# Polymorphism is Set Theoretic, Constructively*

Andrew M. Pitts

Mathematics Division
School of Mathematical and Physical Sciences
University of Sussex
Falmer, Brighton BN1 9QH, England.

**Abstract.** We consider models in toposes of equational theories over the type system consisting of the Girard-Reynolds *polymorphic lambda calculus* augmented with finite product types. The particular notion of model we use is very straightforward, with polymorphic product types and function types both being interpreted in a *standard* way in the topos—the first by internal products and the second by exponentiation. We show that every *hyperdoctrine model* of a polymorphic lambda theory can be fully embedded in such a *topos model*, the topos constructed being simply a functor category. There are precise correspondences between polymorphic lambda theories and their hyperdoctrine models, and between toposes and theories in *higher order intuitionistic predicate logic*. So we can conclude that every theory of the first kind can be interpreted in a theory of the second kind in such a way that the polymorphic types are interpreted in a standard way, but so that up to provability in the higher order theory, they have exactly the same closed terms as before. A simple corollary of this full embedding result is the completeness of topos models : for each polymorphic lambda theory there is a topos model whose valid equations are exactly those derivable in the theory.

* ~~To appear~~ In: *Proceedings of the Summer Conference on Category Theory and Computer Science, Edinburgh, September 1987*, Springer Lecture Notes in Computer Science, No. 283, pp12-39.

# 0. Introduction

The title of this paper ought really to be "Polymorphism *can* be set theoretic, constructively", but the obvious reference to Reynolds' paper "Polymorphism is not set theoretic" [R2] was too tempting. The purpose of the paper is to prove particular kinds of completeness and full embedding theorems for the polymorphic lambda calculus. In order to explain the theorems we need to review some of the recent history of type polymorphism:

In *loc.cit.*, it is shown that the standard interpretation of the first order typed lambda calculus in the category of sets cannot be extended to a model of the second order typed lambda calculus. Reynolds works with quite a general notion of what constitutes a second order model extending a standard first order model, but the naive idea would be to have a set $U$ of sets closed under finite cartesian products, exponentiation and $U$-indexed cartesian products: the types would be interpreted by sets which are elements of $U$, with function types and polymorphic product types *both* being interpreted in a standard way using set-theoretic exponentiation and product. Putting this in more category-theoretic terms, one would have a small category **U** and a full and faithful functor from **U** to the category of sets, $G$:**U**$\hookrightarrow$**Set**, so that **U** has and $G$ preserves finite products, exponentiation and products indexed by the set $U$ of objects of **U**. A simple cardinality argument shows that *there are no non-trivial such* **U**,$G$ (in the sense that any such $U$ can only contain sets with at most one element). If one drops the requirement of having the structure preserving full embedding into **Set**, there are of course many, highly non-trivial, small cartesian closed categories **U** with the above limit closure properties — for example Lambek's *C-monoids*, studied in Part I of [LS], correspond to cartesian closed **U** with just two objects (a terminal object $\top$ and an object $X \cong (X \times X) \cong (X \to X)$), and which therefore certainly have $U$-indexed products. All such **U** give rise to models of the second order typed lambda calculus, but not ones extending the standard set theoretic interpretation of the first order calculus.

Recent work of Hyland, Moggi, Robinson, Rosolini [HRR], Carboni, Freyd and Scedrov [CFS] shows that this non-existence of standard models is due to the non-constructive nature of the category of sets. They have demonstrated that it is possible for elementary toposes (which in general model a particular kind of higher order constructive logic — see [LS]) to contain *internal full subcategories* whch are closed under the operations of taking *any* limits or colimits in the topos (and so in particular also are closed under taking exponentials since these are given by powers, i.e. by internal products). The particular topos they consider is Hyland's *effective topos* [H] and its internal small full subcategory of *subcountable* objects, i.e. those objects which are the quotient of the natural numbers object by a partial equivalence relation. (The related internal subcategory given by $\neg\neg$-*closed* partial equivalence relations also has these strong closure properties in the effective topos; the objects in this subcategory have been termed the *modest sets* by Dana Scott. See also [FS].)

Thus the effective topos is a model of higher order constructive logic in which there is a "set" $U$ of "sets" whose elements are closed under all the operations (indeed, far more) needed to model the second order typed lambda calculus in a completely standard way: the polymorphic types (of one free type variable) are interpreted as arbitrary $U$-indexed collections of "sets" in $U$; function types are interpreted as full function exponentials in the topos; and the product type of a polymorphic type is interpreted as the actual product in the topos of the corresponding indexed collection (which product is again in $U$ despite the size of the indexing "set"). There is no contradiction with Reynolds' result since in the effective topos (and in toposes in general) the classically valid Law of Excluded Middle needed to carry out his argument, is not valid.

Generalizing the naive idea of a standard set theoretic model of the second order typed lambda calculus discussed above, by a *topos model* we shall mean the following: a topos **E** equipped with an internal category **U** and a full and faithful diagram $G$ of type **U** in **E**, so that **U** has and $G$ preserves finite products, exponentials and products indexed by the object of objects of **U**. (We will spell out explicitly what such data amount to in section 3.) Much as indicated above, these properties of the internal subcategory are just what are needed to interpret the language of the second order typed lambda calculus (by which we shall mean Girard's system $F$ [G1, G2] augmented with finite product types: see section 1), in such a way that the equations of $\beta$-conversion, $\eta$-conversion and surjective pairing and the $\xi$-rules of extensionality are always satisfied. Conversely we shall prove that there are enough topos models so that *the only equational consequences of these axioms and rules are those which are satisfied by all topos models.*

More generally one can consider *$2T\lambda C$-theories* $\mathbf{T}=(L,A)$, consisting of a suitable language $L$ (with symbols for constant types, type constructors and constant terms) and a set $A$ of equations between terms in the second order typed lambda calculus built up from the language. A *topos model* of **T** will then mean a topos model in the sense of the previous paragraph, together with an interpretation of the language with the property that all the equations of $A$ are satisfied. (Precise definitions of the notions of theory and satisfaction are given in sections 1 and 2.) We prove:

**Theorem B. (Completeness of topos models.)** *Let* $\mathbf{T}=(L,A)$ *be a $2T\lambda C$-theory. The equations between second order typed lambda calculus terms built up from $L$ which are provable consequences in equational logic of the axioms $A$ of* **T** *together with the axioms of $\beta$-conversion, $\eta$-conversion, surjective pairing and the $\xi$-rules, are just those satisfied by* *all topos models of* **T**. *In fact for each* **T**, *there is a* single *topos model whose valid equations are exactly those derivable in* **T**.

In fact our proof of this theorem is via an even stronger result which shows there are very many topos models:

**Theorem A. (Full embedding in topos models.)** *Every* hyperdoctrine model *of the second order typed lambda calculus fully embeds into a topos model.*

The category theoretic notion of *hyperdoctrine* was introduced by Lawvere [L] in his seminal work on the connexions between categories and logic. The correspondence between theories in a higher order (extensional) typed lambda calculus and an appropriate kind of hyperdoctrine has been developed by R.A.G.Seely [Se]. This correspondence readily specializes to one between $2T\lambda C$-theories and the kind of hyperdoctrines described in section 2 below. The passage from theory to hyperdoctrine is the familiar term-model type construction in categorical logic of organizing the syntax, suitably quotiented by provability in the theory, into a category with suitable extra categorical structure or properties. In this way Theorem B is deduced from Theorem A.

The proof of Theorem A proceeds by first using the *Grothendieck construction* to obtain the total category of the fibration corresponding to the given hyperdoctrine; and then one takes the functor category of contravariant set-valued functors on this total category. Thus the topos constructed is simply a *presheaf topos* and the original hyperdoctrine embeds into it essentially via the Yoneda embedding. The method of proof is therefore very similar to that used by Dana Scott [Sc1] in showing that a model of the untyped $\lambda$-calculus can be realized as a reflexive "set" inside a presheaf topos. The fact that the Yoneda embedding preserves any existing exponentials is a key ingredient of his proof, and here we use this and more — namely that the Yoneda embedding preserves any existing *local* exponentials and instances of right adjoints to pulling back. Part of the data which specifies the original hyperdoctrine naturally gives rise to an internal full subcategory in the constructed presheaf topos. The main part of the proof of Theorem A then resides in showing that this has the requisite limit closure properties.

The significance of Theorem A is that, *as long as our arguments can be carried out in intuitionistic higher order logic* (for a description of which see [LS, Part I] for example), *then we may reason about the second order typed lambda calculus as though polymorphic types were suitably indexed collections of sets, product types of polymorphic types were actual cartesian products of sets, function types were full exponentials of sets, etc.* Our method of proof in fact extends to the case of *higher order* typed lambda calculus (optionally augmented with *sum types* of polymorphic types). The corresponding kind of hyperdoctrine has a cartesian closed base category (optionally augmented with stable left adjoints to substitution along projections): see [Se]. The topos model constructed in the proof of Theorem A will in this case have the further property that the internal subcategory is closed in the topos under products indexed by any finite type on its object of objects (and optionally closed under similarly indexed coproducts). It is important to note that these closure properties are still much less than those enjoyed by the internal categories of the effective topos mentioned above: they are closed with respect to *all* internal limits and (hence) colimits. The question of whether Theorems A and B can be strengthened by restricting to such topos models is not addressed here.

Having read the above description of what will be proved in this paper, one might wonder why it was not entitled "Polymorphism is topos theoretic". What have these results to do with *constructive set theory* ? The kind of intuitionistic higher order logic which toposes model can certainly be regarded as a constructive theory of sets. But if one wishes to model the *untyped* $(\in,=)$-language of set theory, then elementary toposes in general are capable of modelling only a restricted form of Zermelo-Fraenkel set theory, with *bounded* separation and collection axioms: see [J, Chapter 9]. However, the toposes which arise in the proofs of Theorems A and B are *Grothendieck* toposes, and these M.Fourman [F] has shown admit the interpretation of a full intuitionistic Zermelo-Fraenkel set theory with atoms, *IZFA*. (Indeed, the toposes considered here are just *presheaf* toposes, for which the Fourman interpretation of *IZFA* takes a particularly simple form, as Dana Scott shows in [Sc2].) Thus Theorems A and B could be rephased in terms of models of the second order typed lambda calculus in *IZFA*. (It is interesting to note that even though Hyland's effective topos is definitely not a Grothendieck topos, nevertheless it does support an interpretation of *IZFA* — one that is entirely analogous to the standard Heyting-valued models. The resulting model of *IZFA* is presented by C.McCarty in [McC].)

# Acknowledgements

I have had many helpful discussions about type polymorphism with J.M.E.Hyland, E.Robinson, G.Rosolini and R.A.G.Seely (who also suggested the title of this paper). My understanding of fibred category theory in general and internal full subcategories in particular has been very greatly enhanced by the work of J.Bénabou in this area. This research was partly carried out while visiting the Centre en Etudes Catégoriques in Montréal. I also gratefully acknowledge the financial support of the Royal Society.

# 1 Second order equational theories

In this section we review the syntax of equational theories in a system, $2T\lambda C$, of *second order typed lambda calculus*. The type system is that of the Girard-Reynolds polymorphic lambda calculus [G1, G2, R1, R2], augmented with terminal and binary product types. (There are several good reasons for including these extra, first order types, not least of which is their presence in cartesian closed categories.) The equational logic we consider over the type system contains axioms for $\beta$-conversion, $\eta$-conversion, extensionality and surjective pairing.

**1.1. Types.** Let us fix a countably infinite set $TV$ whose elements we shall call *type variables* and denote by $X,Y,Z,\ldots$ . Given a multiset $ar:C\longrightarrow\mathbb{N}$, the set $Types(C)$ of $2T\lambda C$ *types over* $C$ (denoted $\Phi,\Psi,\ldots$) is defined inductively by the clauses given below; simultaneously we define the finite subset $FTV(\Phi)\subseteq TV(C)$ of *free type variables* of a type $\Phi$:

- if $X\in TV$, then $X\in Types(C)$ and $FTV(X)=\{X\}$;
- if $F\in C$ with $ar(F)=n$ say, and if $\Phi_0,\ldots,\Phi_{n-1}\in Types(C)$, then $F(\Phi_0,\ldots,\Phi_{n-1})\in Types(C)$ and $FTV(F(\Phi_0,\ldots,\Phi_{n-1})=FTV(\Phi_0)\cup\ldots\cup FTV(\Phi_{n-1})$;
- $\top\in Types(C)$ and $FTV(\top)=\emptyset$;
- if $\Phi,\Psi\in Types(C)$, then for $*=\times$ or $\rightarrow$, $(\Phi*\Psi)\in Types(C)$ and $FTV(\Phi*\Psi)=FTV(\Phi)\cup FTV(\Psi)$;
- if $\Phi\in Types(C)$ and $X\in TV$, then $\prod X.\Phi\in Types(C)$ and $FTV(\prod X.\Phi)=FTV(\Phi)\setminus\{X\}$.

If $F\in C$ and $ar(F)=n$, we shall call $F$ a *type constructor symbol of arity* $n$; in the special when case $n=0$, $F$ is a *type constant* and we write $F$ rather than $F()$ for the corresponding type. A type $\Phi$ is *closed* if $FTV(\Phi)=\emptyset$. The type variables occurring in a type $\Phi$ which are not in $FTV(\Phi)$ are the *bound* type variables of $\Phi$ ($X$ is bound in $\prod X.\Phi$). As usual, we identify types up to $\alpha$-equivalence, i.e. up to change of bound type variables. If $\Phi,\Psi\in Types(C)$ and $X\in FTV(\Phi)$, then $[X:=\Psi]\Phi$ will denote the result (defined up to $\alpha$-equivalence) of substituting $\Psi$ for $X$ throughout $\Phi$, avoiding variable capture.

**1.2. Terms.** Fix a countably infinite set $IV$ of *individual variables* $x,y,z,\ldots$ . Elements of the product $IV\times Types(C)$ will called *typed* individual variables and a typical such pair will be denoted $x^{\Phi}$. By a *language* $L$ we will mean a pair $L=(C,K)$, where $C$ is a multiset of type constructor symbols and $K$ is a set of *individual constants* $(a,b,c,\ldots)$ together with a specification of their types $(a{:}\Phi,\ldots)$, which we will always assume to be closed types. Given such an $L$, the set $Terms(L)$ of *$2T\lambda C$ terms* over $L$ (denoted $s,t,\ldots$) is defined inductively by the clauses given below; simultaneously we define the type of each term (denoted $s{:}\Phi$) and the finite subsets $FIV(s)$, $FTV(s)$ of *free individual and type variables* of a term $s$:

- if $x^{\Phi}\in IV\times Types(C)$, then $x^{\Phi}\in Terms(L)$ with $x^{\Phi}{:}\Phi$, $FIV(x^{\Phi})=\{x^{\Phi}\}$ and $FTV(x^{\Phi})=FTV(\Phi)$;
- if $a\in K$, then $a\in Terms(L)$ with closed type as specified and $FIV(a)=FTV(a)=\emptyset$;
- $\langle\rangle\in Terms(L)$ with $\langle\rangle{:}\mathrm{T}$ and $FIV(\langle\rangle)=FTV(\langle\rangle)=\emptyset$;
- $\mathsf{Fst}\in Terms(L)$ with $\mathsf{Fst}{:}\prod X.\prod Y.(X\times Y)\to X$ and $FIV(\mathsf{Fst})=FTV(\mathsf{Fst})=\emptyset$;
- $\mathsf{Snd}\in Terms(L)$ with $\mathsf{Snd}{:}\prod X.\prod Y.(X\times Y)\to Y$ and $FIV(\mathsf{Snd})=FTV(\mathsf{Snd})=\emptyset$;
- if $s{:}\Phi$ and $t{:}\Psi$, then $\langle s,t\rangle{:}\Phi\times\Psi$ with $FIV(st)=FIV(s)\cup FIV(t)$ and $FTV(st)=FTV(s)\cup FTV(t)$;
- if $s{:}(\Phi\to\Psi)$ and $t{:}\Phi$, then $st{:}\Psi$ with $FIV(st)=FIV(s)\cup FIV(t)$ and $FTV(st)=FTV(s)\cup FTV(t)$;
- if $s{:}\prod X.\Phi$ and $\Psi\in Types(C)$, then $s\Psi{:}[X{:=}\Psi]\Phi$ with $FIV(s\Psi)=FIV(s)$ and $FTV(s\Psi)=FTV(s)\cup FTV(\Psi)$;
- if $s{:}\Psi$, then $\lambda x^{\Phi}.s:\Phi\to\Psi$ with $FIV(\lambda x^{\Phi}.s)=FIV(s)\setminus\{x^{\Phi}\}$ and $FTV(\lambda x^{\Phi}.s)=FTV(\Phi)\cup FTV(s)$;
- if $s{:}\Phi$, then $\lambda X.s:\prod X.\Phi$ with $FIV(\lambda X.s)=FIV(s)$ and $FTV(\lambda X.s)=FTV(s)\setminus\{X\}$.

A term $s$ is *closed* if $FIV(s)=\emptyset$. The individual and type variables occurring in $s$ but not in the sets $FIV(s)$, $FTV(s)$ are the *bound* variables of $s$ ($x^{\Phi}$ is bound in $\lambda x^{\Phi}.s$ and $X$ is bound in $\lambda X.s$). As for types, we identify terms up to $\alpha$-equivalence, i.e. up to change of bound variables. If $s\in Terms(L)$ and $\Psi\in Types(C)$, then $[X{:=}\Psi]s$ denotes the result of substituting $\Psi$ for a type variable $X$ throughout $s$, avoiding variable capture; evidently $[X{:=}\Psi]s:[X{:=}\Psi]\Phi$ when $s{:}\Phi$. Similarly, if $s{:}\Phi$ and $t{:}\Psi$, then $[x^{\Psi}{:=}t]s{:}\Phi$ denotes the result of substituting $t$ for the individual variable $x^{\Psi}$ throughout $s$, avoiding variable capture.

**1.3. Theories.** Given a language $L=(C,K)$ as above, we consider judgements which assert the equality of two $2T\lambda C$ terms of equal type. Specifically, an *equality judgement* will take the form:

$$\vdash_{\underline{X},\underline{x}} s=t{:}\Phi$$

where $\underline{X}=X_0,\ldots,X_{n-1}$ is a finite list of distinct type variables, $\underline{x}=x_0^{\Phi_0},\ldots,x_{m-1}^{\Phi_{m-1}}$ is a finite list of distinct individual variables of various types over $C$, $s$ and $t$ are terms over $L$ of type $\Phi$ whose free individual variables are contained in $\underline{x}$, and such that the free type variables of $s$, $t$, $\Phi$ and the $\Phi_i$ are contained in $\underline{X}$. Then a *$2T\lambda C$-theory* **T** is specified by a language $L$ and a set $A$ of equality judgements over $L$, called the *axioms of* **T**.

**1.4. Equational logic over *2TλC*.** We next give the basic logical axioms and rules for deriving equality judgements in *2TλC*. In particular, if $\mathbf{T}=(L,A)$ is a *2TλC*-theory, then the ***theorems*** of $\mathbf{T}$ comprise the least set of equality judgements over $L$ containing $A$ and the axioms below, and closed under the rules given below:

- *weakening*
$$\frac{\vdash_{\underline{X},\underline{x}} s=t:\Phi}{\vdash_{\underline{X}',\underline{x}'} s=t:\Phi}\quad (\underline{X}\subseteq\underline{X}' \text{ and } \underline{x}\subseteq\underline{x}')$$

- *reflexivity*
$$\vdash_{\underline{X},\underline{x}} s=s:\Phi$$

- *symmetry*
$$\frac{\vdash_{\underline{X},\underline{x}} s=t:\Phi}{\vdash_{\underline{X},\underline{x}} t=s:\Phi}$$

- *transitivity*
$$\frac{\vdash_{\underline{X},\underline{x}} r=s:\Phi \qquad \vdash_{\underline{X},\underline{x}} s=t:\Phi}{\vdash_{\underline{X},\underline{x}} r=t:\Phi}$$

- *terminal*
$$\vdash_{(),()} t=():\top$$

- *pairing*
$$\frac{\vdash_{\underline{X},\underline{x}} s=s':\Phi \qquad \vdash_{\underline{X},\underline{x}} t=t':\Psi}{\vdash_{\underline{X},\underline{x}} (s,t)=(s',t'):\Phi\times\Psi}$$

- *projections*
$$\vdash_{\underline{X},\underline{x}} \mathrm{Fst}\,\Phi\Psi(s,t)=s:\Phi \;, \qquad \vdash_{\underline{X},\underline{x}} \mathrm{Snd}\,\Phi\Psi(s,t)=t:\Psi$$

- *surjectivity*
$$\vdash_{\underline{X},\underline{x}} (\mathrm{Fst}\,\Phi\Psi r,\mathrm{Snd}\,\Phi\Psi r)=r:\Phi\times\Psi$$

- *individual ξ*
$$\frac{\vdash_{\underline{X},\underline{x},x^{\Phi}} t=t':\Psi}{\vdash_{\underline{X},\underline{x}} \lambda x^{\Phi}.t=\lambda x^{\Phi}.t':\Phi\to\Psi}\quad (x^{\Phi}\notin\underline{x})$$

- *individual application*
$$\frac{\vdash_{\underline{X},\underline{x}} s=s':\Phi\to\Psi \qquad \vdash_{\underline{X},\underline{x}} t=t':\Phi}{\vdash_{\underline{X},\underline{x}} st=s't':\Psi}$$

- *individual β*
$$\vdash_{\underline{X},\underline{x}} (\lambda x^{\Phi}.t)s=[x:=s]t:\Psi$$

- *individual η*
$$\vdash_{\underline{X},\underline{x}} \lambda x^{\Phi}.(tx)=t:\Phi\to\Psi \qquad (x^{\Phi}\notin FIV(t))$$

- *type ξ*
$$\frac{\vdash_{\underline{X},X,\underline{x}} t=t':\Phi}{\vdash_{\underline{X},\underline{x}} \lambda X.t=\lambda X.t'}\quad (X\notin\underline{X})$$

- *type application*
$$\frac{\vdash_{\underline{X},\underline{x}} s=s':\prod X.\Phi}{\vdash_{\underline{X},\underline{x}} s\Psi=s'\Psi:[X:=\Psi]\Phi}$$

- *type β*
$$\vdash_{\underline{X},\underline{x}} (\lambda X.t)\Psi=[X:=\Psi]t:[X:=\Psi]\Phi$$

- *type η*
$$\vdash_{\underline{X},\underline{x}} \lambda X.(tX)=t:\prod X.\Phi \quad (X\notin\underline{X}) \;.$$

We shall write $\mathbf{T}\vdash_{\underline{X},\underline{x}} s=t:\Phi$ to indicate that the equality judgement $\vdash_{\underline{X},\underline{x}} s=t:\Phi$ can be derived as a theorem of the *2TλC*-theory $\mathbf{T}$. The reader will see from the above axioms and rules that we are considering *extensional* theories with *surjective pairing* operations. The first order part of such theories and their relationship to carstesian closed categories are exposed in [LS, Part I]. Seely [Se] considers the case of similar theories in full higher order typed lambda calculus, except that he also allows a second kind of equality judgement — namely the assertion that two *types* are (extensionally) equal. Influenced by Bénabou's observations in [B] on the rôle of equality between objects (≡types) in category theory, we have specifically left out this form of judgement as inappropriate to the intuitive notion of type we have. Rather, in the system we are considering, one can assert that particular terms give an *isomorphism* between particular types.

We conclude this section with some technical remarks about the precise form of the equality judgements in *2TλC*-theories. Because the underlying language of a theory can introduce arbitrary constant types, it is perfectly possible for a type to possess no closed terms, i.e. to be uninhabited. Accordingly the logical system presented above is not equivalent to one in which the judgements are not tagged with a list $\underline{x}$ of free individual variables. On the other hand, since there are always closed types (e.g. the terminal type $\top$), it is not essential that the judgements be tagged with a list $\underline{X}$ of free type variables: using the rules (*type* $\xi$) and (*type* $\beta$), one has that $\mathbf{T}\vdash_{\underline{X},\underline{x}} s{=}t{:}\Phi$ iff $\mathbf{T}\vdash_{\underline{X}',\underline{x}} s{=}t{:}\Phi$, where $\underline{X}'$ comprises *exactly* those free type variables mentioned in $\underline{x}$, $s$, $t$, and $\Phi$, and hence need not be given explicitly. However, the chosen form of judgement is convenient for giving the category theoretic semantics of *2TλC*, to which we now turn.

# 2 Hyperdoctrine models

In this section we review those parts of [Se] concerned with the connexion between *2TλC*-theories and a particular variety of *hyperdoctrine*. The general notion of a hyperdoctrine was introduced by Lawvere in his seminal work on the connexions between category theory and logic; see [L] for example. It has proved to be a very flexible and useful concept, not least in making precise the fundamental observation that *theories* in many different kinds of logic can be specified in a syntax-free way as *models* of various kinds of category theoretic structure, which structure can often be viewed as a particular kind of hyperdoctrine. We refer the reader to [P] and the references there for more information about hyperdoctrines in general. The particular kind of hyperdoctrine we will be concerned with here is a special case of that considered by Seely in [Se] and called a "PL category" there. He shows that there is a correspondence between these PL categories and equational theories over the higher order typed lambda calculus: this correspondence readily specializes to one between the *2TλC*-theories defined in the previous section and the kind of hyperdoctrine defined below. We begin by fixing some notation:

**2.1. Cartesian closed categories.** If a category **C** has a given terminal object, it will be denoted by $\top$. For any object $A$ in **C**, the unique morphism from $A$ to $\top$ will be denoted $A{:}A\longrightarrow\top$, whereas the identity morphism on $A$ will be denoted $1_A$. If **C** has given binary products, the product of objects $A$ and $B$ in **C** will be denoted by

$$A\xleftarrow{\pi_1}A\times B\xrightarrow{\pi_2}B\ .$$

For any morphisms $f{:}C\longrightarrow A$, $g{:}C\longrightarrow B$ in **C**, $\langle f,g\rangle{:}C\longrightarrow A\times B$ will denote the unique morphism with $\pi_1\circ\langle f,g\rangle=f$ and $\pi_2\circ\langle f,g\rangle=g$. For such a category **C**, the exponential of objects $A$ and $B$, if it exists, will be denoted $A\rightarrow B$ and the accompanying evaluation morphism by

$$ev{:}(A\rightarrow B)\times A\longrightarrow B.$$

Thus for a morphism $f{:}C\times A\longrightarrow B$, there is a unique morphism $\overline{f}{:}C\longrightarrow(A\rightarrow B)$ with $ev\circ(\overline{f}\times 1_A)=f$. As usual, a category is called a *cartesian closed category*, or more briefly a *ccc*, if it has a terminal object, binary products and exponentials.

We shall need to consider two slightly different notions of morphism between ccc's, the second a special case of the first. As usual, a functor $F{:}\mathbf{C}\longrightarrow\mathbf{D}$ is said to *preserve terminal objects* if whenever $T$ is terminal in **C**, then $F(T)$ is terminal in **D**. Similarly $F$ *preserves binary products* if whenever $A\xleftarrow{p}P\xrightarrow{q}B$ is a product diagram for $A$ and $B$ in **C**, then $F(A)\xleftarrow{Fp}F(P)\xrightarrow{Fq}F(B)$ is one for $F(A)$ and $F(B)$ in **D**. Such a functor also *preserves exponentials* if whenever $E$ is the exponential of $B$ by $A$ in **C** (via an evaluation morphism $P\xrightarrow{e}B$ with product diagram $E\xleftarrow{p}P\xrightarrow{q}A$), then $F(E)$ is the exponential of $F(B)$ by $F(A)$ in **D** (via $Fe$, $Fp$ and $Fq$). Then if **C** and **D** are ccc's, a functor $F{:}\mathbf{C}\longrightarrow\mathbf{D}$ will be called a *morphism of ccc's* if it preserves the terminal object, binary products and exponentials. $F$ is a *strict morphism of ccc's* if furthermore it sends the given terminal object, binary products (and projections) and exponentials (and evaluation morphisms) in **C** to the given ones in **D**. Thus for a strict morphism one has for example, that $\langle F1_A,F1_B\rangle$ is the identity on $F(A\times B)=F(A)\times F(B)$, whereas for a morphism one has only that $\langle F1_A,F1_B\rangle$ is an isomorphism.

The category of small ccc's and strict morphisms of ccc's will be denoted **Ccc**. The forgetful functor to the category of sets which takes a small ccc to its underlying set of objects will be denoted $ob{:}\mathbf{Ccc}\longrightarrow\mathbf{Set}$.

**2.2. Definition.** A *2T$\lambda$C-hyperdoctrine* **P** is specified by:

(*i*) a small category $|\mathbf{P}|$ with terminal object and binary products;

(*ii*) a distinguished object $U$ in $|\mathbf{P}|$ which generates the other objects, $I$, via finite products, i.e. each $I$ is $U^n$ for some $n\in\mathbb{N}$ (including the case $n=0$, when $U^0=\top$);

(*iii*) a contravariant functor $|\mathbf{P}|^{op}\longrightarrow\mathbf{Ccc}$, such that the composition $|\mathbf{P}|^{op}\longrightarrow\mathbf{Ccc}\xrightarrow{ob}\mathbf{Set}$ is the representable functor $|\mathbf{P}|(-,U)$; the ccc assigned to an object $I$ in $|\mathbf{P}|$ by the functor will be denoted simply by $\mathbf{P}(I,U)$, and the strict ccc morphism assigned to $\alpha{:}I\longrightarrow J$ in $|\mathbf{P}|$ will be denoted $\alpha^*{:}\mathbf{P}(J,U)\longrightarrow\mathbf{P}(I,U)$;

(*iv*) for each object $I$ of $|\mathbf{P}|$, a functor $\prod_I:\mathbf{P}(I\times U,U)\longrightarrow\mathbf{P}(I,U)$ right adjoint to the functor $\pi_I^*:\mathbf{P}(I,U)\longrightarrow\mathbf{P}(I\times U,U)$, such that these functors $\prod_I$ are natural in $I$, i.e. for any $\alpha:I\longrightarrow J$ $\alpha^*\circ\prod_J=\prod_I\circ(\alpha\times 1_U)^*$.

$|\mathbf{P}|$ is the *base* category of the $2T\lambda C$-hyperdoctrine $\mathbf{P}$, and the ccc $\mathbf{P}(I,U)$ is the *fibre* over an object $I$ in the base. Note that the identity on $U$, regarded as on object of $\mathbf{P}(U,U)$, is a *generic* object sense that any object $A$ in any fibre $\mathbf{P}(I,U)$ is equal to $\alpha^*(1_U)$ for some $\alpha:I\longrightarrow U$ (namely $\alpha=A$). Note also that for *any* object $U^n$ of $|\mathbf{P}|$, $\pi_I^*:\mathbf{P}(-,U)\longrightarrow\mathbf{P}(-\times U^n,U)$ has a natural right adjoint, *viz* $\prod_{(-)}\circ\prod_{(-)\times U}\circ\cdots\circ\prod_{(-)\times U^{n-1}}$.

The cartesian closed structure of any $\mathbf{P}(I,U)$ will be denoted by $\top_I$ $\times_I$ and $\rightarrow_I$. In particular when $I=\top$ we get:

$$\ulcorner\top\urcorner =_{def} \top_\top:\top\longrightarrow U; \tag{2.1}$$

and when $I=U^2$ we get:

$$\ulcorner\times\urcorner =_{def}(\pi_1)\times_{U^2}(\pi_2):U\times U\longrightarrow U \quad\text{and}\quad \ulcorner\rightarrow\urcorner =_{def}(\pi_1)\rightarrow_{U^2}(\pi_2):U\times U\longrightarrow U, \tag{2.2}$$

where the $\pi_i\in\mathbf{P}(U^2,U)$ are the projections. The fact that the operations $(-)^*$ are strict ccc morphisms implies that for any object $I$ of $|\mathbf{P}|$

$$\top_I=\ulcorner\top\urcorner\circ I, \tag{2.3}$$

and for any $A,B\in\mathbf{P}(I,U)$

$$A\times_I B=\ulcorner\times\urcorner\circ\langle A,B\rangle \quad\text{and}\quad A\rightarrow_I B=\ulcorner\rightarrow\urcorner\circ\langle A,B\rangle. \tag{2.4}$$

In $\mathbf{P}(\top,U)$ there are morphisms

$$\ulcorner\mathrm{Fst}\urcorner:\ulcorner\top\urcorner\longrightarrow\prod_\top\prod_{\top\times U}\big((\pi_1)\times_{\top\times U^2}(\pi_2)\rightarrow_{\top\times U^2}(\pi_1)\big), \tag{2.5}$$

$$\ulcorner\mathrm{Snd}\urcorner:\ulcorner\top\urcorner\longrightarrow\prod_\top\prod_{\top\times U}\big((\pi_1)\times_{\top\times U^2}(\pi_2)\rightarrow_{\top\times U^2}(\pi_2)\big) \tag{2.6}$$

obtained from the two projection morphisms for $(\pi_1)\times_{\top\times U^2}(\pi_2)$ by transposing across the exponential adjunction and across the adjunctions for $\prod$. Finally, note that for any $I\in|\mathbf{P}|$, $A\in\mathbf{P}(I\times U,U)$ and $B\in\mathbf{P}(I,U)$, there is an internal product projection morphism

$$\pi_B:\prod_I A\longrightarrow A\circ\langle 1_I,B\rangle \tag{2.7}$$

in $\mathbf{P}(I,U)$ given by $\pi_B=\langle 1_I,B\rangle^*(\epsilon_A)$ where $\epsilon_A:\pi_I^*(\prod_I A)\longrightarrow A$ is the counit of the adjunction $\pi_I^*\dashv\prod_I$ at $A$.

The theory of hyperdoctrines is part of the wider theory of *fibred* categories. (See [B] and the references there.) The following remarks on particular aspects of Definition 2.2 assume some familiarity with this theory:

**2.3. Remark.** The kind of hyperdoctrine defined above is "stricter" than usual, in the sense that all the parts of the structure which are normally asked to commute up to (canonical) isomorphisms, are here asked to commute up to equality; in particular $\mathbf{P}$ determines a fibration over $|\mathbf{P}|$ which is *split* (and has various other properties). A laxer notion still suitable for the semantics of $2T\lambda C$-theories, would be that of a category $\mathbf{P}$, fibred over a base category $|\mathbf{P}|$ (with finite products), with stably cartesian closed fibres, containing an object $G$ which is generic in the sense that any other object can be obtained up to isomorphism from $G$ by change of base along a cartesian morphism, and such that **P** has $U$-indexed products where $U \in |\mathbf{P}|$ is the object underlying $G$. In fact it is the case that every such **P** is equivalent over $|\mathbf{P}|$ to a $2T\lambda C$-hyperdoctrine, and the latter notion is often more convenient to work with. However there is one aspect of Definition 2.2 which is not mere convenience, but crucial for the construction of a topos model from a $2T\lambda C$-hyperdoctrine to be given in section 4. This is the requirement that the right adjoint functor $\prod_I$ be natural in $I$ rather than just pseudo-natural. (As noted by Coquand and Ehrhard [CE], Seely's notion of "PL category" is "strict" in all respects except this one, and in giving an equational presentation of the theory of PL categories they add in the naturality of the $\prod$-functors.)

Next we indicate how a $2T\lambda C$-hyperdoctrine **P** provides a semantics for $2T\lambda C$-theories. Roughly speaking, *the denotations of types are objects in the fibres of* **P** (which are particular morphisms in $|\mathbf{P}|$) *and the denotations of terms are morphisms in the fibres.* There is a third syntactic category, variously called "kinds" or "orders", which we have not so far mentioned since for the second order typed lambda calculus the kinds are just finite powers of the basic kind "*Type*". *The denotations of kinds are objects in* $|\mathbf{P}|$, and in particular "*Type*" *is denoted by* $U$. To make these statements more precise, we first have to specify an interpretation of the underlying language of the types and terms and take account of the free type and individual variables they may have:

**2.4. Structures and models.** Let **P** be a $2T\lambda C$-hyperdoctrine and $ar: C \longrightarrow \mathbb{N}$ some multiset of type constructor symbols. A *$C$-structure* $M$ in **P** is specified by giving, for each $F \in C$, an element $MF \in \mathbf{P}(U^{ar(F)}, U)$. Then for a type $\Phi$ over $C$ and a finite list $\underline{X} = X_0, \ldots, X_{n-1}$ of type variables containing the free type varables of $\Phi$, define

$$[\![\Phi, \underline{X}]\!]_M \in \mathbf{P}(U^n, U)$$

by structural recursion:

- $[\![X_i, \underline{X}]\!]_M = \pi_i$, the $i^{\text{th}}$ projection morphism;
- $[\![F(\Phi_0, \ldots, \Phi_{m-1})]\!]_M = MF \circ \langle [\![\Phi_0, \underline{X}]\!]_M, \ldots, [\![\Phi_{m-1}, \underline{X}]\!]_M \rangle$;
- $[\![\top, \underline{X}]\!]_M = \top_{U^n}$;
- $[\![\Phi * \Psi, \underline{X}]\!]_M = [\![\Phi, \underline{X}]\!]_M *_{U^n} [\![\Psi, \underline{X}]\!]_M$ , where $* = \times$ or $\rightarrow$;
- $[\![\prod X.\Phi, \underline{X}]\!]_M = \prod_{U^n} [\![\Phi, \underline{X}, X]\!]_M$.

Note that if $\Phi$ is a *closed* type, then we can take $\underline{X} = \langle\rangle$ the empty list, and get $[\![\Phi, \langle\rangle]\!]_M$ in $\mathbf{P}(\top, U)$: *we will write* $[\![\Phi]\!]_M$ *for* $[\![\Phi, \langle\rangle]\!]_M$. By structural induction, one can show that *substitution of types for type variables is interpreted using composition in* $|\mathbf{P}|$:

$$[\![ [X := \Psi]\Phi, \underline{X}]\!]_M = [\![\Phi, \underline{X}, X]\!]_M \circ \langle 1_{U^n}, [\![\Psi, \underline{X}]\!]_M \rangle. \tag{2.8}$$

Similarly, if $X \notin FTV(\Phi)$, then

$$[\![\Phi, \underline{X}, X]\!]_M = \pi_1^* [\![\Phi, \underline{X}]\!]_M \text{ , where } \pi_1 : U^n \times U \longrightarrow U^n. \tag{2.9}$$

If now $L=(C,K)$ is a language (as defined in 1.2), an *L-structure* $M$ in $\mathbf{P}$ is specified by a $C$-structure together with, for each $a{:}\Phi$ in $K$, a global element of the object $[\![\Phi]\!]_M$, i.e. a morphism $Ma{:}\top_\top \longrightarrow [\![\Phi]\!]_M$ in $\mathbf{P}(\top,U)$. Then for a $2T\lambda C$ term $t$ over $L$ of type $\Phi$ say, together with a finite list $\underline{x}=x_0^{\Phi_0},\ldots,x_{m-1}^{\Phi_{m-1}}$ of distinct individual variables containing $FIV(t)$ and with a finite list $\underline{X}=X_0,\ldots,X_{n-1}$ of distinct type variables containing $FTV(t)$ and $FTV(\Phi_0),\ldots,FTV(\Phi_{m-1})$, define

$$[\![t,\underline{x},\underline{X}]\!]_M : [\![\Phi_0,\underline{X}]\!]_M\times\cdots\times[\![\Phi_{m-1},\underline{X}]\!]_M \longrightarrow [\![\Phi,\underline{X}]\!]_M$$

in $\mathbf{P}(U^n,U)$ by structural recursion:

- $[\![x_i^{\Phi_i},\underline{x},\underline{X}]\!]_M=\pi_i$, the $i^{\text{th}}$ projection morphism;
- $[\![\langle\rangle,\underline{x},\underline{X}]\!]_M=[\![\Phi_0,\underline{X}]\!]_M\times\cdots\times[\![\Phi_{m-1},\underline{X}]\!]_M$, the unique morphism to $\top_{U^n}$;
- $[\![\mathsf{Fst},\underline{x},\underline{X}]\!]_M=(U^n)^*(\ulcorner\mathsf{Fst}\urcorner)\circ([\![\Phi_0,\underline{X}]\!]_M\times\cdots\times[\![\Phi_{m-1},\underline{X}]\!]_M)$, with $\ulcorner\mathsf{Fst}\urcorner$ as in (2.5);
- $[\![\mathsf{Snd},\underline{x},\underline{X}]\!]_M=(U^n)^*(\ulcorner\mathsf{Snd}\urcorner)\circ([\![\Phi_0,\underline{X}]\!]_M\times\cdots\times[\![\Phi_{m-1},\underline{X}]\!]_M)$, with $\ulcorner\mathsf{Snd}\urcorner$ as in (2.6);
- $[\![a,\underline{x},\underline{X}]\!]_M=(U^n)^*(Ma)\circ([\![\Phi_0,\underline{X}]\!]_M\times\cdots\times[\![\Phi_{m-1},\underline{X}]\!]_M)$, for $a\in K$;
- $[\![\langle s,t\rangle,\underline{x},\underline{X}]\!]_M=\langle[\![s,\underline{x},\underline{X}]\!]_M,[\![t,\underline{x},\underline{X}]\!]_M\rangle$;
- $[\![st,\underline{x},\underline{X}]\!]_M=ev\circ\langle[\![s,\underline{x},\underline{X}]\!]_M,[\![t,\underline{x},\underline{X}]\!]_M\rangle$;
- $[\![s\Psi,\underline{x},\underline{X}]\!]_M=\pi_{[\![\Psi,\underline{X}]\!]_M}\circ[\![s,\underline{x},\underline{X}]\!]_M$, with $\pi_{[\![\Psi,\underline{X}]\!]_M}$ as in (2.7) (and using (2.8));
- $[\![\lambda x^\Phi.s,\underline{x},\underline{X}]\!]_M=\overline{f}$, the exponential transpose of $f=[\![s,\underline{x},x,\underline{X}]\!]_M$;
- $[\![\lambda X.s,\underline{x},\underline{X}]\!]_M=\overline{g}$, the transpose across the adjunction $\pi_1^*\dashv\prod_{U^n}$ of
$$g=[\![s,\underline{x},\underline{X},X]\!]_M : \pi_1^*\Big(\prod_{i<m}[\![\Phi_i,\underline{X}]\!]_M\Big)=\prod_{i<m}[\![\Phi_i,\underline{X},X]\!]_M\longrightarrow[\![\Phi,\underline{X},X]\!]_M$$

(where the last clause makes use of (2.9)).

We shall say that the $L$-structure $M$ *satisfies* an equality judgement $\vdash_{\underline{X},\underline{x}} s=t{:}\Phi$ and write

$$M\models_{\underline{X},\underline{x}} s=t{:}\Phi\ ,$$

if $[\![s,\underline{x},\underline{X}]\!]_M=[\![t,\underline{x},\underline{X}]\!]_M$ in $\mathbf{P}(U^n,U)$. Then if $\mathbf{T}=(L,A)$ is a $2T\lambda C$-theory, we will say that an $L$-structure $M$ is a *model* of $\mathbf{T}$ if it satisfies all of the equality judgements which comprise the set $A$ of axioms of $\mathbf{T}$. One easily proves:

**2.5. Soundness Lemma.** *The above definition of satisfaction of an equality judgement by a structure in a $2T\lambda C$-hyperdoctrine is sound for the equational logic of* 1.4. *Thus if* $\mathbf{T}$ *is a $2T\lambda C$-theory and $M$ is a model of* $\mathbf{T}$ *in a $2T\lambda C$-hyperdoctrine* $\mathbf{P}$, *then for any equality judgement one has: if* $\mathbf{T}\vdash_{\underline{X},\underline{x}} s=t{:}\Phi$ *then* $M\models_{\underline{X},\underline{x}} s=t{:}\Phi$ .

□

**2.6. Classifying hyperdoctrines and generic models.** If $\mathbf{T}=(L,A)$ is a $2T\lambda C$-theory, then the $2T\lambda C$ types and terms over $L$ can be used to construct a $2T\lambda C$-hyperdoctrine, called the *classifying* hyperdoctrine of $\mathbf{T}$, denoted $\mathbf{P_T}$ and defined as follows:

The objects of $|\mathbf{P_T}|$ are in bijection with the natural numbers: the object corresponding to $n\in\mathbb{N}$ will be denoted $U^n$, and we will write $U$ for $U^1$ and $\top$ for $U^0$.

Morphisms $U^n\longrightarrow U^m$ in $|\mathbf{P_T}|$ are $m$-tuples of morphisms $U^n\longrightarrow U$; and the latter are equivalence classes $[\Phi,\underline{X}]$ of pairs $(\Phi,\underline{X})$, where $\Phi$ is a type, $\underline{X}$ is a finite list of distinct type variables containing $FTV(\Phi)$ and the equivalence relation on such pairs is that of $\alpha$-equivalence, i.e. $(\Phi,\underline{X})$ is equivalent to $([\underline{X}:=\underline{X}']\Phi,\underline{X}')$. Composition in $|\mathbf{P_T}|$ is given by substitution of types for type variables.

The objects of each fibre $\mathbf{P_T}(U^n,U)$ are necessarily the morphisms $U^n \longrightarrow U$ in $|\mathbf{P_T}|$. Given two such objects, $[\Phi,\underline{X}]$ and $[\Psi,\underline{Y}]$ say, a morphism $[\Phi,\underline{X}]\longrightarrow[\Psi,\underline{Y}]$ in $\mathbf{P_T}(U^n,U)$ is an equivalence class $[t,x^{\Phi},\underline{X},\underline{Y}]$, where $t{:}\Psi$ is a term with $FIV(t)\subseteq\{x^{\Phi}\}$, $FTV(t)\subseteq\underline{X},\underline{Y}$ (we assume $\underline{X}$ distinct from $\underline{Y}$), and where $(t_1,x_1^{\Phi},\underline{X}_1,\underline{Y}_1)$ is equivalent to $(t_2,x_2^{\Phi},\underline{X}_2,\underline{Y}_2)$ iff $\mathbf{T}\vdash_{\langle\rangle,\langle\rangle} \lambda\underline{X}_1\underline{Y}_1.\lambda x_1^{\Phi}.t_1 = \lambda\underline{X}_2\underline{Y}_2.\lambda x_2^{\Phi}.t_2 : \prod\underline{X}\underline{Y}.\Phi\to\Psi$.

Composition is given by substitution of terms for individual variables and the functors $(-)^*$ between the fibres are given by substitution of types for type variables in both types and terms. The terminal object is $[\top,\underline{X}]$; the product of $[\Phi,\underline{X}]$ and $[\Psi,\underline{X}]$ is $[\Phi\times\Psi,\underline{X}]$; and their exponential is $[\Phi\to\Psi,\underline{X}]$.

The right adjoint to $\pi_1^*{:}\mathbf{P_T}(U^n,U)\longrightarrow\mathbf{P_T}(U^n\times U,U)$ sends $[\Phi,\underline{X},X]$ to $[\prod X.\Phi,\underline{X}]$ and $[t,x^{\Phi},\underline{X},X]$ to $[\lambda X.([x^{\Phi}{:=}z\,X]t),z^{\prod X.\Phi},\underline{X}]$.

The verification that the above recipe does give a $2T\lambda C$-hyperdoctrine is a straightforward exercise. (See also [Se, Proposition 4.6] for the full higher order case and [LS, I.11] for the first order case.) Almost tautologically $\mathbf{P_T}$ contains a model of $\mathbf{T}$, which we shall call the *generic* model of $\mathbf{T}$ and denote by $I_{\mathbf{T}}$. The underlying $L$-structure of $I_{\mathbf{T}}$ sends a type constructor $F$ of arity $n$ to $[F(\underline{X}),\underline{X}]{:}U^n\longrightarrow U$ and an individual constant $a$ of (closed) type $\Phi$ to $[a,x^{\top},\langle\rangle]{:}[\top,\langle\rangle]\longrightarrow[\Phi,\langle\rangle]$. It follows from the definitions of $\mathbf{P_T}$ and $I_{\mathbf{T}}$ that:

$$\mathbf{T}\vdash_{\underline{X},\underline{x}} s=t{:}\Phi \text{ iff } I_{\mathbf{T}}\models_{\underline{X},\underline{x}} s=t{:}\Phi \; . \qquad (2.10)$$

In particular, the notion of a model in a $2T\lambda C$-hyperdoctrine is complete for the equational logic of 1.4:

*An equality judgement is a theorem of a $2T\lambda C$-theory iff it is satisfied by all models of the theory in $2T\lambda C$-hyperdoctrines (iff it is satisfied by the generic model).*

However, the generic model $I_{\mathbf{T}}$ has a much stronger property than just (2.10), namely:

*Any model of* **T** *in any $2T\lambda C$-hyperdoctrine* **P** *can be obtained up to isomorphism as the image of the generic model $I_{\mathbf{T}}$ along an essentially unique morphism of $2T\lambda C$-hyperdoctrines $\mathbf{P_T}\longrightarrow\mathbf{P}$.*

(A 2-hyperdoctrine morphism $F{:}\mathbf{P}\longrightarrow\mathbf{Q}$ is specified by a finite product preserving functor $|F|{:}|\mathbf{P}|\longrightarrow|\mathbf{Q}|$ sending the $U\in|\mathbf{P}|$ to the $U\in|\mathbf{Q}|$, and by a natural transformation $F_{(-)}{:}\mathbf{P}(-,U)\longrightarrow\mathbf{Q}(|F|(-),U)$ whose component functors are ccc morphisms commuting with the right adjoints $\prod_I$.)

Thus $\mathbf{P_T}$ is the $2T\lambda C$-hyperdoctrine "freely generated" by the $2T\lambda C$-theory $\mathbf{T}$. Conversely, up to equivalence every $2T\lambda C$-hyperdoctrine $\mathbf{P}$ can be presented as the classifying hyperdoctrine of some $2T\lambda C$-theory $\mathbf{T}$: for the set $C$ of type constructor symbols, for each arity $n$ take one symbol $\ulcorner F\urcorner$ for each morphism $F{:}U^n\longrightarrow U$ in $|\mathbf{P}|$, giving an evident $C$-structure $M{:}\ulcorner F\urcorner\mapsto F$; for the set of individual constants $K$, for each closed type $\Phi$ over

$C$ and each $a{:}\mathrm{T}_{\mathrm{T}} \longrightarrow M\Phi$ in $\mathbf{P}(\mathrm{T},U)$ take a symbol $\ulcorner a \urcorner{:}\Phi$, giving a language $L=(C,K)$ and an evident $L$-structure $M$ in $\mathbf{P}$ with $M(\ulcorner a \urcorner)=a$; then let $\mathbf{T}$ be the $2T\lambda C$-theory $(L,A)$, where $A$ consists of all the equality judgements which are satisfied by the structure $M$. Evidently $M$ is a model of $\mathbf{T}$, and one can show that the corresponding $2T\lambda C$-hyperdoctrine morphism $\mathbf{P}_{\mathbf{T}} \longrightarrow \mathbf{P}$ is an equivalence.

**2.7. Summary.** The notion of satisfaction of an equality judgement by a structure in a $2T\lambda C$-hyperdoctrine is both sound and complete for equational logic over the second order typed lambda calculus. More importantly, the classifying hyperdoctrine construction sets up a correspondence between $2T\lambda C$-theories and $2T\lambda C$-hyperdoctrines which allows us to view the latter notion as a "presentation-free" version of the former. Something has definitely been achieved in this transfer from theories to hyperdoctrines, since $2T\lambda C$-hyperdoctrines are quite elementary kinds of structure: indeed they are models of a particular, essentially equational theory and are thus amenable to study using algebraic and category-theoretic techniques. Thus in proving the completeness and full embedding theorems of section 4, we will be working not with $2T\lambda C$-theories, but with the hyperdoctrines to which they correspond and applying category theoretic constructions to these. The equational aspect of (higher order) hyperdoctrines is emphasised by Coquand and Ehrhard [CE], and of course the first order part of this (the equational presentation of the notion of ccc) lies at the heart of recent work of Cousineau, Curien and Mauny [CCM, Cu].

# 3 Topos models

In this section we assume some familiarity with the theory of toposes and particularly with the use of higher order intuitionistic predicate logic to describe properties of and make constructions in a topos via its internal language: see [J, section 5.4] and [LS, Part II]. Also implicit to the material in this section is the use of fibred (and indexed) categories over a particular topos to provide an *elementary theory of categories* (both large and small) *relative to the topos.* The parts of this (not yet fully developed) theory we shall need are sufficiently simple for them to be given explicitly in terms of more "traditional" topos theory (by which we mean [J] up to, but not including its appendix). [PS] provides a lot of material on indexed category theory, and a taste of the wider aspects of the theory can be got from [B].

The reason why we have to go slightly beyond the traditional internal logic of a topos is simple: the syntax of $2T\lambda C$-theories invovles *variable types* and to model this in a topos $\mathbf{E}$ we will consider the *generalized objects* of $\mathbf{E}$. To understand what is meant by this, recall first the notion of a generalized element of an object $B$ in **E**: for each object $I$ of **E**, the *generalized elements of* $B$ *at stage* $I$ are the morphisms $b:I \longrightarrow B$ in **E**, and they are used to give the Kripke-Joyal forcing semantics of intuitionistic higher order predicate logic in toposes (see [LS, II.8]); the *global* elements of $B$ are morphisms $\top \longrightarrow B$, but in general these are insufficient for detecting properties of the object $B$. Moving up a level, the collection of objects of **E** (global objects, one might call them), are generally insufficient for detecting all the properties of **E** as a constructive universe of "sets". Instead one must consider, for each object $I$, the *generalized objects of* **E** *at stage* $I$, which are by definition morphisms $p:E \longrightarrow I$ in **E** with codomain $I$. The idea behind this definition is that when **E**=**Set**, such $p:E \longrightarrow I$ correspond precisely to $I$-indexed collections of sets, $(E_i | i \in I)$ (by defining $E_i = \{e \in E \mid p(e) = i\}$), just as generalized elements $b:I \longrightarrow B$ correspond to $I$-indexed collections of elements of $B$, $(b(i) \mid i \in I)$.

Just as the generalized elements of $B$ at stage $I$ form a set $\mathbf{E}(I,B)$, so the generalized objects of **E** at stage $I$ form a category (a topos in fact), namely the *slice* category $\mathbf{E}/I$ whose morphisms are commutative triangles over $I$. We begin by fixing some notation for these:

**3.1. Slice categories.** Let **C** be a category. For each object $I$ of **C**, let $\Sigma_I : \mathbf{C}/I \longrightarrow \mathbf{C}$ denote the *forgetful functor from the slice category* $\mathbf{C}/I$: thus a typical object of $\mathbf{C}/I$ is a morphism in **C** of the form $p:\Sigma_I(p) \longrightarrow I$ and a typical morphism $f:p \longrightarrow q$ in $\mathbf{C}/I$ is given by a morphism $f = \Sigma_I(f):\Sigma_I(p) \longrightarrow \Sigma_I(q)$ in **C** satisfying $q \circ f = p$. The identity on $I$ in **C** is a terminal object in $\mathbf{C}/I$. The binary product of $p$ and $q$ in $\mathbf{C}/I$ will be denoted $p \times_I q$ if it exists, in which case it is necessarily given by a pullback square in **C**:

$$\begin{array}{ccc} \Sigma_I(p \times_I q) & \longrightarrow & \Sigma_I(q) \\ \downarrow & & \downarrow{\scriptstyle q} \\ \Sigma_I(p) & \underset{p}{\longrightarrow} & I \end{array} .$$

Similarly, the exponential of $p$ and $q$ in $\mathbf{C}/I$ will be denoted $p \rightarrow_I q$ if it exists and called a *local exponential in* **C**. A morphism $\alpha:I \longrightarrow J$ in **C** will be called *squarable* if the pullback along $\alpha$ of any morphism with codomain $J$ exists in **C**. For such an $\alpha$, the operation of pulling back along $\alpha$ gives a functor between slice categories, which will be denoted

$$\alpha^* : \mathbf{C}/J \longrightarrow \mathbf{C}/I \ .$$

If the right adjoint to $\alpha^*$ exists at an object $p$ of $\mathbf{C}/I$, it will be denoted $\Pi_\alpha(p)$: there is thus a morphism $\varepsilon:\alpha^*(\Pi_\alpha(p)) \longrightarrow p$ in $\mathbf{C}/I$ with the universal property that for any $f:\alpha^*(q) \longrightarrow p$ there is a unique $\overline{f}:q \longrightarrow \Pi_\alpha(p)$ in $\mathbf{C}/J$ with $f = \varepsilon \circ \alpha^*(\overline{f})$. Of course when **C** is a topos it has all pullbacks, local exponentials and right adjoints to pulling back along a morphism (see [J, 1.4]); in particular, when **C**=**Set** $p:E \longrightarrow I$, and $q:F \longrightarrow I$, then for each $i \in I$

$$(\Sigma_I(p \rightarrow_I q))_i = E_i \rightarrow F_i \ ,$$ an exponential in **Set**,

and for each $j \in J$

$(\Sigma_J(\Pi_\alpha p))_j = \prod\{E_i \mid \alpha(i)=j\}$ , a cartesian product in **Set**.

If translated into intuitionistic higher order predicate logic, these formulas remain true for the internal logic of an arbitrary topos **E**. Thus for $i$ a variable of type $I$, let $E_i$ be an abbreviation for $\{e{:}E \mid p(e)=i\}$, etc; then **E** satisfies:

$$(\Sigma_I(p \to_I q))_i \cong \{r{:}\Omega^{E\times F} \mid \text{"}r \text{ is the graph of a function from } E_i \text{ to } F_i\text{"}\} \tag{3.1}$$

$$(\Sigma_J(\Pi_\alpha p))_j \cong \{r{:}\Omega^{I\times E} \mid \text{"}r \text{ is the graph of a partial function from } \{i{:}I \mid \alpha(i)=j\} \text{ to } E\text{" and } \forall i{:}I(\alpha(i)=j \Rightarrow r(i)\in E_i)\} \tag{3.2}$$

**3.2. Internal full subcategories.** Consider the following construction:

> Starting with a set $U$ and a $U$-indexed collection of sets $(G_u \mid u\in U)$, form the small category **U** with underlying set of objects $U$, with hom sets $\mathbf{U}(u,v) = \mathbf{Set}(G_u,G_v)$ and with composition and identities inherited from **Set**; by construction, the assignment $u \mapsto G_u$ extends to a diagram of type **U** in **Set** which is a full and faithful functor $G{:}\mathbf{U} \hookrightarrow \mathbf{Set}$.

Note that specifying the initial data for this construction is equivalent to giving a single morphism $\tau{:}G \longrightarrow U$ in **Set**, where $G$ is the disjoint union $\bigcup\{\{u\}\times G_u \mid u\in U\}$ and $\tau{:}(u,x)\mapsto u$ (so that $G_u \cong \tau^{-1}\{u\}$). Now starting with any topos **E** and a morphism $\tau{:}G\longrightarrow U$ in **E**, it is possible to carry out the analogue of the above construction, obtaining an internal category object **U** in **E** together with a full and faithful internal diagram $G$ of type **U** in **E**. This is achieved as follows:

Pull back $\tau$ along the two projections $\pi_i{:}U\times U \longrightarrow U$, to get $\pi_0^*(\tau) = \tau\times 1_U{:}G\times U \longrightarrow U\times U$ and $\pi_1^*(\tau) = 1_U\times\tau{:}U\times G\longrightarrow U\times U$ in $\mathbf{E}/U^2$. Now form the local exponential $\pi_0^*(\tau)\to_{U^2}\pi_1^*(\tau)$ and suppose it is given by the morphism $\langle d_0,d_1\rangle{:}U_1\longrightarrow U\times U$ in **E**; then in the internal language of **E** one has $(G\times U)_{\langle u,v\rangle}\cong G_u$ and $(U\times G)_{\langle u,v\rangle}\cong G_v$, so that by (3.1) $(U_1)_{\langle u,v\rangle}\cong G_u\to G_v$ (where $u$ and $v$ are variables of type $U$). It follows that

$$d_0,d_1{:}U_1 \rightrightarrows U$$

is the underlying graph of a category object **U** in **E** (with composition in **U** being given by the composition morphism $((G_v\to G_w)\times(G_u\to G_v)\longrightarrow(G_u\to G_w) \mid u,v,w{:}U)$ in $\mathbf{E}/U^3$). Then $\tau{:}G\longrightarrow U$ becomes an internal diagram of type **U** via the action $((U_1)_{\langle u,v\rangle}\times G_u \longrightarrow G_v \mid u,v{:}U)$ which (using the above identifications of internal fibres) is given by the evaluation morphism $ev{:}(\pi_0^*(\tau)\to_{U^2}\pi_1^*(\tau))\times_{U^2}\pi_0^*(\tau)\longrightarrow\pi_1^*(\tau)$ in $\mathbf{E}/U^2$. The transpose $((U_1)_{\langle u,v\rangle}\longrightarrow(G_u\to G_v) \mid u,v\in U)$ of this action across the exponential adjunction, gives the effect of $G$ on morphisms of **U**; and by definition of $U_1$, this is an isomorphism — and hence $G$ is full and faithful. **U** will be called the *internal full subcategory of* **E** *determined by* $\tau{:}G\longrightarrow U$, and the diagram $G\in\mathbf{E}^{\mathbf{U}}$ the *inclusion of* **U** *into* **E**.

We are interested in such internal full subcategories which are *closed in* **E** *under certain internal products*. In order to understand precisely what is meant by this, it is important first to note that our terminology is slightly misleading, since an internal full subcategory is not a generalized collection of generalized objects of **E** (a concept which can be formalized by Bénabou's notion in [B] of "definable class of objects" of a topos), but rather a single object whose variable elements $u$ *name* variable objects $G_u=\{g:G \mid \tau(g)=u\}$. Then if we assert for example, that **U** is *closed in* **E** *under binary products*, we mean that there is a morphism $\ulcorner\times\urcorner:U^2\longrightarrow U$ and a pullback square in **E** of the form:

$$\begin{array}{ccc} \Sigma_{U^2}(\pi_0^*(\tau)\times_{U^2}\pi_1^*(\tau)) & \dashrightarrow & G \\ {\scriptstyle \pi_0^*(\tau)\times_{U^2}\pi_1^*(\tau)}\Big\downarrow & & \Big\downarrow{\scriptstyle \tau} \\ U^2 & \xrightarrow[\ulcorner\times\urcorner]{} & U\ . \end{array} \qquad (3.3)$$

Since the pullback of $\tau$ along $\ulcorner\times\urcorner$ has internal fibres $(G_{\ulcorner\times\urcorner(u,v)} \mid u,v:U)$, the above pullback square furnishes an internal family of isomorphisms $G_{\ulcorner\times\urcorner(u,v)}\cong G_u\times G_v$ $(u,v:U)$, so that $\ulcorner\times\urcorner(u,v)$ names the product of the objects named by $u$ and $v$. Similarly, **U** *is closed in* **E** *under exponentiation* if there is a morphism $\ulcorner\rightarrow\urcorner:U^2\longrightarrow U$ and a pullback square of the form:

$$\begin{array}{ccc} \Sigma_{U^2}(\pi_0(\tau)\rightarrow_{U^2}\pi_1(\tau)) & \dashrightarrow & G \\ {\scriptstyle \pi_0^*(\tau)\rightarrow_{U^2}\pi_1^*(\tau)}\Big\downarrow & & \Big\downarrow{\scriptstyle \tau} \\ U^2 & \xrightarrow[\ulcorner\rightarrow\urcorner]{} & U\ . \end{array} \qquad (3.4)$$

**U** *contains the terminal object* of **E** if there is a morphism $\ulcorner\top\urcorner:\top\longrightarrow U$ and a pullback square of the form:

$$\begin{array}{ccc} \top & \dashrightarrow & G \\ {\scriptstyle 1_\top}\Big\downarrow & & \Big\downarrow{\scriptstyle \tau} \\ \top & \xrightarrow[\ulcorner\top\urcorner]{} & U\ . \end{array} \qquad (3.5)$$

Finally, **U** *is closed in* **E** *under $U$-indexed products* if there is a morphism $\ulcorner\Pi\urcorner:(U\rightarrow U)\longrightarrow U$ and a pullback of the form:

$$\begin{array}{ccc} \Sigma_{U\rightarrow U}(\Pi_p ev^*(\tau)) & \dashrightarrow & G \\ {\scriptstyle \Pi_p ev^*(\tau)}\Big\downarrow & & \Big\downarrow{\scriptstyle \tau} \\ U\rightarrow U & \xrightarrow[\ulcorner\Pi\urcorner]{} & U \end{array} \qquad (3.6)$$

where $p:(U\rightarrow U)\times U\longrightarrow(U\rightarrow U)$ is the first projection and $ev:(U\rightarrow U)\times U\longrightarrow U$ is the evaluation morphism. Using (*3.2*), one sees that the effect of this last pullback square is to provide an internal family of isomorphisms $G_{\ulcorner\Pi\urcorner(f)}\cong\prod\{G_{f(u)} \mid u:U\}$ indexed by $f$ of type $U\rightarrow U$: thus $\ulcorner\Pi\urcorner(f)$ names the product of the $U$-indexed family of objects named by $f(u)$ for $u:U$. Collecting together the data required to specify such an internal full subcategory, we arrive at the following:

**3.3. Definition.** A *topos model* of the second order typed lambda calculus is given by an elementary topos **E** together with morphisms $\tau:G\longrightarrow U$, $\ulcorner\top\urcorner:\top\longrightarrow U$, $\ulcorner\times\urcorner,\ulcorner\rightarrow\urcorner:U^2\longrightarrow U$ and $\ulcorner\Pi\urcorner:(U\rightarrow U)\longrightarrow U$ in **E** and pullback squares of the form (*3.3*),(*3.4*),(*3.5*) and (*3.6*). **E** will be called the *ambient topos* of the model; $\mathbf{U}$ will denote the internal full subcategory of $\mathbf{E}$ determined by $\tau$.

Topos models *are* models of the second order typed lambda calculus because:

**3.4. Lemma.** *Every topos model of the second order typed lambda calculus gives rise to a $2T\lambda C$-hyperdoctrine* $\mathbf{P}$ *via the category valued hom functor* $\mathbf{E}(-,\mathbf{U})$.

**Proof.** Since the internal full subcategory $\mathbf{U}$ is closed in $\mathbf{E}$ under finite products and exponentials, $\mathbf{U}$ is automatically a ccc object in $\mathbf{E}$. Consequently its generalized elements at any stage $I$ form a ccc (with set of objects $\mathbf{E}(I,U)$ and set of morphisms $\mathbf{E}(I,U_1)$); and for any $\alpha:I\longrightarrow J$, the operation $\alpha^*$ gives a strict morphism of ccc's. Thus $\mathbf{E}(-,\mathbf{U})$ is a functor $\mathbf{E}^{op}\longrightarrow\mathbf{Ccc}$. Letting $|\mathbf{P}|$ be the full subcategory of $\mathbf{E}$ whose objects are the finite powers of $U$, we can restrict this functor to get $\mathbf{P}=\mathbf{E}(-,\mathbf{U}):|\mathbf{P}|^{op}\longrightarrow\mathbf{Ccc}$ satisfying parts $(i)$, $(ii)$ and $(iii)$ of Definition 2.2. For part $(iv)$ of the definition, we use the fact that $\mathbf{U}$ is closed in $\mathbf{E}$ under $U$-indexed products. Thus $\prod_I:\mathbf{P}(I\times U,U)\longrightarrow\mathbf{P}(I,U)$ is defined on objects by sending $u:I\times U\longrightarrow U$ to $\prod_I(u)=\ulcorner\Pi\urcorner\circ\overline{u}$ where $\overline{u}:I\longrightarrow(U\to U)$ is the exponential transpose of $u$. Similarly, $\prod_I$ is defined on morphisms by sending $m:I\times U\longrightarrow U_1$ to $\prod_I(m)=\ulcorner\Pi_1\urcorner\circ\overline{m}$, where $\ulcorner\Pi_1\urcorner:(U\to U_1)\longrightarrow U_1$ sends $U$-indexed collections of morphisms in $\mathbf{U}$ to their product in $\mathbf{E}$, (which again lies in $\mathbf{U}$ since it is a *full* sucategory and closed under such products): in other words, $\ulcorner\Pi_1\urcorner$ is uniquely defined by requiring $d_i\circ\ulcorner\Pi_1\urcorner=\ulcorner\Pi\urcorner\circ(1_U\to d_i)$ for $i=0,1$ and (using the internal language of $\mathbf{E}$) further requiring that for $m:U\to U_1$

$$\ulcorner\Pi_1\urcorner(m):G_{\ulcorner\Pi\urcorner(d_0m)}\to G_{\ulcorner\Pi\urcorner(d_1m)}$$

correspond under the isomorphisms

$$G_{\ulcorner\Pi\urcorner(d_im)}\cong\prod\{G_{d_im(u)}\mid u:U\}\qquad\qquad(i=0,1)$$

to $\prod\{m(u)\mid u:U\}$. Evidently this definition of $\prod_I:\mathbf{P}(I\times U,U)\longrightarrow\mathbf{P}(I,U)$ is natural in $I$. That it gives a right adjoint to $\pi_I^*$ follows from the fact that $\pi_I^*:\mathbf{E}(-,\mathbf{U})\longrightarrow\mathbf{E}(-\times U,\mathbf{U})\cong\mathbf{E}(-,\mathbf{U}^U)$ is the representable functor induced by the diagonal internal functor $\overline{\pi}_I:\mathbf{U}\longrightarrow\mathbf{U}^U$ — and taking $U$-indexed products is right adjoint to this.

□

Since each topos model $\mathbf{E},\mathbf{U}$ determines a $2T\lambda C$-hyperdoctrine, it provides a semantics for the types and terms of the second order typed lambda calculus, with the types being denoted by particular morphisms in $\mathbf{E}$ with codomain $U$ and the terms by morphisms with codomain $U_1$. In particular the definitions in 2.4 specialize to give us the notion of a *topos model of a $2T\lambda C$-theory.*

As we explained in the previous section, arbitrary $2T\lambda C$-hyperdoctrines provide a semantics which is completely general. (Perhaps neutral is a better word.) In contrast, topos models embody a very particular and apparently naive idea of the meaning of the various symbols of the calculus: A topos is itself a model of higher order intuitionistic predicate logic; in a topos model of $2T\lambda C$, the closed types get interpreted as elements of a family $U$ of (names of) "sets" in this logic, and more generally the polymorphic (i.e. non-closed) types are given by *arbitrary* functions from this family to itself; furthermore, $\times$ is interpreted by taking actual *cartesian products*, $\rightarrow$ by taking *full function exponentials* and $\prod$ by taking products of "sets" in the family indexed by the elements of $U$. Thus in this notion of model we allow ourselves to work in a "non-standard" universe of sets but make up for this by insisting that the type forming operations be interpreted in a completely standard way. This is in contrast to more traditional formulations of the notion of model of polymorphic lambda calculus (such as in [BM] or [R2]) which are couched in classical set theory, but allow some of the operations ($\prod$ in particular) to be non-standard. The intersection of these two approaches is trivial: when **E** = **Set**, a topos model is given by a set $U$ of sets closed under finite products, exponentiation and $U$-indexed products; and a simple cardinality argument (using the principles of classical logic !) shows that any such $U$ must have as elements only sets with at most one element. There are therefore no non-trivial topos models of $2T\lambda C$ when the ambient topos is the category of sets.

However the more liberal nature of the internal logic of toposes in general means that non-trivial examples of Definition 3.3 do exist. The first such is due to Hyland and Moggi, with **E** being Hyland's effective topos [H], $U$ the object of $\neg\neg$-closed partial equivalence relations on the natural number object $N$ and $\tau:G \longrightarrow U$ having internal fibres $G_u = Eu/u|_{Eu}$, where $Eu = \{n:N \mid u(n,n)\}$ and $u|_{Eu}$ is the equivalence relation obtained by restriction. Dana Scott has dubbed the (generalized) elements of $U$ the *modest sets*: they have far greater closure properties than those required by Definition 3.3 — as Hyland, Robinson and Rosolini show in [HRR], the limit or colimt of *any* internal diagram of modest sets is again modest. (The object of *all* partial equivalence relations on $N$ in the effective topos enjoys similar properties.) There are in all likelyhood very many interesting topos models of $2T\lambda C$ and related type theories. In the next section we will show how to manufacture topos models from $2T\lambda C$-hyperdoctrines, so that in particular there are *enough* such models to distinguish the theorems of a $2T\lambda C$-theory from the non-theorems.

# 4 Full embedding and completeness

In this section we prove the main result of the paper by showing how each $2T\lambda C$-hyperdoctrine can be expanded to a topos model. To do so we will employ two standard category theoretic tools, namely the *Grothendieck construction* of a split fibration from a category-valued functor and the *Yoneda embedding* of a category into a topos of presheaves.

**4.1. The Grothendieck construction.** Suppose that **B** is a category and that $\mathbf{P}:\mathbf{B}^{op}\longrightarrow\mathbf{Cat}$ is a contravariant functor from **B** into the category of small categories. Construct a new category $Gr(\mathbf{P})$ and a functor $P:Gr(\mathbf{P})\longrightarrow\mathbf{B}$ from **P** as follows:

> The objects of $Gr(\mathbf{P})$ are pairs $(I,A)$, where $I$ is an object of **B** and $A$ an object of $\mathbf{P}(I)$. Given two such objects $(I,A),(J,B)$, the morphisms $(I,A)\longrightarrow(J,B)$ in $Gr(\mathbf{P})$ are given by pairs $(\alpha,f)$, where $\alpha:I\longrightarrow J$ in **B** and $f:A\longrightarrow\mathbf{P}(\alpha)B$ in $\mathbf{P}(I)$. The composition of $(\alpha,f):(I,A)\longrightarrow(J,B)$ and $(\beta,g):(J,B)\longrightarrow(K,C)$ in $Gr(\mathbf{P})$ is $(\beta\circ\alpha,\mathbf{P}(\alpha)(g)\circ f)$. The identity on $(I,A)$ is $(1_I,1_A)$. The functor $P:Gr(\mathbf{P})\longrightarrow\mathbf{B}$ sends an object $(I,A)$ to $I$ and a morphism $(\alpha,f)$ to $\alpha$.

The kind of functors into **B** which arise in this way can be characterized by category theoretic properties: they are the *split fibrations* over **B**. More generally, the construction can be applied to *pseudofunctors* $\mathbf{B}^{op}\longrightarrow\mathbf{Cat}$, setting up a correspondence between these and *cloven fibrations* over **B**. We shall not need to use these concepts explicitly here, and refer the interested reader to [Gr].

Now suppose that **P** is a $2T\lambda C$-hyperdoctrine. Recalling the notation of Definition 2.2, we can apply the Grothendieck construction to the composition of $\mathbf{P}(-,U):|\mathbf{P}|^{op}\longrightarrow\mathbf{Ccc}$ with the forgetful functor $\mathbf{Ccc}\longrightarrow\mathbf{Cat}$; we will denote the resulting category and functor simply by:

$$P:Gr(\mathbf{P})\longrightarrow|\mathbf{P}|\ .$$

Because of the special nature of **P**, note that the objects of $Gr(\mathbf{P})$ are given just by morphisms $A:I\longrightarrow U$ in $|\mathbf{P}|$ with codomain $U$, and that a morphism $(A:I\longrightarrow U)\longrightarrow(B:J\longrightarrow U)$ in $Gr(\mathbf{P})$ is given by a pair $(\alpha,f)$, where $\alpha:I\longrightarrow J$ in $|\mathbf{P}|$ and $f:A\longrightarrow\alpha^*(B)=(B\circ\alpha)$ in $\mathbf{P}(I,U)$. The following properties of $Gr(\mathbf{P})$ and $P$ are easily verified:

**4.2. Lemma.** *Let* **P** *be a $2T\lambda C$-hyperdoctrine. Then:*

(*i*) *$Gr(\mathbf{P})$ has and $P$ preserves finite products. The terminal object in $Gr(\mathbf{P})$ is $(\top_T:\top\longrightarrow U)$ and the binary product of $(A:I\longrightarrow U)$ and $(B:J\longrightarrow U)$ is $(\pi_1^*(A)\times_{I\times J}\pi_2^*(B):I\times J\longrightarrow U)$ (where the $\pi_i$ are the product projections for $I\times J$ in $|\mathbf{P}|$).*

(*ii*) *$P$ has a full and faithful right adjoint $T:|\mathbf{P}|\longrightarrow Gr(\mathbf{P})$, given on objects by sending $I$ to the terminal object in the fibre over $I$, i.e. $T(I)=(\top_I:I\longrightarrow U)$.*

□

We now focus attention on a particular morphism in $Gr(\mathbf{P})$, namely:

$$t=_{def}(1_U,1_U):(1_U:U\longrightarrow U)\longrightarrow(\top_U:U\longrightarrow U)=T(U). \qquad (4.1)$$

(Our convention of using the same letter to denote both an object in a category and its associated morphism to the terminal object has become rather confusing at this point: in (*4.1*) the first "$1_U$" denotes the identity on $U$ in $|\mathbf{P}|$, whereas the second denotes the unique morphism in the ccc $\mathbf{P}(U,U)$ from the object $1_U$ to the terminal object.) Although $Gr(\mathbf{P})$ does not have all pullbacks, the morphism $t$ is squarable, i.e. the pullback of it along any morphism with codomain $T(U)$ exists. This is because such a morphism

$(\alpha,f):(A:I\longrightarrow U)\longrightarrow T(U)$ in $Gr(\mathbf{P})$ necessarily has $f=A:A\longrightarrow \mathrm{T}_I$, and then

$$\begin{array}{ccc}
(A\times_I\alpha:I\longrightarrow U) & \xrightarrow{(\alpha,\pi_2)} & (1_U:U\longrightarrow U) \\
{\scriptstyle (1_I,\pi_1)}\downarrow & & \downarrow{\scriptstyle t} \\
(A:I\longrightarrow U) & \xrightarrow[(\alpha,A)]{} & T(U)
\end{array}$$

is easily verified to be a pullback square in $Gr(\mathbf{P})$. Let $\Sigma$ denote the collection of morphisms in $Gr(\mathbf{P})$ which can be obtained by pullback from $t$: thus $(\alpha,f)$ is in $\Sigma$ iff $\alpha$ is an isomorphism and $f$ a product projection. This class $\Sigma$ inherits good properties from $\mathbf{P}$ with respect to local exponentiation and right adjoints to pulling back (*cf.* 3.1):

**4.3. Lemma.** *If* $\mathbf{P}$ *is a* $2T\lambda C$*-hyperdoctrine and* $\Sigma\subseteq mor(Gr(\mathbf{P}))$ *defined as above, then:*

(*i*) *Any morphism in* $\Sigma$ *is squarable.*

(*ii*) *For any* $p:Y\longrightarrow X$ and $q:Z\longrightarrow X$ *in* $\Sigma$, *their local exponential* $p\rightarrow_X q$ *exists in* $Gr(\mathbf{P})$ *and is an element of* $\Sigma$.

(*iii*) *If* $r:W\longrightarrow X\times T(U)$ *is in* $\Sigma$, *then* $\Pi_{\pi_1}(r)$, *the right adjoint at* $r$ *to the pullback functor* $\pi_1^*:Gr(\mathbf{P})/X\rightarrow Gr(\mathbf{P})/X\times T(U)$, *exists and is an element of* $\Sigma$.

**Proof.** Since all the morphisms in $\Sigma$ are obtained by pullback, (*i*) follows immediately from the fact that a pullback of a pullback is a pullback. Thus for any $(1_J,\pi_1):(B\times_J C:J\longrightarrow U)$ in $\Sigma$ and any $(\alpha,f):(A:I\longrightarrow U)\longrightarrow(B:J\longrightarrow U)$ in $Gr(\mathbf{P})$, the pullback $(\alpha,f)^*(1_J,\pi_1)$ is

$$(1_I,\pi_1):(A\times_I\alpha^*C:I\longrightarrow U)\longrightarrow(A:I\longrightarrow U) \qquad (4.2)$$

For (*ii*), suppose that $X=(A:I\longrightarrow U)$, $Y=(A\times_I B:I\longrightarrow U)$, $Z=(A\times_I C:I\longrightarrow U)$, $p=(1_I,\pi_1:A\times_I B\longrightarrow A)$ and $q=(1_I,\pi_1:A\times_I C\longrightarrow A)$. Then for any morphism in $Gr(\mathbf{P})$ with codomain $X$, $(\alpha,f):(D:J\longrightarrow U)\longrightarrow X$ say, it is easy to see that $(\alpha,f)\times_X p$ is

$$(\alpha,f\circ\pi_1):(D\times_J B\alpha:J\longrightarrow U)\longrightarrow X.$$

Hence specifying a morphism $(\alpha,f)\times_X p\longrightarrow q$ in $Gr(\mathbf{P})/X$ amounts to giving a morphism $D\times_J B\alpha\longrightarrow C\alpha$ in $\mathbf{P}(J,U)$; and transposing across the exponential adjunction, this amounts to giving a morphism $D\longrightarrow(B\alpha\rightarrow_J C\alpha)=\alpha^*(B\rightarrow_I C)$. It follows from this that we can take $p\rightarrow_X q$ to be

$$(1_I,\pi_1):(A\times_I(B\rightarrow_I C):I\longrightarrow U)\longrightarrow X, \qquad (4.3)$$

since morphisms in $Gr(\mathbf{P})/X$ into this from $(\alpha,f)$ are also specified by morphisms $D\longrightarrow\alpha^*(B\rightarrow_I C)$. Note that as required, (*4.3*) is in $\Sigma$.

Similar arguments show that for (*iii*) we can take $\Pi_{\pi_1}(r)$ to be

$$(1_I,\pi_1):(A\times_I\textstyle\prod_I(B):I\longrightarrow U)\longrightarrow X, \qquad (4.4)$$

when $X=(A:I\longrightarrow U)$, $W=(\pi_1^*(A)\times_{I\times U}B:I\times U\longrightarrow U)$ and $r=(1_{I\times U},\pi_1):W\longrightarrow(\pi_1^*(A):I\times U\longrightarrow U)\cong X\times T(U)$.

□

Although $Gr(\mathbf{P})$ is far from being a topos (or even locally cartesian closed), Lemma 4.3 means that the construction given in 3.2 (which uses pullbacks and local exponentials) can be carried out starting with the morphism $t$ in $Gr(\mathbf{P})$, to yield an internal full subcategory there with good closure properties. To actually get an internal full subcategory in a topos, we will use:

**4.4. The Yoneda embedding.** If $\mathbf{C}$ is a small category, $[\mathbf{C}^{op},\mathbf{Set}]$ will denote the *topos of presheaves* on $\mathbf{C}$, i.e the category of contravariant, set-valued functors on $\mathbf{C}$ and natural transformations between such. $H:\mathbf{C}\hookrightarrow[\mathbf{C}^{op},\mathbf{Set}]$ will denote the *Yoneda embedding* — the full and faithful functor sending an object $I$ in $\mathbf{C}$ to the hom functor $H(I)=\mathbf{C}(-,I)$, and sending a morphism $\alpha:I\longrightarrow J$ to the natural transformation $H(\alpha)=\alpha^*:\mathbf{C}(-,J)\longrightarrow\mathbf{C}(-,I)$ whose components are given by precomposition with $\alpha$. We will need some properties of the Yoneda embedding under taking slice categories, and for this it is convenient to work with an equivalent version of the topos $[\mathbf{C}^{op},\mathbf{Set}]$ in terms of "discrete fibrations":

Since each set is a discrete category, any functor $X:\mathbf{C}^{op}\longrightarrow\mathbf{Set}$ can be regarded as category-valued and hence one can apply the Grothendieck construction of 4.1 to it to obtain $P:Gr(X)\longrightarrow\mathbf{C}$. $Gr(X)$ is often called the *category of elements of* $X$, since its objects are pairs $(I,x)$ with $I\in ob\mathbf{C}$ and $x\in X(I)$ (and its morphisms $(I,x)\longrightarrow(J,y)$ are just those $\alpha:I\longrightarrow J$ in $\mathbf{C}$ with $X(\alpha)(y)=x$). The functors $P:\mathbf{X}\longrightarrow\mathbf{C}$ that arise in this way are the *discrete fibrations*, which by definition are those with

$$\begin{array}{ccc} mor(\mathbf{X}) & \xrightarrow{cod} & ob(\mathbf{X}) \\ {\scriptstyle mor(P)}\downarrow & & \downarrow{\scriptstyle ob(P)} \\ mor(\mathbf{C}) & \xrightarrow[cod]{} & ob(\mathbf{C}) \end{array}$$

a pullback square in **Set**. More precisely, the Grothendieck construction (together with a similar construction on natural transformations) gives a functor $Gr:[\mathbf{C}^{op},\mathbf{Set}]\longrightarrow\mathbf{Cat}/\mathbf{C}$ which is an equivalence of categories between $[\mathbf{C}^{op},\mathbf{Set}]$ and $\mathbf{Dfib(C)}$, the full subcategory of $\mathbf{Cat}/\mathbf{C}$ whose objects are discrete fibrations. Under this equivalence the Yoneda embedding $H:\mathbf{C}\hookrightarrow[\mathbf{C}^{op},\mathbf{Set}]$ is identified with the functor $\mathbf{C}\longrightarrow\mathbf{Dfib(C)}$ which sends an object $I$ of $\mathbf{C}$ to the functor $\Sigma_I:\mathbf{C}/I\longrightarrow\mathbf{C}$ of 3.1. Now simple properties of pullbacks imply that for functors $P:\mathbf{X}\longrightarrow\mathbf{C}$ and $Q:\mathbf{Y}\longrightarrow\mathbf{X}$, if $P$ is a discrete fibration, then $Q\circ P$ is one iff $Q$ is; consequently $\mathbf{Dfib(C)}/P\cong\mathbf{Dfib(X)}$. Hence for any $X:\mathbf{C}^{op}\longrightarrow\mathbf{Set}$, there is an equivalence of categories:

$$[\mathbf{C}^{op},\mathbf{Set}]/X\simeq[(GrX)^{op},\mathbf{Set}]\,. \qquad (4.5)$$

In particular, when $X$ is a hom functor $H(I)$, (4.5) becomes:

$$[\mathbf{C}^{op},\mathbf{Set}]/H(I)\simeq[(\mathbf{C}/I)^{op},\mathbf{Set}]\,, \qquad (4.6)$$

and under this equivalence, the Yoneda embedding for $\mathbf{C}/I$ is identified with the functor $\mathbf{C}/I\longrightarrow[\mathbf{C}^{op},\mathbf{Set}]/H(I)$ which is "apply $H$ to morphisms".

It is well known that $H$ preserves any limits which exist in $\mathbf{C}$. It is also the case that $H$ preserves any existing exponentials — a fact put to good use by Dana Scott in [Sc1]. In fact something more general is true:

**4.5. Lemma.** *The Yoneda embedding* $H{:}\mathbf{C}\hookrightarrow[\mathbf{C}^{op},\mathbf{Set}]$ *preserves any local exponentials and instances of right adjoints to pulling back* ($\Pi$-*functors*) *that happen to exist in* $\mathbf{C}$.

**Proof.** Local exponentials are definable in terms of $\Pi$-functors: if ($p$ is squarable and) $p\to_I q$ exists, then so does $\Pi_p(p^*(q))$ and they are canonically isomorphic. So it suffices to show that $H$ preserves any instances of $\Pi$-functors. Thus given $\alpha{:}I\longrightarrow J$ and $p{:}E\longrightarrow I$ in $\mathbf{C}$, if $\Pi_\alpha(p)$ exists, with counit morphism $\varepsilon{:}\alpha^*(\Pi_\alpha(p))\longrightarrow p$ say, then we must show that transposing the morphism

$$(H\alpha)^*(H(\Pi_\alpha p))\cong H(\alpha^*(\Pi_\alpha p))\xrightarrow{H\varepsilon} H(p)$$

across the adjunction $(H\alpha)^*\dashv\Pi_{H\alpha}$ gives an isomorphism $H(\Pi_\alpha p)\cong\Pi_{H\alpha}(Hp)$ in $[\mathbf{C}^{op},\mathbf{Set}]/H(J)$. Transferring the problem to $[(\mathbf{C}/J)^{op},\mathbf{Set}]$ via the equivalence (4.6), we can calculate that for any $q$ in $\mathbf{C}/J$:

$$\begin{aligned}
H(\Pi_\alpha p)(q) &\cong [(\mathbf{C}/J)^{op},\mathbf{Set}](H(q),H(\Pi_\alpha p)) && \text{(Yoneda lemma)}\\
&\cong (\mathbf{C}/J)(q,\Pi_\alpha p) && \text{($H$ full and faithful)}\\
&\cong (\mathbf{C}/I)(\alpha^*(q),p) && \text{($\alpha^*$ left adjoint to $\Pi_\alpha$)}\\
&\cong [(\mathbf{C}/I)^{op},\mathbf{Set}]((H\alpha)^*(Hq),Hp) && \text{($H$ full, faithful and pullback preserving)}\\
&\cong [(\mathbf{C}/J)^{op},\mathbf{Set}](Hq,\Pi_{H\alpha}(Hp)) && \text{($(H\alpha)^*$ left adjoint to $\Pi_{H\alpha}$)}\\
&\cong (\Pi_{H\alpha}(Hp))(q)\,. && \text{(Yoneda lemma)}
\end{aligned}$$

These isomorphisms are natural in $q$ and give the required isomorphism $H(\Pi_\alpha p)\cong\Pi_{H\alpha}(Hp)$.

□

We can now state and prove our main result:

**4.6. Theorem. (Full embedding in topos models.)** *Let* $\mathbf{P}$ *be a* $2T\lambda C$*-hyperdoctrine. Apply the Grothendieck construction to* $\mathbf{P}$ *to obtain the category* $Gr(\mathbf{P})$, *containing the morphism* $t{:}(1_U{:}U\to U)\longrightarrow T(U)$ *of* (4.1). *Then the internal full subcategory* $\mathbf{U}$ *of the topos of presheaves* $[(Gr\mathbf{P})^{op},\mathbf{Set}]$ *determined by* $H(t)$ *is a topos model of the second order typed lambda calculus. Moreover, the* $2T\lambda C$*-hyperdoctrine* $\mathbf{P}$ *is fully embedded in the* $2T\lambda C$*-hyperdoctrine determined by this topos model, in the sense that there is a full, faithful and finite product preserving functor*

$$|\mathbf{P}|^{op}\xhookrightarrow{\;T\;}Gr(\mathbf{P})^{op}\xhookrightarrow{\;H\;}[(Gr\mathbf{P})^{op},\mathbf{Set}]$$

*and a natural isomorphism* $\mathbf{P}(-,U)\cong[(Gr\mathbf{P})^{op},\mathbf{Set}](HT(-),\mathbf{U})$.

**Proof.** Let us write $G$ for the object $(1_U{:}U\to U)$ in $Gr(\mathbf{P})$. Referring to Definition 3.3, to see that $H(t){:}H(G)\longrightarrow HT(U)$ determines a topos model, we have to produce morphisms $\ulcorner\mathrm{T}\urcorner{:}\mathrm{T}\longrightarrow HT(U)$, $\ulcorner\times\urcorner,\ulcorner\to\urcorner{:}HT(U)^2\longrightarrow HT(U)$ and $\ulcorner\Pi\urcorner{:}(HT(U)\to HT(U))\longrightarrow HT(U)$ in $[(Gr\mathbf{P})^{op},\mathbf{Set}]$, together with corresponding pullback squares of the form (3.3), (3.4), (3.5) and (3.6). The first three morphisms are rather easy to produce, since they already exist at the level of $|\mathbf{P}|$:

Let $\ulcorner\mathrm{T}\urcorner{:}\mathrm{T}\longrightarrow U$ and $\ulcorner\times\urcorner,\ulcorner\to\urcorner{:}U^2\longrightarrow U$ in $|\mathbf{P}|$ be as in (2.1) and (2.2). Now in $Gr(\mathbf{P})$, by (4.2) we have for the two product projections $\pi_i{:}TU^2\longrightarrow TU$ ($i=0,1$) that:

$\pi_i^*(t) = (1_{U^2},\pi_i):(\pi_i:U^2 \longrightarrow U) \longrightarrow TU^2$ .

Hence by (*4.2*) again

$$\pi_0^*(t)\times_{TU^2}\pi_1^*(t) = (1_{U^2},\pi_0\times_{U^2}\pi_1) = (1_{U^2},\ulcorner\times\urcorner) = T(\ulcorner\times\urcorner)^*(t) \tag{4.7}$$

in $Gr(\mathbf{P})/TU^2$, giving a pullback square in $Gr(\mathbf{P})$ of the form (*3.3*). Similarly, using (*4.3*) we have:

$$\pi_0^*(t)\rightarrow_{TU^2}\pi_1^*(t) = (1_{U^2},\pi_0\rightarrow_{U^2}\pi_1) = (1_{U^2},\ulcorner\rightarrow\urcorner) = T(\ulcorner\rightarrow\urcorner)^*(t) \tag{4.8}$$

giving the required pullback of the form (*3.4*); and by (*4.2*) again,

$$1_{T(\top)} = (1_\top,\top_\top) = (1_\top,\ulcorner\top\urcorner) = T(\ulcorner\top\urcorner)^*(t) \tag{4.9}$$

giving the pullback of the form (*3.5*) in $Gr(\mathbf{P})$. Then applying $H$ to these pullback squares and using the result of Lemma 4.5 that $H$ preserves the (pullbacks and) local exponentials involved in them, we get the required pullbacks in $[(Gr\mathbf{P})^{op},\mathbf{Set}]$ for the morphisms

$$\top \cong HT(\top)\xrightarrow{HT(\ulcorner\top\urcorner)} HT(U) \text{ and } HT(U)^2\cong HT(U^2)\xrightarrow{HT(\ulcorner*\urcorner)} HT(U) \quad (\text{where } * = \times \text{ or } \rightarrow).$$

To produce a morphism $\ulcorner\Pi\urcorner:(HT(U)\rightarrow HT(U))\longrightarrow HT(U)$ and a pullback of the form (*3.6*) is more complicated, since $|\mathbf{P}|$ is not a ccc and the exponential $HT(U)\rightarrow HT(U)$ is not a representable presheaf. However, consider the following calculation:

$$\begin{aligned}
(HT(U)\rightarrow HT(U))(-) &\cong [(Gr\mathbf{P})^{op},\mathbf{Set}](H(-),HT(U)\rightarrow HT(U)) && \text{(Yoneda lemma)}\\
&\cong [(Gr\mathbf{P})^{op},\mathbf{Set}](H(-)\times HT(U),HT(U)) && \text{(definition of exponential)}\\
&\cong [(Gr\mathbf{P})^{op},\mathbf{Set}](H(-\times TU),H(TU)) && (H \text{ preserves products})\\
&\cong Gr(\mathbf{P})(-\times TU,TU) && (H \text{ full and faithful})\\
&\cong |\mathbf{P}|(P(-\times TU),U) && (P \text{ left adjoint to } T)\\
&\cong |\mathbf{P}|(P(-)\times U,U). && (P \text{ preserves products and } PT=1)
\end{aligned}$$

We can therefore identify the exponential $HT(U)\rightarrow HT(U)$ with the functor $|\mathbf{P}|(P(-)\times U,U):Gr(\mathbf{P})^{op}\longrightarrow\mathbf{Set}$. Similarly, the hom functor $HT(U)$ can itself be identified with $|\mathbf{P}|(P(-),U):Gr(\mathbf{P})^{op}\longrightarrow\mathbf{Set}$. But then part (*iv*) of Definition 2.2 gives a natural transformation $\prod_{P(-)}$: $|\mathbf{P}|(P(-)\times U,U)\longrightarrow|\mathbf{P}|(P(-),U)$ and we define $\ulcorner\rightarrow\urcorner:(HT(U)\rightarrow HT(U))\longrightarrow HT(U)$ to be this. To get a pullback square of the form (*3.6*) as well, we have to show that $(\ulcorner\rightarrow\urcorner)^*(Ht)\cong\Pi_p ev^*(Ht)$ in $[(Gr\mathbf{P})^{op},\mathbf{Set}]/(HT(U)\rightarrow HT(U))$. Since by (*4.5*) the latter is equivalent to the topos of presheaves on $Gr(|\mathbf{P}|(P(-)\times U,U))$, it is sufficient to exhibit bijections

$$[(Gr\mathbf{P})^{op},\mathbf{Set}]/(HT(U)\rightarrow HT(U))(\alpha,(\ulcorner\rightarrow\urcorner)^*(Ht))\cong[(Gr\mathbf{P})^{op},\mathbf{Set}]/(HT(U)\rightarrow HT(U))(\alpha,\Pi_p ev^*(Ht))$$

natural in $\alpha:H(X)\longrightarrow(HT(U)\rightarrow HT(U))$ and then apply the Yoneda lemma. But if $\hat{\alpha}:P(X)\times U\longrightarrow U$ corresponds to $\alpha$ under the identifications made above, then the definition of $\ulcorner\rightarrow\urcorner$ gives:

$$\ulcorner\rightarrow\urcorner\circ\alpha = H\left(\overline{\textstyle\prod_{PX}\hat{\alpha}}\right), \tag{4.10}$$

where $\overline{\prod_{PX}\hat{\alpha}}$ is the transpose of $\prod_{PX}\hat{\alpha}:P(X)\longrightarrow U$ across the adjunction $P\dashv T$. Now $\hat{\alpha}$ also corresponds to a morphism $X\times T(U)\longrightarrow T(U)$ in $Gr(\mathbf{P})$ — call it $a$, say: then by (*4.4*)

$$\left(\overline{\textstyle\prod_{PX}\hat{\alpha}}\right)^*(t) = \Pi_{\pi_1}(a^*(t)) \tag{4.11}$$

with $\pi_1$ the first projection $X\times T(U)\longrightarrow X$. Now apply $H$ to (*4.11*) and use (*4.10*), plus the fact that $H$ preseves finite limits and $\Pi$-functors (Lemma 4.5) and sends $a$ to $\overline{\alpha}$, the exponential transpose of $\alpha$; we get:

$$(\ulcorner\rightarrow\urcorner\circ\alpha)^*(Ht)\cong\Pi_p\overline{\alpha}^*(Ht) \qquad (4.12)$$

in $[(Gr\mathbf{P})^{op},\mathbf{Set}]/H(X)$, with $p$ the first projection $H(X)\times HT(U)\longrightarrow H(X)$. Since $\overline{\alpha}=ev\circ(\alpha\times 1_{HX})$ and (by standard properties of $\Pi$-functors) $\Pi_p\circ(\alpha\times 1_{HX})^*\cong\alpha^*\circ\Pi_p$, (4.12) gives:

$$\alpha^*((\ulcorner\rightarrow\urcorner)^*(Ht))\cong\alpha^*(\Pi_p ev^*(Ht)) .$$

Using this isomorphism we get a bijection between morphisms $\alpha\longrightarrow(\ulcorner\rightarrow\urcorner)^*(Ht)$ and morphisms $\alpha\longrightarrow\Pi_p ev^*(Ht)$ in the slice category $[(Gr\mathbf{P})^{op},\mathbf{Set}]/(HT(U)\rightarrow HT(U))$. The constructions performed to get this bijection are evidently natural in $\alpha$: so as remarked above, we can infer that they are induced by an isomorphism $(\ulcorner\rightarrow\urcorner)^*(Ht)\cong\Pi_p ev^*(Ht)$, as required.

This establishes that we have a topos model $\mathbf{U}$ in $[(Gr\mathbf{P})^{op},\mathbf{Set}]$. Since by 4.2(*ii*) and standard properties of the Yoneda embedding, $HT$ *is* full, faithful and finite product preserving, to complete the proof of Theorem 4.6 we just have to exhibit the natural isomorphism:

$$\mathbf{P}(-,U)\cong[(Gr\mathbf{P})^{op},\mathbf{Set}](HT(-),\mathbf{U}) . \qquad (4.13)$$

Our calculations in the earlier part of this proof imply that the category object $\mathbf{U}$ is the image under $H$ of the category object $\mathbf{V}$ in $Gr(\mathbf{P})$ with underlying graph:

$$(U^2,\ulcorner\rightarrow\urcorner)\ \underset{d_1}{\overset{d_0}{\rightrightarrows}}\ (U,\top_U)=T(U) ,$$

where $d_i=(\pi_i,\ulcorner\rightarrow\urcorner)$ $(i=0,1)$. Thus for (4.13) it suffices to give a natural isomorphism

$$v:\mathbf{P}(-,U)\cong Gr(\mathbf{P})(T(-),\mathbf{V}) .$$

For each $I\in|\mathbf{P}|$, define $v$ on objects $A\in|\mathbf{P}|(I,U)$ by $v_I(A)=T(A)$; since $T$ is full and faithful, this gives a bijection. To define $v$ on morphisms, given $A,B:I\longrightarrow U$ in $|\mathbf{P}|$, note that morphisms $\langle TA,TB\rangle\longrightarrow\langle d_0,d_1\rangle$ in $Gr(\mathbf{P})/TU^2$ are specified by morphisms $\top_I\longrightarrow\langle A,B\rangle^*(\ulcorner\rightarrow\urcorner)=A\rightarrow_I B$ in $\mathbf{P}(I,U)$, which correspond under the exponential adjunction to morphisms $A\longrightarrow B$ in $\mathbf{P}(I,U)$: then define $v_I$ on morphisms by sending $f:A\longrightarrow B$ to the morphism $\langle TA,TB\rangle\longrightarrow\langle d_0,d_1\rangle$ over $TU^2$ specified by the exponential transpose of $f$. Routine calculations show that this recipe makes $v_I$ into a functor $\mathbf{P}(I,U)\longrightarrow Gr(\mathbf{P})(TI,\mathbf{V})$, and its construction is evidently natural in $I$. It is a bijection on objects and on hom sets, and hence is an isomorphism of categories (and also of $2T\lambda C$-hyperdoctrines therefore). This completes the proof of the theorem.

□

We can apply Theorem 4.6 to the classifying hyperdoctrine $\mathbf{P_T}$ of a $2T\lambda C$-theory $\mathbf{T}$ (*cf.* 2.6). $\mathbf{P_T}$ contains the generic model $I_\mathbf{T}$ and from (2.10) we have that:

$$\mathbf{T}\vdash_{X,\underline{x}} s=t:\Phi \text{ iff } I_\mathbf{T}\models_{X,\underline{x}} s=t:\Phi .$$

Transporting $I_\mathbf{T}$ along the isomorphism $\mathbf{P_T}(-,U)\cong[(Gr\mathbf{P_T})^{op},\mathbf{Set}](HT(-),\mathbf{U})$ of Theorem 4.6, gives a topos model of $\mathbf{T}$ with the same property. So we deduce:

**4.7. Corollary. (Completeness of topos models.)** *Let* **T** *be a* $2T\lambda C$*-theory. Then the theorems of* **T** *are precisely those equality judgements in the language of* **T** *which are satisfied by all topos models of* **T**; *in fact there is a single topos model whose true equality judgements are exactly the theorems of* **T**.

□

The logical significance of Theorem 4.6 is greater than just the above corollary. Recall the corresponence, reviewed in section 2, between theories in the second order typed lambda calculus and hyperdoctrines; recall also the correspondence, mentioned at the beginning of section 3, between theories in *higher order intuitionistic predicate logic* (or *HOL*, for short) and elementary toposes. In the light of these correspondences, we may rephrase the full embedding theorem as follows:

*Each* $2T\lambda C$*-theory can be interpreted in a theory in HOL so that:*

(*i*) *the types of the* $2T\lambda C$*-theory have a* standard *interpretation in the HOL-theory* (*i.e.* $\top$ *is terminal,* $\times, \prod$ *are products and* $\to$ *is exponentiation*);

(*ii*) *any two closed terms of the* $2T\lambda C$*-theory which can be proved to be equal in the HOL-theory are already provably equal in the* $2T\lambda C$*-theory*;

(*iii*) *any closed term of a type coming from the* $2T\lambda C$*-theory which can be proved to exist using HOL is provably equal to a* $2T\lambda C$*-term.*

**4.8. Extensions.** We conclude by mentioning that the method of proving the full embedding result also suffices to prove similar results for related type theories. In particular theories in the full *higher order typed lambda calculus* (Girard's system $F_\omega$ augmented with finite product types) correspond to the kind of hyperdoctrine where $|\mathbf{P}|$ is required to be cartesian closed (and for any object $V$ of $|\mathbf{P}|$, $\pi_1^*:\mathbf{P}(-,U)\longrightarrow\mathbf{P}(-\times V,U)$ is required to have a natural right adjoint). Since the Yoneda embedding preserves exponentials, the internal full subcategory **U** constructed in $[(Gr\mathbf{P})^{op},\mathbf{Set}]$ will now also be closed in the topos under internal products indexed by any object in the sub-ccc of the topos generated by the object of objects of **U**: indeed, as remarked in the proof of Theorem 4.6, the proof that **U** has the required closure properties is easier when $|\mathbf{P}|$ is a ccc. One could also augment the type theory with polymorphic sums, arriving at Seely's notion [Se] of a "PL-theory" and the corresponding kind of hyperdoctrine: the extra stucture is modelled by left adjoints to $(-)^*$ functors, and these are carried through into the topos model.

# References

[B] J.Bénabou, *Fibred categories and the foundations of naive category theory*, J. Symbolic Logic 50(1985) 10-37.

[BM] K.B.Bruce and A.R.Meyer, *The semantics of second order polymorphic lambda calculus.* In: G.Kahn *et al* (eds), *Semantics of Data Types,* Lecture Notes in Computer Science No.173 (Springer-Verlag, Berlin - Heidelberg - New York - Tokyo, 1984) pp 131-144.

[CCM] G.Cousineau, P.L.Curien and M.Mauny, *The categorical abstract machine.* In: J.P.Jouannaud (ed.), *Functional Programming Languages and Computer Architecture,* Lecture Notes in Computer Science No.201 (Springer-Verlag, Berlin - Heidelberg - New York - Tokyo, 1985) pp 50-64.

[CE] T.Coquand and T.Ehrhard, *An equational presentation of higher order logic.* In this volume.

[CFS] A.Carboni, P.J.Freyd and A.Scedrov, in preparation.

[Cu] P.L.Curien, *Categorical combinators, sequential algorithms and functional programming,* Research Notes in Theoretical Computer Science (Pitman, London, 1986).

[F] M.Fourman, *Sheaf models for set theory,* J. Pure Appl. Algebra 19(1980) 91-101.

[FS] P.J.Freyd and A.Scedrov, *Some semantic aspects of polymorphic lambda calculus.* In: *Proc. 2nd Annual Symp. on Logic in Computer Science* (IEEE, 1987) pp 315-319.

[G1] J.-Y.Girard, *Interprétation fonctionelle et elimination des coupures dans l'arithmétique d'ordre supérieur,* Thèse de Doctorat d'Etat (Paris, 1972).

[G2] J.-Y.Girard, *The system F of variable types, fifteen years later,* Theoretical Computer Science 45(1986) 159-192.

[Gr] J.W.Gray, *Fibred and cofibred categories.* In: S.Eilenberg *et al* (eds), *Proc. LaJolla Conference on Categorical Algebra* (Springer-Verlag, Berlin - Heidelberg, 1966) pp 21-83.

[H] J.M.E.Hyland, *The effective topos.* In: A.S.Troelstra and D.vanDalen (eds), *The L.E.J.Brouwer Centenary Symposium* (North-Holland, Amsterdam, 1982) pp 165-216.

[HRR] J.M.E.Hyland, E.Robinson and G.Rosolini, *Discrete objects in the effective topos.* To appear.

[J] P.T.Johnstone, *Topos Theory,* L.M.S. Monographs No.10 (Academic Press, London, 1977).

[LS] J.Lambek and P.J.Scott, *Introduction to Higher Order Categorical Logic,* Cambridge Studies in Advanced Mathematics No.7 (Cambridge University Press, 1986).

[L] F.W.Lawvere, *Equality in hyperdoctrines and comprehension schema as an adjoint functor.* In: A.Heller (ed.), *Proc. New York Symp. on Applications of Categorical Algebra* (Amer. Math. Soc., Providence, 1970) pp 1-14.

[McC] D.C.McCarty, *Realizability and recursive set theory,* Annals Pure Appl. Logic 32(1986) 153-183.

[P] A.Poigné, *Tutorial on category theory and logic.* In: D.Pitt *et al* (eds), *Category theory and computer programming,* Lecture Notes in Computer Science No.240 (Springer-Verlag, Berlin - Heidelberg - New York, 1986) pp 103-142.

[PS] R.Paré and D.Schumacher, *Abstract families and the adjoint functor theorems.* In: P.T.Johnstone and R.Paré (eds), *Indexed Categories and their Applications,* Lecture Notes in Mathematics No.661 (Springer-Verlag, Berlin - Heidelberg - New York, 1978) pp 1-125.

[R1] J.C.Reynolds, *Types, abstraction and parametric polymorphism,* Information Processing 83(1983) 513-523.

[R2] J.C.Reynolds, *Polymorphism is not set-theoretic.* In: G.Kahn *et al* (eds), *Semantics of Data Types,* Lecture Notes in Computer Science Vol.173 (Springer-Verlag, Berlin - Heidelberg - New York - Tokyo, 1984) pp 145-156.

[Sc1] D.S.Scott, *Relating theories of the lambda calculus.* In: J.R.Hindley and J.P.Seldin (eds), *To H.B.Curry: essays on combinatory logic, lambda calculus and formalism*

(Academic Press, London, 1980) pp 403-450.

[Sc2] D.S.Scott, *The presheaf model for set theory*. University of Oxford, February 1980. Manuscript.

[Se] R.A.G.Seely, *Categorical semantics for higher order polymorphic lambda calculus*, J. Symbolic Logic, to appear.