrow & \swarrow \\
 & \mathcal{C}(FUFB, FB) &
 \end{array}$$

commutes. Let $m: FB' \rightarrow FB$ be $h(\text{id}_{FB})$. The commutation implies that $m \circ Ff = \epsilon FB$. This shows that the diagram

$$\begin{array}{ccccc}
 B' & \xrightarrow{\eta B'} & UFB' & \xrightarrow{Um} & UFB \\
 f \uparrow & & \searrow UFf & & \swarrow U\epsilon FB \\
 UFB & \xrightarrow{\eta UFB} & FUFB & &
 \end{array}$$

commutes. Since $U\epsilon FB \circ \eta UFB = \text{id}$, we see that f is a split monomorphism. But we began with $f \in \mathcal{E}$ and thus f is an isomorphism, contrary to its construction. This establishes the proposition. \square

Proof of the theorem: Every algebra is the target of a U -split epimorphism whose source is a free model. Thus if \mathcal{D} contains every free model, it contains every model. Hence if \mathcal{D} is a proper subcategory, there is some free algebra FB that it does not contain.

Each instance of such a free algebra not in \mathcal{D} gives a Horn and the class of all of them determines an HSP subcategory $\mathcal{C}_1 \subseteq \mathcal{C}$ such that $\mathcal{D} \subseteq \mathcal{C}_1$. It follows from Theorem 4.2 that \mathcal{C}_1 is tripleable over \mathcal{B} and it is obvious that \mathcal{D} is an HSP subcategory of \mathcal{C}_1 . If the inclusion $\mathcal{D} \subseteq \mathcal{C}_1$ is proper, then we may repeat the construction to get a subcategory $\mathcal{C}_2 \subseteq \mathcal{C}_1$ defined by a class of Horns of \mathcal{C} . In this way, we get a descending ordinal sequence of HSP subcategories. At limit ordinals, we take the intersection of the subcategories. Let $U^\lambda: \mathcal{C}_\lambda \rightarrow \mathcal{B}$ be the underlying functor and $F_\lambda: \mathcal{B} \rightarrow \mathcal{C}_\lambda$ be its left adjoint.

Each of these subcategories is an HSP and therefore \mathcal{E} -reflective subcategory of all the preceding ones. Thus we have for all pairs of ordinals $\kappa > \lambda$ an inclusion $I_\lambda^\kappa: \mathcal{C}_\kappa \rightarrow \mathcal{C}_\lambda$ with a left adjoint $L_\kappa^\lambda: \mathcal{C}_\lambda \rightarrow \mathcal{C}_\kappa$ such that the adjunction morphism $\alpha_\kappa^\lambda: \text{id} \rightarrow I_\kappa^\lambda L_\lambda^\kappa$ is in \mathcal{E} . For a fixed object C of \mathcal{C} , the class $\{\alpha_\lambda^0 C: C \rightarrow I_0^\lambda L_\lambda^0 C\}$ is only a set because of well-co-poweredness hypothesis. Hence there is a λ such that for all $\mu > \lambda$, there is a $\kappa > \mu$ for which α_λ^κ is an isomorphism. That is to say, at least one of the objects in the quotient lattice of C appears cofinally in the sequence. From $\alpha_\kappa^\lambda = \alpha_\kappa^\mu \circ \alpha_\mu^\lambda$ we conclude that α_λ^μ is a monomorphism and hence an isomorphism for all $\mu > \lambda$. This means that the $\alpha_{\lambda+1}^\lambda C$ are isomorphisms for all sufficiently large λ , depending on C .

Since $U^\lambda I_\lambda^\kappa = U^\kappa$, we conclude that $F_\kappa = L_\kappa^\lambda F_\lambda$. Suppose that C is in every \mathcal{C}_λ . Choose λ sufficiently large that $\alpha_{\lambda+1}^\lambda F_\lambda U^\lambda C$ is an isomorphism. If the object C is not in \mathcal{D} , then neither is $F_\lambda U^\lambda C$. Thus $\alpha_{\lambda+1}^\lambda F_\lambda U^\lambda C$ is not an isomorphism, contradicting the choice of λ . This shows that C is in \mathcal{C}_λ and establishes the theorem. \square

5.6 The case of regular epis. There is one case that is of special interest. This is the case that for each object B of \mathcal{B} , the arrow $FB \rightarrow ILFB$ is a U -regular epimorphism (meaning that U applied to it gives a regular epimorphism).

5.7 Theorem. *Suppose that $\mathcal{D} \subseteq \mathcal{C}$ is an HSP subcategory of the category of algebras with the property that for each object of B , the adjunction $FB \rightarrow ILFB$ is both a regular epimorphism and a U -regular epimorphism. Then \mathcal{D} is the category of models of the Horns satisfied by \mathcal{D} .*

In other words, in this case, we don't have to go to a sequence of such subcategories; a single step is sufficient. Before beginning the proof, we require a diagram chasing lemma.

5.8 Lemma. *Let*

$$\begin{array}{ccccc}
 A & \rightrightarrows & B & \longrightarrow & C \\
 \downarrow & & \downarrow & & \downarrow \\
 A' & \rightrightarrows & B' & \longrightarrow & C'
 \end{array}$$

be a diagram whose rows are coequalizers, whose squares serially commute and whose left hand vertical arrow is an epimorphism. Then the right hand square is a pushout.

Proof. The dual is a simple diagram chase using elements. \square

Proof of the theorem We construct the Horns by factoring the arrow $UFB \rightarrow UILFB$ which we are supposing to be a regular epi and hence in the epi part of any factorization system. Let $C \rightrightarrows FB$ be the kernel pair of $FB \rightarrow ILFB$. Then UC is the kernel pair of $UFB \rightarrow UILFB$, which is a regular epi. It follows from standard category theory that $UC \rightrightarrows UFB \rightarrow UILFB$ is a coequalizer. But free functors preserve coequalizers, so that $FUC \rightrightarrows FUFB \rightarrow FUILFB$ is a coequalizer as well. We can now apply the lemma to the diagram

$$\begin{array}{ccccc}
 FUC & \rightrightarrows & FUFB & \longrightarrow & FUILFB \\
 \downarrow & & \downarrow & & \downarrow \\
 C & \rightrightarrows & FB & \longrightarrow & UILFB
 \end{array}$$

to conclude that

$$\begin{array}{ccc}
 FUFB & \longrightarrow & FUILFB \\
 \downarrow & & \downarrow \\
 FB & \longrightarrow & ILFB
 \end{array}$$

is a pushout. Now consider the way in which a Horn is constructed over B . We begin by factoring the arrow $UFB \rightarrow UILFB$. But is already a regular epi, hence cannot be further factored. Thus the B' of the construction is $UILFB$ and the Horn is

$$\begin{array}{ccc}
 FUFB & \longrightarrow & FB \\
 \downarrow & & \\
 FUILFB & &
 \end{array}$$

whose pushout is $ILFB$. From this it follows from the definition of what it means to satisfy the Horn that for any object D that does satisfy it, every map $FB \rightarrow D$ induces a unique map $ILFB \rightarrow D$. But this is possible if and only if D is already in \mathcal{D} . \square

5.9 Extremal monics. There is one important special case that should be mentioned. In the case that \mathcal{M} is the class of extremal monomorphisms it will not be necessary to show that an HSP category is closed under all the extremal monics, provided \mathcal{B} is complete. We begin with a well-known property of extremal monics.

5.10 Proposition. *Suppose \mathcal{B} is a complete category in which every class of subobjects of an object has an intersection. Let \mathcal{M} be a class of monics in \mathcal{B} that includes all regular monics and is closed under composition and all intersections. Then \mathcal{M} includes all extremal monics.*

Note that an accessible category is well-powered, so the intersection condition follows from completeness.

Proof. Let $f: A \rightarrow B$ be an extremal monic not in \mathcal{M} , if possible. Let C be the intersection of all its \mathcal{M} subobjects of B that include A . Since \mathcal{M} is closed under arbitrary intersection, C is an \mathcal{M} subobject of A . Also C can have no proper \mathcal{M} subobject that includes A , since that would be an \mathcal{M} subobject of B . If $A \rightarrow C$ were an epimorphism, it would be both epic and extremal monic and hence an isomorphism and hence a regular monic. Thus there are two arrows $g \neq h: C \rightarrow D$ for some D such that $g \circ f = h \circ f$. The equalizer of g and h is then a proper \mathcal{M} subobject of C that includes A , a contradiction. \square

5.11 Corollary. *Under the same hypotheses, in order that a reflective subcategory $\mathcal{D} \subseteq \mathcal{C}$ be an HSP subcategory for the class of extremal monics, it is sufficient that it be closed under U -split epics, regular monics and products.*

Proof. Let \mathcal{M} be the class defined by $f: A \rightarrow B$ in \mathcal{M} if f is monic and if B in \mathcal{D} , then A in \mathcal{D} . Of course, this means that \mathcal{M} contains every subobject of an object not in \mathcal{D} , but that is of no significance. It is immediate that \mathcal{M} is closed under composition and intersection. Thus \mathcal{M} contains all extremal subobjects. In particular, \mathcal{D} is closed under extremal subobjects. \square

6 Examples

6.1 Example. We see how this works in the category of posets. There are two distinct factorization systems (and others that do not appear to be interesting for this purpose). One is regular-epi/mono and the other is epi/regular-mono. If we take a theory whose operations are defined on all tuples, then the resultant triple will preserve regular epimorphisms and the regular image of a morphism of algebras will be an algebra for the theory. It follows that the conditions of Theorem 5.7 are satisfied for the regular-epi/mono factorization system. Of course, this system is not really very interesting. The only conditions that can be imposed are equational.

Much more interesting is the case of epi/regular-mono factorization. The reason is that such arrows as $1 + 1 \rightarrow \mathbf{2}$ is an epi, but not a regular epi. A Horn based on that epi is a way of imposing an order relation among terms. We have already seen this in the case of the theory with one $\mathbf{2}$ -ary operation. Here is another interesting example.

Begin with the theory with one ω -ary operation, call it \lim . Thus to every ω indexed sequence $x_0 \leq x_1 \leq \dots$ there is an element we denote $\lim x_i$. We now want to

say that for all j , we have $x_j \leq \lim x_i$. This is easily done by using the Horn

$$\begin{array}{ccc}
 F\mathbf{2} & & F\omega \\
 \swarrow Ff & & \nearrow g \\
 & F(1+1) &
 \end{array}$$

where $f: 1+1 \rightarrow \mathbf{2}$ is the inclusion and $g: F(1+1) \rightarrow F(\omega)$ takes the first generator of $F(1+1)$ to the integer j and the second generator to the element $\lim x_i$. It should be understood that in the free algebra, every subsequence of ω has a different “limit”, including even the constant ones. Now an arrow $F(\omega) \rightarrow C$ is uniquely determined by a sequence $x_0 \leq x_1 \leq \dots$ of elements of C . The composite with g takes such a sequence to the map $F(1+1) \rightarrow C$ that takes the two generators to x_j and $\lim x_i$. This factors through $\mathbf{2}$ if and only if $x_j \leq \lim x_i$. A model satisfies this condition for all j if and only if $\lim x_i$ is an upper bound for the sequence.

It is made into a least upper bound in a similar way. Take the Horn

$$\begin{array}{ccc}
 F\mathbf{2} & & F(\omega+1) \\
 \swarrow Ff & & \nearrow h \\
 & F(1+1) &
 \end{array}$$

where h is the arrow that takes the first generator to $\lim x_i$ and the second to ω . A morphism $F(\omega+1) \rightarrow C$ is determined by a sequence $x_0 \leq x_1 \leq \dots$ and an upper bound x_ω . The Horn is satisfied for this sequence if and only if $\lim x_i \leq x_\omega$. Since x_ω is an arbitrary upper bound, the Horn is satisfied if and only if $\lim x_i$ precedes *every* upper bound of the sequence. Thus all these conditions are satisfied if and only if $\lim x_i$ is the least upper bound for the sequence. The category of algebras for this theory is precisely the category of ω -CPOs.

Here is another example based on posets. Take the sketch with one everywhere defined binary operation. This means that the sketch consists of one arrow $\sigma: 1 \rightarrow 1+1$. A model is a poset together with a binary operation that is not assumed to satisfy any condition. Let $f: 1+1 \rightarrow \mathbf{2}$ denote the inclusion. Now look at the Horn

$$\begin{array}{ccc}
 \mathbf{2} & & \mathbf{2} + \mathbf{2} \\
 \swarrow f & & \nearrow \\
 & 1+1 &
 \end{array}$$

The arrow going up the right is $\langle \sigma \circ \langle 0, 0 \rangle, \sigma \circ \langle 1, 1 \rangle \rangle$. On a model P it takes $((x_0, x_1), (y_0, y_1))$ such that $x_0 \leq x_1$ and $y_0 \leq y_1$ to $(\sigma(x_0, y_0), \sigma(y_0, y_1))$ and P satisfies the Horn if and

only if for all such 4-tuples, we have $\sigma(x_0, y_0) \leq \sigma(x_1, y_1)$. Thus a model P satisfies the Horn if and only if it satisfies the Horn

$$x_0 \leq x_1 \wedge y_0 \leq y_1 \Rightarrow \sigma(x_0, y_0) \leq \sigma(x_1, y_1)$$

6.2 Example. A nice base category is the category of graphs, \mathbf{Gph} . Let us denote by \mathbf{n} the graph with n objects $0, 1, \dots, n-1$ and one arrow $i-1 \rightarrow i$ for $i = 1, \dots, n-1$. Thus $\mathbf{1}$ is the graph with one object and no morphisms and $\mathbf{2}$ is the graph with two objects and one arrow between them. Note that $\mathbf{1}$ is not the terminal graph; that is the graph with one object and one endomorphism. An operation $\mathbf{2} \rightarrow \mathbf{3}$ in a theory is modeled by a function $\mathbf{Gph}(\mathbf{3}, G) \rightarrow \mathbf{Gph}(\mathbf{2}, G)$ on a graph \mathbf{Gph} . If this is forced to be associative, it determines a composition law among arrows on a graph. An operation $\emptyset \rightarrow \mathbf{1}$ gives an object in a model and an operation $\emptyset \rightarrow \mathbf{2}$ gives an arrow. With the proper equations, this gives identity arrows and so we get a theory on graphs whose models are the category of small categories.

With a little more work, we can, for example create a theory whose models are categories with terminal objects. Begin by adding to the sketch described above an operation $\emptyset \rightarrow \mathbf{1}$. This is a constant of type object, intended to represent the terminal object. We then add an operation of type $\mathbf{1} \rightarrow \mathbf{2}$. This assigns to each object an arrow. By a suitable equation, we can make the domain of that arrow be the object in question and the codomain be the intended terminal object. The result will be a theory whose models are categories with a specified weak terminal object.

Now consider the full subcategory consisting of those small categories for which the specified object is actually terminal. It is clear that this is closed under products and subobjects. (Note: the inclusion of a subobject must take the specified weak terminal object of the one category to that of the second. If the latter is actually terminal, so is the former.) Finally, we must show that this subcategory is closed under the formation of quotients that split at the level of underlying graph.

So suppose $\phi: \mathcal{A} \rightarrow \mathcal{B}$ is a functor such that there is graph morphism $\psi: \mathcal{B} \rightarrow \mathcal{A}$ such that $\phi \circ \psi = \text{id}_{\mathcal{B}}$. We will simplify somewhat by assuming that \mathcal{B} is a subgraph of \mathcal{A} and ψ is the inclusion.

By elaborating this kind of construction, we can construct more complicated theories based on graphs whose algebras are categories with equalizers, or with all finite limits or toposes. This has been carried in detail in Burroni [1981], (see also Lambek [1982] and Lambek and Scott [1987] for a somewhat easier exposition of some of the details).

6.3 Example. We will give an outline of how the category of H -spaces is the category of models of a theory based on the category of topological spaces. It is clear that construction can be elaborated to give A_n or E_n spaces for n an integer or infinity.

First off, we need a multiplication, which is an operation of type $\mathbf{2} \rightarrow \mathbf{1}$, these being the discrete spaces with 2 and 1 elements, respectively. Then a unit of type 0

$\rightarrow 1$ and equations to make its value be a left and right unit. Up to this point, there is nothing to make these operations continuous. The algebras for what we have are simply topological spaces with a multiplication and unit on its underlying set. Morphisms are continuous functions that preserve the operations. Call this category \mathcal{C} and let \mathcal{D} be the full subcategory of those models on which the multiplication is continuous. Let $U: \mathcal{C} \rightarrow \text{Top}$ be the underlying functor to topological spaces.

6.4 Proposition. *The full subcategory \mathcal{D} is closed under products, extremal subobjects and U -split quotients.*

Proof. The closure under limits is obvious. Suppose $Y \subseteq X$ is an extremal submodel where the multiplication on X is continuous. The topology on Y is that a subset U is open if and only if it has the form $Y \cap V$ where V is an open subset of X . Let $\mu_X: X^2 \rightarrow X$ and $\mu_Y: Y^2 \rightarrow Y$ be the operations on X and Y , respectively. Then

$$\mu_Y^{-1}(U) = \mu_Y^{-1}(V \cap Y) = \mu_X^{-1}(V) \cap Y^2$$

which is the intersection with Y^2 of an open set in X^2 . This shows that the multiplication in Y is continuous.

Suppose now that $f: X \twoheadrightarrow Y$ is a map of models that is a retraction to a subspace in Top and that the multiplication in X is continuous. Let μ_X and μ_Y be the multiplication as above. For U open in Y , $f^{-1}(U)$ is open in X and hence $\mu_X^{-1}f^{-1}(U) = (f^2)^{-1}\mu_Y^{-1}(U)$ is open in X^2 . But then so is

$$\mu_Y^{-1}(U) = Y^2 \cap (f^2)^{-1}\mu_Y^{-1}(U)$$

□

6.5 Corollary. *Under the same hypotheses, \mathcal{D} is an HSP subcategory of \mathcal{C} for the epi/extremal-monic factorization system.*

Proof. The only thing left to show is that the subcategory is reflective. But epis are surjective, hence from which it follows that if $X \twoheadrightarrow Y$, then $X^2 \twoheadrightarrow Y^2$. From this it is easy to see that the extremal image of a morphism of models is a model. Also the category is well-co-powered. Thus the solution set for some model X consists of all the epimorphic images of X for which the multiplication is continuous. □

The next step is to add the homotopy. This is an operation of type $3 \rightarrow [0, 1]$ that assigns to each 3-tuple (x, y, z) of elements of a model X a homotopy between $\mu(x, \mu(y, z))$ and $\mu(\mu(x, y), z)$. The equations for this are immediate. Thus the category of H -spaces (and maps that preserve the operation and the homotopy) is the category of models of a theory on Top .

6.6 Example. In this example, we take for base category the category whose objects are metric spaces (that allow points to be infinitely far apart) and morphisms are functions that do not increase distance. This category is complete and cocomplete and it is not hard to show it is locally \aleph_1 -presentable. In fact, using Theorem 2.5.1 of [Makkai & Paré] with P the set of one and two point metric spaces gives a quick proof it is accessible, but doesn't give us the cardinal. Actually a representative sample of the metric spaces that have at most countably many points are an \aleph_1 -accessible generating set. (Interestingly, the semi-metric spaces—distinct points are allowed to be at zero distance—even form an \aleph_0 -accessible category. You can take for generators all the finite semi-metric spaces.) Let S and S^+ denote, respectively, the spaces consisting of all the $1/n$, $n \in \mathbf{N}$ and that space together with 0. Then a sketch built with the single operation of type $S \rightarrow S^+$ assigns to each sequence of type $\{1/n\}$ (that is each sequence of elements x_1, x_2, \dots , such that $d(x_n, x_m) \leq |1/n - 1/m|$) a convergent sequence of the same type. With appropriate equations, we can force the non-limit terms of the second sequence to be the same as those of the first. Thus a model will satisfy this equation if and only if every such sequence converges. it is easy to see that this is so if and only if the space is complete. Thus the complete metric spaces are an HSP subcategory of the models of this operation.

6.7 Example. Consider now the category \mathcal{N} whose objects are sets with a real value function, say $\| - \|: X \rightarrow \mathbf{R}$. We will call this function a norm and define an arrow in this category to be a function that doesn't increase norms. We call this the category of normed sets. For $\lambda \in \mathbf{R}$, let $G(\lambda)$ denote the set with one element of norm λ . The set $\mathcal{N}(G(\lambda), X)$ is the subset of X consisting of those elements whose norm is at most λ . Now build a sketch that has, for each $\lambda_1 < 1/2$ and $\lambda_2 < 1/2$ an operation of type

$$G(\lambda_1 + \lambda_2) \rightarrow G(\lambda_1) + G(\lambda_2)$$

for each $\lambda < 1/2$ an operation $(-)^{-1}: G(\lambda) \rightarrow G(\lambda)$ and an operation $\iota: \emptyset \rightarrow G(0)$ and operations $G(\lambda) \rightarrow G(\lambda)$, defined for $\lambda < 1/2$. A model is a normed set with a set of partial multiplication, defined when both arguments have norm less than $1/2$ and a set of inverses defined on the same range. In addition there is a constant of norm 0. The next thing is to add equations that say that if $\|x\| < 1/2$ and $\|y\| < 1/2$, then all the multiplications defined for x and y are the same. With similar equations for the unary operation, which we will denote $()^{-1}$ we get a set with a partial binary operation we denote $*$, defined for all x, y both of norm less than $1/2$, an inverse in the same range and a constant of norm 0. Notice that we have insured that $\|x * y\| \leq \|x\| + \|y\|$ and that $\|x^{-1}\| \leq \|x\|$. We next impose the equation that for all $\lambda < 1/2$, for all x with $\|x\| < \lambda$, we have $x * \iota = \iota * x = x$ and $x * x^{-1} = x^{-1} * x = \iota$ and for all λ, μ, ν all less than $1/4$ and all x, y, z of norm λ, μ, ν , respectively, we have $x * (y * z) = (x * y) * z$. We call the resultant category the category of *normed local groups*.

Any sufficiently small neighborhood of the identity in a Lie groups, for example, will give a normed local group with an appropriate norm.

6.8 Example. A *coherence space* [Girard, 1989] is a poset with a bottom in which the down segment of each element is a complete atomic boolean algebra (CABA) and in which each set of elements of which any two have a common upper bound has an upper bound (and therefore a sup, since the down segment of that sup is a CABA). It is well-known that the property of being a CABA can be defined equationally among complete boolean algebras; the equations are the infinitary distributive laws. A morphism of coherence spaces is an order preserving function that preserves the structure of the CABAs. More precisely, if C and C' are coherence spaces, let $x\downarrow$ denote the down segment of the elements that are below x (including x). Then a morphism is an order preserving $f: C \rightarrow C'$ such that for each $x \in C$, the restriction of f induces a complete boolean homomorphism from $x\downarrow \rightarrow f(x)\downarrow$.

This can be made into the category of models of a theory on poset as follows. For every set X , let X^\top denote the poset consisting of the discrete set X plus one more element above every element of X . There is a nullary operation that chooses a bottom element and equations that make it be the bottom. For each set X there is an X^\top -ary operation that is thought of as taking a set with an upper bound to its sup and another taking a set with an upper bound to its inf. A 1^\top -ary operation takes a pair $y \leq x$ to the relative complement of y in $x\downarrow$. Equations will force these operations to have the appropriate properties to be the sup, inf and relative complement. So far, a model will be a poset with bottom in which down segments are CABAs.

For each set X let \widehat{X} denote a set that is the union of all the subsets of X with at most two elements, ordered by inclusion. We add to the theory above an \widehat{X} -ary operation for each X that is thought of as taking a set of elements, each pair of which has a common upper bound, to a common upper bound. Equations will guarantee that the value of the operation is an upper bound for the set.

This example illustrates as well as any the power of building a theory using a sketch of it. In this case, the sketch happens not to be finite, or even small, but it is nonetheless very simple and is clearly based on the way we perceive the theory.

7 Related work

There is an unpublished paper of Kelly and Power [to appear] which deals with related material. Their hypotheses are quite different since they deal with finitary theories, but their categories enriched over closed categories. I don't think that finitariness is essential for their results (although some adjoints won't automatically exist but will have to be assumed) and the results of this paper can no doubt be generalized to the enriched context. But their emphasis is on presenting every tripleable category

as an equational subcategory of a category of algebras for some operations. In terms of the results in this paper, what they prove is that if you put in enough operations, then equations are enough. By contrast, we are interested in how to describe all HSP subcategories of a category of algebras already given.

Nonetheless their results do show that in the locally presentable case, every theory can be presented with operations and equations. One way to see this is to take the graph underlying a small generating subcategory of the Kleisli category as operations. Then the equations are the commutative diagrams.

On the other hand, this doesn't appear to work with topological spaces (Example 6.3) because you would almost surely need a proper class of operations that is made manageable only by the equations.

Another related work is that of Andr eka and N emeti [1982]. This is a generalization in a completely different direction. The base category remains sets, but they are interested in SP subcategories and in partial algebras and similar things. They give several references to further related work.

It seems likely that both of these papers illustrate useful ways in which the ideas of the current paper might be extended. It can be done over a \mathcal{V} -based category for a closed category \mathcal{V} and there may well be an interesting theory for SP subcategories and for partial algebras.

References

- H. Andr eka and I. N emeti, *Generalization of the concept of variety and quasivariety*. Dissertationes Math. **CCIV** (1982).
- M. Barr, *Accessible categories and models of linear logic*. J. Pure Applied Algebra, **69** (1991), 219-232.
- M. Barr, *HSP type theorems in the category of posets*. To appear in S. Brookes, ed., Proc. Fifth Conference in Mathematical Foundations of Programming Semantics.
- M. Barr and C. Wells, **Toposes, Triples and Theories**. Grundlehren der math. Wissenschaften 278, Springer-Verlag, Berlin, Heidelberg, New York, 1985.
- G. Birkhoff, *On the structure of abstract algebra*. Proc. Cambridge Phil. Soc. **31** (1935), 433-454.
- A. Burroni, *Alg ebres graphiques*. Cahiers de Topologie et G eom etrie Diff erentielle, **22** (1981), 249-265.
- J.-Y. Girard, Y. Lafont, P. Taylor, *Proofs and Types*. Cambridge University Press, 1989.
- G. M. Kelly and A. J. Power, *algebraic structure in the enriched context*. To appear.

- H. Kleisli, Every standard construction is induced by a pair of adjoint functors. Proc. Amer. Math. Soc. 16, 544-546 (1965).
- J. Lambek, *Toposes are monadic over categories*. Lecture Notes in Mathematics **962** (1982), Springer-Verlag, Berlin, Heidelberg, New York, 153–166.
- J. Lambek and P. J. Scott, **Introduction to Higher Order Categorical Logic**. Cambridge studies in advanced mathematics, **7** (1986), Cambridge University Press.
- F. E. J. Linton, *An outline of functorial semantics*. Lecture Notes in Mathematics **80** (1969), Springer-Verlag, Berlin, Heidelberg, New York, 7–53.
- F. E. J. Linton. Mimeographed notes, 1974.
- M. Makkai and R. Paré, **Accessible Categories**. Contemporary Mathematics, **104**, Amer. Math. Soc. 1990.