GENERALISED ALGEBRAIC THEORIES

AND

CONTEXTUAL CATEGORIES

by

J.W.Cartmell

Submitted for the degree of D.Phil.

Oxford University

1978

Acknowledgement

I am indebted to Professor Dana Scott for supervision, criticism and direction.

CONTENTS

		PAGE NO.
PREFACE		
CHAPTER 1 GENERALISED ALGEBRAIC THEORIES		1.1
1.1	Introduction.	1.2
1.2	Examples of theories.	1.7
1.3	Predicates as types.	1.13
1.4	Essentially algebraic theories and categories With finite limits.	1.16
1.5	The extra generality of the algebraic semantics.	1.19
1.6	The formal definition.	1.22
1.7	The substitution lemma and other lemmas.	1.28
1.8	Informal syntax.	1.39
1.9	Models and homorphisms.	1.45
1.10	A list of theories.	1.49
1.11	Interpretations.	1.5Ø
1.12	Contexts and realisations.	1.58
1.13	Intended identity of denotation.	1.6Ø
1.14	The Category GAT.	1.69
CHAPTER 2 CONTEXTUAL CATEGORIES		
2.1	Algebraic semantics.	2.1
2.2	Definition and examples.	2.3
2.3	Notation and basic lemmas.	2.12
2.4	Contextual categories = generalised algebraic theories.	2.24
2.5	Functorial Semantics, Universal algebra.	2.77

PAGE NO.

CHAPTER 3 THE ALGEBRAIC SEMANTICS OF MARTIN-LOF TYPE THEORY 3.1

3.1	Disjoint unions and singleton object.	3.2
3.2	Categories with attributes.	3.14
3.3	Categories and fibrations.	3.32
3.4	Martin-Lof type theory.	3.35
3.5	Limit spaces - a model of M-L type theory.	3.5Ø

PREFACE

A notion of equational theory is introduced; more general than previous notions, equal in descriptive power to the essentially algebraic theories of Freyd [5] , and hence to the logic of left exact categories, we call the theories generalised algebraic. The extra generality of these equational theories is achieved by the introduction of sort structures more general than those usually considered in that sorts may denote sets as is usual or else they may denote families of sets, families of families of sets and the like. This acceptance of variable types at the level of syntax (the idea and the form of syntax is taken directly from Martin-Löf type theory) makes the theories particularly suited to the description of the structures that occur in category theory. The basic example being the theory of categories, in which Ob appears as a sort to be interpreted as a set where as Hom appears as a sort to be interpreted as a family of sets indexed by Ob X Ob. Hom(x,y) appears in the syntax as a variable type.

The definition of the most general or algebraic semantics for generalised algebraic theories necessitates the introduction of the notion of a contextual category. So called because we shall see that the objects of a contextual category should be thought of as contexts.

The theory of contextual categories is seen as an algebraic description of the structure imposed on certain classes of term and type expressions by the operation of substitution of correctly typed terms for variables. Now this is something one might also say of the theory of

categories. However the theory of contextual categories captures the structure of substitution at work in a more general situation, it is the structure of substitution as found in the generalised algebraic theories but not in algebraic theories, as found originally in Martin-Lof type theory but not in theories of the typed λ- calculus.

It is proved that the category of contextual categories is equivalent to the category of generalised algebraic theories and equivalence classes of interpretations. Thus we say that we have the most general possible semantics. This result is a generalisation of the result implicit in Lawvere [ ] that the old syntactic notion of algebraic theory (i.e. one sorted equational) and Lawvere's algebraic notion are both one and the same (i.e. equivalent categories).

This thesis developed from the desire to develop the model theory of Martin-Lof type theory. The model theory rests on the notions of generalised algebraic theory and contextual category. It is only in these terms that we can define the notion model of Martin-Lof type theory. We also give the definition of model for a strengthened version of Martin-Lof type theory, this definition can be reworked algebraically into a hyperdoctrinal format. We briefly describe a new model of the type theory in which types are interpreted as limit spaces.

The model theory of the strengthened version of Martin-Lof type theory is a generalisation of the well known correspondence of the typed λ-calculus with cartesian closed categories.

CHAPTER 1

GENERALISED ALGEBRAIC THEORIES

The purpose of this chapter is to describe and to formally define the notion of generalised algebraic theory. It is hoped that it will be clear from the description that (i) the notion is a natural one formalising actual mathematical language and that (ii) the notion is a simple generalisation of the notion of a many sorted algebraic theory. Though (ii) tends to be obscured by the form of the chosen syntax no doubt the choice is correct.

The formal definition is given in §1.6. Most of the material that follows §1.6 is in preparation for Chapter Two, §1.8 is partially in digression and partially to explain some of the informal syntax that is used in the early sections of this Chapter.

1.1 Introduction

The notion of generalised algebraic theory is a generalisation of the notion of many sorted algebraic theory in just the following manner. Whereas the sorts of a many sorted algebraic theory are constant types in the sense that they are to be interpreted as sets the sort of a generalised algebraic theory need not all be constant types some of them may be nominated to be variable types in which case they are to be interpreted as families of sets. The type or types on which the variation of a variable type depends must always be specified.

Thus a generalised algebraic theory consists of (i) a set of sorts, each with a specified role either as a constant type or else as a variable type varying in some way, (ii) a set of operator symbols, each with its argument types and its value type specified, (the value type may vary as the argument varies), (iii) a set of axioms. Each axiom must be an identity between similar well formed expressions. either between terms of the same possible varying type or else between type expressions.

The theory of categories is a good example. The sort symbols we shall call Ob and Hom, the operator symbols id and o.

Ob is a constant type. Hom is a symbol for a variable type depending twice on Ob. That is to say that if t_1 and t_2 are both terms of type Ob then Hom (t_1,t_2) is a type. In particular if x and y are both variables of type Ob then Hom (x,y) is a type.

The operator symbol id has one argument type, namely Ob. The value type of id varies as the argument varies, for if x is a variable of type Ob then id(x) is of type Hom(x,x).

Not all the argument types of o are constant. If x,y and z are variables of type Ob, if f is a variable of type Hom(x,y) and if g is a variable of type Hom(y,z), then o(f,g) is a term of type Hom(x,z).

One way of setting up the syntax to deal with variables would be to assume that for every type Δ we had a supply V_Δ of variables of type Δ. However this method would lead to complications. Instead we assume just one set V of variables and then repeatedly assign types to variables as required. In a particular context the assertion or assumption that the variable x is of type Δ is written shorthand as $x \in \Delta$. More generally, the assertion that an expression t is a term of type Δ will be written as $t \in \Delta$. If the term t has variables $x_1, \ldots x_n$ occuring within it then it will only make sense to assert $t \in \Delta$ under an assumption that $x_1, \ldots x_n$ are variables of particular types. The complete assertion will be of the form: if x_1 is a variable of type $\Delta_1, \ldots$ and if x_n is a variable of type Δ_n then t is a term of type Δ. This complete assertion we write shorthand as

$$\frac{x_1 \in \Delta_1, x_2 \in \Delta_2, \ldots x_n \in \Delta_n}{t \in \Delta}$$

or else as $x_1 \in \Delta_1, x_2 \in \Delta_2, \ldots x_n \in \Delta_n : t \in \Delta$.

Similarly $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ is used to assert that if x_1 is a variable of type $\Delta_1, \ldots$ if x_n is a variable of type Δ_n, then Δ is a type.

These shorthands of the forms $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ we call rules. They serve to express which expressions of a given language are well formed as terms or as types. We work with these rules as units rather than with the basic expressions. For example, in the formal definition instead of defining the notions of well formed term and well formed type we define inductively a set of rules, to be called the derivable rules, which express the well formed types, the well formed terms their types.

The axioms of a theory are also written as rules. Instead of the more usual $\forall x_1 \in \Delta_1, \forall x_2 \in \Delta_2, \ldots \forall x_n \in \Delta_n, t_1 = t_2$ we write $\dfrac{x_1 \in \Delta_1, x_2 \in \Delta_2, \ldots x_n \in \Delta_n}{t_1 = t_2}$

There again, we might just write $t_I = t_2$, whenever $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$.

For example the theory of categories has as axioms the following:

$o(id(x),f) = f$, whenever $x,y \in Ob$ and $f \in Hom(x,y)$.

$o(f,id(y)) = f$, whenever $x,y \in Ob$ and $f \in Hom(x,y)$.

$o(o(f,g),h) = o(f,o(g,h))$, whenever $w,x,y,z \in Ob$, $f \in Hom(w,x)$, $g \in Hom(x,y)$ and $h \in Hom(y,z)$.

A theory is presented by specifiying the language and by listing the axioms. The language is specified by listing the symbols and by specifying the role that each symbol plays within the language either as a sort symbol of some kind or as a particularly typed operator symbol. The role that a symbol plays can always be specified by way of the assertion of a single rule. In the case of a sort symbol A there is a rule of the form $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$ that will correctly specify over what types A is dependent. In the case of an operator symbol f a rule of the form $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{f(x_1, \ldots x_n) \in \Delta}$ suffices to specify of what types its arguments are to be and of what type its values will be. In either case we call the symbol the introductory rule associated with the symbol.

For example the sort Hom of the theory of categories has introductory rule $x \in Ob, y \in Ob$: Hom(x,y) is a type. The symbol id has introductory rule $x \in Ob$: id(x) $\in$ Hom(x,x).

Finally, then, every theory is presented as a set of symbols each with associated introductory rule and a set of axioms. And of course everything must be well formed, but we leave all that until we give the formal definition in § 1.6.

The theory of categories nowlooks like this:

Symbol.	Introductory Rule.
Ob	Ob is a type.
Hom	$x,y \in Ob$: Hom(x,y) is a type.
o	$x,y,z \in Ob$, $f \in Hom(x,y), g \in Hom(y,z)$: $o(f,g) \in Hom(x,z)$.
id	$x \in Ob$: $id(x) \in Hom(x,x)$.

Axioms.

$o(id(x),f) = f$, whenever $x,y \in Ob$ and $f \in Hom(x,y)$.

$o(f,id(y)) = f$, whenever $x,y \in Ob$ and $f \in Hom(x,y)$.

$o(o(f,g),h) = o(f,o(g,h))$, whever $w,x,y,z \in Ob$,

$f \in Hom(w,x), g \in Hom(x,y)$ and $h \in Hom(y,z)$.

Whenever we speak of a model of a theory U, without qualification, then we shall mean a model in the usual sense, that is where type symbols are interpreted as sets, symbols for families of types are interpreted as families of sets, operator symbols are interpreted at operators and so on. Later we shall be interpreting theories in algebraic structures, in which case type symbols will be interpreted as objects within a structure rather than as sets.

1.2 Examples of Theories

The first example is a theory which can be called the theory of families of elements of families of sets:

Symbol	Introductory Rule
A	A is a type
B	For $x \in A$: $B(x)$ is a type
b	For $x \in A$: $b(x) \in B(x)$

Axioms - None

A model of this theory will consist of a set, a family indexed by this set and a distinguished element of each set in this family; which is to say that a model will consist of a set indexed family of elements of a family of sets. We are not sure of the notation that we should be using but if we denote the interpretation of a symbol in a model $\mathfrak{M}$ by that symbol superscripted by $\mathfrak{M}$ then a model $\mathfrak{M}$ consists of i. a set $A^{\mathfrak{M}}$, ii. an $A^{\mathfrak{M}}$ - indexed family of elements $b^{\mathfrak{M}}$ of the family of sets $B^{\mathfrak{M}}$.

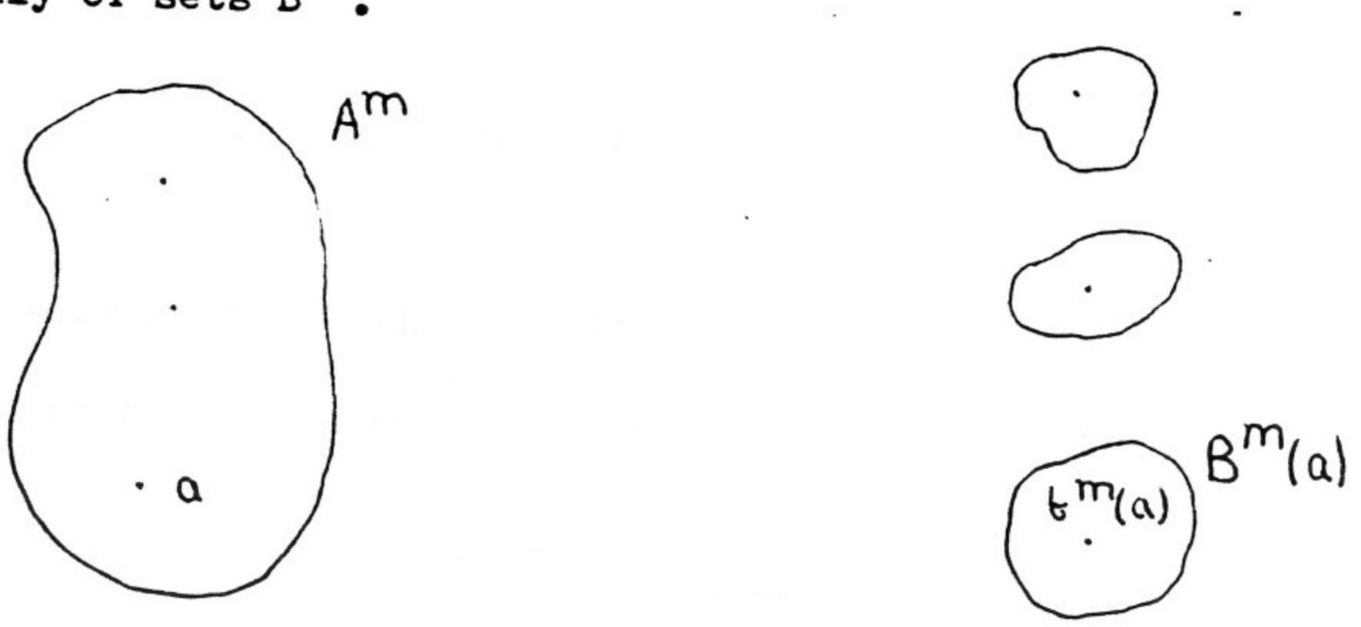


Fig. 1. - for every element a of the set $A^{\mathfrak{M}}$ we have i. a set $B^{\mathfrak{M}}(a)$ and ii. an element $b^{\mathfrak{M}}(a)$ of the set $B^{\mathfrak{M}}(a)$.

If both $\mathfrak{M}$ and $\mathfrak{M}'$ are models of this theory then a homomorphism $f : \mathfrak{M} \longrightarrow \mathfrak{M}'$ consists of a function $f_A : A^{\mathfrak{M}} \longrightarrow A^{\mathfrak{M}'}$ and an $A^{\mathfrak{M}}$ - indexed family of functions f_B such that for every $a \in A^{\mathfrak{M}}$, $f_B(a) : B^{\mathfrak{M}}(a) \longrightarrow B^{\mathfrak{M}'}(f_A(a))$ and such that for every $a \in A^{\mathfrak{M}}$, $f_B(a)(b^{\mathfrak{M}}(a)) = b^{\mathfrak{M}'}(f_A(a))$.

Alternatively we can say that a homomorphism consists of a function $f_A : A^{\mathfrak{M}} \longrightarrow A^{\mathfrak{M}'}$ and an operator f_B such that for every $a \in A^{\mathfrak{M}}$, for every $b \in B^{\mathfrak{M}}(a)$, $f_B(a,b) \in B^{\mathfrak{M}'}(f_A(a))$ and satisfying $f_B(a,b^{\mathfrak{M}}(a)) = b^{\mathfrak{M}'}(f_A(a))$, whenever $a \in A^{\mathfrak{M}}$. Now this means that there is a generalised algebraic theory whose models are just homomorphisms between the models of the given theory (in fact this is quite generally the case). This theory of homomorphisms can be presented as follows:

The theory of families of elements of families of sets in the language $\langle A,B,b \rangle$ + the same theory in the language $\langle A',B',b' \rangle$ +

Symbol.	Introductory Rule
f_A	For $x \in A$: $f_A(x) \in A'$.
f_B	For $x \in A$, for $y \in B(x)$: $f_B(x,y) \in B'(f_A(x))$.

Axiom.

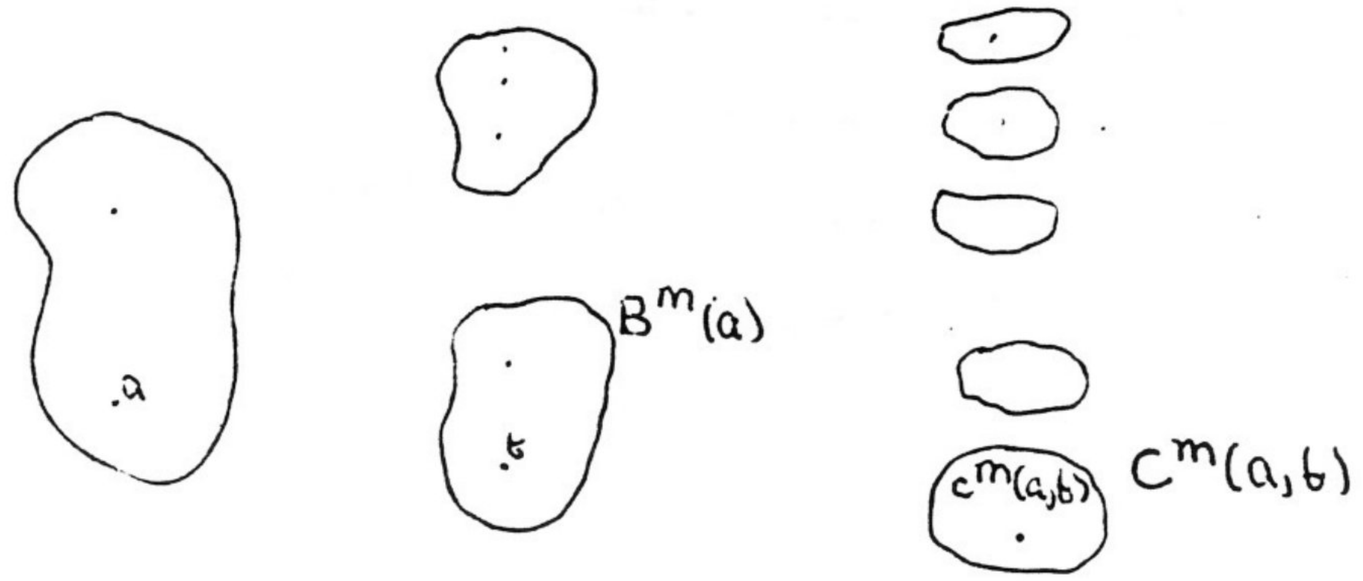
$f_B(x,b(x)) = b'(f_A(x))$, whenever $x \in A$.

An example similar to the first example we call the theory of families of families of elements of families of families of sets:

Symbol.	Introductory Rule
A	A is a type
B	For $x \in A$: $B(x)$ is a type
C	For $x \in A$, for $y \in B(x)$: $C(x,y)$ is a type
c	For $x \in A$, for $y \in B(x)$: $c(x,y) \in C(x,y)$

Axioms - None.

Suppose that $\mathfrak{m}$ is a model of this theory. Then $A^{\mathfrak{m}}$ is a set. For every element a of the set $A^{\mathfrak{m}}$ we have a set $B^{\mathfrak{m}}(a)$ and for every element b of the set $B^{\mathfrak{m}}(a)$ we have a set $C^{\mathfrak{m}}(a,b)$ and an element $c^{\mathfrak{m}}(a,b)$ of the set $C^{\mathfrak{m}}(a,b)$.



Now for every element a of $A^{\mathfrak{m}}$, $\lambda b.C^{\mathfrak{m}}(a,b)$ is a $B^{\mathfrak{m}}$ - indexed family of sets. Thus $\lambda a.\lambda b.C^{\mathfrak{m}}(a,b)$, ie. $C^{\mathfrak{m}}$, is an $A^{\mathfrak{m}}$ - indexed family of families of sets. Similarly $c^{\mathfrak{m}}$ is an $A^{\mathfrak{m}}$ - indexed family of families of elements.

Note that in the presentation of this theory no harm is done if we replace the introductory rule for C by the rule:-
for $x \in A$, for $y \in B(x)$: C(y) is a type, this rule having the same meaning as the given rule. The expression C(x,y) in the given rule depends explicitly on x and y. We say that the expression C(y) in the alternative rule depends implicitly on x by virtue of its explicit dependence on y and by virtue of the dependence of y on x. In the alternative version of the theory we say that a variable has been omitted. This is one way in which a theory may be informally presented. We use this method and another in presenting the next theory - the theory of trees.

The theory of trees has countably many sort symbols, no operator symbols and no axioms. However, we chose to write the theory informally with just two sort symbols, one of these symbols doing the work that in a formal presentation would be shared among countably many distinct symbols.

Symbol	Introductory Rule
S_1	S_1 is a type
S	For $x_1 \in S_1$: $S(x_1)$ is a type
S	For $x_1 \in S_1$, for $x_2 \in S(x_1)$: $S(x_2)$ is a type
.	.
.	.
.	.
.	.
S	For $x_1 \in S_1$, for $x_2 \in S(x_1)$, ... for $x_n \in S(x_{n-1})$: $S(x_n)$ is a type
.	.
.	.
.	.
.	.
.	.

Axioms - None.

S_1, then, is a symbol denoting the set of nodes at the base of the tree. If x is any node of the tree then S(x) is the set of nodes immediately above x in the tree, that is to say the set of successor nodes to x. In a formal presentation of this theory there would be symbols $S_1, S_2, S_3, \ldots$ and the symbol S_{n+1} would be introduced by the rule

$$\frac{x_1 \in S_1, x_2 \in S_2(x_1), \ldots x_n \in S_n(x_1, \ldots x_{n-1})}{S_{n+1}(x_1, \ldots x_n) \text{ is a type}}$$

We use the same methods in presenting the theory of functors informally. The theory of functors consists of the theory of categories in the language $\langle Ob, Hom, id, o\rangle$ + the theory of categories in the language $\langle Ob', Hom, id, o\rangle$ (and at this point we have used the same three symbols Hom, id and o in new roles) +

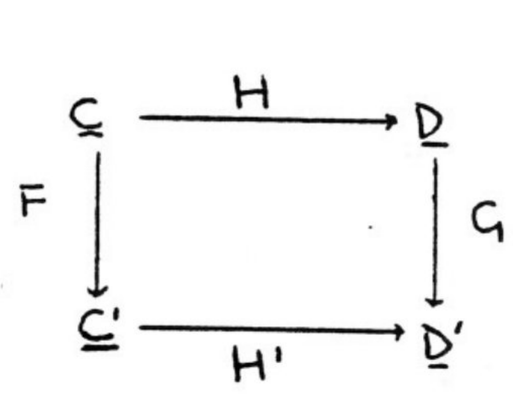
Symbol.	Introductory Rule
F	For $x \in Ob$: $F(x) \in Ob'$
F	For $x, y \in Ob$, for $f \in Hom(x,y)$: $F(f) \in Hom(F(x), F(y))$

Axioms.

$F(id(x)) = id(F(x))$, whenever $x \in Ob$.

$F(o(f,g)) = o(F(f), F(g))$, whenever $x,y,z \in Ob, f \in Hom(x,y)$ and $g \in Hom(y,z)$.

A model of this theory is just a functor. The category of models is the category $\underline{Cat}^2$, which is to say that if $F : \underline{C} \longrightarrow \underline{C}'$ is a functor and if $G : \underline{D} \longrightarrow \underline{D}'$ is a functor then a homomorphism from F to G consists of a pair of functors $\langle H, H'\rangle$ such that $H : \underline{C} \longrightarrow \underline{D}$ and $H' : \underline{C}' \longrightarrow \underline{D}'$ and such that



commutes.

The final example is to indicate one way of axiomatising the disjoint union of a family of types.

If U is a theory which includes a type symbol A and a symbol B for an A-indexed family of types then U can be extended by three operator symbols, three axioms and one type symbol $\sum_A B$ in such a way that i. every model $\mathfrak{M}$ of U uniquely extends to a model of the extended theory and ii. every model $\mathfrak{M}$ of the extended theory interprets the symbol $\sum_A B$ by the set $\{\langle a,b\rangle \mid a \in A^{\mathfrak{M}} \text{ and } b \in B^{\mathfrak{M}}(a)\}$, that is to say as the disjoint union of the family of sets interpreting B. The extended theory is taken to be U +

Symbol.	Introductory Rule
$\sum_A B$	$\sum_A B$ is a type
P_1	For $z \in \sum_A B$: $P_1(z) \in A$
P_2	For $z \in \sum_A B$: $P_2(z) \in B(P_1(z))$
Pr	For $x \in A$, for $y \in B(x)$: $Pr(x,y) \in \sum_A B$

Axioms

$Pr(P_1(z),P_2(z)) = z$, whenever $z \in \sum_A B$.

$P_1(Pr(x,y)) = x$, whenever $x \in A$ and $y \in B(x)$.

$P_2(Pr(x,y)) = y$, whenever $x \in A$ and $y \in B(x)$.

In future we might refer to an extension of a theory by symbols for disjoint unions of specified families of types.

1.3 Predicates as types

It is possible to introduce sort symbols into a generalised algebraic theory and then axiomatise them in such a way as they are effectively predicate symbols. In this way any theory of predicate calculus all of whose axioms are of the form $\forall \bar{x}(\varphi_1 \wedge \varphi_2 \ldots \wedge \varphi_n \longrightarrow \psi)$, where $\varphi_1, \ldots \varphi_n$ and ψ are all atomic, can be expressed as generalised algebraic. Let us call such an axiom a universal condition.

We do not work with relations directly but rather with their characteristic families. If R is an n-ary relation on a set A then its characteristic family is the family $\lambda a_1 . \lambda a_2 \ldots \lambda a_n . P(a_1, \ldots a_n)$, where $P(a_1, \ldots a_n) = \{\emptyset\}$ if $R(a_1, \ldots a_n)$ and $P(a_1, \ldots a_n) = \emptyset$ otherwise.

The following theory indicates how an n-ary predicate symbol may be introduced into a theory. The given theory has as models just characteristic families of n-ary relations on a set.

Symbol.	Introductory Rule
A	A is a type
P	For $x_1, \ldots x_n \in A$: $P(x_1, \ldots x_n)$ is a type

Axiom

$y_1 = y_2$, whenever $x_1, \ldots x_n \in A$ and $y_1, y_2 \in P(x_1, \ldots x_n)$.

It remains to show how universal conditions may be expressed as generalised algebraic. We distinguish three forms that such a condition might take. The first case is when each of $\varphi_1,\ldots\varphi_n$ and ψ are instances of a predicate other than the equality predicate. In this case $\forall\bar{x}(\varphi_1\wedge\ldots\varphi_n\longrightarrow\psi)$ cannot be expressed as an axiom but it can be expressed merely by the introduction of a new operator symbol. For example the transitivity of a binary predicate P is expressed by the introduction of a new operator symbol t by the rule :- for $x_1,x_2,x_3\in A$, for $y_1\in P(x_1,x_2)$ and for $y_2\in P(x_2,x_3)$: $t(y_1,y_2)\in P(x_1,x_3)$. The point is that once P is interpreted then t is interpretable in at most one way and then only in case the predicate is transitive.

The second case to consider is the case where each of $\varphi_1,\ldots\varphi_n$ are instances of a predicate other than the equality predicate and ψ is an instance of the equality predicate. In this case $\forall\bar{x}(\varphi_1\wedge\ldots\wedge\varphi_n\longrightarrow\psi)$ can be expressed as an axiom of the theory. For example the anti-symetry of a binary predicate P can be expressed by the axiom :- $x_1=x_2$, whenever $x_1,x_2\in A$ and $y_1\in P(x_1,x_2), y_2\in P(x_2,x_1)$.

Lastly we must consider the case when one of the φi's is an instance of the equality predicate. In this case a new binary predicate must be added to the language and axiomatised to be the equality predicate. The axiom $\forall\bar{x}(\varphi_1\wedge\ldots\varphi_n\longrightarrow\psi)$ can then be dealt with by one of the first two cases. The following theory indicates the way in which the binary predicate Eq can be added to a theory and axiomatised to be the equality predicate.

Symbol.	Introductory Rule
A	A is a type
Eq	For $x_1, x_2 \in A$: $Eq(x_1, x_2)$ is a type
r	For $x \in A$: $r(x) \in Eq(x,x)$

Axioms

$y_1 = y_2$, whenever $x_1, x_2 \in A$ and $y_1, y_2 \in Eq(x_1, x_2)$

$x_1 = x_2$, whenever $x_1, x_2 \in A$ and $y \in Eq(x_1, x_2)$.

One final example. The theory of a 1-1 function is the theory of equality in the language $\langle B, Eq, r \rangle$ +

Symbol.	Introductory Rule
A	A is a type
f	For $x \in A$: $f(x) \in B$.

Axioms

$x_1 = x_2$, whenever $x_1, x_2 \in A$ and $y \in Eq(f(x_1), f(x_2))$.

1.4 Essentially Algebraic Theories and Categories with Finite Limits

The essentially algebraic theories of Freyd [5] can be seen to have the same descriptive power as generalised algebraic theories, at least as far as the usual set valued models are concerned. In this section we look at the relationship between these two notions and also at the relationship between essentially algebraic theories and categories with all finite limits. In the next section we point out the way in which generalised algebraic is a more general notion than essentially algebraic.

Essentially algebraic theories are introduced and very briefly discussed in Freyd [5]; they are many sorted partial algebraic theories such that the domain of every partial operation is specified as the extension of some conjunction of identities between terms compounded from previously introduced operators.

Thus the theory of categories as an essentially algebraic theory has two sorts, Ob and morph, three total operations, dom : Morph $\longrightarrow$ Ob, cod : Morph $\longrightarrow$ Ob and id : Ob $\longrightarrow$ Morph, and one binary partial operation o form Morph x Morph to Morph whose domain is specified by asserting that o(x,y) is defined iff cod(x) = dom(y).

In order to write an essentially algebraic theory as generalised algebraic, all the equality predicates used in defining domains of partial operations must be introduced. For example if f is to be a partial C-valued function defined on $\{x \in A \mid t_1 = t_2\}$, where for $x \in A : t_1 \in B$ and for $x \in A : t_2 \subset B$, then the equality predicate on B must be introduced and axiomatised. Then f can be introduced by

the rule for $x \in A$, for $y \in Eq(t_1,t_2)$: $f(x,y) \in C$.

In this way every essentially algebraic theory can be rewritten as generalised algebraic. The converse is also the case, at least in so far as that to every generalised algebraic theory there corresponds an essentially algebraic theory with the same category of models. This is the case because of the equivalence between A-indexed families of sets and morphisms in the category Set with codomain A. This equivalence holds for any set A and is given by the following

1. If $\{B(a) | a \in A\}$ is an A-indexed family of sets then proj : $\dot{\cup} B(a) \longrightarrow A$ is a morphism of Set with codomain A (remember that $\dot{\cup} B(a) = \{\langle a,b\rangle \mid a \in A, b \in B(a)\}$) .

2. If $f : A' \longrightarrow A$ is a map in Set with codomain A then $\{ f^{-1}(a) \mid a \in A \}$ is an A-indexed family of sets.

1. and 2. establish an isomorphism between the class of A-indexed families of sets and the class of functions with codomain A. Thus, if in a generalised algebraic theory there is a sort symbol B introduced as an A-indexed family of types then in the corresponding essentially algebraic theory there is introduced a new sort symbol A' and a map $\rho : A' \longrightarrow A$.

The notion of an essentially algebraic theory can be seen as a notion of type theory in which the only type forming principles are for the formation of product types and for the formation of types of the form $\{x \in \Delta \mid t_1 = t_2\}$, were Δ is a type and t_1 and t_2 are terms of the same type. Now if we think of the objects of an arbitrary category .

as types then to have these two type forming principles is just to have finite products and equalisers of pairs. Since a category with finite products and equalisers of pairs is precisely a category with finite limits, the notions of essentially algebraic theory and category with finite limits are closely connected.

In fact for every essentially algebraic theory U there is a category with finite limits C(U) such that the category of models of U is equivalent to the category LEX(C(U),Set) of all finite limit preserving functors from C(U) to Set, with all natural transformations between them as morphisms.

This is the content of a remark made by Lawvere, pages 8-9 of Lawvere [17], though the remark does not actually use the term essentially algebraic.

1.5 The Extra Generality of the Algebraic Semantics

One of the advantages of generalised algebraic over essentially algebraic is to be found in the syntax particularly with regard to the presentation of theories. In presenting theories as essentially algebraic there is a coding process in that, in general, families of sets indexed by a set are represented by functions with codomain that set. On the otherhand in presenting a theory as generalised algebraic there need be no such coding. This distinction whereby families in a generalised algebraic theory can have a life of their own goes through into the algebraic semantics. The algebraic semantics of generalised algebraic theories is more general than any possible such semantics for essentially algebraic theories. There are perfectly coherent interpretations of generalised algebraic theories into structures in which the elements of the structure that are there to interpret type-indexed families of types are distinct from the elements that are there to interpret functions with codomain. This can never be the case with essentially algebraic theories because already in the syntax families of types are coded as functions with codomain.

The notion of type is adequately captured by the notion of object of category. However having decided to think of the objects of a particular category $\mathcal{C}$ as types and in particular having decided to think of an actual object A of $\mathcal{C}$ as a type then it is incorrect then to suppose that the notion of A-indexed family of types should always be taken to be the notion morphism of $\mathcal{C}$ with codomain A. This is just one possibility. In general, there will be other possibilities some

of which may be more attractive. The example that we have in mind is when $\mathcal{C}$ is taken to be the category Cat of all (small) categories. Now the question is what shall we chose to mean by category indexed family of categories? In particular if A is a category then what shall we mean by an A-indexed family of categories? Well, what we would like to mean by that is any functor $B:A \longrightarrow Cat$. This is not the same as taking it to mean a morphism of Cat with codomain A. The idea that a family of categories indexed by the category A should just be a functor $B:A \longrightarrow Cat$ arises because there is a *category* of all (small) *categories* just as the fact that there is a *class* U of all sets (= small *classes*) leads to the definition of a family of sets indexed by a set A as a function $B:A \longrightarrow U$.

A functor $B:A \longrightarrow Cat$ can be thought of as a structure of the general kind (for example take sort symbols ObA, HomA, ObB and HomB introduced by rules ObA is a type; for $x,y \in ObA$:HomA(x,y) is a type; for $x \in ObA$:ObB(x) is a type; for $x \in ObA$, for $y,z \in ObB(x)$:HomB(x,y,z) is a type). It follows that there is a category of category indexed families of categories and structure preserving homomorphisms (it turns out that a homomorphism from $B:A \longrightarrow Cat$ to $B':A' \longrightarrow Cat$ is describable just as a pair F,N wnere $F:A \longrightarrow A'$ is a functor and $N:B \longrightarrow FoB'$ is a natural transformation). But now there follows the notion of a category indexed family of category indexed families of categories. This procedure can be iterated. We get a huge structure of categories, category indexed families of categories, category indexed families of category indexed families of categories and so on. It is a structure into which generalised algebraic theories can be interpreted - by

interpreting types as categories, type indexed families of types as category indexed families of categories and so on.

A model of the theory of families of elements of families of sets in this structure will consist of category $\underline{A}$, a functor $B:\underline{A}\longrightarrow\underline{Cat}$, for each $a\in|\underline{A}|$, an object $b(a)$ of $B(\underline{a})$ and for each $f:a\longrightarrow a'$ in $\underline{A}$, a morphism $b(f)$: $B(f)(b(a))\longrightarrow b(a')$ in $B(a')$. Such that $b(id(a))=id(b(a))$, for all objects a of A and such that $B(f')(b(f))ob(f') = b(fof')$, whenever $\multimap f\rightarrow\multimap f'\rightarrow$ in $\underline{A}$.

1.6 The Formal Definition

We must insist that the introductory rule for a symbol into into a language be well formed. In order that we may say what it is for a rule to be well formed we require a notion of derivability. Since the notion of derivability depends upon the introductory rules there is a difficulty in giving the formal definition.

The difficulty is that we need knowledge of the derived rules of a theory when we are still in the process of defining the possible languages in which the theory may be written. We choose to overcome the difficulty by leaving aside the question of wellformedness until we have available the complete set of derived rules of the theory. For this reason the theories that are admitted by the definition below may not be well formed; we call them pretheories and accept that they might make little sense. Later we shall define a theory to be a well formed pretheory.

We assume throughout that we have a set V of variables which has countably many distinct members. We begin by giving a definition of rule, more precisely a definition of rule of the alphabet W. The definition is crude in that most of the permitted rules are meaningless in all circumstances; it does suffice, though, for the purpose of turning rules into objects. Suppose then that W is a set. We consider the set W to be an alphabet and its elements to be symbols. The following definition is relative to W(and, of course, it is relative to the set of variables V, but V will remain fixed throughout).

The set of <u>expressions</u> is defined inductively in such a way that every expression is a finite sequence of elements of $W \cup V \cup \{(\} \cup \{)\} \cup \{,\}$ by the clauses: 1. If $x \in V$ then x is an expression. 2. If $L \in W$ then L is an expression. 3. If $L \in W$ and $e_1, \ldots e_n$ are expressions then $L(e_1, \ldots e_n)$ is an expression.

A <u>premise</u> is defined to be any finite sequence of elements of Vx the set of expressions. The empty sequence is included as a premise, called funnily enough, the empty premise. The premise determined by $((x_1, \Delta_1), \ldots (x_n, \Delta_n))$ is written as $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$.

A <u>T-conclusion</u> is determined by a single expression Δ and is written as 'Δ is a type'.

An <u>$\in$-conclusion</u> is determined by a pair of expressions (t, Δ) and is written as '$t \in \Delta$'.

A <u>T=-conclusion</u> is determined by a pair of expressions (Δ, Δ') and is written as '$\Delta = \Delta'$'.

An <u>$\in$ =-conclusion</u> is determined by a triple of expressions (t, t', Δ) and is written as '$t = t' \in \Delta$'.

A <u>rule</u> is determined by a premise P and a conclusion C and is written as $\frac{P}{C}$. A rule is said to be a T-rule, an $\in$-rule, a T=rule or an $\in$=rule according as to the form of its conclusion.

If $\Delta, t_1, \ldots t_n$ are expressions and if $x_1, \ldots x_n$ are distinct variables then the expression $\Delta[t_1|x_1, \ldots t_n|x_n]$ is that expression which results from simultaneously replacing every occurence of the

variables $x_1,\ldots x_n$ in the expression Δ by $t_1,\ldots t_n$. Please note that $\Delta[t_1|x_1,\ldots t_n|x_n]$ and $\Delta[t_1|x_1]\ldots[t_n|x_n]$ are not usually the same, indeed $\Delta[t_1|x_1,t_2|x_2]$ and $\Delta[t_1|x_1][t_2|x_2]$ are distinct whenever x_1 appears in Δ, x_2 appears in t_1 and t_2 is distinct from x_2.
We can now give the main definitions.

Definition 1. A <u>pretheory</u> consists of 1. a set S, called the set of sort symbols. 2. A set Z called the set of operator symbols. 3. To each sort symbol A, an associated rule of the alphabet $S \cup Z$, called the introductory rule for A and of the form $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{A(x_1,\ldots x_n) \text{ is a type}}$ for some $n \geq o$. 4. To each operator symbol F, an associated rule of the alphabet $S \cup Z$ called the introductory rule for F and of the form $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{F(x_1,\ldots x_n) \in \Delta}$, for some $n \geq o$. 5. A set of axioms. Each axiom is either a T=rule or an $\in$=rule of the alphabet.

Taken together, definition 2(a) and 2(b) define the derived rules of a pretheory. The definition is of an inductive nature.

Definition 2(a). If U is a pretheory then (i) a <u>context</u> is a premise $x_1 \in \Delta_1,\ldots x_n \in \Delta_n$ such that the rule $\dfrac{x_1 \in \Delta_1,\ldots x_{n-1} \in \Delta_{n-1}}{\Delta_n \text{ is a type}}$ is a derived rule of U. (ii) (a) The rule $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ is <u>wellformed</u> iff $x_1 \in \Delta_1,\ldots x_n \in \Delta_n$ is a context. (b) The rule $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{t \in \Delta}$ is <u>wellformed</u> iff $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ is a derived rule of U.

(d) The rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t=t' \in \Delta}$ is wellformed iff $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t' \in \Delta}$ are both derived rules of U.

Definition 2(b). The set of derived rules of U is the set of rules derivable by the following principles of derivation.

LI1. From $\dfrac{P}{\Delta \text{ is a type}}$ derive $\dfrac{P}{\Delta = \Delta}$.

LI2. From $\dfrac{P}{t \in \Delta}$ derive $\dfrac{P}{t=t \in \Delta}$.

LI3. From $\dfrac{P}{\Delta_1 = \Delta_2}$ derive $\dfrac{P}{\Delta_2 = \Delta_1}$.

LI4. From $\dfrac{P}{t_1=t_2 \in \Delta}$ derive $\dfrac{P}{t_2=t_1 \in \Delta}$.

LI5. From $\dfrac{P}{\Delta_1 = \Delta_2}$ and $\dfrac{P}{\Delta_2 = \Delta_3}$ derive $\dfrac{P}{\Delta_1 = \Delta_3}$.

LI6. From $\dfrac{P}{t_1=t_2 \in \Delta}$ and $\dfrac{P}{t_2=t_3 \in \Delta}$ derive $\dfrac{P}{t_1=t_3 \in \Delta}$.

LI7. From $\dfrac{P}{t_1=t_2 \in \Delta_1}$ and $\dfrac{P}{\Delta_1 = \Delta_2}$ derive $\dfrac{P}{t_1=t_2 \in \Delta_2}$.

T1. From $\dfrac{P}{\Delta_1 = \Delta_2}$ and $\dfrac{P}{t \in \Delta_1}$ derive $\dfrac{P}{t \in \Delta_2}$.

CF1. For $n \geqslant 0$, $1 \leqslant i \leqslant n+1$.

From $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ derive $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}$,

providing that x_{n+1} is a variable distinct from all of $x_1, \ldots x_n$.

CF2(a). For every sort symbol A with wellformed introductory rule

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$$, for every context P, from

$$\frac{P}{t_1 \in \Delta_1}\ ,\ \frac{P}{t_2 \in \Delta_2[t_1 \mid x_2]}\ ,\ \ldots\ \frac{P}{t_n \in \Delta_n[t_1 \mid x_1, \ldots t_{n-1} \mid x_{n-1}]}$$ derive

$$\frac{P}{A(t_1, \ldots t_n) \text{ is a type.}}$$

CF2(b). For every operator symbol F with wellformed introductory rule

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{F(x_1, \ldots x_n) \in \Delta}$$, for every context P,

from $\frac{P}{t_1 \in \Delta_1}$, $\frac{P}{t_2 \in \Delta_2[t_1 \mid x_1]}$, ... and $\frac{P}{t_n \in \Delta_n[t_1 \mid x_1, \ldots t_{n-1} \mid x_{n-1}]}$

derive $\frac{P}{F(t_1, \ldots t_n) \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}$.

SI1. If Q is a context then from $\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega = \Omega'}$ and $\frac{Q}{S_1 = S_1' \in \Omega_1}$,

$\frac{Q}{S_2 = S_2' \in \Omega_2[S_1 \mid Y_1]}$, ... $\frac{Q}{S_m = S_m' \in \Omega_m[S_1 \mid Y_1, \ldots S_{m-1} \mid Y_{m-1}]}$ derive $\frac{Q}{\Omega[S_1 \mid Y_1, \ldots S_m \mid Y_m] = \Omega'[S_1 \mid Y_1, \ldots S_m \mid$

SI2. If Q is a context then from $\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{S = S' \in \Omega}$ and $\frac{Q}{S_1 = S_1' \in \Omega_1}$,

$\frac{Q}{S_2 = S_2' \in \Omega_2[S_1 \mid Y_1]}$, ... $\frac{Q}{S_m = S_m' \in \Omega_m[S_1 \mid Y_1, \ldots S_{m-1} \mid Y_{m-1}]}$ derive

$$\frac{Q}{S[S_1 \mid y_1, \ldots S_m \mid y_m] = S'[S_1' \mid y_1, \ldots S_m' \mid y_m] \in \Omega[S_1 \mid y_1, \ldots S_m \mid y_m]}$$.

A1. If $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$ is an axiom then from

$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ and $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta' \text{ is a type}}$ derive $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$.

A2. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}$ is an axiom then from

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t' \in \Delta}$ derive $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}$

Definition 3. A pretheory is <u>wellformed</u> iff all of its introductory rules and axioms are wellformed. A <u>generalised algebraic theory</u> is a wellformed pretheory.

1.7 The Substitution Lemma and Other Lemmas

Each lemma in this section is needed at some later stage. For example the substitution lemma which asserts that the set of derived rules of a theory is closed under the operation of the substitution of correctly typed terms for variables, is needed in the definition of the category of generalised algebraic theories.

Substitution could have been taken as one of the principles of derivation; however to have done this would have hindered the definitions by induction which surround the semantics. Compare with Lambek [19], though of course the problem is Gentzens.

It is assumed throughout that U is some generalised algebraic theory. Let us say that a derived rule of U of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}}$ has the substitution property iff for every context Q of U, whenever $S_1, S_2, \ldots S_m$ are expressions such that $\dfrac{Q}{S_1 \in \Omega_1}$, $\dfrac{Q}{S_2 \in \Omega_2 [S_1 \mid y_1]}$, ... and $\dfrac{Q}{S_m \in \Omega_m [S_1 \mid y_1, \ldots S_{m-1} \mid y_{m-1}]}$ are all derived rules of U then the rule $\dfrac{Q}{\text{Conclusion}\,[S_1 \mid y_1, \ldots S_m \mid y_m]}$ is also a derived rule of U.

We aim to show that all derived rules of U have the substitution property.

We need two preliminary lemmas.

Lemma 1. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}}$ is a derived rule of U then any variables appearing in the conclusion occur among $\{x_1, \ldots x_n\}$.

Proof. By induction on the derivation of rules in U. Look at each principle of derivation in turn and see that it is impossible to use the principle to derive a rule without this property from rules which do have the property. This is very easy to see.

Lemma 2. (i) The premise of a derived rule is a context. (ii) If $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$ is a context then for all i, $1 \leq i \leq n$, the rule

$$\frac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i \text{ is a type.}}$$ is a derived rule.

Proof. (i) is proved by induction on derivations. If each principle of derivation is checked it will be seen that the premise of the derived rule is either a context by hypothesis (CF2, SI1 and SI2), or is a premise of a previously derived rule (LI1,...LI7, T1, R1 and A2), or else satisfies the conditions necessary to be a context (CF1).

(ii) follows from an iteration of (i). If $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$ is a context then $\dfrac{x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}}{\Delta_n \text{ is a type}}$ is a derived rule. Hence by (i), $x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}$ is a context. Continue until you get to $x_1 \in \Delta_1, \ldots x_i \in \Delta_i$ is a context and $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i \text{ is a type}}$ is a derived rule.

The substitution lemma. Every derived rule of the theory U has the substitution property.

Proof. The derived T=rules and $\in$=rules of U have the substitution property because there are principles of derivation SI1 and SI2 which have just that effect.

The proof that T and $\in$-rules of U have the substitution property is by induction on derivations in U. It suffices to show that no principle of derivation by which such rules are derived can be used to derive a rule without the property from rules with the property Thus we just have to check the principles T1, CF1 and CF2. T1. Suppose that both $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta_1 = \Delta_2}$ and $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t \in \Delta_1}$ are derived rules of U which have the substitution property. We must show that $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t \in \Delta_2}$ has the substitution property. So suppose that for each j, $1 \leq j \leq m$, $\dfrac{Q}{s_j \in \Omega_j [s_1 \mid y_1, \ldots s_{j-1} \mid y_{j-1}]}$ is a derived rule of U.

By our first assumption both $\dfrac{Q}{\Delta_1 [s_1 \mid y_1, \ldots s_m \mid y_m] = \Delta_2 [s_1 \mid y_1, \ldots s_m \mid y_m]}$

and $\dfrac{Q}{t [s_1 \mid y_1, \ldots s_m \mid y_m] \in \Delta_1 [s_1 \mid y_1, \ldots s_m \mid y_m]}$ are derived rules of U.

Thus, by an application of T1, So is

$\dfrac{Q}{t [s_1 \mid y_1, \ldots s_m \mid y_m] \in \Delta_2 [s_1 \mid y_1, \ldots s_m \mid y_m]}$ a derived rule of U.

Hence $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t \in \Delta_2}$ has the substitution property.

CF1. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a derived rule of U having the substitution property. We must show that $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}$ has the substitution property. So suppose that for each j, $1 \leq j \leq n+1$, $\dfrac{Q}{s_j \in \Delta_j [s_1 \mid x_1, \ldots s_{j-1} \mid x_{j-1}]}$ is a derived rule of U. By lemmas 1

and 2, $x_{j+1},\ldots x_{n+1}$ do not occur in Δ_j. Hence
$\Delta_j [s_1 \mid x_1,\ldots s_{j-1} \mid x_{j-1}] = \Delta_j [s_1 \mid x_1,\ldots s_{n+1} \mid x_{n+1}]$. Thus

$$\frac{Q}{s_i \in \Delta_i [s_1 \mid x_1,\ldots s_{n+1} \mid x_{n+1}]}$$ is a derived rule of U. Which is to say

$$\frac{Q}{(x_i \in \Delta_i)[s_1 \mid x_1,\ldots s_{n+1} \mid x_{n+1}]}$$ is a derived rule of U. Thus

$\underline{x_1 \in \Delta_1,\ldots x_{n+1} \in \Delta_{n+1}}$ has the substitution property. CF2 (a).
$x_i \in \Delta_i$

(CF2(b) is very similar and we shall not bother checking it).

Suppose that A is a sort symbol of U introduced by the rule
$\frac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{A(x_1,\ldots x_n) \text{ is a type}}$. Suppose that for each i, $1 \leq i \leq n$,

$$\frac{y_1 \in \Omega_1,\ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1,\ldots t_{i-1} \mid x_{i-1}]}$$

is a derived rule of U and has the substitution property. We must show that $\frac{y_1 \in \Omega_1,\ldots y_m \in \Omega_m}{A(t_1,\ldots t_n) \text{ is a type}}$ has the substitution property. So suppose that for each j, $1 \leq j \leq m$, the rule $\frac{Q}{S_j \in \Omega_j [s_1 \mid y_1,\ldots s_{j-1} \mid y_{j-1}]}$

is a derived rule of U. Then because for each i, $1 \leq i \leq n$,

$\underline{y_1 \in \Omega_1,\ldots y_m \in \Omega_m}$ has the substitution property and beacuse
$t_i \in \Delta_i [t_1 \mid x_1,\ldots t_{i-1} \mid x_{i-1}]$ it follows from lemma 1 that

$$\Delta_i [t_1 \mid x_1,\ldots t_{i-1} \mid x_{i-1}] [s_1 \mid y_1,\ldots s_m \mid y_m]$$
$$= \Delta_i [t_1 [s_1 \mid y_1,\ldots s_m \mid y_m] \mid x_1,\ldots t_{i-1} [s_1 \mid y_1,\ldots s_m \mid y_m] \mid x_{i-1}],$$

so it is the case that for each i, $1 \leq i \leq n$,

$$\frac{Q}{t_i[S_1 \mid y_1,\ldots S_m \mid y_m] \in \Delta_i[t_1[S_1 \mid y_1,\ldots S_m \mid y_m] \mid x_1,\ldots \quad / \quad t_{i-1}[S_1 \mid y_1,\ldots S_m \mid y_m] \mid x\text{-}1]}$$

is a derived rule of U. Thus, by an application of CF2(a), the

rule $\dfrac{Q}{A(t_1[S_1 \mid y_1,\ldots S_m \mid y_m],\ldots t_n[S_1 \mid y_1,\ldots S_m \mid y_m] \text{ is a type}}$

is a derived rule of U. Which is to say that the rule

$\dfrac{Q}{A(t_1,\ldots t_n)[S_1 \mid y_1,\ldots S_m \mid y_m] \text{ is a type}}$ is a derived rule of U.

Thus $\dfrac{y_1 \in \Omega_1,\ldots y_m \in \Omega_m}{A(t_1,\ldots t_n) \text{ is a type}}$ has the substitution property.

<u>Corollary</u>. (Change of Variables) If $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\text{Conclusion}}$ is a derived

rule of U and if $y_1,\ldots y_n$ is a sequence of distinct variables then

$$\frac{y_1 \in \Delta_1, y_2 \in \Delta_2[y_1 \mid x_1],\ldots y_n \in \Delta_n[y_1 \mid x_1,\ldots y_{n-1} \mid x_{n-1}]}{\text{Conclusion}\,[y_1\ x_1,\ldots y_n \mid x_n]}$$

is a derived rule of U.

<u>Proof</u>. The proof is by induction on n. If n=o then there is nothing to prove. If the result holds for n then it holds for n+1 as follows.

Suppose $\dfrac{x_1 \in \Delta_1,\ldots x_{n+1} \in \Delta_{n+1}}{\text{Conclusion}}$ is a derived rule of U and suppose that

$y_1,\ldots y_{n+1}$ is a sequence of distinct variables. Then by lemma 2 the rule

$\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a derived rule of U. Hence by the inductive

hypothesis so to is the rule $\dfrac{y_1 \in \Delta_1,\ldots y_n \in \Delta_n[y_1 \mid x_1,\ldots y_{n-1} \mid x_{n-1}]}{\Delta_{n+1}[y_1 \mid x_1,\ldots y_n \mid x_n] \text{ is a type}}$.

By applying CF1, the rule $\dfrac{y_1 \in \Delta_1, \ldots y_{n+1} \in \Delta_{n+1}[y_1 \mid x_1, \ldots y_n \mid x_n]}{y_i \in \Delta_i[y_1 \mid x_1, \ldots y_{i-1} \mid x_{i-1}]}$

is a derived rule of U, for each i, $1 \leqslant i \leqslant n+1$. Therefore by the substitution lemma and since $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\text{Conclusion}}$ is a derived rule of U we can conclude that $\dfrac{y_1 \in \Delta_1, \ldots y_{n+1} \in \Delta_{n+1}[y_1 \mid x_1, \ldots y_n \mid x_n]}{\text{Conclusion}\ [y_1 \mid x_1, \ldots y_{n+1} \mid x_{n+1}]}$ is a derived rule of U. The result holds for n+1. Hence the result holds for all $n \geqslant o$.

Lemma Every derived rule of a theory is wellformed.

Proof By induction on derivations in the theory U. We check each principle in turn, showing that all rules derived from wellformed rules are wellformed. LI1-LI7 and T1 are very easy to check. CF1. We must show that if $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a wellformed derived rule of U and if x_{n+1} is a variable distinct from $x_1, \ldots x_n$ then for each i, $1 \leqslant i \leqslant n$ the rule $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}$ is wellformed. That is we must show that $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}$

This is the case because, as above, for each j, $1 \leqslant j \leqslant i$, the rule $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_j \in \Delta_j}$ is derivable and because by lemma 2. (ii) the rule $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i \text{ is a type}}$ is derived rule. Using the substitution lemma, the rule $\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}$ is a derived rule.

CF2(a). Follows immediately from lemma 2. CF2(b), SI1 and SI2. Follow immediately from the substitution lemma. A1 and A2. These state that an axiom is a derived rule only if it is wellformed.

The Derivation Lemma. (a). Every derived T-rule of the theory U is of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{A(t_1, \ldots t_n) \text{ is a type}}$ for some sort symbol A of U with introductory rule of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$ and for some expressions $t_1, \ldots t_n$ such that for each i, $1 \leq i \leq n$, the rule $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}$ is derived rule of U. (b) Every derived $\in$-rule of U is either of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{y_j \in \Omega}$ for some j, $1 \leq j \leq m$, and for some Λ such that $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega_j = \Lambda}$ is a derived rule of U, or else is of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{f(t_1, \ldots t_n) \in \Omega}$ for some operator symbol f of U with introductory rule of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{f(x_1, \ldots x_n) \in \Delta}$ and for some expressions $t_1, \ldots t_n$, such that for each i, $1 \leq i \leq n$, $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}$ is a derived rule of U and such that $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] = \Omega}$ is a derived rule of U.

Proof. (a) Simply because the only principle of derivation that enables us to derive T-rules is principle CF2(a). (b) The principes which allow us to derive $\in$-rules are principles T1, CF1 and CF2(b).

If an $\in$-rule is derived by CF1 then it is immediately of the first of the two forms stated above, if it is derived by CF2(b) then it is immediately of the second form. It remains to consider the case of an $\in$-rule derived by T1.

First suppose that a rule $\dfrac{P}{t \in \Omega}$ is derived by T1 from $\dfrac{P}{t \in \Omega'}$ and $\dfrac{P}{\Omega = \Omega'}$ and also suppose that the-rule $\dfrac{P}{t \in \Omega'}$ is derived by T1 from some $\dfrac{P}{t \in \Omega''}$ and $\dfrac{P}{\Omega' = \Omega''}$. In this situation the rule $\dfrac{P}{t \in \Omega}$ could have been derived directly by T1 from $\dfrac{P}{t \in \Omega''}$ and $\dfrac{P}{\Omega = \Omega''}$, thus missing out a double application of T1. It follows that if a rule $\dfrac{P}{t \in \Omega}$ is derivable by an application of T1 then it is derivable by an application of T1 to some rules $\dfrac{P}{t \in \Omega'}$ and $\dfrac{P}{\Omega = \Omega'}$ such that the rule $\dfrac{P}{t \in \Omega'}$ is derivable by CF1 or CF2(a). It then follows that $\dfrac{P}{t \in \Omega}$ is of one of the two forms stated above.

<u>Corollary</u>. If $\dfrac{P}{t \in \Omega}$ and $\dfrac{P}{t \in \Omega'}$ are both derived rules of U then so to is $\dfrac{P}{\Omega = \Omega'}$.

The next lemma might indicate an alternative inductive definition of the notion of generalised algebraic theory.

If we say that U′ is a theory extending U then it is meant that all

the introductory rules and axioms of U are included among the introductory rules and axioms of U′. In particular every symbol of U is a symbol of U′.

An extension U′ of U is said to be a <u>simple extension</u> of U iff all of the introductory rules and axioms of U′ are well formed wrt U. For example, the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ of U′ is wellformed wrt U iff $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$ is a context of U.

<u>Lemma 3</u>. If U′ is a theory extending the theory U then there exists a sequence of theories $U_o, U_1, U_2, \ldots\ldots$ such that for each $i \geq 0$, U_{i+1} is a simple extension of U_i and such that $U_o = U$ and $\bigcup_{i \geq 0} U_i = U'$.

Proof. U_o is defined to be U. U_{i+1} is defined to be the simple extension of U_i given by all those symbols of U′ whose indroductory rules are wellformed wrt U_i and all those axioms of U′ which are wellformed wrt U_i. The only problem is to show that every symbol and axiom of U′ is eventually in U_i for some i. We just have to show that every introductory rule and axiom of U′ is wellformed wrt U_i for some i.

Suppose then that R is an introductory rule or an axiom of U′. Because R is wellformed wrt U′ it must be wellformed wrt some finite number k of introductory rules and axioms of U′. We show by induction on k that R is wellformed wrt U_k.

If $k=o$ then R is wellformed wrt $U=U_o$.

If $k>o$ Suppose S is one of those k introductory rules and axioms from which R can be shown to be wellformed. In any derivation we only use

an axiom or an introductory rule after it has been shown to be well formed; in particular since S is used in showing that R is wellformed there must be rules capable of showing that S is wellformed among the k rules that can be used to show R is wellformed. Thus it can be shown that S is wellformed from some number p of rules and axioms of U' where p is strictly smaller than k. By the inductive hypothesis S is wellformed wrt U_p. Thus S is an introductory rule or an axiom of U_k. This is the case for any of those k introductory rules and axioms of U' which can be used to show that R is wellformed. Thus R is well formed wrt U_k.

Lemma 4. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}}$

are both derived rules of U, if z is a variable distinct from $x_1, \ldots x_n, y_1, \ldots y_m$ then $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta, y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}}$

is a derived rule of U.

Proof. By induction on m. If m=o then from $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$

we can derive $\dfrac{}{x_i \in \Delta_i}$

for each i, $1 \leqslant i \leqslant n$. Since $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}}$ is a derived

rule by the substitution lemma so too is $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta}{\text{Conclusion}}$.

If $m \geqslant o$. Then since $x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y_1 \in \Omega_1, \ldots y_m \in \Omega_m$ is a context, $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y_1 \in \Omega_1, \ldots y_{m-1} \in \Omega_{m-1}}{\Omega_m \text{ is a type}}$ is a derived rule.

Thus by the inductive hypothesis $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta, y_1 \in \Omega_1, \ldots y_{m-1} \in \Omega_{m-1}}{\Omega_m \text{ is a type}}$ $E\Omega_{m-}$

is a derived rule. Thus $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta, y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{x_i \in \Delta_i}$,

$1 \leqslant i \leqslant n$, and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta, y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{y_j \in \Omega_j}$, $1 \leqslant j \leqslant m$, are

derived rules. By the substitution lemma

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Delta, y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}}$ is a derived rule.

1.8 Informal Syntax

There is a discrepancy between the syntax adopted in the formal definition of §1.6 and the syntax used in informally presenting theories in other sections. We say that we have a formal syntax and an informal syntax. The informal syntax is the syntax that is used in practice. In a particular case it provides an adequate and unambiguous language for the description of the structure involved. Since the informal syntax varies non uniformly from one particular theory to another it is impossible to give a direct description of the informal syntax. In this section we classify the discrepancies that occur between formal and informal. We also state a general problem and provide a partial solution. In this section we have in mind a very practical approach to mathematical syntax.

In the first place, the actual forms of the rules are of no consequence. Thus the form that is used in the formal syntax, that is $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}}$, has the advantage of alienation and is not significantly different from any of the other forms which have to some extent the advantage of naturality, forms such as

$x_1 \in \Delta_1, \ldots x_n \in \Delta_n$: Conclusion,

for $x_1 \in \Delta_1$, for $x_2 \in \Delta_2, \ldots$ and for $x_n \in \Delta_n$: Conclusion,

Conclusion, whenever $x_1 \in \Delta_1, x_2 \in \Delta_2, \ldots$ and $x_n \in \Delta_n$,

and such that, wherever possible, repetitions of expressions in the premise are avoided by writing $x_k, x_{k+1}, \ldots x_{k+c} \in \Delta$ instead of $x_k \in \Delta, \ldots x_{k+c} \in \Delta$.

There are two significant differences between the formal and the informal. The first of these is the omission of some of the

variables which would formally have to appear in the term that introduces a symbol into a theory, and the subsequent omission of terms from the argument places of the symbol as it appears in the derived rules of the theory. Thus in the theory of categories the term o(f,g) occurs in the introductory rule for o instead of the term o(x,y,z,f,g); the axioms differ accordingly.

The second difference between formal and informal is that informally one symbol can be made to do the work that several symbols would have to do formally. As examples of symbols that have to do the work of several we have the symbol S of the theory of trees and the symbols id,o,Hom and F of the theory of functors (with reference to the presentations of these theories in §1.2).

If we arbitrarily rewrite a formal theory by these two methods, that is if we omit certain variables from certain introductory rules, altering the derived rules accordingly, and if we replace certain collections of symbols by single symbols, then ambiguities may or may not arise. There is ambiguity just when two formally distinct derived rules are rewritten as indentical, for this would mean that there was an informal rule which had two meanings, was ambiguous. So the problem is - in what ways can we rewrite a given formal theory without ambiguities arisisng? The answer is that it depends on the theory in question. The best general answer that we can give consists of a condition that the omission of variables must respect if ambiguities are not to arise. This condition objectifies the dropping of the variables x,y,z from the term o(x,y,z,f,g) in the introductory rule for o and the wrongheadedness of dropping f or g or both from this same term. Intuitively, o(f,g) depends explicitly on f and g, f depends explicitly on x and on y

because $f \in Hom(x,y)$, g depends explicitly on y and on z because $g \in Hom(y,z)$, thus o(f,g) depends implicitly on all of the variables x,y,z,f and g that occur in the premise of the introductory rule, so that although the variables x,y and z no longer appear explicitly in o(f,g) there is still an implicit dependence of o(f,g) on each of x,y, and z. The condition on an introductory rule which is necessary if ambiguities are not to arise is that all variables occuring in the premise must occur implicitly in the conclusion. This condition we can call the condition of implicit occurence. It is necessary but not sufficient, as we shall show.

The definition of implicit occurence must be given inductively. Suppose that P is the premise $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$, suppose C is a conclusion and that $1 < i \leq n$, then we say that the variable

x_i occurs implicitly in the conclusion C wrt the premise P

iff either x_i actually occurs in C (in which case we also say x_i occurs explicitly) or if for some $j > i$, x_i appears in Δ_j and x_j occurs implicitly in C wrt P.

Lemma If ambiguity is not to occur in an informal theory then whenever L is a symbol introduced by the rule $x_1 \in \Delta_1, \ldots x_n \in \Delta_n : C$ then each of $x_1, \ldots x_n$ must occur implicitly in C (the condition of implicit occurence).

Proof. Assume that we have a theory and an informal presentation of that theory in which there is a symbol L introduced by a rule that does not satisfy the condition of implicit occurence. We shall suppose that L is an operator symbol for if otherwise and L is a sort symbol then the argument is the same. Suppose that $n \geq 1$ and that $1 \leq j_1 < j_2 \ldots \leq j_r$ suppose that L is introduced by the rule $x_1 \in \Delta_1, \ldots x_n \in \Delta_n : L(x_{j_1}, \ldots x_{j_r}) \in \Delta$. We are

assuming that not all of $x_1,\ldots x_n$ occur implicitly in $L(x_{j_1},\ldots x_{j_r})\in\Delta$ wrt the premise $x_1\in\Delta_1,\ldots x_n\in\Delta_n$. We show that there are two derived rules which are distinct in the more formal syntax, that is when L is introduced by $x_1\in\Delta_1,\ldots x_n\in\Delta_n: L(x_1,\ldots x_n)\in\Delta$, but which are indistinct in the informal syntax.

Let $x_1',\ldots x_n'$ be a sequence of variables each one of which is distinct from each one of $x_1,\ldots x_n$. Let $J=\{j \mid 1\leq j\leq n$ and x_j does not occur implicitly in $L(x_{j_1},\ldots x_{j_r})\in\Delta\}$. Let $y_1,\ldots y_n$ be the sequence of variables given by $y_j=x_j$ if $j\in J$, $y_j=x_j'$ if $j\notin J$.

Suppose that $j\notin J$ and $j'\in J$, then x_j occurs implicitly in $L(x_{j_1},\ldots x_{j_r})\in\Delta$ whereas $x_{j'}$ does not, hence $x_{j'}$ does not occur in Δ_j. Thus if $j\notin J$ then $\Delta_j[y_1 \mid x_1,\ldots y_{j-1} \mid x_{j-1}]$ is Δ_j. Also note that $\Delta[y_1 \mid x_1,\ldots y_n \mid x_n]$ is Δ, since if $j\in J$ then x_j does not occur in Δ.

By the change of variables lemma the rule

$y_1\in\Delta_1,\ldots y_n\in\Delta_n[y_1 \mid x_1,\ldots y_{n-1} \mid x_{n-1}]$: $L(y_1,\ldots y_n)\in\Delta[y_1 \mid x_1,\ldots y_n \mid x_n]$ is a derived rule. By the preceeding paragraph this rule is just the rule $a_1,\ldots a_n: L(y_1,\ldots y_n)\in\Delta$, where a_j is $x_j\in\Delta_j$ when $j\notin J$ and a_j is $x_j\in\Delta_j[y_1 \mid x_1,\ldots y_{j-1} \mid x_{j-1}]$ when $j\in J$. By lemma 4 of §1.7 if we extend the premise of this rule by inserting the clause $x_j\in\Delta_j$ after the clause a_j whenever $j\in J$ then the new rule is still derivable. If we call this extended premise Q then the rule Q: $L(y_1,\ldots y_n)\in\Delta$ is a derived rule.

Since Q is a context extending the context $x_1\in\Delta_1,\ldots x_n\in\Delta_n$ and using lemma 4 of §1.7 we deduce that Q: $L(x_1,\ldots x_n)\in\Delta$ is a derived rule

Now $Q{:}L(y_1,\ldots y_n)\in\Delta$ and $Q{:}L(x_1,\ldots x_n)\in\Delta$ are distinct since we have assumed J to be non empty. In the informal syntax both these rules become $Q{:}L(x_{j_1},\ldots x_{j_r})\in\Delta$.
Hence informally the rule $Q{:}L(x_{j_1},\ldots x_{j_r})\in\Delta$ is ambiguous.

Finally we show that the condition of implicit occurence is not sufficient to assure that ambiguities do not arise when variables are omitted.

Consider the following theory:

Symbol	Introductory Rule
A	A is a type.
B	For $x\in A{:}B(x)$ is a type.
C	For $x\in A$, for $y\in B(x){:}C(x,y)$ is a type.
D	For $y\in B(a_1)$, for $z\in C(a_1,y){:}D(y,z)$ is a type.
a_1	$a_1\in A$.
a_2	$a_2\in A$.

Axioms.

$B(a_1)=B(a_2)$.

See that in this theory the rules $y\in B(a_1){:}C(a_1,y)$ is a type and $y\in B(a_1){:}C(a_2,y)$ is a type are both derivable.

The theory might be rewritten informally by introducing the symbol C by the rule for $x\in A$, for $y\in B(x){:}C(y)$ is a type, clearly the condition of implicit occurence is respected. However the two rules $y\in B(a_1){:}C(a_1y)$ is a type and $y\in B(a_1){:}C(a_2,y)$ is a type are rewritten as the rule $y\in B(a_1){:}C(y)$ is a type in this informal syntax.

Thus though the condition of implicit occurence is respected ambiguity still arises.

We have included the symbol D in the theory to illustrate that an ambiguous rule in a theory easily leads to an ambiguous theory. Informally the symbol D is now introduced by the rule for $y \in B(a_1)$, for $z \in C(y):D(y,z)$ is a type; the formal theory can no longer be recovered from its informal presentation since this presentation could equally, well be the informal presentation of the theory that differs from the given theory in that the symbol D is introduced by the rule

for $y \in B(a_1)$, for $z \in C(a_2,y):D(y,z)$ is a type.

1.9 Models and Homomorphisms

We neglect the formal definitions of model and homomorphism. It should be quite clear what we mean by model. We have a few words to say about homomorphisms but first we wish to develop some notations which relate to the matter but are actually of more importance in the next chapter.

We begin rather distantly with trees. The theory of trees was given in §1.2 but from now on we want all our trees to be trees with a unique least element. If we come across a tree which does not have a unique least element then we quickly adjoin a new element 1 beneath all the other elements. If Θ is a tree and if A is a node of the tree then we say that $A \in \Theta$, so confusing the tree with its set of nodes. If we wish to assert that B is a node of the tree Θ and that B succeeds A then we just say that $A \lhd B$ in Θ. The least node of Θ is alway denoted 1. Thus if A is any node of the tree Θ distinct from 1 then there exists a unique $n \geqslant o$, there exists uniquely $A_1, \ldots A_n$ such that $1 \lhd A_1 \lhd A_2 \ldots A_n \lhd A$ in Θ.

We are interested in trees because for any theory U, the set of contexts of U is structured as a tree. The least element of the tree is the empty context $\langle\rangle$. For any $n \geqslant 1$ the predecessor of the node $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is the node $\langle x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1} \rangle$.

We wish to identify a large tree of sets, families of sets, families of families of sets and so on. It looks as if we should call it the family tree. But we won't. Anyway it is first necessary to consider the notation that we use for families.

If A is a set and if for every element $a \in A$, B(a) is a set, then the corresponding A-indexed family of sets is denoted $\lambda a \in A.B(a)$ or just as $\lambda a.B(a)$. Now suppose that in this situation we also have a set C(a,b) for every $a \in A$ and for every $b \in B(a)$. Now if a is an element of A then $\lambda b \in B(a).C(a,b)$ is a B(a) indexed family of sets. Thus $\lambda a \in A.\lambda\, b \in B(a).C(a,b)$ is an A-indexed family of families of sets. We can continue in this way. The whole collection of sets, families of set, families of families of sets and so on is structured as a (large) tree. We call it the tree of families. The next thing to do is to turn the notation about. If A_1 is a set and if A_2 is an A_1-indexed family of sets then we write $A_2(a_1)$ for the value of the family at an element $a_1 \in A_1$. We do the same for families of families and so on. In general if $1 \lhd A_1 \lhd A_2 \ldots \lhd A_n \lhd A$ in the tree of families then for any $a_1 \in A_1, \ldots$ for any $a_n \in A_n(a_1, \ldots a_{n-1})$, $A(a_1, \ldots a_n)$ is a set.

Lastly we wish to be precise about the term operator. If $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in the tree of families and if for any $a_1 \in A_1, \ldots$ for any $a_n \in A_n(a_1, \ldots a_{n-1})$, $f(a_1, \ldots a_n) \in A(a_1, \ldots a_n)$ then we say that $\lambda a_1 \in A_1 \ldots \lambda a_n \in A_n.f(a_1, \ldots a_n)$ is an operator at A. Thus for any node A of the tree of families there is a set of operators at A. If we turn the notation about then we always write $g(a_1, \ldots a_n)$ for the value of the operator g at arguments $a_1, \ldots a_n$. If $1 \lhd A_1 \ldots A_n \lhd A$ in the tree of families, if g is an operator at A then the <u>status</u> of the operator g is given by the rule for every $a_1 \in A_1, \ldots$ for every $a_n \in A_n(a_1, \ldots a_{n-1})$, $g(a_1, \ldots a_n) \in A(a_1, \ldots a_n)$. For example if $\underline{C}$ is a real live category then id is an operator whose status is given by for every $a \in |\underline{C}|$, $id(a) \in Hom_{\underline{C}}(a,a)$. Alternivel id is an operator at $\lambda a \in |\underline{C}|.Hom_{\underline{C}}(a,a)$.

Now we return to the question of models of a theory. In the first place it is the derived rules of a theory that are interpreted in a model, not the expressions. If U is a theory with model M and if R is a derived T-rule of the theory, say $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$,

if we let R_i be the rule $\frac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i \text{ is a type}}$, whenever $1 \leq i \leq n$,

if the interpretation of a rule within the model M is written as that rule superscripted by m. Then R_1^m is a set, R_2^m is an R_1^m- indexed family of sets and in general $1 \triangleleft R_1^m \triangleleft R_2^m \ldots \triangleleft R_n^m \triangleleft R^m$ in the tree of families.

Further if R_t is a derived rule of U of the form $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$

then the interpretation R_t^m of the rule R_t by the model M is an operator at R^m.

We note that if M and M′ are both models of U and if f: M $\longrightarrow$ M′ is a homomorphism then for every derived rule R of U of the form $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ there is an operator f_R whose status is given by for $a_1 \in R_1^m, \ldots$ for $a_n \in R_n^m(a_1, \ldots a_{n-1})$, for $a \in R^m(a_1, \ldots a_n)$, $f_R(a_1, \ldots a_n, a) \in R^{m'}(f_{R_1}(a), \ldots f_{R_n}(a_1, \ldots a_n), f_R(a_1, \ldots a_n, a))$. It is required that for all derived rules R_t of the form $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$

it is the case that for all $a_1 \in R_1^m, \ldots$ for all $a_n \in R_n^m(a_1, \ldots a_{n-1})$, $f_R(a_1, \ldots a_n, R_t^m(a_1, \ldots a_n)) = R_t^{m'}(f_{R_1}(a_1), \ldots f_{R_n}(a_1, \ldots a_n))$. In fact if this condition holds for all $F(x_1, \ldots x_n)$, F an operator symbol, then it will hold for all t. The only requirement of a homomorphism that

might appear unusual is the requirement that whenever $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$ is a derived rule of U then $f_R = f_{R'}$ where R is the rule $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ and where, R' is the rule $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta' \text{ is a type}}$.

In the definition of homomorphism we just ask for a family $\lambda L \in \{\text{sort symbols of } U\}.f_L$ of operators of the correct status and then define the f_R by induction (in this notation $f_{\text{introductory rule of } L} = f_L$). And we require just that the two conditions mentioned above are satisfied by the f_R.

With homomorphism so defined the notion reduces to the usual model theoretic notion in the special case of the universal conditional theories expressed as generalised algebraic as in §1.3.

1.10 A Short list of theories

We list some generalised algebraic theories and pay particular attention to the sort structures.

Firstly, extending the theory of categories by operator symbols and axioms alone (thus there are no additional sort symbols) are the theories of categories with finite products, cartesian closed categories, categories with finite coproducts, monoidal categories, closed categories, additive categories, U -categories for a given algebraic theory U , groupoids, preorders, partial orders, lattices and monads. Extending the theory of categories by just the equality predicate for morphisms, new operator symbols and new axioms are the theories of categories with equalisers of pairs, categories with finite limits, abelian categories and topoi.

Extending the theory of functors by just operator symbols and axioms are the theories of natural transformations, adjoint pairs and equivalences. In more diverse sort structures we have the theories of n-categories for fixed n, multicategories, category valued presheaves on an arbitrary category, category valued presheaves on a given category. The theory of hyperdoctrines extends the theory of category valued presheaves by operator symbols and axioms alone. Later we shall come across the theory of contextual categories, a theory which extends the theory of trees.

1.11 Interpretations

The notion of an interpretation of one generalised algebraic theory in another is defined in such a way that there is a category of generalised algebraic theories and interpretations. It is worth noting of the definition that the induction that occurs is on the expressions of a language rather than on the derived rules and that as such the construction is very simple.

The alphabet in which a theory U is written we denote A_U.

We assume throughout some fixed enumeration $v_1, v_2, v_3, \ldots$ of the set V of variables. We do this because symbols $L \in A_U$ are going to be interpreted by expressions I(L) of a theory U'. We will wish to know which free variables in the expression I(L) correspond to which argument places associated with L. The simplest way this can be done is to assume the enumeration of V and then chose I(L) such that v_1 corresponds to the first argument place of L, v_2 to the second and so on.

We first define the notion of a preinterpretation and then go on to eventually define an interpretation to be a well formed preinterpretation.

Definitions

A Preinterpretation I of the theory U in the theory U' consists just of a function I: $A_U \longrightarrow$ Expressions of U'.

If I is a preinterpretation of U in U' then define the function $\dot{I}$: Expressions of U $\longrightarrow$ Expressions of U' by induction (see the inductive definition of expression in §1.6) with the following clauses:

1. If $x \in V$ the $\dot{I}(x) = x$.

2. If $L \in A_u$ then $\dot{I}(L) = I(L)$.

3. If $L \in A_u$ and $e_1, \ldots e_n$ are expressions then
$\dot{I}(L(e_1, \ldots e_n)) = I(L)[\dot{I}(e_1) \mid v_1, \ldots \dot{I}(e_n) \mid v_n]$.

If I is a preinterpretation of U in U' then define a function $\hat{I}$: Rules of U $\longrightarrow$ Rules of U' by:

1. $$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}\right) = \frac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(\Delta) \text{ is a type}} .$$

2. $$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right) = \frac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t) \in \dot{I}(\Delta)} .$$

3. $$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}\right) = \frac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{I(\Delta) = I(\Delta')} .$$

4. $$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}\right) = \frac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t) = \dot{I}(t') \in \dot{I}(\Delta)} .$$

An <u>interpretation</u> I of U in U' is a preinterpretation I of U in U' such that for all introductory rules and axioms r of U, I(r) is a derived rule of U'.

If I is an interpretation of U in U' then for any derived rule r of U, I(r) is a derived rule of U'. We prove this after proving a preliminary lemma.

<u>Lemma 1</u>. If I is a preinterpretation of U in U' and if e and $d_1, \ldots d_m$ are expressions of A_U then

$$\dot{I}(e[d_1 \mid y_1,\ldots d_m \mid y_m]) = \dot{I}(e)[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m].$$

Proof. By induction on the length of the expression e.

1. If $e = x \in V$ then, if $x = y_i$ for some i, $1 \leq i \leq m$, then $L.H.S = \dot{I}(e_i) = R.H.S.$, otherwise $L.H.S. = x = R.H.S.$

2. If $e = L \in A_U$ then $L.H.S. = I(L) = R.H.S.$

3. If $e = L(e_1,\ldots e_n)$ for some $L \in A_U$ and for some expressions $e_1,\ldots e_n$ of A_U then $\dot{I}(e[d_1 \mid y_1,\ldots d_m \mid y_m]) =$
$\dot{I}(L(e_1[d_1 \mid y_1,\ldots d_m \mid y_m],\ldots e_n[d_1 \mid y_1,\ldots d_m \mid y_m])) =$
$I(L)[\dot{I}(e_1[d_1 \mid y_1,\ldots d_m \mid y_m]) \mid v_1,\ldots \dot{I}(e_n[d_1 \mid y_1,\ldots d_m \mid y_m]) \mid v_n]$.
By the inductive hypothesis $\dot{I}(e_i[d_1 \mid y_1,\ldots d_m \mid y_m]) =$
$\dot{I}(e_i)[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m]$. Hence
$\dot{I}(e[d_1 \mid y_1,\ldots d_m \mid y_m]) = I(L)[\dot{I}(e_1)[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m] \mid v_1,\ldots$
$\dot{I}(e_n)[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m] \mid v_n] = I(L)[\dot{I}(e_1) \mid v_1,\ldots \dot{I}(e_n) \mid v_n]$
$[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m] = \dot{I}(L(e_1,\ldots e_n))[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m]$
$= \dot{I}(e)[\dot{I}(d_1) \mid y_1,\ldots \dot{I}(d_m) \mid y_m]$. As required.

Lemma 2. If I is an interpretation of U in U' then for every derived rule r of U, $\hat{I}(r)$ is a derived rule of U'.

Proof. We wish to show that all derived rules r of U have the property that $\hat{I}(r)$ is derivable. We are given that the axioms have this property so it suffices to show that the principles of derivation transmit the property. That is we should check that each principle of derivation when applied to rules with the property yields a rule with the property. Principles LI1.-7., T1. and CF1. are incredibly easy to check. The principle CF2(b) is similar to the principle CF2(a)

and also principle SI1. is similar to principle SI2.. In view of this we just check CF2(a). and SI2.

CF2(a). Suppose that A is a sort symbol of U introduced by the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$, and that for each i, $1 \leq i \leq n$,

$\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 | x_1, \ldots t_{i-1} | x_{i-1}]}$ is a derived rule of U. Also suppose that

$\hat{I}\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 | x_1, \ldots t_{i-1} | x_{i-1}]}\right)$ is a derivedrule of U'. We wish to show that $\hat{I}\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{A(t_1, \ldots t_n) \text{ is a type}}\right)$ is a derived rule of U'.

Since I is an interpretation, $\hat{I}$ (introductory rule for A) is a derived rule of U'. That is $\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(A(x_1, \ldots x_n)) \text{ is a type}}$ is a derived rule of U'.

By Lemma 1., $\hat{I}\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 | x_1, \ldots t_{i-1} | x_{i-1}]}\right)$

$$= \frac{y_1 \in \dot{I}(\Delta_1), \ldots y_m \in \dot{I}(\Delta_m)}{\dot{I}(t_i) \in \dot{I}(\Delta_i) [\dot{I}(t_1) | x_1, \ldots \dot{I}(t_n) | x_n]}$$

This rule is a derived rule for each i, $1 \leq i \leq n$. Hence by the substitution lemma the rule $\dfrac{y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)}{\dot{I}(A(x_1, \ldots x_n)) [\dot{I}(t_1) | x_1, \ldots \dot{I}(t_n) | x_n]}$ is a derived rule of U

But $\dot{I}(A(x_1,\ldots x_n)) = I(A)[\,x_1 \mid v_1,\ldots x_n \mid v_n\,]$, hence $\dot{I}(A(x_1,\ldots x_n))[\,\dot{I}(t_1)\mid x_1,\ldots \dot{I}(t_n)\mid x_n\,] = I(A)[\,\dot{I}(t_1)\mid v_1,\ldots \dot{I}(t_n)\mid v_n\,] = \dot{I}(A(t_1,\ldots t_n))$. So we have shown that $\hat{I}\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{A(t_1,\ldots t_n) \text{ is a type}}\right)$ is a derived rule of U', as $\hat{I}\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{A(t_1,\ldots t_n) \text{ is a type}}\right)$ is by definition

$$\frac{y_1\in \dot{I}(\Omega_m),\ldots y_m\in \dot{I}(\Omega_m)}{\dot{I}(A(t_1,\ldots t_n)) \text{ is a type}}\ .$$

SI2. Suppose that $\dfrac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{t = t'\in\Delta}$ is a derived rule of U and that for each $i, 1\leq i\leq n$, $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{t_i=t_i'\in\Delta_i[\,t_i' \mid x_1,\ldots t_{i-1}'\mid x_{i-1}]}$ is a derived rule of U. Suppose that $\hat{I}$ applied to each of the rules yields a derivable rule of U'.

By definition $\hat{I}\left(\dfrac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{t = t'\in\Delta}\right) = \dfrac{x_1\in\dot{I}(\Delta_1),\ldots x_n\in\dot{I}(\Delta_n)}{\dot{I}(t) = \dot{I}(t')\in\dot{I}(\Delta)}$,

this latter rule then is a derived rule of U'.

By definition and Lemma 1., $\hat{I}\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{t_i = t_i'\in\Delta_i[\,t_i' \mid x_1,\ldots t_{i-1}' \mid x_{i-1}]}\right)$

$= \dfrac{y_1\in \dot{I}(\Omega_1),\ldots y_m\in \dot{I}(\Omega_m)}{\dot{I}(t_i) = \dot{I}(t_i')\in \dot{I}(\Delta_i)[\,(t_1') \mid x_1,\ldots I(t_{i-1}') \mid x_{i-1}]}$. This latter rule is derivable for each $i, 1\leq i\leq n$.

Hence using principle SI2.(wrt.U'), the rule

$$\frac{y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)}{\dot{I}(t)[\dot{I}(t_1) \mid x_1, \ldots \dot{I}(t_n) \mid x_n] = \dot{I}(t')[\dot{I}(t_1') \mid x_1, \ldots \dot{I}(t_n') \mid x_n] \in \dot{I}(\Delta)[I(t_1') \mid x_1, \ldots I(t_n') \mid x_n]}$$

is a derived rule of U'. By definition and lemma 1. this last rule is just

$$\hat{I}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] = t'[t_1' \mid x_1, \ldots t_n' \mid x_n] \in \Delta[t_1' \mid x_1, \ldots t_n' \mid x_n]}\right).$$

As required, this rule we have shown to be derivable in U'.

If U,U' and U'' are theories and I is a preinterpretation of U in U' and I'is a preinterpretation of U' in U'' then define I'oI to be the preinterpretation of U in U'' given by $(I'oI)(L) = \dot{I}'(I(L))$, for all $L \in A_U$. On the way to showing that there is a category of generalised algebraic theories and interpretations we need the following lemma and corollaries.

Lemma 3. If I and I' are as above then for any expression e of A_U, $(I'oI)^{\cdot}(e) = \dot{I}'(\dot{I}(e))$.

Proof. By induction on the length of e.

1. If $e = x \in V$ then $(I'oI)^{\cdot}(e) = x = \dot{I}'(\dot{I}(e))$.

2. If $e = L \in A_U$ then $(I'oI)^{\cdot}(L) = (I'oI)(L) = \dot{I}'(I(L)) = \dot{I}'(\dot{I}(L))$.

3. If $e = L(e_1, \ldots e_n)$ then $(I'oI)^{\cdot}(e) = (I'oI)(L)[(I'oI)^{\cdot}(e_1) \mid v_1, \ldots (I'oI)^{\cdot}(e_n) \mid v_n] = \dot{I}'(I(L))[\dot{I}'(\dot{I}(e_1)) \mid v_1, \ldots \dot{I}'(I(e_n)) \mid v_n]$, by the inductive hypothesis, $= \dot{I}'(I(L)[\dot{I}(e_1) \mid v_1, \ldots \dot{I}(e_n) \mid v_n]$, by lemma 1, $= \dot{I}'(\dot{I}(L(e_1, \ldots e_n)) = \dot{I}'(\dot{I}(e))$, as required.

Corollary 4 If I an I' are as above then for any derived rule r of $(\widehat{I'oI})(r) = \hat{I}'(\hat{I}(r))$.

Corollary 5 If I is an interpretation of U in U' and I' in an interpretation of U' in U'' then IoI' is an interpretation of U in U''.

We need some identity morphisms. If U is a theory define a preinterpretation id_U of U in U by:

1. If A is a sort symbol of U introduced by $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$,

then define $id_U(A) = A(v_1, \ldots v_n)$.

2. If f is an operator symbol of U introduced by $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta}{f(x_1, \ldots x_n) \in \Delta}$

then define $id_U(f) = f(v_1, \ldots v_n)$.

id_U is a preinterpretation of U in U which is quickly seen to have the property that for all expressions e of A_U $\dot{id}_U(e) = e$. Thus it has the property that for all rules r of U, $\hat{id}_U(r) = r$. Hence id_U is an interpretation.

It is now clear that there is a category of generalised algebraic theories and interpretations.

Any interpretation $I:U \longrightarrow U'$ induces a functor between the categories of algebras, denoted $I:alg: \underline{U'\text{-alg}} \longrightarrow \underline{U\text{-alg}}$. Any funct equivalent to I-alg, for some I, is said to be generalised algebraic.

Examples

1. If the theory U is included in the theory U', that is if every introductory rule and axiom of U is a derived rule of the theory U', then there is a canonical interpretation of U in U'. The corresponding algebraic functor for $\underline{U'\text{-alg}}$ to $\underline{U\text{-alg}}$ is usually called a forgetful functor.

2. If $\underline{C}$ and $\underline{D}$ are categories and $F:\underline{C}\longrightarrow\underline{D}$ is a functor then there is an interpretation I_F of the theory of category valued presheaves on $\underline{C}$ into the theory of category valued presheaves on $\underline{D}$. The induced generalised algebraic functor is the functor $\underline{Cat}^F$: $\underline{Cat}^{\underline{D}^{op}}\longrightarrow\underline{Cat}^{\underline{C}^{op}}$.

3. The functor which takes an adjoint pair to the monad induced by that adjoint pair is generalised algebraic. It is induced by an interpretation of the theory of monads in the theory of adjoint pairs.

4. The functor which takes a category valued presheaf on an arbitrary category to the total category of its fibration is generalised algebraic. However it is not induced by an interpretation of the theory of categories into the theory of category valued presheaves, as such. Rather it is induced by an interpretation into the theory of category valued presheaves extended by symbols for disjoint unions (one for objects, one for morphisms). The extension of a theory by symbols for disjoint unions is discussed in §1.2.

1.12 Contexts amd Realisations

Definition If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ are contexts of a theory U then a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is an m-tuple $\langle t_1, \ldots t_m \rangle$ such that for each $j, 1 \leqslant j \leqslant m$, $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t_j \in \Omega_j [t_1 \mid y_1, \ldots t_{j-1} \mid y_{j-1}]}$

is a derived rule of U.

If U is a generalised algebraic theory then there is a category $\underline{R(U)}$ of contexts and realisations of U. In $\underline{R(U)}$, $\langle t_1, \ldots t_m \rangle : \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \longrightarrow \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ iff $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$. The identity morphism on an object $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ of $\underline{R(U)}$ is the realisation $\langle x_1, \ldots x_n \rangle$ of $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$. Composition in $\underline{R(U)}$ is defined by $\langle t_1, \ldots t_m \rangle \circ \langle S_1, \ldots S_\ell \rangle =$

$\langle S_1 [t_1 \mid y_1, \ldots t_m \mid y_m], \ldots S_\ell [t_1 \mid y_1, \ldots t_m \mid y_m] \rangle$, whenever

$\langle t_1, \ldots t_m \rangle : \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \longrightarrow \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ and

$\langle S_1, \ldots S_\ell \rangle : \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle \longrightarrow \langle z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell \rangle$ in $\underline{R(U)}$.

The set of objects of $\underline{R(U)}$, ie the set of contexts of U, is structured as a tree with least element $\langle\rangle$ (the empty context) and with $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$, $n > o$, preceeded by $\langle x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1} \rangle$. For any $n, m > o$ and for any context $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x_{n+1} \in \Delta_{n+1}, \ldots x_{n+m} \in \Delta_{n+m} \rangle$ of U, $\langle x_1, \ldots x_n \rangle$ is a realisation of $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_{n+m} \in \Delta_{n+m} \rangle$. This map, $\langle x_1, \ldots x_n \rangle : \langle x_1 \in \Delta_1, \ldots x_{n+m} \in \Delta_{n+m} \rangle \longrightarrow$

$\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ in R(U), is denoted $p(\langle x_1 \in \Delta_1, \ldots x_{n+m} \in \Delta_{n+m} \rangle, \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle)$. Thus for any $A, B \in R(U)$ such that $A \leq B$, the morphism $p(B,A)$ is defined and $p(B,A): B \longrightarrow A$ in R(U).

If $A = \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$, $A' = \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$, if $B = \langle y_1 \in \Omega_1, \ldots y_{m+\ell} \in \Omega_{m+\ell} \rangle$ and if $f = \langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$, (in which case

$$\begin{array}{ccc} B & & \\ \downarrow & p(B,A') & \\ A & \xrightarrow[f]{} & A' \end{array}$$

in R(W) then

$$\begin{array}{ccc} f^*B & \xrightarrow{q(f,B)} & B \\ \downarrow p(f^*B,A) & & \downarrow p(B,A') \\ A & \xrightarrow[f]{} & A' \end{array}$$

is a pullback diagram in R( ).

Where $f^*B = \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y_{m+1} \in \Omega_{m+1}[t_1 \mid y_1, \ldots t_m \mid y_m], \ldots$ $y_{m+\ell} \in \Omega_{m+\ell}[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle$ and $q(f,B) = \langle t_1, \ldots t_m, y_{m+1}, \ldots y_{m+\ell} \rangle$.

1.13 Intended identity of denotation

We define an equivalence relation on the derived T and $\in$-rules of a theory U which we call the equivalence relation of intended identity of denotation. The idea is that if $r \equiv r'$ by this relation then any kind of model of U should interpret r and r' as identical objects. Semantically, we should not distinguish between the rule

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}} \quad \text{and the rule} \quad \frac{x_1' \in \Delta_1, \ldots x_n' \in \Delta_n[x_1' \mid x_1, \ldots x_{n-1}' \mid x_{n-1}]}{\Delta[x_1' \mid x_1, \ldots x_n' \mid x_n] \text{ is a type}}$$

because the two rules only differ by the choice of variables. Neither should we distinguish between $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ and $\dfrac{x_1 \in \Delta_1', \ldots x_n \in \Delta_n'}{\Delta' \text{ is a type}}$

if for each $i, 1 \leq i \leq n$, $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Delta_i'}$ is a derived rule and

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$$ is a derived rule.

The relation $\equiv$, of intended identity of denotation, is defined between T-rules of a theory U by $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega_{m+1} \text{ is a type}}$

iff $n = m$ and for each $i, 1 \leq i \leq n+1$, $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$ is a

derived rule of U.

Lemma 1. $\equiv$ is a equivalence relation.

Proof.

1. Reflexiveness follows from the principle of derivation LI1.

2. Symmetry. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{\Omega_{n+1} \text{ is a type}}$

we prove by induction on i that for all $i, 1 \leq i \leq n+1$,

$\dfrac{y_1 \in \Omega_1, \ldots y_{i-1} \in \Omega_{i-1}}{\Omega_i = \Delta_i[\, y_1 \mid x_1, \ldots y_{i-1} \mid x_{i-1}\,]}$ is derivable.

If $i = 1$ then certainly $\Omega_1 = \Delta_1$ is derivable since $\Delta_1 = \Omega_1$ is.

Now if $i > 1$ and if for all j, $1 \leq j < i$, $\dfrac{y_1 \in \Omega_1, \ldots y_{j-1} \in \Omega_{j-1}}{\Omega_j = \Delta_j[\, y_1 \mid x_1, \ldots y_{j-1} \mid x_{j-1}\,]}$

is derivable; then for all such j $\dfrac{y_1 \in \Omega_1, \ldots y_{j-1} \in \Omega_{j-1},\ y_j \in \Omega_j}{y_j \in \Delta_j[\, y_1 \mid x_1, \ldots y_{j-1} \mid x_{j-1}\,]}$

is derivable. Hence for all such j, $\dfrac{y_1 \in \Omega_1, \ldots y_{i-1} \in \Omega_{i-1}}{y_j \in \Delta_j[\, y_1 \mid x_1, \ldots y_{j-1} \mid x_{j-1}\,]}$

is derivable. Since $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Omega_i[\, x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}\,]}$ is derivable,

by the substitution lemma so is $\dfrac{y_1 \in \Omega_1, \ldots y_{i-1} \in \Omega_{i-1}}{\Delta_i[\, y_1 \mid x_1, \ldots y_{i-1} \mid x_{i-1}\,] = \Omega_i}$

(because $\Omega_i[\, x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}\,][\, y_1 \mid x_1, \ldots y_{i-1} \mid x_{i-1}\,] = \Omega_i$).

Hence $\dfrac{y_1 \in \Omega_1, \ldots y_{i-1} \in \Omega_{i-1}}{\Omega_i = \Delta_i[\, y_1 \mid x_1, \ldots y_{i-1} \mid x_{i-1}\,]}$ is derivable as required. It

follows by induction that $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{\Omega_{n+1} \text{ is a type}} \equiv \dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$

3. Transitivity. Assume that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega_{m+1} \text{ is a type}}$

and that $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega_{m+1} \text{ is a type}} \equiv \dfrac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Lambda_{\ell+1} \text{ is a type}}$.

Suppose $1 \leq i \leq n+1$. We have seen above that for each $j, 1 \leq j < i$

$\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{x_j \in \Omega_j[x_1 \mid y_1, \ldots x_{j-1} \mid y_{j-1}]}$ is a derived rule of U. From the substitution lemma and $\dfrac{y_1 \in \Omega_1, \ldots y_{i-1} \in \Omega_{i-1}}{\Omega_i = \Lambda_i[y_1 \mid z_1, \ldots y_{i-1} \mid z_{i-1}]}$ it follows that

$$\frac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}] = \Lambda_i[y_1 \mid z_1, \ldots y_{i-1} \mid z_{i-1}][x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$$

is a derived rule. That is $\dfrac{x_1 \quad _1, \ldots x_{i-1} \quad _{i-1}}{\Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}] = \Lambda_i[x_1 \mid z_1, \ldots x_{i-1} \mid z_{i-1}]}$

is a derived rule. Since $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$

is a derived rule it follows that $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Lambda_i[x_1 \mid z_1, \ldots x_{i-1} \mid z_{i-1}]}$ is a

derived rule. Since this is the case for each $i, 1 \leq i \leq n+1$, it follows that

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}} \equiv \dfrac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Lambda_{\ell+1} \text{ is a type}}$. Which completes the proof that $\equiv$ is a equivalence relation.

The equivalence relation $\equiv$ on T-rules can be used to define an equivalence relation on contexts - simply by

$\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ iff $\dfrac{x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}}{\Delta_n \text{ is a type}}$

$\equiv \dfrac{y_1 \in \Omega_1, \ldots y_{m-1} \in \Omega_{m-1}}{\Omega_m \text{ is a type}}$. We define the equivalence relation $\equiv$

of intended identity of denotation on $\in$-rules of a theory U by

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{S \in \Omega}$ iff $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}} \equiv$

$\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Omega \text{ is a type}}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = S[x_1 \mid y_1, \ldots x_n \mid y_n]}$ is a derived rule

of U. That $\equiv$ is an equivalence relation on $\in$-rules of U is a consequence of the following:

Lemma 2. If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ are contexts of U and $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ then for all derived rules of U of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}}$, the rule

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}[x_1 \mid y_1, \ldots x_n \mid y_n]}$ is a derived rule of U.

Proof. For each $i, 1 \leq i \leq n$, $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$ is a

derived rule of U thus, by the substitution lemma, so is

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_i = \Omega_i[x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$. Since for each $i, 1 \leq i \leq n$,

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{x_i \in \Delta_i}$ is a derived rule of U then so is

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{x_i \in \Omega_i [x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$. Now by the substitution lemma, since

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{x_i \in \Omega_i [x_1 \mid y_1, \ldots x_{i-1} \mid y_{i-1}]}$ is a derived rule, $1 \leqslant i \leqslant n$, and since

$\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{\text{Conclusion}}$ is a derived rule then so is $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}\ [x_1 \mid y_1, \ldots x_n \mid y_n]}$

<u>Corollary</u> $\equiv$ is an equivalence relation on the derived $\in$-rules of U.

<u>Proof</u>.

1. Reflexiveness follows from principle of derivation LI.2.

2. Symmetry. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Omega}$

Then $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle y_1 \in \Omega_1, \ldots y_n \in \Omega_n \rangle$ so by lemma 2,

$\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{t [y_1 \mid x_1, \ldots y_n \mid x_n] = S \in \Delta [y_1 \mid x_1, \ldots y_n \mid x_n]}$ is a derived rule of U.

By wellformedness, since $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Omega}$ and $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Delta [y_1 \mid x_1, \ldots y_n \mid x_n]}$

are derived rules of U so is $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{\Omega = \Delta [y_1 \mid x_1, \ldots y_n \mid x_n]}$.

Thus $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S = t[y_1 \mid x_1, \ldots y_n \mid x_n] \in \Omega}$ is a derived rule of U and

$$\frac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Omega} \equiv \frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} .$$

3. Transitivity. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Omega}$.

and $\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S \in \Omega} \equiv \dfrac{z_1 \in \Lambda_1, \ldots z_n \in \Lambda_n}{r \in \Lambda}$. Then

$\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle y_1 \in \Omega_1, \ldots y_n \in \Omega_n \rangle$ and

$\dfrac{y_1 \in \Omega_1, \ldots y_n \in \Omega_n}{S = r[y_1 \mid z_1, \ldots y_n \mid z_n] \in \Omega}$ is a derived rule of U thus by lemma 2.

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{S[x_1 \mid y_1, \ldots x_n \mid y_n] = r[x_1 \mid z_1, \ldots x_n \mid z_n] \in \Omega[x_1 \mid y_1, \ldots x_n \mid y_n]}$ is

a derived rule of U. Since we also have that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = S[x_1 \mid y_1, \ldots x_n \mid y_n] \in \Delta}$

is a derived rule of U we can conclude that so to is $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = r[x_1 \mid z_1, \ldots x_n \mid z_n] \in \Delta}$. Thus $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \equiv \dfrac{z_1 \in \Lambda_1, \ldots z_n \in \Lambda}{r \in \Lambda}$

We define an equivalence relation $\equiv$ on realisations of U as follows. If $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and if $\langle t'_1, \ldots t'_m \rangle$ is a realisation of $\langle y'_1 \in \Omega'_1, \ldots y'_m \in \Omega'_m \rangle$ wrt $\langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n \rangle$ then $\langle t_1, \ldots t_m \rangle \equiv$

$\langle t'_1,\ldots t'_m\rangle$ iff $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m\rangle \equiv \langle y'_1\in {}'y,\ldots y'_m\in\Omega'_m\rangle$ and

for each j, $1\leq j\leq m$, $\dfrac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{t_j\in\Omega_j[t_1\mid y_1,\ldots t_{j-1}\mid y_{j-1}]} \equiv$

$$\frac{x'_1\in\Omega'_1,\ldots x'_n\in\Omega'_n}{t'_j\in\Omega'_j\,[t'_1\mid y_1,\ldots t'_{j-1}\mid y_{j-1}]}\ .$$

Lemma 3. If $\langle t_1,\ldots t_m\rangle$ is a realisation of $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m\rangle$ wrt $\langle x_1\in\Delta_1,\ldots x_n\in\Delta_n\rangle$ and if $\langle t'_1,\ldots t'_m\rangle$ is a realisation of $\langle y'_1\in\Omega'_1,\ldots y'_m\in\Omega'_m\rangle$ wrt $\langle x'_1\in\Delta'_1,\ldots x'_n\in\Delta'_n\rangle$ and if $\langle t_1,\ldots t_m\rangle \equiv \langle t'_1,\ldots t'_m\rangle$ then (i) if $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Omega \text{ is a type}}$ and $\dfrac{y'_1\in\Omega'_1,\ldots y'_m\in\Omega'_m}{\Omega' \text{ is a type}}$

are derived rules of U s.t. $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m, y\in\Omega\rangle \equiv \langle y'_1\in\Omega'_1,\ldots y'_m\in\Omega'_m, y'\in\Omega'\rangle$ then $\langle x_1\in\Delta_1,\ldots x_n\in\Delta_n, y\in\Omega[t_1\mid y_1,\ldots t_m\mid y_m]\rangle \equiv \langle x'_1\in\Delta'_1,\ldots x'_n\in\Delta'_n, y'\in\Omega'[t'_1\mid y'_1,\ldots t'_m\mid y'_m]$. (ii) If $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{S\in\Omega}$ and $\dfrac{y'_1\in\Omega'_1,\ldots y'_m\in\Omega'_m}{S'\in\Omega'}$ are derived rules of U s.t. $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{S = S'[y_1\mid y'_1,\ldots y_m\mid y'_m]\in\Omega}$ is a derived rule of U then

$$\frac{x_1\in\Omega_1,\ldots x_n\in\Delta_n}{S[t_1\mid y_1,\ldots t_m\mid y_m] = S'[t'_1\mid y'_1,\ldots t'_m\mid y'_m][x_1\mid x'_1,\ldots x_n\mid x'_n]\in \overline{\Omega[t_1\mid y_1,\ldots t_m\mid y_m]}}$$ is a derived rule of U.

Proof Since $\langle t_1,\ldots t_m\rangle \equiv \langle t'_1,\ldots t'_m\rangle$, for each j, $1\leq j\leq m$,

$$\frac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{t_j\in\Omega_j\,[t_1\mid y_1,\ldots t_{j-1}\mid y_{j-1}]} \equiv \frac{x'_1\in\Delta'_1,\ldots x_n\in\Delta'_n}{t'_j\in\Omega'_j\,[t'_1\mid y_1,\ldots t'_{j-1}\mid y_{j-1}]}\ .$$

That is for each $j, 1 \leq j \leq m$, $x_1 \in \Delta_1, \ldots x_n \in \Delta_n$

$$t_j = t'_j [x_1 \mid x'_1, \ldots x_n \mid x'_n] \in \Omega_j [t_1 \mid y_1, \ldots$$

$t_{j-1} \mid y_{j-1}]$ is a derived rule of U. Now (i) follows immediately by use of principle of derivation SI1.. (ii) follows from SI2. using the fact that $S' [y_1 \mid y'_1, \ldots y_m \mid y'_m] [t'_1 [x_1 \mid x'_1, \ldots x_n \mid x'_n] \mid y_1, \ldots$

$t'_m [x_1 \mid x'_1, \ldots x_n \mid x'_n] \mid y_m] = S' [t'_1 \mid y'_1, \ldots t'_m \mid y'_m] [x_1 \mid x'_1, \ldots x_n \mid x'_n]$.

Lemma 4. (i) If $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and if $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n \rangle$ and $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle \equiv \langle y'_1 \in \Omega'_1, \ldots y'_m \in \Omega'_m \rangle$ then there exists a realisation $\langle t'_1, \ldots t'_m \rangle$ of $\langle y'_1 \in \Omega'_1, \ldots y'_m \in \Omega'_m \rangle$ wrt $\langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n \rangle$ such that $\langle t_1, \ldots t_m \rangle \equiv \langle t'_1, \ldots t'_m \rangle$.

(ii) If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta \rangle$ is a context of U and $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n \rangle$ then there exists Δ' s.t. $\langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n, x' \in \Delta' \rangle$ is a context and $\langle x'_1 \in \Delta'_1, \ldots x'_n \in \Delta'_n, x' \in \Delta' \rangle \equiv \langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta \rangle$.

Proof. For (i) take $t'_j = t_j [x'_1 \mid x_1, \ldots x'_{j-1} \mid x_{j-1}]$ and then use lemma 2. Similarly in (ii) take $\Delta' = \Delta [x'_1 \mid x_1, \ldots x'_n \mid x_n]$ and use lemma 2.

The following lemma follows directly from the SI2 principle of derivation.

Lemma 5. If $\langle t_1, \ldots t_{m+1} \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_{m+1} \in \Omega_{m+1} \rangle$

wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$, if $\langle S_1, \ldots S_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and if $\langle S_1, \ldots S_m \rangle \equiv \langle t_1, \ldots t_m \rangle$ then $\langle S_1, \ldots S_m, t_{m+1} \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_{m+1} \in \Omega_{m+1} \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle S_1, \ldots S_m, t_{m+1} \rangle \equiv \langle t_1, \ldots t_{m+1} \rangle$.

1.14 The Category GAT

We can define an equivalence relation $\equiv$ on the set of interpretations of a theory U in a theory U'. If I and J are two such interpretations then define $I \equiv J$ iff for all introductory rules R of U, $\hat{I}(R) \equiv \hat{J}(R)$ in U'.

Lemma 1. If I an J are interpretations of U in U' and $I \equiv J$ then for all derived T and $\in$-rules R of U, $\hat{I}(R) \equiv \hat{J}(R)$.

Proof. By induction on derivations in U. We check that principles T1, CF1, CF2(a) and (b) preserve the property.

T1. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ and $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$

are derived rules of U and that $\hat{I}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right) \equiv$

$\hat{J}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right)$.

Then $\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t) = \dot{J}(t) \in \dot{I}(\Delta)}$ is a derived rule of U' and since

I is an interpretation, so is $\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(\Delta) = \dot{I}(\Delta')}$. Thus

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t) = \dot{J}(t) \in \dot{I}(\Delta')}$ is a derived rule of U'. And as

$\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle \equiv \langle x_1 \in \dot{J}(\Delta_1), \ldots x_n \in \dot{J}(\Delta_n) \rangle$ is the case so

$\hat{I}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta'}\right) \equiv \hat{J}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta'}\right)$

CF1 Suppose $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a derived rule of U and that

$$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}\right) \equiv \hat{J}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}\right).$$ Then

$\langle x_1 \in \dot{I}(\Delta_1), \ldots x_{n+1} \in \dot{I}(\Delta_{n+1})\rangle \equiv \langle x_1 \in \dot{J}(\Delta_1), \ldots x_{n+1} \in \dot{J}(\Delta_{n+1})\rangle$ and

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_{n+1} \in \dot{I}(\Delta_{n+1})}{x_i = x_i \in \dot{I}(\Delta_i)}$ is a derived rule of U', hence

$$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}\right) \equiv \hat{J}\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \quad \Delta_{n+1}}{x_i \in \Delta_i}\right).$$

CF2(b) (Principle CF2(a) will be similar). Suppose that f is an operator symbol of U introduced by $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{f(x_1, \ldots x_n) \in \Delta}$ and

suppose that for each $i, 1 \leq i \leq n$, $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}$ is a

derived rule of U. Suppose that for each $i, 1 \leq i \leq n$,

$$\hat{I}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}\right) \equiv \hat{J}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i \in \Delta_i [t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}\right).$$

Using lemma 1 of §1.10 we see that for each $i, 1 \leq i \leq n$,

$$\frac{y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)}{\dot{I}(t_i) \in \dot{I}(\Delta_i) [\dot{I}(t_1) \mid x_1, \ldots I(t_{i-1}) \mid x_{i-1}]} \equiv$$

$$\frac{y_1 \in \dot{J}(\Delta_1), \ldots y_m \in \dot{J}(\Delta_m)}{\dot{J}(t_i) \in \dot{J}(\Delta_i) [\dot{J}(t_1) \mid x_1, \ldots \dot{J}(t_{i-1}) \mid x_{i-1}]}$$. Also, since I ≡ J,

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(f(x_1, \ldots x_n)) \in \dot{I}(\Delta)} \equiv \dfrac{x_1 \in \dot{J}(\Delta_1), \ldots x_n \in \dot{J}(\Delta_n)}{\dot{J}(f(x_1, \ldots x_1)) \in \dot{J}(\Delta)}$. In particular

$\langle \dot{I}(t_1), \ldots \dot{I}(t_n) \rangle$ is a realisation of $\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle$ wrt $\langle y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m) \rangle$, $\langle \dot{J}(t_1), \ldots \dot{J}(t_n) \rangle$ is a realisation of $\langle x_1 \in \dot{J}(\Delta_1), \ldots x_n \in \dot{J}(\Delta_n) \rangle$ wrt $\langle y_1 \in \dot{J}(\Omega_1), \ldots y_m \in \dot{J}(\Omega_m) \rangle$ and $\langle \dot{I}(t_i), \ldots \dot{I}(t_n) \rangle \equiv \langle \dot{J}(t_1), \ldots \dot{J}(t_n) \rangle$. Thus by lemma 3 (i) of §1.12,

$$\frac{y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)}{\dot{I}(f(x_1, \ldots x_n))[\dot{I}(t_1)|x_1, \ldots \dot{I}(t_m)|x_n] = \dot{J}(f(x_1, \ldots x_n))[\dot{J}(t_1)|x_1, \ldots \dot{J}(t_n)|x_n] \in \dot{J}(\Delta)[\dot{J}(t_1)|x_1 \ldots \dot{J}(t_n)|x_n]}$$

is derived rule of U'.

Using lemma 1 of §1.10, this rule is just

$$\frac{y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)}{\dot{I}(f(t_1, \ldots t_m)) = \dot{J}(f(t_1, \ldots t_m)) \in \dot{I}(\Delta)[t_1|x_1, \ldots t_n|x_n]}$$. Thus

$$\hat{I}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{f(t_1, \ldots t_m) \in \Delta[t_1|x_1, \ldots t_n|x_n]}\right) \equiv \hat{J}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{f(t_1, \ldots t_n) \in \Delta[t_1|x_1, \ldots t_n|x_n]}\right)$$

Corollary 2 If I and J are interpretations of U in U' then $I \equiv J$ iff for all derived $\in$-rules R of U, $\hat{I}(R) \equiv \hat{J}(R)$.

Proof. Use the fact that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ is a derived rule of U iff $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta}{x \in \Delta}$ is a derived rule of U. From

$$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta}{x \in \Delta}\right) \equiv \hat{J}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta}{x \in \Delta}\right)$$ we can deduce that $\hat{I}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}\right) \equiv \hat{J}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}\right)$.

Corollary 3 If I and J are interpretations of U in U' and I' and J' are interpretations of U' in U'' then $I \equiv J$ and $I' \equiv J'$ implies $I'oI \equiv J'oJ$.

We denote the category of generalised algebraic theories and equivalence classes of interpretations by GAT. Composition is defined by $[I'] \circ [I] = [I'oI]$. Well defined by corollary 3.

CHAPTER 2 CONTEXTUAL CATEGORIES

2.1 Algebraic Semantics

In this chapter we show that there is a generalised algebraic theory (the theory of contextual categories) whose category of models (the category of contextual categories) is equivalent to the category GAT of generalised algebraic theories and equivalence classes of interpretations.

How do we interpret this result? Well, there are many other examples of this very strong kind of relationship holding between an algebraic notion of structure and a syntactic notion of theory. The following list is by no means exhaustive:

Syntactic Notion	Algebraic Notion	Reference
Propositional Theory of classical logic.	Boolean Algebra.	
Propositional Theory of Intuitionistic Logic.	Heyting Algebra.	
Single Sorted Algebraic or Equational Theory.	Lawveres Notion of an Algebraic Theory.	Lawvere [11]
Equational Theory in the language of the typed λ-calculus.	Cartesian Closed Category.	Myers [26]
Theory of higher order Intuitionistic Logic.	Topos.	Fourman [
Coherent Theory.	Grothendieck Site.	Reyes [28]

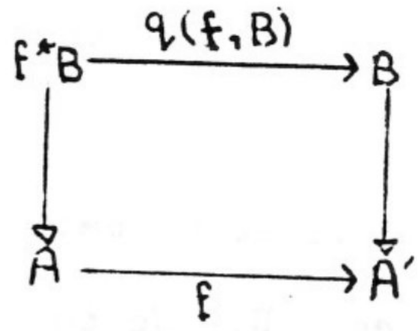
In all these cases there is definable the notion of a model of a given theory in a given structure. In each case the category of syntactic theories and equivalence classes of interpretations is equivalent to the category of algebraic structures. This last property is the important characterising property. It can lead to the view that the theories in syntactic form should be dispensed with entirely and the structures be given the title of theories. This seems wasteful. It is to be preferred that we think of the structures as providing a semantics for the theories, in fact, the most general possible semantics. We shall call it <u>the</u> algebraic semantics. Thus contextual categories are to provide us with the alg semantics of generalised algebraic theories.

In case it should be argued that what we have called the algebrai semantics is really noneother than the interpretations of one theory i considered as a notion of semantics; well we more or less agree, thoug perhaps it is only when such are considered as interpretations into algebraic structures that they can be properly said to constitute a notion of semantics. The important point here,though,is that struct do frequently appear quite independently of theories; thus the notion of model is certainly enriched by the isomorphism between theories and structures because theories which arise first as structures (being defined by something like "the theory that corresponds to this here structure") are usually theories which would not otherwise have occure

2.2 Definition and Examples

Definition. A <u>contextual category</u> consists of

1. A category $\underline{C}$ with terminal object 1.
2. A tree structure on the objects of $\underline{C}$ such that the terminal object 1 is the unique least element of the tree.
3. For all $A,B \in |\underline{C}|$ such that $A \lhd B$ a morphism $p(B)$: $B \longrightarrow A$ in $\underline{C}$. This morphism will also be written just as $B \longrightarrow\!\!\triangleright A$.
4. For all $A,A' \in |\underline{C}|$, for all $f:A \longrightarrow A'$ in $\underline{C}$, for all $B \in |\underline{C}|$ such that $A' \lhd B$, an object f^*B of $\underline{C}$ and a morphism $q(f,B)$: $f^*B \longrightarrow B$ such that $A \lhd f^*B$ and such that the diagram

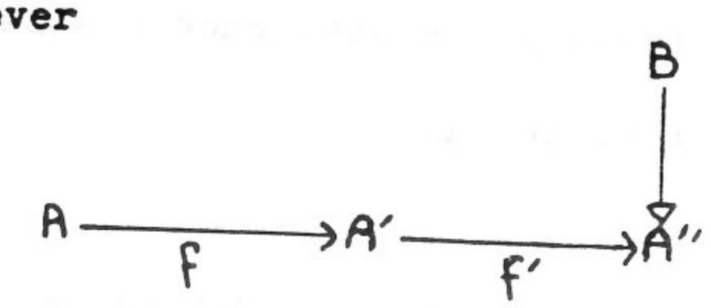


is a pullback diagram in $\underline{C}$.

Such that:

(I) For all $A,B \in |\underline{C}|$ such that $A \lhd B$, $id_A{}^*B = B$ and $q(id_A, B) = id_B$.

(II) Whenever



in $\underline{C}$ then $(ff')^*B = f^*(f'^*B)$ and $q(ff',B) = q(f,f'^*B) \circ q(f',B)$.

We shall see that the objects of a contextual category should be thought of as contexts. Recall that a context is a sequence $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ such that the rule $x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1} :$ Δ_n is a type is a derived rule and such that x_n is a variable distinct from each of $x_1, \ldots x_{n-1}$. The tree structure on the set of contexts of a theory is easily seen. For $n \geqslant 1$, the predecessor of a context $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is the context $\langle x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1} \rangle$. The empty context $\langle\rangle$ is the unique least element of the tree.

The morphisms of a contextual category should be thought of as realisations. Recall that a realisation of a context $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m$ with respect to the context $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is just an m-tuple $\langle t_1, \ldots t_m \rangle$ such that for each j, $1 \leqslant j \leqslant m$, the rule $x_1 \in \Delta_1, \ldots x_n \in \Delta_n :$ $t_j \in \Omega_j [t_1 \mid y_1, \ldots t_{j-1} \mid y_{j-1}]$ is a derived rule. Think of the morphism $f: A \longrightarrow A'$ in a contextual category as being a realisation of the context A' wrt the context A.

In §1.12 we defined the category $\underline{R(U)}$ of contexts and realisations of a theory U, We could go on and show that for any theory U, the category $\underline{R(U)}$ with the pullback structure defined in §1.12 (actually we defined more structure than was necessary) is a contextual category But we do not need this construction. Rather we need the construction of a contextual category $\mathbb{C}(U)$ associated with a theory U as part of the equivalence between contextual categories and generalised algebraic theories. This category $\mathbb{C}(U)$ is a category of equivalence classes of contexts and equivalence classes of realisations of U. We shall describe in some detail.

Recall that in §1.13 we defined an equivalence relation $\equiv$ on derived T and $\in$-rules of a theory U. This equivalence relation we call

the equivalence relation of intended identity of denotation. We used this equivalence relation in defining an equivalence relation $\equiv$ on contexts and realisations: $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m$ iff $\dfrac{x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}}{\Delta \text{ is a type}} \equiv \dfrac{y_1 \in \Omega_1, \ldots y_{m-1} \in \Omega_{m-1}}{\Omega_m \text{ is a type}}$. And if $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots \Delta_n \rangle$ and if $\langle t_1', \ldots t_m' \rangle$ is a realisation of $\langle y_1' \in \Omega_1', \ldots y_m' \in \Omega_m' \rangle$ wrt $\langle x_1' \in \Delta_1', \ldots x_n \in \Delta_n' \rangle$ then $\langle t_1, \ldots t_m \rangle \equiv \langle t_1', \ldots t_m' \rangle$ iff for each j, $1 \leq j \leq m$,

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t_j \in \Omega_j [t_1 \mid y_1, \ldots t_{j-1} \mid y_{j-1}]} \equiv \frac{x_1' \in \Delta_1', \ldots x_n' \in \Delta_n'}{t_j' \in \Omega_j' [t_1' \mid y_1', \ldots t_{j-1}' \mid y_{j-1}']}$$

The category $\mathbb{C}(U)$ is defined to have as objects the equivalence classes of contexts of U and to have as morphisms the equivalence classes of realisations of U. More precisely if $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ are contexts of U then define $\mathrm{Hom}_{\mathbb{C}(U)}([\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle], [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle]) =$

$\{ [\langle t_1, \ldots t_m \rangle] \mid \langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \}$. Hom is well defined just by lemma 4(i) of §1.13.

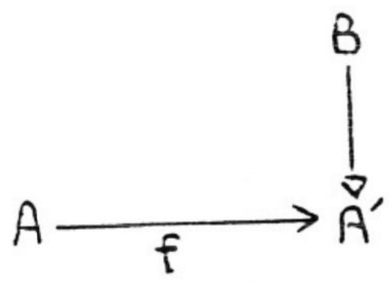
Whenever $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ and whenever $\langle s_1, \ldots s_\ell \rangle$ is a realisation wrt $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ then the composition in $\mathbb{C}(U)$ of $[\langle t_1, \ldots t_m \rangle]$ with $[\langle s_1, \ldots s_\ell \rangle]$ is defined by $[\langle t_1, \ldots t_m \rangle] \circ [\langle s_1, \ldots s_\ell \rangle] = [\langle s_1 [t_1 \mid y_1, \ldots t_m \mid y_m], \ldots s_\ell [t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]$. Composition is well defined, this follows from lemma 3(ii) of §1.13. The identity morphisms in $\mathbb{C}(U)$ are given by $\mathrm{id}_{[\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle]} = [\langle x_1, \ldots x_n \rangle]$. Well definedness is trivial.

The objects of $\mathbb{C}(U)$ are structured as a tree by taking the predecessor of $[<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>]$ to be $[<x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}>]$. The tree structure is well defined because by definition if $<x_1 \in \Delta_1, \ldots x_n \in \Delta_n> \equiv <x_1' \in \Delta_1', \ldots x_n' \in \Delta_n'>$ then $<x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}> \equiv <x_1' \in \Delta_1', \ldots x_{n-1}' \in \Delta_{n-1}'>$.
$[<>]$ is the least element of the tree.

$[<>]$ is a terminal onject of $\mathbb{C}(U)$ because by definition of $\mathrm{Hom}\,\mathbb{C}(U)$, $\mathrm{Hom}\,\mathbb{C}(U)\,([<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>], [<>]) =$
$\{[<>] \mid <>$ is a realisation of $<>$ wrt $<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>\}$ and because by the definition of realisation, $<>$ is a realisation of the context $<>$ wrt the context $<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>$.

If $A \lhd B$ in $\mathbb{C}(U)$, say $A = [<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>]$ and $B = [<x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta>]$, then define $p(B): B \longrightarrow A$ by $p(B) = [<x_1, \ldots x_n>]$.

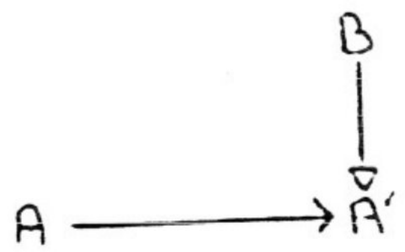
If



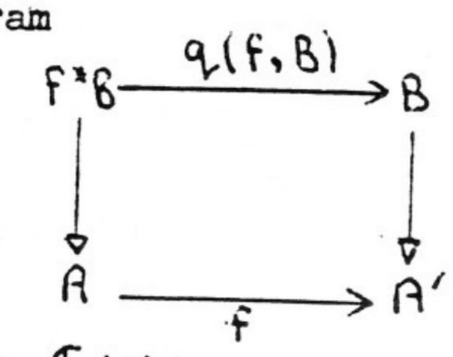
in $\mathbb{C}(U)$, say $A = [<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>]$, $A' = [<y_1 \in \Omega_1, \ldots y_m \in \Omega_m>]$, $B = [<y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega>]$ and $f = [<t_1, \ldots t_m>]$, then define $f^*B = [<x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y \in \Omega[t_1 \mid y_1, \ldots t_m \mid y_m]>]$ and $q(f,B) = [<t_1, \ldots t_m, y>]$. f^*B is well defined, by lemma 3(i) of §1.13

Lemma. $\mathbb{C}(U)$ is a contextual category.

Proof. Firstly we must show that whenever

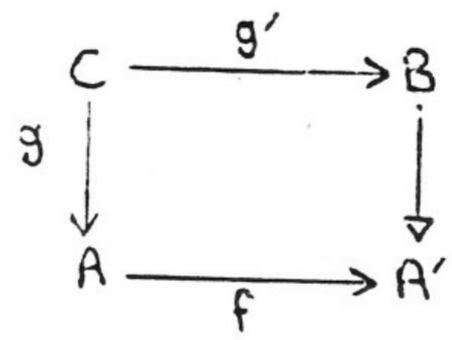


in $\mathbb{C}(U)$ then the diagram



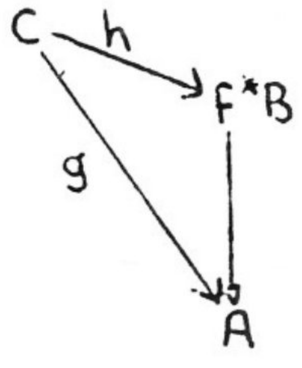
is a pullback diagram in $\mathbb{C}(U)$.

So suppose that $A= [<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>]$, $B= [<y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega>]$ and that $f= [<t_1, \ldots t_m>]$. Suppose also that C is an object of $\mathbb{C}(U)$ and that $g:C \longrightarrow A$ and $g':C \longrightarrow B$ in $\mathbb{C}(U)$ such that the diagram

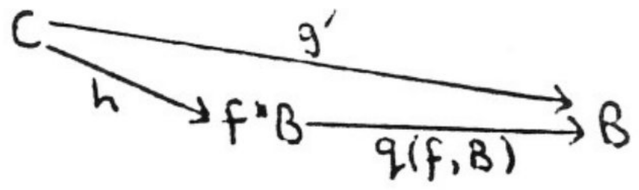


commutes. Call this diagram (I). We can suppose that $C= [<z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell>]$, $g= [<r_1, \ldots r_n>]$, $g'= [<S_1, \ldots S_m, S>]$, where $<r_1, \ldots r_n>$ is some realisation of $<x_1 \in \Delta_1, \ldots x_n \in \Delta_n>$ wrt $<z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell>$ and $<S_1, \ldots S_m, S>$ is a realisation of $<y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega>$ wrt $<z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell>$.

We must show that there exists a unique $h:C \longrightarrow f^*B$ in $\mathbb{C}(U)$ such that diagrams (II) and (III) both commute.



(II)



(III)

I claim that $[<r_1, \ldots r_m, S>]$ is such an h. Since diagram (I) commutes, $[<t_1[r_1 \mid x_1, \ldots r_n \mid x_n], \ldots t_m[r_1 \mid x_1, \ldots r_n \mid x_n]>] = [<S_1, \ldots S_m>]$. Hence for all j, $1 \leq j \leq m$,

$z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell$

$t_j[r_1 \mid x_1, \ldots r_n \mid x_n] \in C[\ldots]$ is a derived r[ule]

of U. Thus as $\langle S_1,\ldots S_m,S\rangle$ is a realisation of $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m,y\in$

$\Omega[t_1[r_1\mid x_1,\ldots r_n\mid x_n]\mid y_1,\ldots t_m[r_1\mid x_1,\ldots r_n\mid x_n]\mid y_m$ the rule $\dfrac{z_1\in\Lambda_1,\ldots z_\rho\in\Lambda}{}$

is a derived rule of U. Hence $\langle r_1,\ldots r_n,S\rangle$ is a realisation of

$\langle x_1\in\Delta_1,\ldots x_n\in\Delta_n,y\in\Omega[t_1\mid y_1,\ldots t_m\mid y_m]\rangle$ wrt

$\langle z_1\in\Lambda_1,\ldots z_\rho\in\Lambda_\rho\rangle$ and thus $[\langle r_1,\ldots r_n,S\rangle]: C\longrightarrow f^*B$ in $\mathbb{C}(U)$.

Setting $h=[\langle r_1,\ldots r_n,S\rangle]$ then (II) commutes because

$[\langle r_1,\ldots r_n,S\rangle]\circ[\langle x_1,\ldots x_n\rangle]=[\langle r_1,\ldots r_n\rangle]$ and (III) commutes

because $[\langle r_1,\ldots r_n,S\rangle]\circ[\langle t_1,\ldots t_m,y\rangle]=[\langle t_1[r_1\mid x_1,\ldots r_n\mid$

$x_n],\ldots t_m[r_1\mid x_1,\ldots r_n\mid x_n],S\rangle=[\langle S_1,\ldots S_m,S\rangle]$. So

$[\langle r_1,\ldots r_m,S\rangle]$ is certainly such an h. To show that it is the unique

such h suppose now that h is an arbitrary morphism $h:C\longrightarrow f^*B$ in

$\mathbb{C}(U)$ such that the diagrams(II) and (III) commute say $h=$

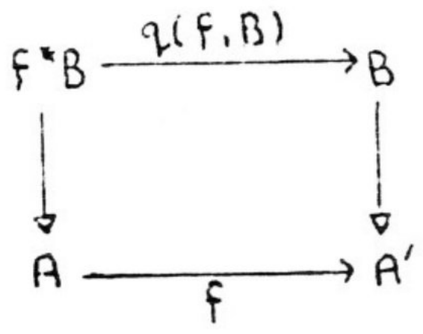
$[\langle r_1',\ldots r_n',S'\rangle]$. Since (II) commutes, for each i, $1\leq i\leq n$, the

rule $\dfrac{z_1\in\Lambda_1,\ldots z_\rho\in\Lambda_\rho}{r_i=r_i'\in\Delta_i[r_1\mid x_1,\ldots r_{i-1}\mid x_{i-1}]}$ is a derived rule of U. Since

(III) commutes, the rule $\dfrac{z_1\in\Lambda_1,\ldots z_\rho\in\Lambda_\rho}{S=S'\in\Omega[S_1\mid y_1,\ldots S_m\mid y_m]}$ is a derived rule

of U. Hence $h=[\langle r_1',\ldots r_n',S'\rangle]=[\langle r_1,\ldots r_n,S\rangle]$. Which completes

the proof that



is a pullback diagram in $\mathbb{C}(U)$.

It remains to show that the axioms (I) and (II) of the definition o

contextual category hold of the structure $\mathbb{C}(U)$. Well, it is easy to se

that (I) must hold because if B= $[\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta \rangle]$ then $\Delta[x_1 \mid x_1, \ldots x_n \mid x_n] = \Delta$. Similarly (III) holds because $\Lambda[S_1 \mid z_1, \ldots S_\varrho \mid z_\varrho][t_1 \mid y_1, \ldots t_m \mid y_m] = \Lambda[S_1[t_1 \mid y_1, \ldots t_m \mid y_m] \mid S_\varrho[t_1 \mid y_1, \ldots t_m \mid y_m] \mid z_\varrho]$, whenever $\langle z_1 \in \Lambda_1, \ldots z_\varrho \in \Lambda_\varrho, z \in \Lambda \rangle$ is a context and $\langle S_1, \ldots S_\varrho \rangle$ is a realisation of $\langle z_1 \in \Lambda_1, \ldots z_\varrho \in \Lambda_\varrho \rangle$ wrt $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ and $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$.

We now turn to the definition of a (large) contextual category, Fam, which plays the same role among contextual categories as does the category Set among categories. Whereas Set is the structured collection of functions so it is that Fam is the structured collection of operators. We must refer back to § 1.9 to the discussion of operators and sets, families of sets, families of families of sets and so

The tree of objects of Fam is the tree of families introduced in §1.9.

Thus it is the tree of sets, families of sets, families of families of sets and so on with a formally adjoined least element 1.

For $n,m \geqslant o$, if $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m$ in Fam then $\mathrm{Hom}_{\mathrm{Fam}}(A_n, B_m)$ $\overline{\mathrm{def}}$ $\{\langle f_1, \ldots f_m \rangle \mid f_1, \ldots f_m$ are operators such that the status of the operator f_j is given by for $a_1 \in A_1$, for $a_2 \in A_2(a_1)$, ... for $a_n \in A_n(a_1, \ldots a_{n-1})$: $f_j(a_1, \ldots a_n) \in B_j(f_1(a_1, \ldots a_n), \ldots f_{j-1}(a_1, \ldots a_n))$

In particular if n=o then we get $\mathrm{Hom}(1, B_m) = \{\langle b_1, \ldots b_m \rangle \mid b_1 \in B_1, b_2 \in B_2(b_1), \ldots$ and $b_m \in B_m(b_1, \ldots b_{m-1})\}$. On the other hand 1 is the terminal object of Fam because $\mathrm{Hom}(A_n, 1) = \{\langle\rangle\}$.

Composition in Fam is defined as follows. If $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$ and $1 \lhd C_1 \ldots \lhd C_\ell$ in Fam and if $\langle f_1, \ldots f_m \rangle : A_n \longrightarrow B_m$, $\langle g_1, \ldots g_\ell \rangle : B_m \longrightarrow C_\ell$ then the composition is given by $\langle f_1, \ldots f_m \rangle \circ \langle g_1, \ldots g_\ell \rangle = \langle h_1, \ldots h_\ell \rangle$, where for each k, $1 \leq k \leq \ell$, h_k is defined by $h_k(a_1, \ldots a_n) = g_k(f_1(a_1, \ldots a_n), \ldots f_m(a_1, \ldots a_n))$, whenever $a_1 \in A_1, \ldots a_n \in A_n(a_1, \ldots a_{n-1})$.

If $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in Fam then $p(A) : A \longrightarrow A_n$ in Fam is given by $p(A) = \langle h_1, \ldots h_n \rangle$, where for each i, $1 \leq i \leq n$, h_i is defined by $h_i(a_1, \ldots a_n, a) = a_i$, whenever $a_1 \in A_1, \ldots a_n \in A_n(a_1, \ldots a_{n-i})$ and $a \in A(a_1, \ldots a_n)$.

If $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m \lhd B$ in Fam and if $\langle f_1, \ldots f_m \rangle : A_n \longrightarrow B_m$ then $\langle f_1, \ldots f_m \rangle * B$ is defined to be the family $\lambda a_1 \in A_1 . \lambda a_2 \in A_2(a_1) \ldots . \lambda a_n \in A_n(a_1, \ldots a_{n-1}) . B(f_1(a_1, \ldots a_n), \ldots f_m(a_1, \ldots a_n))$. In this situation $q(\langle f_1, \ldots f_m \rangle, B) = \langle f_1, \ldots f_m, \gamma \rangle$, where γ is the operator defined by $\gamma(a_1, \ldots a_n, b) = b$, whenever $a_1 \in A_1, \ldots a_n \in A_n(a_1, \ldots a_{n-1})$ and $b \in B(f_1(a_1, \ldots a_n), \ldots f_m(a_1, \ldots a_n))$.

The proof that Fam, so defined, is a contextual category is rather simple. Because the statements asserting the status of operators and families are so similar to the formal rules the proof is similar to the proof that $\mathbb{C}(U)$ is always a contextual category, only easier.

The homomorphisms between contextual categories are called contextual functors. Thus:

Definition. If $\mathbb{C}$ and $\mathbb{C}'$ are contextual categories then a contextual functor $F : \mathbb{C} \longrightarrow \mathbb{C}'$ is a functor $F : \mathbb{C} \longrightarrow \mathbb{C}'$ such that:

1. $F(1)=1$ and if $A \lhd B$ in $\mathfrak{C}$ then $F(A) \lhd F(B)$ in $\mathfrak{C}'$.
2. For all objects A of $\mathfrak{C}$, $F(p(A))=p(F(A))$.
3. For all f and B such that f^*B is defined in $\mathfrak{C}$, $F(f^*B)=F(f)^*F(B)$ and $F(q(f,B))=q(F(f),F(B))$.

The category of contextual categories and contextual functors is denoted Con.

2.3 Notation and Basic Lemmas

If $A \leqslant B$ in the contextual category $\mathbb{C}$ then define the morphism $p(B,A):B \longrightarrow A$ in $\mathbb{C}$, also written just as $B \longrightarrow\!\!\!\!\!\rightarrow A$, by $p(B,A)=p(B)\circ p(X_n)\circ \ldots \circ p(X_1)$, where $X_1,\ldots X_n$ is the unique sequence of objects of $\mathbb{C}$ such that $A \lhd X_1 \ldots \lhd X_n \lhd B$ in $\mathbb{C}$. (in the case $A=B$, $p(B,A)=id_A$).

$$A \xleftarrow{P(X_1)} X_1 \xleftarrow{P(X_2)} X_2 \qquad \ldots . X_n \xleftarrow{P(B)} B$$

The contextual category structure supplies us with pullbacks for any map of the form $p(X)$, these given pullbacks can be pieced together to obtain a pullback for any morphism of the form $p(B,A)$ along any morphism with condomain A. For we have the following very trivial lemma about pullbacks in any category $\underline{C}$.

Lemma In any category $\underline{C}$, (a) if $f:A \longrightarrow B$ in $\underline{C}$ then

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle id(A)}\downarrow & & \downarrow{\scriptstyle id(B)} \\ A & \xrightarrow[f]{} & B \end{array}$$

is a pullback diagram in $\underline{C}$. (b) If

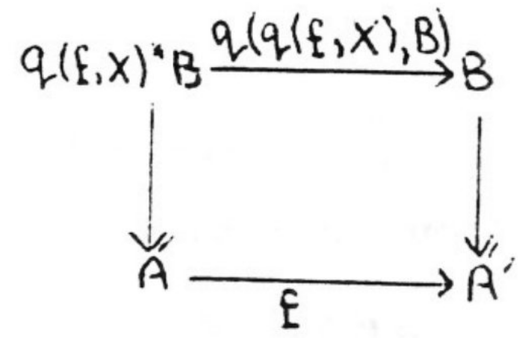
$$\begin{array}{ccc} A_2 & \xrightarrow{f_2} & B_2 \\ {\scriptstyle h_1}\downarrow & & \downarrow{\scriptstyle g_1} \\ A_1 & \xrightarrow[f_1]{} & B_1 \end{array} \qquad \text{and} \qquad \begin{array}{ccc} A_3 & \xrightarrow{f_3} & B_3 \\ {\scriptstyle h_2}\downarrow & & \downarrow{\scriptstyle g_2} \\ A_2 & \xrightarrow[f_2]{} & B_2 \end{array}$$

are pullback diagrams in $\underline{C}$, then so is

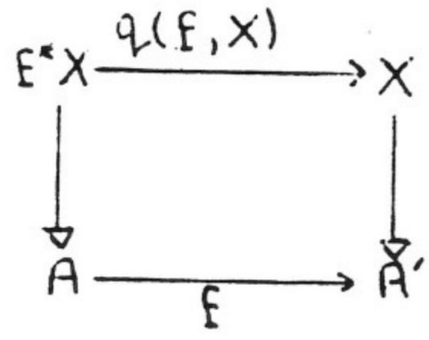
$$\begin{array}{ccc} A_3 & \xrightarrow{f_3} & B_3 \\ {\scriptstyle h_1\circ h_2}\downarrow & & \downarrow{\scriptstyle g_1\circ g_2} \\ A_1 & \xrightarrow[f_1]{} & B_1 \end{array}$$

This means, for example, that if $A' \lhd X \lhd B$ in a contextual category $\mathbb{C}$ a

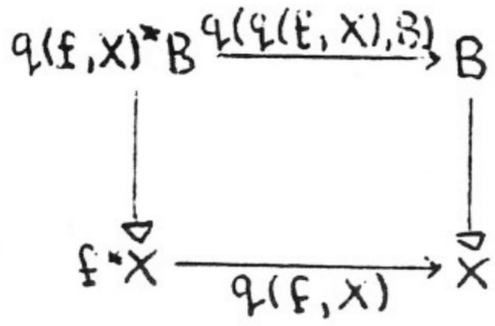
if $f:A \longrightarrow A'$ in $\mathbb{C}$ then the diagram



is a pullback diagram, since both of the two diagrams

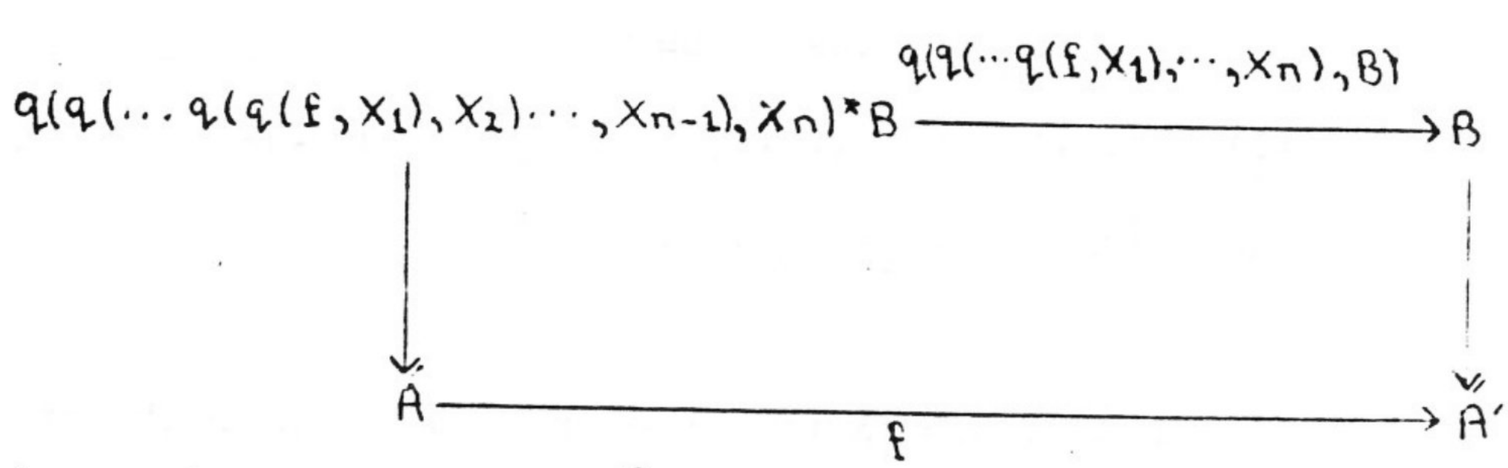


and



are pullback diagrams. So that gives us a canonical pullback for the morphism $B \longrightarrow A'$ along any $f:A \longrightarrow A'$.

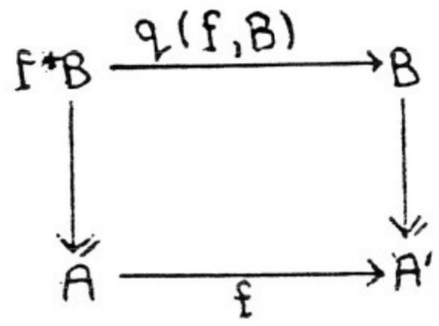
In general, it means that whenever $A' \leq B$ in $\mathbb{C}$ and $f:A \longrightarrow A'$ in $\mathbb{C}$ then we have a canonical pullback for the morphism $p(B,A')$ along f. For by iterating the method used when $A' \lhd X \lhd B$ we see that if $A' \lhd X_1 \ldots \lhd X_n \lhd B$ in $\mathbb{C}$ and if $f:A \longrightarrow A'$ then the diagram



is a pullback diagram in $\mathbb{C}$.

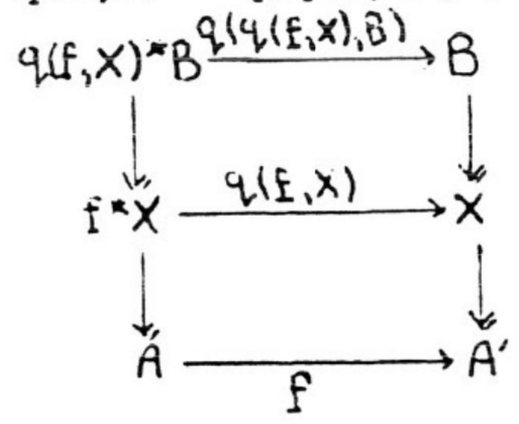
Since these constructed pullbacks form an important part of contextual category structure we would like a simpler notation for them. As no confusion is likely, we extend the * and q notation to cover these new pullback diagrams.

From now on if $f:A \longrightarrow A'$ in $\mathbb{C}$ and $A' \leqslant B$ in $\mathbb{C}$ then the diagram



is the cannonical pullback diagram defined above. The following observation which follows from the way the new pullback diagrams are constructed contains all the information we need to remember about that construction:

Lemma 1. If $f:A \longrightarrow A'$ and $A' \leqslant X \leqslant B$ in the contextual category $\mathbb{C}$ then $f^*B = q(f,X)^*B$ and $q(f,B) = q(q(f,X),B)$.



If U is a generalised algebraic theory and if $f:A \longrightarrow A'$ and $A' \leqslant B$ in the contextual category $\mathbb{C}(U)$ then supposing that

$A = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle]$, $A' = [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle]$,

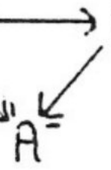
$B = [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y_{m+1} \in \Omega_{m+1}, \ldots y_{m+\varrho} \in \Omega_{m+\varrho} \rangle]$ and

$f = [\langle t_1, \ldots t_m \rangle]$ then $f^*B = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y_{m+1} \in \Omega_{m+1}[t_1 \mid y_1, \ldots t_m \mid y_m] , \ldots y_{m+\varrho} \in \Omega_{m+\varrho}[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]$ and

$q(f,B) = [\langle t_1, \ldots t_m, y_{m+1}, \ldots y_{m+\varrho} \rangle]$.

For $\mathbb{C}$ a contextual category and for A an object of $\mathbb{C}$ we can define a contextual category $\mathbb{C}_A$ whose tree of objects is the tree of those objects of $\mathbb{C}$ which appears above A. The construction is similar to the construction of the comma category $\underline{C}/_A$ of a category $\underline{C}$ at an object A. The similarity is enhanced by the fact that it $f:A \longrightarrow A'$ in $\mathbb{C}$ then pulling back along f induces a contextual functor $\mathbb{C}_f : \mathbb{C}_{A'} \longrightarrow \mathbb{C}_A$.

$\mathbb{C}_A$ is defined by $|\mathbb{C}_A| = \{B \in |\mathbb{C}| \text{ s.t. } A \leq B\}$, $\mathrm{Hom}\,\mathbb{C}_A(B,B') =$

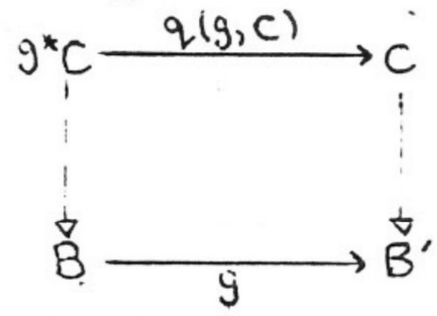
$\{g:B \longrightarrow B' \text{ in } \mathbb{C} \mid$  commutes. $\}$.

We show that $\mathbb{C}_A$ inherits the structure of a contextual category from $\mathbb{C}$. It suffices to show that $\mathbb{C}_A$ is closed under the operations $p, {}^*$ and q of $\mathbb{C}$ and that these operations yield pullback diagrams in $\mathbb{C}_A$.

Lemma. (a). If $A \leq B \lhd C$ in $\mathbb{C}$ then $p(C):C \longrightarrow B$ in $\mathbb{C}_A$.

(b). If $g:B \longrightarrow B'$ in $\mathbb{C}_A$ and $B' \lhd C$ in $\mathbb{C}_A$ then $g^*C \in |\mathbb{C}_A|$ and $q(g,C):\ g^*C \longrightarrow C$ in $\mathbb{C}_A$.

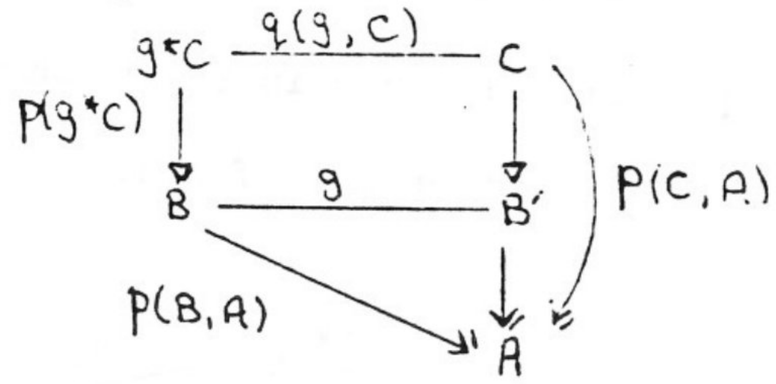
(c). If $g:B \longrightarrow B'$ in $\mathbb{C}_A$ and $B' \lhd C$ in $\mathbb{C}_A$ then



is a pullback diagram in $\mathbb{C}_A$.

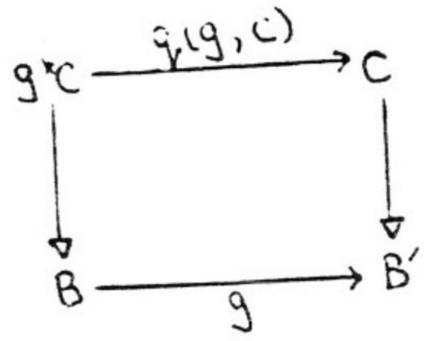
Proof. (a) Trivial - $p(C) \circ p(B,A) = p(C,A)$.

(b) $g^*C \in |\mathbb{C}_A|$ since $A \leq B < g^*C$. $q(g,C):g^*C \longrightarrow C$ in $\mathbb{C}_A$ because $q(g,C):g^*C \longrightarrow C$ in $\mathbb{C}$ and $q(g,C) \circ p(C,A) = q(g,C) \circ p(C) \circ p(B',A) =$ $p(g^*C) \circ g \circ p(B',A) = p(g^*C) \circ p(B,A) = p(g^*C,A)$.



(c). Assume that $D \in |\mathbb{C}_A|$ and that $h_1:D \longrightarrow B$, $h_2:D \longrightarrow C$ in $\mathbb{C}_A$ such that $h_1 \circ g = h_2 \circ p(C)$. We must show that there exists a unique $k:D \longrightarrow g^*C$ in $\mathbb{C}_A$ such that $k \circ p(g^*C) = h_1$ and $k \circ q(f,C) = h_2$.

Since we also have $D \in |\mathbb{C}|$ and $h_1: D \longrightarrow B$, $h_2: D \longrightarrow C$ in $\mathbb{C}$ such that $h_1 \circ g = h_2 \circ p(C)$, and since

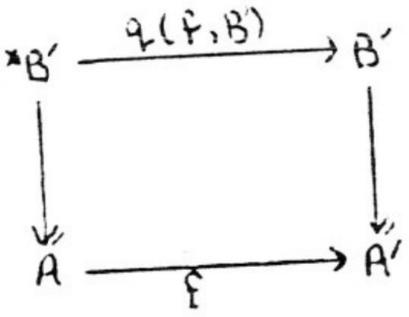


is a pullback diagram in $\mathbb{C}$, it follows that there exists a unique $k: D \longrightarrow g^*C$ in $\mathbb{C}$ such that $k \circ p(g^*C) = h_1$ and $k \circ q(g,C) = h_2$. It suffices to show that k is in $\mathbb{C}_A$, ie. that $k \circ p(g^*C, A) = p(D, A)$. Since h_1 is in $\mathbb{C}_A$ we known that $h_1 \circ p(B,A) = p(D,A)$. Hence $k \circ p(g^*C, A) = k \circ p(g^*C) \circ p(B,A) = h_1 \circ p(B,A) = p(D,A)$.

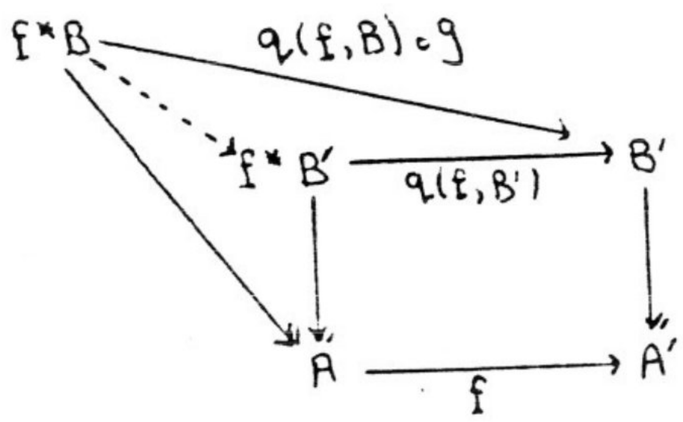
So $\mathbb{C}_A$ is a contextual category. Now if we suppose that $f: A \longrightarrow A'$ in $\mathbb{C}$ then we can define a functor $\mathbb{C}_f: \mathbb{C}_{A'} \longrightarrow \mathbb{C}_A$ by

$$B \xrightarrow{g} B' \qquad \Big|\!\!\Big\rightarrow^{\mathbb{C}_f} \qquad f^*B \xrightarrow{f^*g} f^*B',$$

where for any $g: B \longrightarrow B'$ in $\mathbb{C}_{A'}$, f^*g is the unique morphism from f^*B to f^*B' in $\mathbb{C}_A$ such that $f^*g \circ q(f,B') = q(f,B) \circ g$. Such a morphism exists uniquely in $\mathbb{C}_A$ because



is a pullback diagram in $\mathbb{C}$, because $q(f,B) \circ g \circ p(B',A') = q(f,B) \circ p(B,A) = p(f^*B,A) \circ f$ and because f^*g is in $\mathbb{C}_A$ iff $f^*g \circ p(f^*B',A) = p(f^*B,A)$.



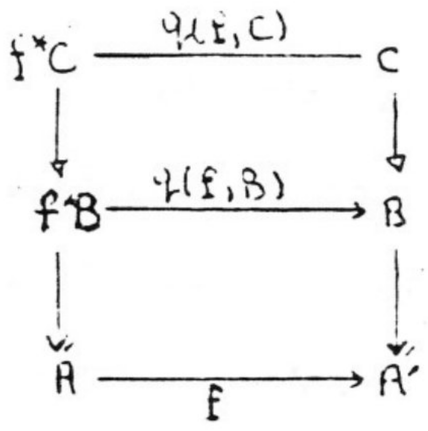
We must check that $\mathbb{C}_f$, so defined, is a functor. So we check that if $g:B \longrightarrow B'$ and $g':B' \longrightarrow B''$ in $\mathbb{C}_A$ then $f^*(g \circ g') = f^*g \circ f^*g'$. Well, $f^*(g \circ g')$ is defined as the unique morphism from f^*B to f^*B'' in $\mathbb{C}_A$ such that $f^*(g \circ g') \circ q(f,B'') = q(f,B) \circ g \circ g'$. But $f^*g \circ f^*g' : f^*B \longrightarrow f^*B''$ in $\mathbb{C}_A$ and satisfies $f^*g \circ f^*g' \circ q(f,B'') = f^*g \circ q(f,B) \circ g' = q(f,B) \circ g \circ g'$. So because of the uniqueness of $f^*(g \circ g')$ we must have $f^*(g \circ g') = f^*g \circ f^*g'$.

In fact, $\mathbb{C}_f$ is a contextual functor:

Lemma. (a). If $A' < B \lhd C$ in $\mathbb{C}$ then $f^*p(C) = p(f^*C)$.

(b). If $g:B \longrightarrow B'$ in $\mathbb{C}_A$ and $B' \leqslant C$ then $f^*g^*C = (f^*g)^*(f^*C)$ and $f^*(q(g,C)) = q(f^*g, f^*C)$.

Proof. (a). $f^*(p(C))$ is defined to be the unique map from f^*C to f^*B in $\mathbb{C}_A$ such that $f^*p(C) \circ q(f,B) = q(f,C) \circ p(C)$. So by uniqueness $p(f^*C) = f^*p(C)$ since $p(f^*C) \circ q(f,B) = q(f,C) \circ p(C)$ follows from lemma 1.



(b). $f^*(g^*C) = q(f,B)^*g^*C$, by lemma 1.

$= (q(f,B) \circ g)^*C$, by axiom (II).

$= (f^*g \circ q(f,B'))^*C$, by def. of f^*g.

$= (f^*g)^*(q(f,B')^*C)$, by axiom (II).

$= (f^*g)^*(f^*C)$, by lemma 1. As required for the first part.

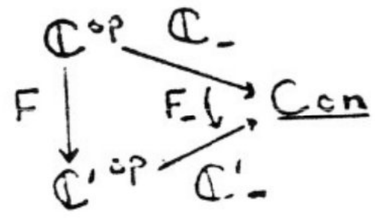
Now $f^*(q(g,C))$ is the unique morphism from f^*g^*C to f^*C in $\mathbb{C}_A$ such that $f^*(q(g,C))\circ q(f,C) = q(f,g^*C)\circ q(g,C)$. That $q(f^*g,f^*C) = f^*(q(g,C))$ follows from uniqueness because

$$\begin{aligned} q(f^*g,f^*C)\circ q(f,C) &= q(f^*g,q(f,B')^*C)\cdot q(q(f,B'),C) \text{ by lemma 1.} \\ &= q(f^*g\circ q(f,B'),C), \text{ by axiom (II).} \\ &= q(q(f,B)\circ g,C), \text{ by definition of } f^*g. \\ &= q(q(f,B),g^*C)\circ q(g,C), \text{ by axiom (II).} \\ &= q(f,g^*C)\circ q(g,C), \text{ by lemma 1.} \end{aligned}$$

Axiom (I) of the theory of contextual categories ensures that if A is an object of the contextual category $\mathbb{C}$ then $\mathbb{C}_{idA} = id\,\mathbb{C}_A$. Axiom (II) ensures that whenever $f:A\longrightarrow A'$ in $\mathbb{C}$ and $f':A'\longrightarrow A''$ in $\mathbb{C}$ then $\mathbb{C}_f\cdot\mathbb{C}_{f'} = \mathbb{C}_{f'\circ f}$. Thus $\mathbb{C}_-:\mathbb{C}^{op}\longrightarrow\underline{Con}$ is a functor.

Definition. If $F:\mathbb{C}\longrightarrow\mathbb{C}'$ is a contextual functor then if A is an object of $\mathbb{C}$ let $F_A:\mathbb{C}_A\longrightarrow\mathbb{C}_{F(A)}$ be the restriction of F to $\mathbb{C}_A$. F_A is, in fact, a contextual functor.

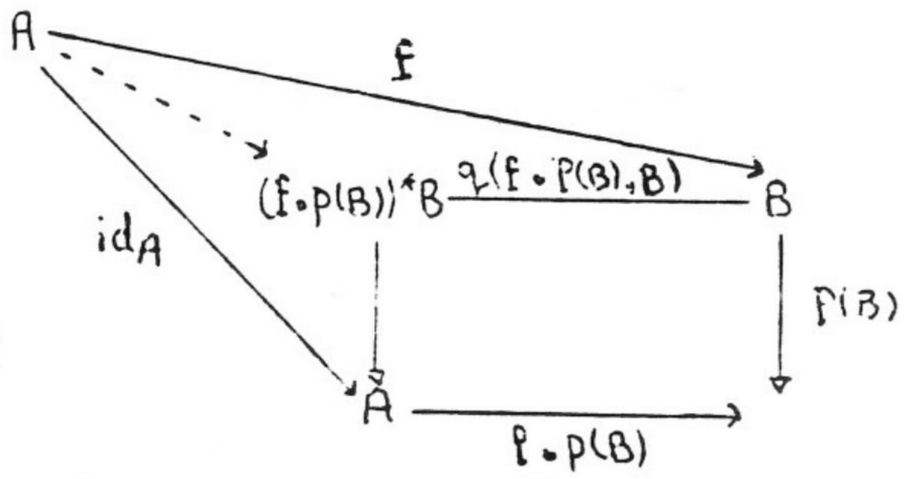
It is amusing to note that whenever $F:\mathbb{C}\longrightarrow\mathbb{C}'$ is a contextual functor then $F_-:\mathbb{C}_-\longrightarrow\mathbb{C}'_{F(-)}$ is a natural transformation. Thus whenever $F:\mathbb{C}\longrightarrow\mathbb{C}'$ is a contextual functor then we have the following diagram in the 2-category of categories.



The remainder of this section is ground work for the proof of the next section of the equivalence between theories and contextual categories. We begin by giving two definitions. With reference to the first of these two definitions we must apologise for the notation for we are not asking

that the morphism 'f' be thought of in any way as a quotation of f, the notation is merely convenient.

Definition. If $A,B \in |\mathbb{C}|$ and $f:A \longrightarrow B$ then 'f' is the unique morphism from A to $(f \circ p(B))^* B$ such that $'f' \circ p((f \circ p(B))^* B) = id_A$ and $'f' \circ q(f \circ p(B), B) = f$.



Definition. If $A \lhd B$ in $\mathbb{C}$ then $Arr_{\mathbb{C}}(B) = \{ f:A \longrightarrow B \mid f \circ p(B) = id_A \}$

Note that for all $A,B \in |\mathbb{C}|$, for all $f:A \longrightarrow B$, $'f' \in Arr_{\mathbb{C}}((f \circ p(B))^* B)$.

Lemma 2. If U is a generalised algebraic theory and if $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$ in $\mathbb{C}(U)$ are given by $A_n = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n >]$ and $B_m = [< y_1 \in \Omega_1, \ldots y_m \in \Omega_m >]$ then

(i) If $f:A_n \longrightarrow B_m$ is given by $f = [< t_1, \ldots t_m >]$ then $(f \circ p(B_m))^* B_m = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, z \in \Omega_m [t_1 \mid y_1, \ldots t_{m-1} \mid y_{m-1}] >]$ and $'f' = [< x_1, \ldots x_n, t_m >]$.

(ii) If $A_n \lhd A$ in $\mathbb{C}(U)$ and $A = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]$ then for any morphism g of $\mathbb{C}(U)$, $g \in Arr_{\mathbb{C}(U)}(A)$ iff g is of the form $[< x_1, \ldots x_n, t >]$ for some t such that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ is a derived rule of U.

(iii) For any i, $1 \leq i \leq n$, $p(A_n, A_i) = [< x_1, \ldots x_i >]$.

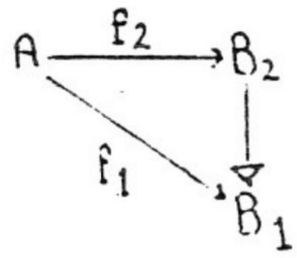
Proof. (i) amd (ii) follow directly from the definition of $\mathbb{C}(U)$.

(iii) The proof is by induction. Certainly $p(A_n, A_n) = [< x_1, \ldots x_n >]$

by definition of $\mathbb{C}(U)$. If we assume that the result holds for i+1, that is if we assume that in $\mathbb{C}(U)$, $p(A_n,A_{i+1}) = [<x_1,\ldots x_{i+1}>]$, then the result follows for i, since $p(A_n,A_i) = p(A_n,A_{i+1})_{\circ}p(A_{i+1}) = [<x_1,\ldots x_{i+1}>] \circ [<x_1,\ldots x_i>] = [<x_1,\ldots x_i>]$.

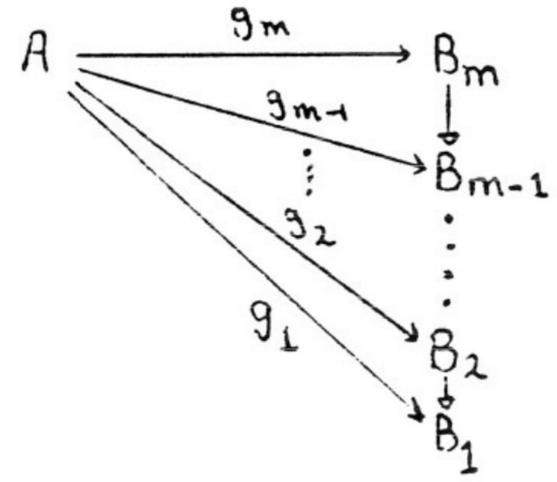
Hence the result holds for all i, $1 \leq i \leq n$.

<u>Lemma 3</u> (i) If



is a commutative diagram in $\mathbb{C}$ then for all $x:X \longrightarrow X'$ in $\mathbb{C}_{B_2}$, $f_2{}^*x = {}'f_2{}'^*f_1{}^*x$.

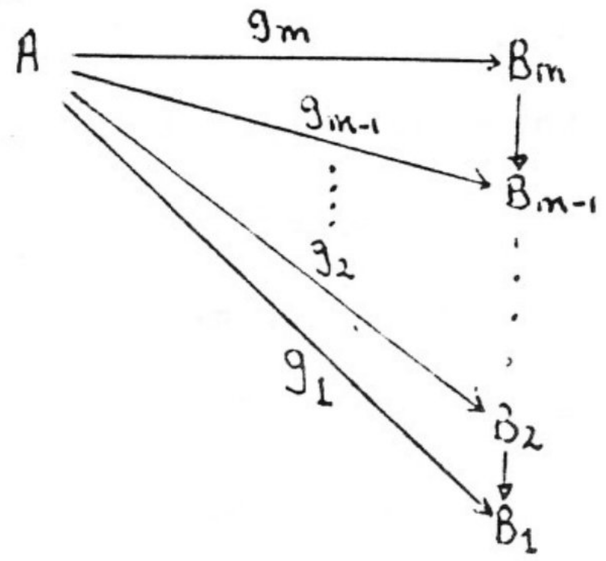
(ii) If $A \in |\mathbb{C}|$ and $1 \lhd B_1 \ldots \lhd B_m$ in $\mathbb{C}$ and for each j, $1 \leq j \leq m$ $g_j:A \longrightarrow B_j$ such that each triangle in the diagram commutes,



then for all $x:X \longrightarrow X'$ in $\mathbb{C}_{B_m}$, ${}'g_m{}'^*\ldots{}'g_1{}'^*p(A,1)^*x = g_m{}^*x$.

(iii) If $A \in |\mathbb{C}|$ and $1 \lhd B_1 .. \lhd B_m$ in $\mathbb{C}$ and for each j, $1 \leq j \leq m$ $\alpha_j \in Arr_{\mathbb{C}}(\alpha_{j-1}{}^*\ldots\alpha_1{}^*p(A,1)^*B_j)$ then there exists a unique sequence of morphisms $g_1,\ldots g_m$ of $\mathbb{C}$ such that for each $j, 1 \leq j \leq m$, $g_j:A \longrightarrow B_j$

such that the diagram



commutes and such that for all j, $1 \leq j \leq m$, $'g_j' = \alpha_j$.

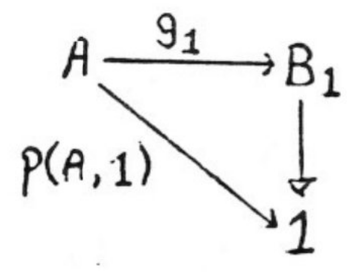
Proof. (i) $f_2^*x = ('f_2' \circ q(f_1, B_2))^*x$, since $f_2 = 'f_2' \circ q(f_1, B_2)$.

$= {}'f_2'^* q(f_1, B_2)^*x$, by axiom (II).

$= 'f_2'^* f_1^* x$, by lemma 1.

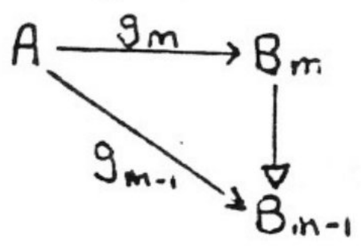
(ii) The proof is by induction on m.

If m = 1 then we have



in $\mathbb{C}$, thus by part (i) $g_1^*x = 'g_1'^* p(A,1)^*x$.

If $m > 1$, if we assume that for all j, $1 \leq j < m$, $g_j^*x = 'g_j'^* \ldots 'g_1'^* p(A,1)^*x$, for all x in $\mathbb{C}_{B_j}$. Then since



in $\mathbb{C}$ thus by part (i) $g_m^*x = 'g_m'^*(g_{m-1}^*x) = 'g_m'^* 'g_{m-1}'^* \ldots 'g_1'^* p(A,1)^*x$, by inductive hypothesis.

(iii) The proof is by induction on m.

If m = 1 then since 'g_1' is required to be α_1 so we must choose g_1 such that $g_1 = \alpha_1 \circ q(p(A,1),B_1)$.

If $m > 1$, and if we assume that $g_1, \ldots g_{m-1}$ are such that $g_j \circ p(B_j) = B_{j-1}$, for all $j, 1 < j \leq m-1$ and such that 'g_j' $= \alpha_j$, for each $j, 1 \leq j < m$. Then by part (ii) $\alpha_m : A \longrightarrow g_{m-1}{}^*B$. So we can and must choose $g_m = \alpha_m \circ q(g_{m-1}, B_m)$.

Lemma 4. (i) If $A \xrightarrow{f} B \xrightarrow{g} C$ in $\mathbb{C}$ then $f^*'g' = 'f \circ g'$.

(ii) If $A \in |\mathbb{C}|$ and $1 \lhd B_1 \ldots \lhd B_m$ in $\mathbb{C}$ and for each $j, 1 \leq j \leq m$, $\beta_j \in Arr_{\mathbb{C}}(\beta_{j-1}{}^* \ldots \beta_1{}^* p(A,1)^* B_j)$ then for each $j, 1 \leq j \leq m$, $\beta_m{}^* \ldots \beta_1{}^* \; p(A,1)^* 'p(B_n, B_j)' = \beta_j$.

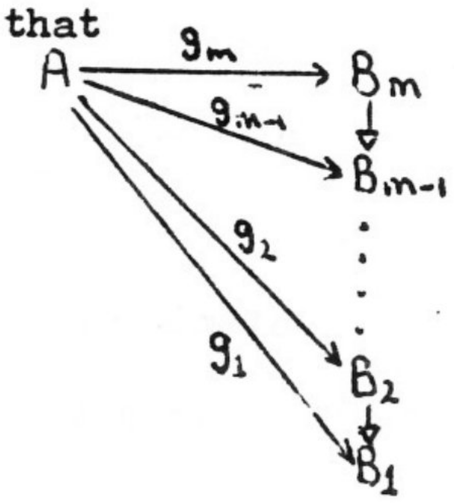
(iii). If $A \in |\mathbb{C}|$, $1 \lhd B_1 \ldots \lhd B_m$ and $1 \lhd C_1 \ldots \lhd C_\ell$ in $\mathbb{C}$, if for all $j, 1 \leq j \leq m$, $\beta_j \in Arr_{\mathbb{C}}(\beta_{j-1}{}^* \ldots \beta_1{}^* p(A,1)^* B_j)$ and for all $k, 1 \leq k \leq \ell$, $\gamma_k \in Arr_{\mathbb{C}}(\gamma_{k-1}{}^* \ldots \gamma_1{}^* p(B_m,1)^* C_k)$, then for all $x : X \longrightarrow X'$ in $\mathbb{C}$, $(\beta_m{}^* \ldots \beta_1{}^* p(A,1)^* \gamma_\ell)^* (\beta_m{}^* \ldots \beta_1{}^* p(A,1)^* \gamma_{\ell-1})^* \ldots (\beta_m{}^* \ldots \beta_1{}^* p(A,1)^* \gamma_1)^* p(A,1)^* x = \beta_m{}^* \ldots \beta_1{}^* p(A,1)^* \gamma_\ell{}^* \ldots \gamma_1{}^* p(B_m,1)^* x$.

Proof (i) '$f \circ g$' is the unique morphism from A to $(f \circ g \circ p(C))^* C$ such that '$f \circ g$'$\circ p((f \circ g \circ p(C))^* C) = id_A$ and '$f \circ g$'$\circ q(f \circ g \circ p(C), C) = f \circ g$. It thus suffices to show that $f^*'g'$ is such a morphism.

By definition of 'g', 'g' : $B \longrightarrow (g \circ p(C))^* C$ such that 'g'$\circ p((g \circ p(C))^* C) = id_B$ and 'g'$\circ q(g \circ p(C), C) = g$. Thus $f^*'g' : A \longrightarrow f^*(g\ p(C))^* C$, that is to say, $f^*'g' : A \longrightarrow (f \circ g \circ p(C))^* C$. Also since 'g'$\circ p((g \circ p(C))^* C = id_B$, $f^*'g' \circ f^* p((g \circ p(C))^* C) = id_A$. But $f^* p((g \circ p(C))^* C) = p(f^*(g \circ p(C))^* C) = p((f \circ g \circ p(C))^* C)$, hence $f^*'g' \circ p((f \circ g \circ p(C))^* C) = id_A$. Which is one property of '$f \circ g$'.

As for the other, from $'g'\circ q(g\circ p(C),C) = g$ deduce that $f\circ'g'\circ q(g\circ p(C),C) = f\circ g$. But $f\circ'g' = f^*'g'\circ q(f\circ(g\circ p(C))^*C)$ and $q(f,(g\circ p(C))^*C)\circ q(g\circ p(C),C) = q(f\circ g\circ p(C),C)$ hence we have $f^*'g'\circ q(f\circ g\circ p(C),C) = f\circ g$. As required. From the uniqueness of $'f\circ g'$ we conclude that $f^*'g' = 'f\circ g'$.

(ii). By lemma 3(iii), there corresponds to $\beta_1,\ldots\beta_m$ a sequence of morphisms $g_1,\ldots g_m$ such that



commutes and such that $'g_j' = \beta_j$.

By lemma 3(ii) $\beta_m^*\ldots\beta_1^*p(A,1)^*'p(B_m,B_j)' = g_m^*'p(B_m,B_j)'$. By part (i) of this lemma $g_m^*'p(B_m,B_j)' = 'g_m\circ p(B_m,B_j)' = \beta_j$ as required.

(iii). Again we use lemma 3(iii). Corresponding to (β_j) $1\leq j\leq m$ we have (g_j) $1\leq j\leq m$ such that $'g_j' = \beta_j$. Corresponding to (γ_k) $1\leq k\leq \ell$ we have (h_k) $1\leq k\leq \ell$ such that $'h_k' = \gamma_k$.

$(\beta_m^*\ldots\beta_1^*p(A,1)^*\gamma_\ell)^*\ \ldots\ \ldots(\beta_m^*\ldots\beta_1^*p(A,1)^*\gamma_1)^*p(A,1)^*x$

$= (g_m^*\gamma_\ell)^*\ldots(g_m^*\gamma_1)^*p(A,1)^*x$, by lemma 3(ii).

$= 'g_m\circ h_\ell'^*\ldots'g_m\circ h_1'^*p(A,1)^*x$, by this lemma part (i).

$= g_m\circ h_\ell^*x$, by lemma 3(ii).

$= g_m^*h_\ell^*x$, by axiom (II).

$= 'g_m'^*\ldots'g_1'^*p(A,1)^*'h_\ell'^*\ldots'h_1'^*p(B_m,1)^*x$, by lemma 3(ii).

$= \beta_m^*\ldots\beta_1^*p(A,1)^*\gamma_\ell^*\ldots\gamma_1^*p(B_m,1)^*x$. As required.

Lemma 5. If $f:A\longrightarrow A'$ and A' B in $\mathbb{C}$ then $'q(f,B)' = 'id_{f^*B}'$.

Proof. By definition of $'id_{f^*B}'$, $'id_{f^*B}'\circ q(p(f^*B),f^*B) = id_{f^*B}$. Hence $'id_{f^*B}'\circ q(p(f^*B),f^*B)\circ q(f,B) = q(f,B)$, that is to say $'id_{f^*B}'\circ q(p(f^*B)\ f,B) = q(f,B)$, that is

$'id_{f^*B}'\circ q(p(f^*B)\ f,B) = q(f,B)$ follows from (II)

2.4 Contextual Categories = Generalised Algebraic Theories

In this section we establish the equivalence between the category Con of contextual categories and the category GAT of generalised algebraic theories. We split the section into subsections as follows: 1. define a functor $\mathbb{C} : \underline{GAT} \longrightarrow \underline{Con}$, 2. define a functor $U : \underline{Con} \longrightarrow \underline{GAT}$, 3. prove that $U \circ \mathbb{C} \simeq id_{\underline{GAT}}$, 4. prove that $C \circ U \simeq id_{\underline{Con}}$.

2.4.1 Definition of $\mathbb{C} : \underline{GAT} \longrightarrow \underline{Con}$. $\mathbb{C}$ has been defined on objects in §2.2, if U is a theory then $\mathbb{C}(U)$ is a contextual category.

If U and U' are theories and $[I] : U \longrightarrow U'$ in GAT the define $\mathbb{C}([I]) : \mathbb{C}(U) \longrightarrow \mathbb{C}(U')$ by

$$
\begin{array}{ccc}
[\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle] & \xrightarrow{\mathbb{C}([I])} & [\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle] \\
\Big\downarrow [\langle t_1, \ldots t_m \rangle] & & \Big\downarrow [\dot{I}(t_1), \ldots \dot{I}(t_m) \rangle] \\
[\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle] & & [\langle y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m) \rangle]
\end{array}
$$

Then $\mathbb{C}$ is well defined on morphisms because by lemma 1 of §1.13 if I and J are interpretations of U in U' and if $I \equiv J$ then for all derived T and $\in$-rules R of U, $\hat{I}(R) \equiv \hat{J}(R)$,

To see that $\mathbb{C}([I])$ is well defined and takes objects and morphisms of $\mathbb{C}(U)$ to objects respectively morphisms of $\mathbb{C}(U')$ we require the following:-

Lemma. If U and U' are generalised algebraic theories and if I is an

interpretation of U in U' then

(i). If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is a U- context then $\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)$ is a U'-context.

(ii). If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle x_1' \in \Delta_1', \ldots x_n' \in \Delta_n' \rangle$ are U-contexts and $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle \equiv \langle x_1' \in \Delta_1', \ldots x_n' \in \Delta_n' \rangle$ then $\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle \equiv \langle x_1' \in I(\Delta_1'), \ldots x_n' \in I(\Delta_n') \rangle$.

(iii). If $\langle t_1, \ldots t_m \rangle$ is a U-realisation of the context $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt the context $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ then $\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle$ is a U'-realisation of $\langle y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m) \rangle$ wrt $\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle$.

(iv). If $\langle t_1, \ldots t_m \rangle$ is a U-realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and if $\langle t_1' \ldots t_m' \rangle$ is a U-realisation of $\langle y_1' \in \Omega_1', \ldots y_m' \in \Omega_m' \rangle$ wrt $\langle x_1' \in \Delta_1', \ldots x_n' \in \Delta_n' \rangle$ and if $\langle t_1, \ldots t_m \rangle \equiv \langle t_1', \ldots t_m' \rangle$ then $\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle \equiv \langle \dot{I}(t_1'), \ldots \dot{I}(t_m') \rangle$.

Proof. (i) Follows immediately from lemma 2 of §1.11.

(ii). For each i, $1 \leq i \leq n$, $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \Delta_i' [x_1 \mid x_1', \ldots x_{i-1} \mid x_{i-1}']}$ is a derived rule of U, hence by lemma 2 of §1.11,

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_{i-1} \in \dot{I}(\Delta_{i-1})}{\dot{I}(\Delta_i) = \dot{I}(\Delta_i' [x_1 \mid x_1', \ldots x_{i-1} \mid x_{i-1}'])}$ is a derived rule of U'. But

by lemma 1 of §1.11, $\dot{I}(\Delta_i [x_1 \mid x_1', \ldots x_{i-1} \mid x_{i-1}']) \equiv \dot{I}(\Delta_i) [x_1 \mid x_1', \ldots x_{i-1} \mid x_{i-1}']$. Thus for each i, $1 \leq i \leq n$,

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_{i-1} \in \dot{I}(\Delta_{i-1})}{\dot{I}(\Delta_i) = \dot{I}(\Delta_i) [x_1 \mid x_1', \ldots x_{i-1} \mid x_{i-1}']}$ is a derived rule of U'. That is

$\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle \equiv \langle x_1' \in \dot{I}(\Delta_1'), \ldots x_n \in \dot{I}(\Delta_n') \rangle$.

(iii) For each j, $1 \leq j \leq m$, $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t_j \in \Omega_j [\, t_1 \mid y_1, \ldots t_{j-1} \mid y_{j-1} \,]}$ is a

derived rule of U, thus by lemmas 1 and 2 of § 1.11,

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t_j) \in \dot{I}(\Omega_j)[\, \dot{I}(t_1) \mid y_1, \ldots \dot{I}(t_{j-1}) \mid y_{j-1} \,]}$ is a derived rule of U'.

Thus $\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle$ is a realisation of $\langle y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m)$

$\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n) \rangle$.

(iv) For each j, $1 \leq j \leq m$, $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t_j = t_j' [\, x_1 \mid x_1', \ldots x_n \mid x_n' \,] \in \Omega_j [\, t_1 \mid y_1, \ldots t_{j-1}$

is a derived rule of U, thus by lemmas 1 and 2 of § 1.11,

$\dfrac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n)}{\dot{I}(t_j) = \dot{I}(t_j')[\, x_1 \mid x_1', \ldots x_n \mid x_n' \,] \in \dot{I}(\Omega_j)[\, \dot{I}(t_1) \mid y_1, \ldots \dot{I}(t_{j-1}) \mid y_{j-1}$

is a derived rule of U'. That is $\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle \equiv \langle \dot{I}(t_1'), \ldots \dot{I}(t_m'$

So $\mathbb{C}([\, I \,])$ is well defined. It remains to show that $\mathbb{C}([\, I \,])$ is a contextual functor:

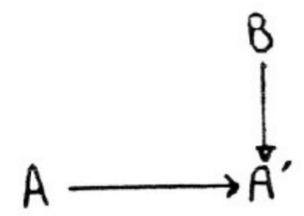
Lemma (i) If $A \in |\mathbb{C}(U)|$ then $\mathbb{C}([\, I \,])(id_A) = id_{\mathbb{C}([\, I \,])(A)}$.

(ii) If $A \xrightarrow{f} B \xrightarrow{g} C$ in $\mathbb{C}(U)$ then $\mathbb{C}([\, I \,])(fog) =$

$\mathbb{C}([\, I \,])(f) \circ \mathbb{C}([\, I \,])(g)$.

(iii) If $A \in |\mathbb{C}|$ and $1 < A$ then $\mathbb{C}([\, I \,])(p(A)) = p(\mathbb{C}([\, I \,])(A))$.

(iv) If



in $\mathbb{C}(U)$ then $\mathbb{C}([\, I \,])(f^*B) = \mathbb{C}([\, I \,])(f)^* \mathbb{C}([\, I \,])(B)$ and

$\mathbb{C}([\, I \,])(q(f,B)) = q(\mathbb{C}([\, I \,])(f), \mathbb{C}([\, I \,])(B))$.

Proof. (i) and (iii) are both remarkably trivial.

(ii) Suppose that $A = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle]$, $B = [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle]$ and $C = [\langle z_1 \in \Lambda_1, \ldots z_\varrho \in \Lambda_\varrho \rangle]$, $f = [\langle t_1, \ldots t_m \rangle]$ and $g = [\langle S_1, \ldots S_\varrho \rangle]$, where $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle S_1, \ldots S_\varrho \rangle$ is a realisation of $\langle z_1 \in \Lambda_1, \ldots z_\varrho \in \Lambda_\varrho \rangle$ wrt $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$. If we use the definition of $\mathbb{C}([I])$, the definition of composition in $\mathbb{C}(U)$ and $\mathbb{C}(U')$, and lemma 1 of §1.11 then:

$$\mathbb{C}([I])(fog) = \mathbb{C}([I])([\langle S_1[t_1 \mid y_1, \ldots t_m \mid y_m], \ldots S_\varrho[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]) = [\langle \dot{I}(S_1[t_1 \mid y_1, \ldots t_m \mid y_m]), \ldots \dot{I}(S_\varrho[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]$$
$$= [\langle \dot{I}(S_1)[\dot{I}(t_1) \mid y_1, \ldots \dot{I}(t_m) \mid y_m], \ldots \dot{I}(S_\varrho)[\dot{I}(t_1) \mid y_1, \ldots \dot{I}(t_m) \mid y_m] \rangle]$$
$$= [\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle] \circ [\langle \dot{I}(S_1), \ldots \dot{I}(S_m) \rangle] = \mathbb{C}([I])(f) \circ \mathbb{C}([I])(g).$$

(iv) Suppose that $A = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle]$, $B = [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega \rangle]$ and $f = [\langle t_1, \ldots t_m \rangle]$, where $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$.

$$\mathbb{C}([I])(f^*B) = \mathbb{C}([I])([\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y \in \Omega[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]) = [\langle x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta_n), y \in \dot{I}(\Omega)[\dot{I}(t_1) \mid y_1, \ldots \dot{I}(t_m) \mid y_m] \rangle]$$
$$= [\langle \dot{I}(t_1), \ldots \dot{I}(t_m) \rangle] * [\langle y_1 \in \dot{I}(\Omega_1), \ldots y_m \in \dot{I}(\Omega_m), y \in \dot{I}(\Omega) \rangle] = \mathbb{C}([I])(f) * \mathbb{C}([I])(B).$$

$$\mathbb{C}([I])(q(f,B)) = \mathbb{C}([I])([\langle t_1, \ldots t_m, y \rangle]) = [\langle \dot{I}(t_1), \ldots \dot{I}(t_m), y \rangle] = q(\mathbb{C}([I])(f), \mathbb{C}([I])(B)).$$

Supposing that I is an interpretation of U in U' and that I' is an interpretation of U' in U", then by lemma 3 of §1.11 for any

expression e of U, $(I' \circ I)\,(e) = \dot{I}'(\dot{I}(e))$. Thus it follows that in this section $\mathbb{C}([\,I'\,]) \circ \mathbb{C}([\,I\,]) = \mathbb{C}([\,I' \circ I\,]) = \mathbb{C}([\,I'\,] \circ [\,I\,])$. So for sure $\mathbb{C}$:GAT $\longrightarrow$ Con is a functor.

2.4.2 The definition of U:Con $\longrightarrow$ GAT. It is convenient to say that an object A of a contextual category $\mathbb{C}$ is trivial just in case A is the least element 1 of $\mathbb{C}$. Similarly we say that a morphism of $\mathbb{C}$ is trivial just in case its codomain is the trivial object.

We begin by describing the functor U:Con $\longrightarrow$ GAT on objects of Con. If $\mathbb{C}$ is a contextual category then $U(\mathbb{C})$ is the generalised algebraic theory described as follows: $U(\mathbb{C})$ has a sort symbol $\overline{A}$ for every non-trivial object A of $\mathbb{C}$. $U(\mathbb{C})$ has an operator symbol $\overline{f}$ for every non-trivial morphism f of $\mathbb{C}$. If $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$ then the introductory rule for $\overline{A}$ in $U(\mathbb{C})$ is

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{A}(x_1, \ldots x_n) \text{ is a type}} \quad .$$

If $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ and $f: A_n \longrightarrow B$ where $1 < B$ then the introductory rule for $\overline{f}$ in $U(\mathbb{C})$ is

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{(f \circ p(B)) * B}(x_1, \ldots x_n)} \quad .$$

The axioms of $U(\mathbb{C})$ arise from three different situations, $U(\mathbb{C})$ has just the following axioms:

(i) For $n \geqslant 0$, $m \geqslant 1$ and $\ell \geqslant 0$, if $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$ and $1 \lhd C_1 \ldots \lhd C_\ell$ in $\mathbb{C}$ and if $f: A_n \longrightarrow B_m$ and $g: B_m \longrightarrow C_\ell$ in $\mathbb{C}$ then $U(\mathbb{C})$ has the axiom

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{f \circ g}(x_1, \ldots x_n) = \overline{g}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f \circ p(B_m, B_{m-1})}(x_1, \ldots x_n), \overline{f}(x_1, \ldots x_n)) \in \overline{(f \circ g \circ p(C_\ell)) * C}\,(x_1, \ldots x_n)} \quad .$$

(ii) For $n > o$, if $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ then for each i, $1 \leq i \leq n$, $U(\mathbb{C})$ h

the axiom $\dfrac{x_1 \in \bar{A}_1, \ldots x_n \in \bar{A}_n(x_1, \ldots x_{n-1})}{\overline{p(A_n, A_i)}(X_1, \ldots x_n) = x_i \in \bar{A}_i(x_1, \ldots x_{i-1})}$

(iii) For $n \geq o$, $m \geq 1$, if $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m \lhd B$ in $\mathbb{C}$ and if $f: A_n \longrightarrow B_m$ then $U(\mathbb{C})$ has the axioms

$$\frac{x_1 \in \bar{A}_1, \ldots x_n \in \bar{A}_n(x_1, \ldots x_{n-1})}{\overline{f^*B}(x_1, \ldots x_n) = \bar{B}(\overline{fop(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{fop(B_m, B_{m-1})}(x_1, \ldots x_n), \bar{f}(x_1, \ldots$$

and $\dfrac{x_1 \in \bar{A}_1, \ldots x_n \in \bar{A}_n(x_1, \ldots x_{n-1}), y \in \overline{f^*B}(x_1, \ldots x_n)}{\overline{q(f,B)}(x_1, \ldots x_n, y) = y \in \overline{f^*B}(x_1, \ldots x_n)}$. This completes the definition of $U(\mathbb{C})$.

As for the action of the functor U on morphisms, if $F: \mathbb{C} \longrightarrow \mathbb{C}'$ is a contextual functor then define a preinterpretation $U(F)$ of $U(\mathbb{C})$ in $U(\mathbb{C}')$ follows: If $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$ then define $U(F)(\bar{A}) = \overline{F(A)}(v_1, \ldots v_n)$, if $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ and $f: A_n \longrightarrow B$ then define $U(F)(\bar{f}) = \overline{F(f)}(v_1, \ldots v_n)$ $v_1, v_2, \ldots$ is supposed to be the standard ennumeration of the set V of variables, see §1.11.

Lemma. $U(F)$ is an interpretation of $U(\mathbb{C})$ in $U(\mathbb{C}')$.

Proof. We have to check that for every introductory rule or axiom R of $U(\mathbb{C})$, the rule $\widehat{U(F)}(R)$ is a derived rule of $U(\mathbb{C}')$. But it happens that in each case $\widehat{U(F)}(R)$ is actually an introductory rule or an axiom of $U(\mathbb{C}$ and thus a derived rule. Thus there is little work to be done.

For example, if $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$, so that $U(\mathbb{C})$ has the introductory rule $\dfrac{x_1 \in \bar{A}_1, \ldots x_n \in \bar{A}_n(x_1, \ldots x_{n-1})}{\bar{\ }(x_1, \ldots x_n) \text{ is a typ}}$ then

$$\widehat{I(F)}\left(\frac{x_1 \in \bar{A}_1, \ldots x_n \in \bar{A}_n(x_1, \ldots x_{n-1})}{\bar{A}(x_1, \ldots x_n) \text{ is a type}}\right) =$$

$$\frac{x_1 \in U\dot{(}F)(\bar{A}_1), \ldots x_n \in U\dot{(}F)(\bar{A}_n(x_1, \ldots x_{n-1}))}{U\dot{(}F)(\bar{A}(x_1, \ldots x_n)) \text{ is a type}} =$$

$$\frac{x_1 \in U(F)(\overline{A_1}), \ldots x_n \quad U(F)(\overline{A_n})[x_1 | v_1, \ldots x_{n-1} | v_{n-1}]}{U(F)(\bar{A})[x_1 | v_1, \ldots x_n | v_n] \text{ is a type}} =$$

$$\frac{x_1 \in \overline{F(A_1)}, \ldots x_n \in \overline{F(A_n)}(x_1, \ldots x_{n-1})}{\overline{F(A)}(x_1, \ldots x_n) \text{ is a type}}$$

which is of course the introductory rule for $\overline{F(A)}$ in $U(\mathfrak{C}')$.

If $F: \mathfrak{C} \longrightarrow \mathfrak{C}'$ and $F': \mathfrak{C}' \longrightarrow \mathfrak{C}''$ in Con then $U(F'oF) = U(F')oU(F)$. This is because for any symbol $\bar{L}$ of $U(\mathfrak{C})$ and for appropriate n,

$U(F'oF)(L) = \overline{F'(F(L))}(v_1, \ldots v_n) = U\dot{(}F')(\overline{F(L)}(v_1, \ldots v_n)) = U\dot{(}F')(U(F)(\bar{L})) =$ $(U(F')oU(F))(\bar{L})$. Thus we have defined a functor from the category Con to the category of generalised algebraic theories and interpretations. By taking the value of U at F to be $[U(F)]$ we get a functor $U: Con \longrightarrow$

2.4.2 The proof that $Uo\mathfrak{C} \simeq id_{GAT}$. For every generalised algebraic theory U we define an interpretation ϕ_U of U in $U(\mathfrak{C}(U))$. We show that $[\phi_-]$ (ie $\lambda U \in |GAT| . [\phi_U]$) is a natural transformation $[\phi_-] : id_{GAT} \longrightarrow Uo\mathfrak{C} : GAT \longrightarrow GAT$. For every theory U we define an interpretation ψ of $U(\mathfrak{C}(U))$ in U and show that $[\phi_U]o[\psi_U] = id_{U(\mathfrak{C}(U))}$ and that $[\psi_U]o[\phi_U] = id_U$.

If U is a theory then the preinterpretation ϕ_U of U in $U(\mathfrak{C}(U))$ is defined as follows: If A is a sort symbol of U introduced by

the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{A(x_1, \ldots x_n) \text{ is a type}}$ then define $\varphi_U(A)$

$= \overline{[< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in A(x_1, \ldots x_n) >]}\,(v_1, \ldots v_n)$. If f is an operator symbol of U introduced by the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{f(x_1, \ldots x_n) \in \Delta}$ then define

$$\varphi_U(f) = \overline{[< x_1, \ldots x_n, f(x_1, \ldots x_n) >]}\ (v_1, \ldots v_n).$$

When it is unlikely to lead to misunderstanding then we drop the subscript U from φ_U. We wish to show that for any theory U, the preinterpretation φ of U in $U(\mathbb{C}(U))$ is actually an interpretation. This requires a long string of lemmas. We do in fact show that for any derived rule of U of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$, $\hat{\varphi}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}\right) =$

$$\frac{v_1 \in \overline{[< x_1 \in \Delta_1 >]}, \ldots v_n \in \overline{[< x_1 \in \Delta_1, \ldots x_n \in \Delta_n >]}\,(v_1, \ldots v_{n-1})}{\overline{[< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]}\,(v_1, \ldots v_n) \text{ is a type}}$$

and we show that for any derived rule of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$,

$$\hat{\varphi}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right) \equiv$$

$$\frac{v_1 \in \overline{[< x_1 \in \Delta_1 >]}, \ldots v_n \in \overline{[< x_1 \in \Delta_1, \ldots x_n \in \Delta_n >]}\,(v_1, \ldots v_n)}{\overline{[< x_1, \ldots x_n, t >]}\,(v_1, \ldots v_n) \in \overline{[< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]}\,(v_1, \ldots v_n)}$$

From which it follows of course that φ is an interpretation.

Lemma 1. If $\mathbb{C}$ is a contextual category and if for some $n \geqslant 0$ and $m \geqslant 1$, $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}$ then the rule

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{B}_m(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f \circ p(B_m, B_{n-1})}(x_1, \ldots x_n))}$$

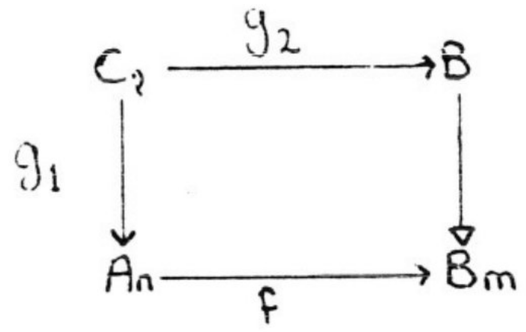
is

a derived rule of $U(\mathbb{C})$.

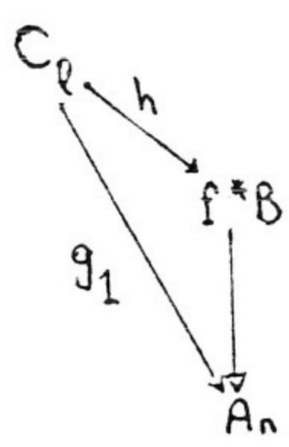
Proof. The rule can be derived by the principle T1 from the introductory rule of $\overline{f}$ and the axiom of kind (iii) which is

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{f \circ p(B_m)^* B_m}(x_1, \ldots x_n) = \overline{B_m}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f \circ p(B_m, B_{m-1})}(x_1, \ldots x_n))}$$

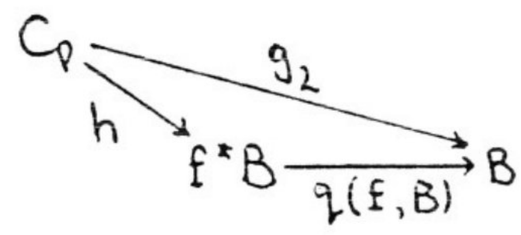
Lemma 2. If $\mathbb{C}$ is a contextual category and if for some $n \geqslant 1$, $m \geqslant 1$ and $\ell \geqslant 0$, $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m \lhd B$, $1 \lhd C_1 \ldots C_\ell$, $f: A_n \longrightarrow B_m$, $g_1: C_\ell \longrightarrow A_n$ and $g_2: C_\ell \longrightarrow B$ in $\mathbb{C}$ such that the diagram



commutes, and if $h: C_\ell \longrightarrow f^*B$ is the unique morphism such that



and



commute

then the rule $$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{h}(z_1, \ldots z_\ell) = \overline{g_2}(z_1, \ldots z_\ell) \in (\overline{g_1 \circ f})^* B(z_1, \ldots z_\ell)}$$

is a derived rule of $U(\mathbb{C})$.

Proof. From Lemma 1 it follows that for each i, $1 \leq i \leq n$ the rule

$$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{hop(f^*B, A_i)}(z_1, \ldots z_\ell) \in \overline{A_i}(\overline{hop(f^*B, A_1)}(z_1, \ldots z_\ell), \ldots \overline{hop(f^*B, A_{i-1})}(z_1, \ldots z_\ell))}$$

is a derived rule of $U(\mathbb{C})$. By the same lemma so is

$$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{h}(z_1, \ldots z_\ell) \in \overline{f^*B}(\overline{hop(f^*B, A_1)}(z_1, \ldots z_\ell), \ldots \overline{hop(f^*B, A_n)}(z_1, \ldots z_\ell))}$$

. Hence we can substitute this n+1-tuple of terms into the axiom

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}), y \in \overline{f^*B}(x_1, \ldots x_n)}{\overline{q(f,B)}(x_1, \ldots x_n, y) \in \overline{f^*B}(x_1, \ldots x_n)}$$

$$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{q(f,B)}(\overline{hop(f^*B, A_1)}(z_1, \ldots z_\ell), \ldots \overline{hop(f^*B, A_n)}(z_1, \ldots z_\ell), \overline{h}(z_1, \ldots z_\ell) = \overline{h}(z_1, \ldots z$$

is a derived rule of $U(\mathbb{C})$.

But since $hoq(f,B) = g_2$ (and also using $g_2 op(B, B_m) = g_1 of$) the rule

$$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{g_2}(z_1, \ldots z_\ell) = \overline{q(f,B)}(\overline{hop(f^*B, A_1)}(z_1, \ldots z_\ell), \ldots \overline{hop(f^*B, A_n)}(z_1, \ldots z_\ell), \overline{h}(z_1, \ldots z_\ell)) \in \overline{(g_1 of)^*B}(z_1, \ldots z_\ell)}$$

is an axiom of $U(\mathbb{C})$. Thus by transitivity of $=$

$$\frac{z_1 \in \overline{C_1}, \ldots z_\ell \in \overline{C_\ell}(z_1, \ldots z_{\ell-1})}{\overline{h}(z_1, \ldots z_\ell) = \overline{g_2}(z_1, \ldots z_\ell) \in (\overline{g_1 of)^*B}(z_1, \ldots z_\ell)}$$

is a derived rule of $U(\mathbb{C})$.

As a special case of lemma 2 we have:

Corollary 3. If $\mathbb{C}$ is a contextual category and if for some $n \geq 0$, $1 \lhd A_1 \ldots \lhd A_n$ and $f: A_n \longrightarrow B$ in $\mathbb{C}$ then the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{'f'}(x_1, \ldots x_n) = \overline{f}(x_1, \ldots x_n) \in (\overline{f \circ p(B))^* B}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathbb{C})$.

If we use the description lemma 2 of §2.3 of the contextual category $\mathbb{C}(U)$ associated with a theory U then corollary 3 can be rewritten as follows:

Corollary 4. If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ and $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega \rangle$ are contexts of a theory U and if $\langle t_1, \ldots t_m, t \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ then the rule

$$\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{f}(v_1, \ldots v_n) = \overline{g}(v_1, \ldots v_n) \in \overline{C}(v_1, \ldots v_n)}$$

is a derived rule of $U(\mathbb{C}(U))$,

where $A_i = [\langle x_1 \in \Delta_1, \ldots x_i \in \Delta_i \rangle]$, $f = [\langle t_1, \ldots t_m, t \rangle]$, $g = [\langle x_1, \ldots x_n, t \rangle]$ and $C = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y \in \Omega[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]$

Lemma 5. If $\mathbb{C}$ is a contextual category and for some $n \geqslant 0$ ad $m \geqslant 1$, $1 \lhd A_1 \ldots \lhd A_n, 1 \lhd B_1 \ldots \lhd B_m \lhd B$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}$, then (i) the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f^*B}(x_1, \ldots x_n) = \overline{B}(\overline{'f \circ p(B_m, B_1)'}(x_1, \ldots x_n), \ldots \overline{'f \circ p(B_m, B_{m-1})'}(x_1, \ldots x_n), \overline{'f'}(x_1, \ldots x_n))}$$

is a derived rule of $U(\mathbb{C})$. (ii) If also $g \in \mathrm{Arr}_{\mathbb{C}}(B)$ then the rule

$$\frac{x_1 \in \overline{A}_1, \ldots x_n \in \overline{A}_n(x_1, \ldots x_{n-1})}{\overline{'f \circ g'}(x_1, \ldots x_n) = \overline{'g'}(\overline{'f \circ p(B_m, B_1)'}(x_1, \ldots x_n), \ldots \overline{'f \circ p(B_m, B_{m-1})'}(x_1, \ldots x_n), \overline{'f'}(x_1, \ldots x_n)) \in \overline{f^*B}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathbb{C})$.

Proof. Both (i) and (ii) follow from corollary 3. For example (ii) follows from the axiom

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{fog}(x_1, \ldots x_n) = \overline{g}(\overline{fop(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{fop(B_m, B_{m-1})}(x_1, \ldots x_n), \overline{f}(x_1, \ldots x_n}$$

since by corollary 3 we have as derived rules of $U(\mathbb{C})$ the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{'fog'}(x_1, \ldots x_n) = \overline{fog}(x_1, \ldots x_n) \in \overline{f^*B}(x_1, \ldots x_n)}$$

(since $fogop(B) = fid_B = f$) and for each j, $1 \leq j \leq m$, the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{'fop(B_m, B_j)'}(x_1, \ldots x_n) = \overline{fop(B_m, B_j)}(x_1, \ldots x_n) \in \overline{(fop(B_m, B_{j-1}))^*B_j}(x_1, \ldots x_n)}$$

□

In particular if $\mathbb{C} = \mathbb{C}(U)$ then we get the following:

Corollary 6. If U is a theory, if for some $n \geq 0$, $m \geq 1$, $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$, $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega \rangle$ are contexts of U and $\langle t_1, \ldots t_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ then (i) the rule

$$\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{C}(v_1, \ldots v_n) = \overline{B}(\overline{g_1}(v_1, \ldots v_n), \ldots \overline{g_m}(v_1, \ldots v_n))}$$

is a derived rule of $U(\mathbb{C}(U))$, where $A_i = [\langle x_1 \in \Delta_1, \ldots x_i \in \Delta_i \rangle]$,
$C = [\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n, y \in \Omega[t_1 \mid y_1, \ldots t_m \mid y_m] \rangle]$,
$B = [\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m, y \in \Omega \rangle]$ and $g_j = [\langle x_1, \ldots x_n, t_j \rangle]$.

(ii). If $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t \in \Omega}$ is a derived rule of U the the rule

$$\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(V_1, \ldots V_{n-1})}{\overline{f}(v_1, \ldots V_n) = \overline{h}(g_1(v_1, \ldots v_n), \ldots \overline{g_m}(v_1, \ldots v_n)) \in \overline{C}(v_1, \ldots v_n)}$$

is a derived rule of $U(\mathfrak{C}(U))$, where $f = [< x_1, \ldots x_n, t[t_1 \mid y_1, \ldots t_m \mid y_m] >]$ and $h = [< y_1, \ldots y_m, t >]$.

Lemma 7. If U is a theory then (i) for every derived T-rule $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$ of U the rule $\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{A}(x_1, \ldots x_m) = \dot{\varphi}(\Delta)}$

is a derived rule of $U(\mathfrak{C}(U))$, where $A_i = [< x_1 \in \Delta_1, \ldots x_i \in \Delta_i >]$ and $A = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]$. (ii) For every derived $\in$-rule $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ of U the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{[< x_1, \ldots x_n, t >]}(x_1, \ldots x_n) = \dot{\varphi}(t) \in \overline{A}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathfrak{C}(U))$.

Proof. The proof is by induction on derivations in U. We wish to show that all the derived T and $\in$-rules of U have a certain property so we just show that any rule that is derived from rules that have the propert must itself have the property. We must check the principles of derivation T1, CF1 and CF2.

T1. Suppose that we derive the rule $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta'}$ from the rules $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ and $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$, and suppose also that $\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ has the property, which is to say suppose that

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{[< x_1, \ldots x_n, t >]}(x_1, \ldots x_n) = \dot{\varphi}(t) \in \overline{A}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathfrak{C}(U))$. We wish to show that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta'}$

has the property i.e. that

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{[< x_1, \ldots x_n, t >]}(x_1, \ldots x_n) = \dot{\varphi}(t) \in \overline{A'}(x_1, \ldots x_n)}$$

is a derived rule

of $U(\mathfrak{C}(U))$, where $A' = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta' >]$. But

of course it is, because $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$ is a derived rule implies A=A'.

CF1. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a derived rule of U such that

$\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{A_{n+1}}(x_1, \ldots x_n) = \dot{\varphi}(\Delta_{n+1})}$ is a derived rule of $U(\mathfrak{C}(U))$. We

must show that for each i, $1 \leq i \leq n+1$, the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_{n+1} \in \overline{A_{n+1}}(x_1, \ldots x_n)}{\overline{[< x_1, \ldots x_{n+1}, x_i >]}(x_1, \ldots x_{n+1}) = x_i \in \overline{C_i}(x_1, \ldots x_{n+1})}$$

is a derived rule of $U(\mathfrak{C}(U))$, where $C_i = [< x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}, y \in \Delta_i >]$

This follows because $< x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1} >$ and $< x_1 \in \Delta_1, \ldots x_i \in \Delta_i >$

are contexts of U and $< x_1, \ldots x_i >$ is a realisation of $< x_1 \in \Delta_1, \ldots x_i \in \Delta_i >$

wrt $< x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1} >$, thus by corollary 4 the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_{n+1} \in \overline{A_{n+1}}(x_1, \ldots x_n)}{\overline{[< x_1, \ldots x_i >]}(x_1, \ldots x_{n+1}) = \overline{[< x_1, \ldots x_{n+1}, x_i >]}(x_1, \ldots x_{n+1}) \in \overline{C}(x_1, \ldots x_{n+1}}$$

is a derived rule of $U(\mathbb{C}(U))$; and because, since $p(A_{n+1},A_i)$ in $\mathbb{C}(U)$: just $\langle x_1,\ldots x_i\rangle$ (lemma 2 of § 2.3), $U(\mathbb{C}(U))$ has the axiom

$$\frac{x_1 \in \overline{A_1},\ldots x_{n+1} \in \overline{A_{n+1}}(x_1,\ldots x_n)}{\overline{[\langle x_1,\ldots x_i\rangle]}(x_1,\ldots x_n) = x_i \in \overline{A_i}(x_i,\ldots x_n)} \quad .$$

CF2(a). Suppose that B is a sort symbol of U introduced by

$$\frac{y_1 \in \Omega_1,\ldots y_m \in \Omega_m}{B(y_1,\ldots y_m) \text{ is a type.}}$$

Suppose that for each j, $1 \leq j \leq m$, the rule $\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{t_j \in \Omega_j[t_1 \mid y_1,\ldots t_{j-1} \mid y_{j-1}}$

is a derived rule of U with the property, i.e. such that the rule

$$\frac{x_1 \in \overline{A_1},\ldots x_n \in \overline{A_n}(x_1,\ldots x_{n-1})}{\overline{[\langle x_1,\ldots x_n,t_j\rangle]}(x_1,\ldots x_n) = \dot{\Phi}(t_j) \in \overline{Q_j}(x_1,\ldots x_n)}$$

is a derived rule of $U(\mathbb{C}(U))$ where $Q_j = [\langle x_1 \in \Delta_1,\ldots x_n \in \Delta_n, y_j \in \Omega_j[t_1 \mid y_1,\ldots t_{j-1} \mid y_{j-1}]\rangle]$. We wish to show the rule

$\dfrac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{B(t_1,\ldots t_m) \text{ is a type}}$ has the property, i.e that the rule

$\dfrac{x_1 \in \overline{A_1},\ldots x_n \in \overline{A_n}(x_1,\ldots x_{n-1})}{\overline{C}(x_1,\ldots x_n) = \dot{\Phi}(\mathcal{B}t_1,\ldots t_m))}$ is a derived rule of $U(\mathbb{C}(U))$, where

$C = [\langle x_1 \in \Delta_1,\ldots x_n \in \Delta_n, z \in B(t_1,\ldots t_m)\rangle]$.

Let $L = [\langle y_1 \in \Omega_1,\ldots y_m \in \Omega_m, y \in B(y_1,\ldots y_m)\rangle]$.

By corollary 6(i), $\dfrac{x_1 \in \overline{A_1},\ldots x_n \in \overline{A_n}(x_1,\ldots x_{n-1})}{\overline{Q_j}(x_1,\ldots x_n) = \overline{B_j}(g_i(x_1,\ldots x_n),\ldots \overline{g_{j-1}}(x_1,\ldots x_n))}$

is a derived rule of $U(\mathfrak{C}(U))$ for each j, $1 \leq j \leq m$, where

$B_j = \lceil\langle y_1 \in \Omega_1, \ldots y_j \in \Omega_j \rangle\rceil$, and $g_j = \lceil\langle x_1, \ldots x_n, t_j \rangle\rceil$.

Hence for each j, $1 \leq j \leq m$, the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{g_j}(x_1, \ldots x_n) = \dot{\varphi}(t_j) \in \overline{B_j}(\overline{g_1}(x_1, \ldots x_n), \ldots \overline{g_{j-1}}(x_1, \ldots x_n))}$$

is a derived rule of $U(\mathfrak{C}(U))$, and from the introductory rule for $\overline{L}$ in $U(\mathfrak{C}(U))$, which is $\dfrac{y_1 \in \overline{B_1}, \ldots y_m \in \overline{B_m}(y_1, \ldots y_{m-1})}{\overline{L}(y_1, \ldots y_m) \text{ is a type}}$

we get $$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{L}(\overline{g_1}(x_1, \ldots x_n), \ldots \overline{g_m}(x_1, \ldots)) = \overline{L}(\dot{\varphi}(t_1), \ldots \dot{\varphi}(t_m))}$$

as a derived rule of $U(\mathfrak{C}(U))$. But again by 6(i), the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{L}(\overline{g_1}(x_1, \ldots x_n), \ldots \overline{g_m}(x_1, \ldots x_n)) = \overline{C}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathfrak{C}(U))$. Thus the rule

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{C}(x_1, \ldots x_n) = \overline{L}(\dot{\varphi}(t_1), \ldots \dot{\varphi}(t_n))}$$ is a derived rule of $U(\mathfrak{C}(U))$,

which is just whats wanted since $\dot{\varphi}(C(t_1, \ldots t_m)) = \overline{L}(\dot{\varphi}(t_1), \ldots \dot{\varphi}(t_m))$.

CF2(b). Very similar to CF2(a), uses corollary 6(i) and (ii).

Corollary 8. For every theory U, φ_U is an interpretation of U in $U(\mathfrak{C}(U))$.

Proof. It suffices to show that for any derived rule R of U, $\hat{\Phi}(R)$ is a derived rule of $U(\mathbb{C}(U))$. We check for each of the four forms separately.

1. The T-rules. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}$ is a derived rule of U

then by definition $\hat{\Phi}\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1} \text{ is a type}}\right) = \dfrac{x_1 \in \dot{\Phi}(\Delta_1), \ldots x_n \in \dot{\Phi}(\Delta_n)}{\dot{\Phi}(\Delta_{n+1}) \text{ is a type}}$

By lemma 7, for each i, $1 \leq i \leq n+1$, the rule

$\dfrac{x_1 \in \overline{A_1}, \ldots x_{i-1} \in \overline{A_{i-1}}(x_1, \ldots x_{i-1})}{\overline{A_i}(x_1, \ldots x_{i-1}) = \dot{\Phi}(\Delta_i)}$ is a derived rule of

$U(\mathbb{C}(U))$. Hence for each i, $1 \leq i \leq n+1$, the rule

$\dfrac{x_1 \in \dot{\Phi}(\Delta_1), \ldots x_{i-1} \in \dot{\Phi}(\Delta_{i-1})}{\overline{A_i}(x_1, \ldots x_{i-1}) = \dot{\Phi}(\Delta_i)}$ is a derived rule of

$U(\mathbb{C}(U))$ (argue by induction). In particular $\dfrac{x_1 \in \dot{\Phi}(\Delta_1), \ldots x_n \in \dot{\Phi}(\Delta_n)}{\overline{A_{n+1}}(x_1, \ldots x_n) = \dot{\Phi}(\Delta_{n+1})}$

is a derived rule of $U(\mathbb{C}(U))$. Thus because of wellformedness of derived rules (see § 1.7) we must have $\dfrac{x_1 \in \dot{\Phi}(\Delta_1), \ldots x_n \in \dot{\Phi}(\Delta_n)}{\dot{\Phi}(\Delta_{n+1}) \text{ is a type}}$

as a derived rule of $U(\mathbb{C}(U))$.

2. The $\in$-rules. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ is a derived rule

of U. By wellformedness and by part 1. above,

$$\frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\dot{\varphi}(\Delta) = \overline{A}(x_1, \ldots x_n)} \qquad \text{and} \qquad \frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_{i-1} \in \dot{\varphi}(\Delta_{i-1})}{\dot{\varphi}(\Delta_i) = \overline{A_i}(x_i, \ldots x_{i-1})}$$

$1 \leq i \leq n$, are derived rules of $U(\mathfrak{C}(U))$.

By lemma 7, $$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{\lceil < x_1, \ldots x_n, t > \rceil}(x_1, \ldots x_n) = \bar{\varphi}(t) \in \overline{A}(x_1, \ldots x_n)}$$

is a derived rule of $U(\mathfrak{C}(U))$. Thus so is

$$\frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \bar{\varphi}(\Delta_n)}{\lceil < x_1, \ldots x_n, t > \rceil\ (x_1, \ldots x_n) = \dot{\varphi}(t) \in \dot{\varphi}(\Delta)}$$ a derived rule of $U(\mathfrak{C}(U))$.

Hence by wellformedness the rule $$\frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\dot{\varphi}(t) \in \dot{\varphi}(\Delta)}$$

is a derived rule of $U(\mathfrak{C}(U))$.

3. The T=rules. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$ is a derived rule of U then

$\dfrac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\dot{\varphi}(\Delta) = \dot{\varphi}(\Delta')}$ is a derived rule of $U(\mathfrak{C}(U))$ because by

lemma 7 and 1. above, the rules $$\frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\overline{A}(x_1, \ldots x_n) = \dot{\varphi}(\Delta)} \qquad \text{and}$$

$$\frac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\overline{A}'(x_1, \ldots x_n) = \dot{\varphi}(\Delta')}$$ are derived rules of $U(\mathfrak{C}(U))$, where of course $A = \lceil < x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta > \rceil = \lceil < x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x' \in \Delta' > \rceil = A'$.

4. Similarly if $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}$ is a derived rule of U then

$\dfrac{x_1 \in \dot{\varphi}(\Delta_1), \ldots x_n \in \dot{\varphi}(\Delta_n)}{\dot{\varphi}(t) = \dot{\varphi}(t') \in \dot{\varphi}(\Delta)}$ is a derived rule of $U(\mathbb{C}(U))$.

Recap: We are attempting to show that the functor $id_{\underline{GAT}} : \underline{GAT} \longrightarrow \underline{GAT}$ is isomorphic to the functor $U \circ \mathbb{C} : \underline{GAT} \longrightarrow \underline{GAT}$. So far we have defined an interpretation φ_U of U in $U(\mathbb{C}(U))$ for every theory U. Thus for every $U \in |\underline{GAT}|$, $[\varphi_U] : U \longrightarrow U(\mathbb{C}(U))$ is a morphism of $\underline{GAT}$. It remains to show that for every $U \in |\underline{GAT}|$, φ_U is an isomorphism and that $[\varphi_-]$ is a natural transformation, $[\varphi_-] : id_{\underline{GAT}} \longrightarrow U \circ \mathbb{C}$. It is understood that we write $[\varphi_-]$ for what otherwise might be written as $\lambda U \in |\underline{GAT}| . [\varphi_U]$.

Lemma 9. $[\varphi_-]$ is a natural transformation, $[\varphi_-] : id_{\underline{GAT}} \longrightarrow U \circ \mathbb{C}$.

Proof. We must show that whenever U and U' are theories and I is an interpretation of U in U' then the diagram

$$\begin{array}{ccc} U & \xrightarrow{[\varphi_U]} & U(\mathbb{C}(U)) \\ {\scriptstyle [I]}\downarrow & & \downarrow{\scriptstyle U(\mathbb{C}([I]))} \\ U' & \xrightarrow[[\varphi_{U'}]]{} & U(\mathbb{C}(U')) \end{array}$$

commutes in $\underline{GAT}$.

Suppose that we have such an I. By corollary 2 of §1.14 it suffices to show that for any derived $\in$-rule R of U,

$\widehat{U(\mathbb{C}([I]))}(\hat{\varphi}_U(R)) = \hat{\varphi}_{U'}(\hat{I}(R))$.

Suppose then that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ is a derived rule of U. By

lemma 7, $$\hat{\Phi}_U\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right) \equiv \frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{A}(x_1, \ldots x_n)}$$

where $A_i = [< x_1 \in \Delta_1, \ldots x_i \in \Delta_i >]$, $A = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]$

and $f = [< x_1, \ldots x_n, t >]$. Thus by definition of the functor U,

$$\widehat{U(\mathbb{C}([I]))}\left(\hat{\Phi}_U\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right)\right) \equiv$$

$$\frac{x_1 \in \overline{\mathbb{C}([I])(A_1)}, \ldots x_n \in \overline{\mathbb{C}([I])(A_n)}(x_1, \ldots x_{n-1})}{\overline{\mathbb{C}([I])(f)}(x_1, \ldots x_n) \in \overline{\mathbb{C}([I])(A)}(x_1, \ldots x_n)} .$$

On the otherhand, $$\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right) = \frac{x_1 \in \dot{I}(\Delta_1), \ldots x_n \in \dot{I}(\Delta}{\dot{I}(t) \in \dot{I}(\Delta)}$$

Thus by lemma 7, $$\hat{\Phi}_{U'}\left(\hat{I}\left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}\right)\right) \equiv$$

$$\frac{x_1 \in \overline{B_1}, \ldots x_n \in \overline{B_n}(x_1, \ldots x_{n-1})}{\overline{g}(x_1, \ldots x_n) \in \overline{B}(x_1, \ldots x_n)}$$ where

$B_i = [< x_1 \in \dot{I}(\Delta_1), \ldots x_i \in \dot{I}(\Delta_i) >]$, $B = [< x_1 \in \dot{I}(\Delta_1), \ldots$

$x_n \in \dot{I}(\Delta_n), x \in \dot{I}(\Delta) >]$ and $g = [< x_1, \ldots x_n, \dot{I}(t) >]$ But by definition of $\mathbb{C}([I])$, $\mathbb{C}([I])(A_i) = B_i$, $\mathbb{C}([I])(A) = B$ and $\mathbb{C}([I])(f) = g$. Hence $\widehat{U(\mathbb{C}([I]))}(\hat{\Phi}_U(R)) = \hat{\Phi}_{U'}(\hat{I}(R))$, as required.

It remains to show that for any generalised algebraic theory U, the morphism $[\Phi_U]$ of the category GAT is an isomorphism. It suffices to define an interpretation Ψ_U of $U(\mathbb{C}(U))$ in U and to show that $\Psi_U \circ \Phi_U = id_U$ and $\Phi_U \circ \Psi_U = id_{U(\mathbb{C}(U))}$.

Let U be a theory. We define a preinterpretation ψ_U of $U(\mathbb{C}(U))$ in U. The preinterpretation ψ_U is defined on sort symbols $\overline{A}$ of U by choosing an element $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta, v_{n+1} \in \Delta \rangle$ of the equivalence class A and by defining $\psi_U(\overline{A}) = \Delta$. To simplify matters we can make the choices in such a way that if $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}(U)$ and if $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n \rangle$ is chosen to represent A_n then a context of the form $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Delta \rangle$ is chosen to represent A. This is always possible by virtue of corollary 2(b) of § 2.2.

If $f: A_n \longrightarrow B$ in $\mathbb{C}(U)$ and B is non-trivial then $\psi_U(\overline{f})$ is defined by choosing an element $\langle t_1, \ldots t_m, t \rangle$ of the equivalence class f and by defining $\psi_U(\overline{f}) = t$. However we choose the representation $\langle t_1, \ldots t_m, t \rangle$ of f in such a way that if $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n \rangle$ represents A_n and if $\langle v_1 \in \Omega_1, \ldots v_m \in \Omega_m, v_{m+1} \in \Omega \rangle$ represents B then $\langle t_1, \ldots t_m, t \rangle$ is a realisation of $\langle v_1 \in \Omega_1, \ldots v_m \in \Omega_m, v_{m+1} \in \Omega \rangle$ wrt $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta$
This is possible by lemma 4(i) of §1.13.
Moreover to simplify matters the choices are made in such a way that if $f: A_n \longrightarrow B$ and if $f \circ p(B)$ (assuming it is non-trivial) is represented by $\langle t_1, \ldots t_m \rangle$ then f is represented by $\langle t_1, \ldots t_m, t \rangle$, for some t.
This is possible by lemma 5 of §1.13.

Lemma 10. ψ_U is an interpretation of $U(\mathbb{C}(U))$ in U.

Proof. We must check that all the introductory rules and axioms of $U(\mathbb{C}(U))$ are mapped by ψ_U to derived rules of U. We just check two cases the other cases are just as simple to check.

1. Suppose $\overline{A}$ is a sort symbol of $U(\mathbb{C}(U))$. Say $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}(U)$, so that $\overline{A}$ has the introductory rule $\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{A}(x_1, \ldots x_n) \text{ is a type}}$

Suppose that A has been represented by $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Delta \rangle$, then $\hat{\psi}_U \left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{A}(x_1, \ldots x_n) \text{ is a type}} \right) =$

$$\frac{x_1 \in \Delta_1, x_2 \in \Delta_2[x_1 | v_1], \ldots x_n \in \Delta_n[x_1 | v_1, \ldots x_{n-1} | v_{n-1}]}{\Delta[x_1 | v_1, \ldots x_n | v_n] \text{ is a type}}$$

and this rule is a derived rule of U by the change of variable lemma of §1.7 because the rule $\dfrac{v_1 \in \Delta_1, \ldots v_n \in \Delta_n}{\Delta \text{ is a type}}$ is a derived rule of U.

2. If $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m \lhd B$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}(U)$ so that $U(\mathbb{C}(U))$ has the axiom R, where R =

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f^*B}(x_1, \ldots x_n) = \overline{B}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f \circ p(B_m, B_{m-1})}(x_1, \ldots x_n), \overline{f}(x_1, \ldots x_n)}$$

Suppose that A_n has been represented by $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n \rangle$, B has been represented by $\langle v_1 \in \Omega_1, \ldots v_m \in \Omega_m, v_{m+1} \in \Omega \rangle$, f^*B has been represented by $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Delta \rangle$ and f has been represented by $\langle t_1, \ldots t_m \rangle$. Then $\psi(R) =$

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n[x_1 | v_1, \ldots x_{n-1} | v_{n-1}]}{\Delta[x_1 | v_1, \ldots x_n | v_n] = \Omega[t_1[x_1 | v_1, \ldots x_n | v_n] | v_1, \ldots t_m[x_1 | v_1, \ldots x_n | v_n] | v_m]}$$

$$= \frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n [\, x_1 \mid v_1, \ldots x_{n-1} \mid v_{n-1}]}{\Delta [\, x_1 \mid v_1, \ldots x_n \mid v_n] = \Omega [t_1 \mid v_1, \ldots t_m \mid v_m] [\, x_1 \mid v_1, \ldots x_n \mid v_n]}$$

which is a derived rule of U by the changeof variable lemma since

$\dfrac{v_1 \in \Delta_1, \ldots v_n \in \Delta_n}{\Delta = \Omega [\, t_1 \mid v_1, \ldots t_m \mid v_m]}$ is a derived rule of U because

$[< v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Delta >] = f^{*}B = [< v_1 \in \Delta_1, \ldots v_n \in \Delta_n,$
$v_{n+1} \in \Omega [\, t_1 \mid v_1, \ldots t_m \mid v_m] >]$.

Lemma. $\psi_U \circ \varphi_U \equiv id_U$.

Proof. Use corollary 2 of §1.14. Suppose that $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$

is a derived rule of U. Let $A_i = [< x_1 \in \Delta_1, \ldots x_i \in \Delta_i >]$ and let $A = [< x_1 \in \Delta_1, \ldots x_n \in \Delta_n, x \in \Delta >]$. By lemma 7 of this section,

$$\hat{\varphi}_U \left(\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \right) \equiv \frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{[< x_1, \ldots x_n, t >]}(x_1, \ldots x_n) \in \overline{A}(x_1, \ldots x_n)} .$$

Therefore $\hat{\psi}_U \left(\hat{\varphi}_U \left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta} \right) \right) \equiv$

$$\hat{\psi}_U \left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{[< x_1, \ldots x_n, t >]}(x_1, \ldots x_n) \in \overline{A}(x_1, \ldots x_n)} \right)$$

and it follows from the definition of ψ_U that this rule is equivalent to

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$.

Lemma. $\varphi_U \circ \psi_U \equiv id_{U(\mathfrak{C}(U))}$.

Proof. Suppose that $1 \lhd A_1 \ldots A_n \lhd A$ in $\mathbb{C}(U)$, so that $\overline{A}$ is a sort symbol of $U(\mathbb{C}(U))$ introduced by the rule

$$\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{A}(v_1, \ldots v_n) \text{ is a type}}$$

If A has been represented by $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Delta \rangle$ then

$$\hat{\phi}_u\left(\hat{\psi}_u\left(\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{A}(v_1, \ldots v_n) \text{ is a type}}\right)\right) = \hat{\phi}_u\left(\frac{v_1 \in \Delta_1, \ldots v_n \in \Delta_n}{\Delta \text{ is a type}}\right) \equiv$$

$$\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{A}(v_1, \ldots v_n) \text{ is a type}} \quad \text{, by lemma 7.}$$

If also $1 \lhd B_1 \ldots \lhd B_m$ in $\mathbb{C}(U)$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}(U)$. If B_m has been represented by $\langle v_1 \in \Omega_1, \ldots v_m \in \Omega_m \rangle$, if f has been represented by $\langle t_1, \ldots t_m \rangle$ and if $(f \circ p(B_m)) * B_m$ has been represented by $\langle v_1 \in \Delta_1, \ldots v_n \in \Delta_n, v_{n+1} \in \Omega \rangle$ then

$$\hat{\phi}_u\left(\hat{\psi}_u\left(\frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{f}(v_1, \ldots v_n) \in \overline{(f \circ p(B_m)) * B_m}(v_1, \ldots v_n)}\right)\right) = \hat{\phi}_u\left(\frac{v_1 \in \Delta_1, \ldots v_n \in \Delta_r}{t_m \in \Omega}\right.$$

$$\equiv \frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{[\langle v_1, \ldots v_n, t_m \rangle]}(v_1, \ldots v_n) \in \overline{(f \circ p(B_m)) * B_m}(v_1, \ldots v_n)}$$

by lemma 7 of this section,

$$\equiv \frac{v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1})}{\overline{[\langle t_1, \ldots t_m \rangle]}(v_1, \ldots v_n) \in \overline{f \circ p(B_m)) * B_m}(v_1, \ldots v_n)}$$

by corollary 4 of this section.

This completes the proof that $\phi_u \circ \psi_u$ and $id_{U(\mathbb{C}(U))}$ agree up to equivalence on the introductory rules of $U(\mathbb{C}(U))$.

2.4.4 The proof that $\mathbb{C}oU \simeq id_{\underline{Con}}$. We define a natural transformation $\eta: id_{\underline{Con}} \longrightarrow \mathbb{C}oU$. That is we define for each contextual category $\mathbb{C}$, a contextual functor $\eta_{\mathbb{C}}: \mathbb{C} \longrightarrow \mathbb{C}(U(\mathbb{C}))$, such that if $F: \mathbb{C} \longrightarrow \mathbb{C}'$ is a contextual functor then the diagram

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{\eta_{\mathbb{C}}} & \mathbb{C}(U(\mathbb{C})) \\ {\scriptstyle F}\downarrow & & \downarrow{\scriptstyle \mathbb{C}(U(F))} \\ \mathbb{C}' & \xrightarrow[\eta_{\mathbb{C}'}]{} & \mathbb{C}(U(\mathbb{C}')) \end{array}$$

commutes in $\underline{Con}$. Eventually we show that η is an isomorphism, that is that for each contextual category $\mathbb{C}$, $\eta_{\mathbb{C}}$ is an isomorphism. Then η is the required isomorphism between $id_{\underline{Con}}$ and $\mathbb{C}oU$.

If $\mathbb{C}$ is a contextual category then $\eta_{\mathbb{C}}$ is defined on the trivial objects and on the trivial morphisms of $\mathbb{C}$ in the trivial manner, that is by $\eta_{\mathbb{C}}(1) = 1$ and $\eta_{\mathbb{C}}(p(A,1)) = p(\eta_{\mathbb{C}}(A),1)$. $\eta_{\mathbb{C}}$ is defined on the non-trivial objects and morphisms of $\mathbb{C}$ as follows:

If $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$ then $\eta_{\mathbb{C}}(A) = [\langle v_1 \in \overline{A_1}, \ldots v_n \in \overline{A_n}(v_1, \ldots v_{n-1}), v_{n+1} \in \overline{A}(v_1, \ldots v_n) \rangle]$.

If $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}$ then

$$\eta_{\mathbb{C}}(f) = [\langle \overline{f \circ p(B_m, B_1)}(v_1, \ldots v_n), \ldots \overline{f \circ p(B_m, B_{m-1})}(v_1, \ldots v_n), \overline{f}(v_1, \ldots v_n) \rangle$$

Lemma. If $\mathbb{C}$ is a contextual category then $\eta_{\mathbb{C}}: \mathbb{C} \longrightarrow \mathbb{C}(U(\mathbb{C}))$ is a contextual functor.

Proof. (i). $\eta_{\mathbb{C}}$ preserves identity morphisms because if $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ then $\eta_{\mathbb{C}}(id_A) =$

$[\langle \overline{p(A_n, A_1)}(v_1, \ldots v_n), \ldots \overline{p(A_n, A_{n-1})}(v_1, \ldots v_n), \overline{id_{A_n}}(v_1, \ldots v_n) \rangle]$, by

def. of $\eta_{\mathbb{C}}$ = $[\langle v_1,\ldots v_n\rangle]$, because of axioms of $U(\mathbb{C})$ to that effect,

= $^{id}\ [\langle v_1 \in \overline{A_1},\ldots v_n \in \overline{A_n}(v_1,\ldots v_{n-1})\rangle]$.

(ii). $\eta_{\mathbb{C}}$ preserves composition because if $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$, $1 \lhd C_1 \ldots \lhd C_\ell$ and $f: A_n \longrightarrow B_m$, $g: B_m \longrightarrow C_\ell$ in $\mathbb{C}$ then $\eta(f \circ g) =$

$[\langle \overline{f \circ g \circ p(C_\ell, C_1)}(v_1,\ldots v_n),\ldots \overline{f \circ g}(v_1,\ldots v_n)\rangle]$, by def. of $\eta_{\mathbb{C}}$.

= $[\langle \overline{g \circ p(C_\ell, C_1)}(\overline{f \circ p(B_m, B_1)}(v_1,\ldots v_n),\ldots \bar{f}(v_1,\ldots v_n)),\ldots$

$\bar{g}(\overline{f \circ p(B_m, B_1)}(v_1,\ldots v_n),\ldots \bar{f}(v_1,\ldots v_n))\rangle]$, because of axioms to that effect in $U(\mathbb{C})$.

= $[\langle \overline{f \circ p(B_m, B_1)}(v_1,\ldots v_n),\ldots \bar{f}(v_1,\ldots v_n)\rangle] \circ [\langle \overline{g \circ p(C_\ell, C_1)}(v_1,\ldots v_m),\ldots$

$\bar{g}(v_1,\ldots v_m)\rangle]$ = $\eta_{\mathbb{C}}(f) \circ \eta_{\mathbb{C}}(g)$, as required.

(iii). If $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ then $\eta_{\mathbb{C}}(p(A_n))$ =

$[\langle \overline{p(A_n, A_1)}(v_1,\ldots v_n),\ldots \overline{p(A_n, A_{n-1})}(v_1,\ldots v_n)\rangle]$ = $[\langle v_1,\ldots v_{n-1}\rangle]$

$p(\eta_{\mathbb{C}}(A_n))$.

(iv). If $f: A_n \longrightarrow B_m$ in $\mathbb{C}$ where $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m \lhd B$ in $\mathbb{C}$ then $\eta_{\mathbb{C}}(f^*B)$ = $[\langle v_1 \in \overline{A_1},\ldots v_n \in \overline{A_n}(v_1,\ldots v_{n-1}), v_{n+1} \in \overline{f^*B}(v_1,\ldots v_n)\rangle]$

by def. of $\eta_{\mathbb{C}}$.

= $[\langle v_1 \in \overline{A_1},\ldots v_n \in \overline{A_n}(v_1,\ldots v_{n-1}), v_{n+1} \in \overline{B}(\overline{f \circ p(B_m, B_1)}(v_1,\ldots v_n),\ldots$

$\bar{f}(v_1,\ldots v_n))\rangle]$, because there is an axiom to that effect in $U(\mathbb{C})$.

= $[\langle \overline{f \circ p(B_m, B_1)}(v_1,\ldots v_n),\ldots \bar{f}(v_1,\ldots v_n)\rangle] * [\langle v_1 \in \overline{B_1},\ldots v_m \in \overline{B_m}(v_1,\ldots$

$v_{m-1}), v_{m+1} \in \overline{B}(v_1,\ldots v_m)\rangle]$, by def. of $\mathbb{C}(U(\mathbb{C}))$.

= $\eta_{\mathbb{C}}(f) * \eta_{\mathbb{C}}(B)$, As required.

Also in this situation, $\eta_{\mathfrak{C}}(q(f,B)) = [< \overline{q(f,B)\circ p(B,B_1)}(V_1,\ldots V_{n+1}),.. \overline{q(f,B)}(V_1,\ldots V_{n+1}) >] = [< \overline{p(f^*B)\circ f\circ p(B_m,B_1)}(V_1,\ldots V_{n+1}),\ldots \overline{p(f^*B)\circ f}(V_1,\ldots V_{n+1}),\overline{q(f,B)}(V_1,\ldots V_{n+1}) >] = [< \overline{f\circ p(B_m,B_1)}(V_1,\ldots V_n), \ldots \bar{f}(V_1,\ldots V_n),V_{n+1}>] = q(\eta_{\mathfrak{C}}(f),\eta_{\mathfrak{C}}(B))$.

Lemma. If $F:\mathfrak{C}\longrightarrow\mathfrak{C}'$ in Con then the diagram

$$\begin{array}{ccc} \mathfrak{C} & \xrightarrow{\eta_{\mathfrak{C}}} & \mathfrak{C}(U(\mathfrak{C})) \\ F\downarrow & & \downarrow \mathfrak{C}(U(F)) \\ \mathfrak{C}' & \xrightarrow[\eta_{\mathfrak{C}'}]{} & \mathfrak{C}(U(\mathfrak{C}')) \end{array}$$

commutes.

Proof. If $f:A_n\longrightarrow B_m$, where $1\lhd A_1\ldots\lhd A_n, 1\lhd B_1\ldots\lhd B_m$ in $\mathfrak{C}$, then $\eta_{\mathfrak{C}}(f) = [< \overline{f\circ p(B_m,B_1)}(V_1,\ldots V_n),\ldots \bar{f}(V_1,\ldots V_n) >]$. Thus $\mathfrak{C}(U(F))(\eta_{\mathfrak{C}}(f)) =$ $[< \overline{F(f\circ p(B_m,B_1)}(V_1,\ldots V_n),\ldots \overline{F(f)}(V_1,\ldots V_n) >] = [< \overline{F(f)\circ p(F(B_m)}, \overline{F(B_1))}(V_1,\ldots V_n),\ldots \overline{F(f)}(V_1,\ldots V_n) >] = \eta_{\mathfrak{C}'}(F(f))$.

So we have a natural transformation $\eta: id_{\underline{Con}}\longrightarrow \mathfrak{C}\circ U$. We wish to show that for each $\mathfrak{C}$, $\eta_{\mathfrak{C}}$ is an isomorphism in Con. Unfortunately this turns out to be rather tricky. We have to define a contextual functor $\xi_{\mathfrak{C}}: \mathfrak{C}(U(\mathfrak{C}))\longrightarrow\mathfrak{C}$ and show that $\xi_{\mathfrak{C}}\circ\eta_{\mathfrak{C}} = id_{\mathfrak{C}}$ and $\eta_{\mathfrak{C}}\circ\xi_{\mathfrak{C}} = id_{\mathfrak{C}(U(\mathfrak{C}))}$. The procedure that we adopt in defining $\xi_{\mathfrak{C}}$ is to define a function J from derived T and $\in$-rules of $U(\mathfrak{C})$ to objects respectively morphisms of $\mathfrak{C}$. We show that J induces an equivalence preserving map from contexts and realisations of $U(\mathfrak{C})$ to

objects respectively morphisms of $\mathbb{C}$. Thus we get a map from objects and morphisms of $\mathbb{C}(U(\mathbb{C}))$ to objects respectively morphisms of $\mathbb{C}$, we show that this map is a left and right inverse to $\eta_{\mathbb{C}}$.

Initially J is defined just as a partial function from the derived T and $\in$-rules of $U(\mathbb{C})$ to the objects respectively morphisms of $\mathbb{C}$, though eventually we show that J is total.

Consider the forms that the derived T and $\in$-rules of $U(\mathbb{C})$ can take. By the derivation lemma of §1.7 every derived T-rule of $U(\mathbb{C})$ is of the form $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{A}(t_1, \ldots t_n) \text{ is a type}}$ for some object A of $\mathbb{C}$ such that $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$ and such that for each $i, 1 \leqslant i \leqslant n$, $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{t_i \in \overline{A_i}(t_1, \ldots t_{i-1})}$ is a derived rule of $U(\mathbb{C})$. By the same lemma any derived $\in$-rule of $U(\mathbb{C})$ is either of the form $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{x_j \in \Lambda}$ or else is of the form $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{f}(t_1, \ldots t_m) \in \Delta}$ for some morphism $f: A_n \longrightarrow B$ where $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$, such that for each $i, 1 \leqslant i \leqslant n$, the rule $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{t_i \in \overline{A_i}(t_1, \ldots t_{i-1})}$ is a derived rule of $U(\mathbb{C})$ and such that $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{B}(t_1, \ldots t_m) = \Delta}$ is a derived rule of $U(\mathbb{C})$.

Bearing this in mind the function J from derived T and ϵ-rules of $U(\mathbb{C})$ to the objects and morphisms of $\mathbb{C}$ is defined inductively as follows:

$$J\left(\frac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{A}(t_1, \ldots t_n) \text{ is a type}}\right) = J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Delta_m}), 1)^* A$$

$$J\left(\frac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{f}(t_1, \ldots t_n) \in \Delta}\right) = J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Delta_m}), 1)^* \text{'f'}.$$

$$J\left(\frac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{x_j \in \Lambda}\right) = \text{'}p(J(R_{\Delta_m}), J(R_{\Delta_j}))\text{'}.$$ Where for each $i, 1 \leq i \leq n$, R_{t_i} is the derived rule $\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{t_i \in \overline{A_i}(t_1, \ldots t_{i-1})}$

and where for each $j, 1 \leq j \leq m$, R_{Δ_j} is the derived rule

$$\frac{x_1 \in \Delta_1, \ldots x_{j-1} \in \Delta_{j-1}}{\Delta_j \text{ is a type}}.$$

If R is a derived rule of $U(\mathbb{C})$ and if J(R) is defined then J(R) is an object of $\mathbb{C}$. If R is a derived ϵ-rule of $U(\mathbb{C})$ and if J(R) is defined then J(R) is a morphism of $\mathbb{C}$.

$J\left(\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{A}(t_1, \ldots t_n) \text{ is a type}}\right)$ can fail to be defined either because one or more of the $J(R_{t_i})$'s is not defined or else because the composite is not defined in $\mathbb{C}$. For example, $J\left(\dfrac{x_1 \in \Delta_1, \ldots x_m \in \Delta_m}{\overline{A}(t_1, \ldots t_n) \text{ is a type}}\right)$ will certainly not be defined unless $J(R_{t_1})$ is a morphism of $\mathbb{C}$ whose codomai

occurs lower in the tree structure of $\mathbb{C}$ than the object $p(J(R_{\Delta_m}),1)^*A$, for otherwise $J(R_{t_1})^* p(J(R_{\Delta_m}),1)^*A$ will not be defined and thus $J(R_{t_n})^*\ldots J(R_{t_1})^*p(J(R_{\Delta_m}),1)^*A$ will not be defined.

We wish to show that J(R) is defined for all derived T and $\in$ -rules R of $U(\mathbb{C})$. This is going to require a proof by induction on derivations in $U(\mathbb{C})$. It turns out that the inductive hypothesis that we must use is rather complicated. This is because as we proceed to prove that J is defined on the derived rules R of $U(\mathbb{C})$ we must keep an eye on the behaviour of J on substitution instances of R, otherwise the induction does not go through. If we call the inductive hypothesis H then H is a possible property of derived rules of $U(\mathbb{C})$. That is for any derived rule R of U either H(R) is or is not the case. Of course we aim to show that H(R) is always the case.

H is defined inductively. The definition of H(R) depends on which kind of rule R is, thus there are four cases to consider:

Case1. T-rules. If R_Δ is a derived T-rule of $U(\mathbb{C})$ of the form $\dfrac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{\Delta \text{ is a type}}$ then $H(R_\Delta)$ is equivalent to 1(a) and 1(b) and 1(c), which are as follows:

1(a). If $n\geqslant 1$, that is if the premise of R_Δ is not the empty premise, then $H(R_{\Delta_n})$, where R_{Δ_n} is the rule $\dfrac{x_1\in\Delta_1,\ldots x_{n-1}\in\Delta_{n-1}}{\Delta_n \text{ is a type}}$

1(b). $J(R_\Delta)$ is defined and if $n\geqslant 1$ the $J(R_{\Delta_n}) \lhd J(R_\Delta)$ in $\mathbb{C}$. If $n = o$ then $1 \lhd J(R_\Delta)$ in $\mathbb{C}$.

1(c). $J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n]\ \text{ is a type}}\right)$

is defined and is equal to $J(R_{t_n})^*\ldots J(R_{t_1})^*p(J(R_{\Omega_m}),1)^*J(R_\Delta)$, where $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m\rangle$ is a context of $U(\mathbb{C})$ and whenever $\langle t_1,\ldots t_n\rangle$ a realisation of $\langle x_1\in\Delta_1,\ldots x_n\in\Delta_n\rangle$ wrt $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m\rangle$ with the property that for each $i, 1\leqslant i\leqslant n$,

$$J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]\ \text{ is a type}}\right)$$

is defined and $J(R_{t_i})$ is defined and $J(R_{t_i})\in$

$$\mathrm{Arr}_{\mathbb{C}}\left(J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]\ \text{ is a type}}\right)\right).$$

Where R_{t_i} is the rule $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{t_i\in\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]}$ and R_{Ω_m} is the rule $\dfrac{y_1\in\Omega_1,\ldots y_{m-1}\in\Omega_{m-1}}{\Omega_m\ \text{is a type}}$. (In future we use the convention that R_t denotes a rule of the form $\dfrac{-\in-,\ldots-\in-}{t\in-}$.

It will then usually be quite clear which rule R_t denotes and we will not have to mention it. Similarly, unless we say otherwise R_{Δ_i} denotes the rule $\dfrac{x_1\in\Delta_1,\ldots x_{i-1}\in\Delta_{i-1}}{\Delta_i\ \text{is a type}}$.)

Case 2. $\in$-rules. If R_t is a derived $\in$-rule of $U(\mathbb{C})$ of the form $\dfrac{x_1\in\Delta_1,\ldots x_n\in\Delta_n}{t\in\Delta}$ then $H(R_t)$ is equivalent to 2(a) and 2(b) and 2(c) which are as follows:

2(a). $H(R_\Delta)$. (it is hoped that it is understood that R_Δ is the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$!!)

2(b). $J(R_t)$, is defined and $J(R_t) \in Arr_{\mathbb{C}}(J(R_\Delta))$.

2(c). $$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right)$$

is defined and is equal to $J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_t)$, whenever $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ is a context of $U(\mathbb{C})$ and whenever $\langle t_1, \ldots t_n \rangle$ is a realisation of $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ wrt $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ with the property that for all $i, 1 \leq i \leq n$,

$$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}] \text{ is a type}}\right)$$

is defined and $J(R_{t_i})$ is defined and $J(R_{t_i}) \in$

$$Arr_{\mathbb{C}}\left(J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}] \text{ is a type}}\right)\right).$$

Case 3. T=rules. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$ is a derived rule of $U(\mathbb{C})$

then $H\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}\right)$ is equivalent to $H(R_\Delta)$ and $H(R_{\Delta'})$

and $J(R_\Delta) = J(R_{\Delta'})$.

Case 4. $\in$=rules. If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}$ is a derived rule of $U(\mathbb{C})$

then $H\left(\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = t' \in \Delta}\right)$ is equivalent to $H(R_t)$ and $H(R_{t'})$

and $J(R_t) = J(R_{t'})$.

This completes the definition of the inductive hypothesis H. We still need two lemmas before we can proceed with the induction.

Lemma 11. If R is a derived rule of $U(\mathfrak{C})$ of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\text{Conclusion}}$ such that H(R), If $\langle t_1, \ldots t_n \rangle$ is a realisation of $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ wrt $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ such that for each i, $1 \leq i \leq n$, $H(R_{t_i})$ then also

$$H\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\text{Conclusion}\,[\,t_1 \mid x_1, \ldots t_n \mid x_n\,]}\right)$$

Proof. By induction on n. We suppose that the result holds for all rules R' of the form $\dfrac{x'_1 \in \Delta'_1, \ldots x'_{n'} \in \Delta'_{n'}}{\text{Conclusion}'}$ for $n' < n$.

We show that the result holds of R, we must treat each of the four kinds of rules separately.

Case 1. R is a T-rule, say $R = R_\Delta = \dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta \text{ is a type}}$. We must show that H holds of the rule $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[\,t_1 \mid x_1, \ldots t_n \mid x_n\,] \text{ is a type}}$

that is we must show that 1(a), 1(b) and 1(c) hold of this rule.

1(a). $H(R_{\Omega_m})$ is the case because $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_1 \in \Delta_1}\right)$ is the case.

1(b). Since for each i,1 ⩽ i ⩽ n $H(R_{t_i})$, so for each i,1 ⩽ i ⩽ n, $J(R_{t_i})$ is defined and $J(R_{t_i}) \in Arr_{\mathbb{C}}\, \bar{J}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}] \text{ is a type}}\right)$

Since $H(R_\Delta)$, so R_Δ has property 1(c). Thus

$\bar{J}\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$ is defined and is equal to

$J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_\Delta)$. We wish to show that $J(R_{\Omega_m}) \lhd$

$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$ thus we must show that

$J(R_{\Omega_m}) \lhd J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_\Delta)$.

$H(R_\Delta)$ implies that for each i,1⩽ i⩽ n, $H(R_{\Delta_i})$. Thus, just as above, for each i,1⩽ i⩽ n, $J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}] \text{ is a type}}\right)$

is defined and is equal to $J(R_{t_{i-1}})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_{\Delta_i})$.

The situation in $\mathbb{C}$ then is that $1 \lhd J(R_{\Delta_1}) \ldots \lhd J(R_{\Delta_n}) \lhd J(R_\Delta)$ and for each i,1⩽ i⩽ n, $J(R_{t_i}) \in Arr_{\mathbb{C}}(J(R_{t_{i-1}})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_{\Delta_i}))$. Hence $J(R_{\Omega_m}) \lhd J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_\Delta)$ in $\mathbb{C}$. The situation is as in lemma 3 of §2.3.

1(c). Suppose that $\langle S_1, \ldots S_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in \Omega_m \rangle$ wrt $\langle z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell \rangle$ and suppose that for each j,1⩽ j⩽ m, $J(R_{S_j})$ is defined and $J(R_{S_j}) \in$

$Arr_{\mathbb{C}}\left(J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Omega_j[S_1 \mid y_1, \ldots S_{j-1} \mid y_{j-1}] \text{ is a type}}\right)\right)$.

We must show that $J\left(\dfrac{z_1\in\Lambda_1,\ldots z_\ell\in\Lambda_\ell}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n][S_1\mid y_1,\ldots S_m\mid y_m]}\right)$

is defined and is equal to $J(R_{S_m})^*\ldots J(R_{S_1})^*p(J(R_{\Lambda_\ell}),1)^*$

$J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n]\quad\text{is a type}}\right)$.

By the induction we can assume the corresponding result for the rule

$\dfrac{x_1\in\Delta_1,\ldots x_{i-1}\in\Delta_{i-1}}{\Delta_i\text{ is a type}}$, whenever $i\leq n$. That is we can assume that fo

each $i, 1\leq i\leq n$, $J\left(\dfrac{z_1\in\Lambda_1,\ldots z_\ell\in\Lambda_\ell}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}][S_1\mid y_1,\ldots S_m\mid y_m]\text{is}}\right)$

is defined and is equal to $J(R_{S_m})^*\ldots J(R_{S_1})^*p(J(R_{\Delta_i}),1)^*$

$J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]\quad\text{is a type}}\right)$. Call this the inductive

assumption.

Since for each $i, 1\leq i\leq n$, $H(R_{t_i})$, so for each $i, 1\leq i\leq n$,

$J\left(\dfrac{z_1\in\Lambda_1,\ldots z_\ell\in\Lambda_\ell}{t_i[S_1\mid y_1,\ldots S_m\mid y_m]\in\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}][S_1\mid y_1,\ldots S_m\mid}\right)$

is defined and is equal to $J(R_{S_m})^*\ldots J(R_{S_1})^*p(J(R_{\Lambda_\ell}),1)^*J(R_{t_i})$. Since

$J(R_{t_i})\in Arr_{\mathbb{C}}J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]\text{ is a type}}\right)$, so

$J\left(\dfrac{z_1\in\Lambda_1,\ldots z_\ell\in\Lambda_\ell}{t_i[S_1\mid y_1,\ldots S_m\mid y_m]\in\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}][S_1\mid y_1,\ldots S_m\mid}\right)$

$\in Arr_{\mathbb{C}}(J(R_S)^*\ldots J(R_S)^*p(J(R\quad),1)^*J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_i[t_1\mid x_1,\ldots t_{i-1}\mid x_{i-1}]i}\right.$

Now we use the inductive assumption to get, for each $i, 1 \leq i \leq n$,

$$J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{t_i[S_1 \mid y_1, \ldots S_m \mid y_m] \in \Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}][S_1 \mid y_1, \ldots S_m \mid y}\right.$$

$$\left. \in Arr_{\mathbb{C}}\left(J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}][S_1 \mid y_1, \ldots S_m \mid y_m]}\right)\right)\right).$$

Let $\sigma_i = t_i[S_1 \mid y_1, \ldots S_m \mid y_m]$. We have shown that $\langle \sigma_1, \ldots \sigma_n \rangle$ i
a realisation of $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ wrt $\langle z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell \rangle$
such that for each $i, 1 \leq i \leq n$, $J(R_{\sigma_i})$ is defined and $J(R_{\sigma_i}) \in$

$$Arr_{\mathbb{C}}\left(J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta_i[\sigma_1 \mid x_1, \ldots \sigma_{i-1} \mid x_{i-1}] \text{ is a type}}\right)\right).$$

Since $H(R_\Delta)$ we can conclude that $J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \text{ is a type}}\right)$
is defined and is equal to $J(R_{\sigma_n}) * \ldots J(R_{\sigma_1}) * p(J(R_{\Lambda_\ell}), 1) * J(R_\Delta)$.

Since $H(R_{t_i})$, so $J(R_{\sigma_i}) = J(R_{S_m}) * \ldots J(R_{S_1}) * p(J(R_{\Lambda_\ell}), 1) * J(R_{t_i})$.

Hence $J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \text{ is a type}}\right) = (J(R_{S_m}) * \ldots J(R_{S_1}) *$

$p(J(R_{\Lambda_\ell}), 1) * J(R_{t_n})) * \ldots (J(R_{S_m}) * \ldots J(R_{S_1}) * p(J(R_{\Lambda_\ell}), 1) J(R_{t_1}) * p(J(R_{\Lambda_\ell}), 1)$

$J(R_\Delta) = J(R_{S_m}) * \ldots J(R_{S_1}) * p(J(R_{\Lambda_\ell}), 1) * J(R_{t_n}) * \ldots J(R_{t_1}) * p(J(R_{\Omega_m}), 1) * J(R_\Delta)$

by lemma 4(iii) of §2.3. Thus $J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n][S_1 \mid y_1, \ldots S_m \mid}\right.$

$$= J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \text{ is a type}}\right) =$$

$J(R_{S_m}) * \ldots J(R_{S_1}) * p(J(R_{\Lambda_\ell}), 1) * J(R_{t_n}) * \ldots J(R_{t_1}) * p(J(R_{\Omega_m}), 1) * J(R_\Delta) =$

$J(R_{S_m}) * \ldots J(R_{S_1}) * p(J(R_{\Lambda_\ell}), 1) *$ $J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$,

as required, since $H(R_\Delta)$.

Case 2. R is an $\in$-rule, say $R = R_t = \dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$

We assume $H(R_t)$ and we must show that

$$H\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right),$$

that is we must show that 2(a), 2(b) and 2(c) hold of the rule

$$\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}.$$

2(a). It follows from Case 1 that $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right.$

2(b). Because $H(R_t)$ it follows that

$$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right)$$

is defined and is equal to $J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}), 1)^* J(R_J)$ because $H(R_t)$ it follows that $J(R_t) \in \mathrm{Arr}_{\mathbb{C}}(J(R_\Delta))$. Hence

$$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right) \in$$

$\mathrm{Arr}_{\mathbb{C}}(J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}), 1)^* J(R_\Delta))$. But since $H(R_\Delta)$,

$J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}), 1)^* J(R_\Delta) = J\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a typ}}\right.$

thus $J\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right) \in$

$$Arr_{\mathbb{C}}\left(J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)\right) \quad \text{as required.}$$

2(c). Suppose that $\langle S_1, \ldots S_m \rangle$ is a realisation of $\langle y_1 \in \Omega_1, \ldots y_m \in$ wrt $\langle z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell \rangle$ such that for each $j, 1 \leq j \leq m$, $J(R_{S_j})$ is defined and $J(R_{S_j}) \in Arr_{\mathbb{C}}\left(J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Omega_j[S_1 \mid y_1, \ldots S_{j-1} \mid y_{j-1}] \text{ is a type}}\right)\right)$

If we let $\sigma_i = t_1[S_1 \mid y_1, \ldots S_m \mid y_m]$ for each $i, 1 \leq i \leq n$, then $t[t_1 \mid x_1, \ldots t_n \mid x_n][S_1 \mid y_1, \ldots S_m \mid y_m] = t[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n]$ and $\Delta[t_1 \mid x_1, \ldots t_n \mid x_n][S_1 \mid y_1, \ldots S_m \mid y_m] = \Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n]$.

Thus we just have to show that $J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{t[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \in \Delta[\sigma_1 \mid x_1, \ldots \sigma$

is defined and is equal to $J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_\ell}), 1)^*$

$$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right) .$$

As in case 1, $\langle \sigma_1, \ldots \sigma_n \rangle$ is a realisation of $\langle x_1 \in \Delta_1, \ldots x_n \in$ wrt $\langle z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell \rangle$ such that for each $i, 1 \leq i \leq n$, $J(R_{\sigma_i})$ is defined and $J(R_{\sigma_i}) \in Arr\left(J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{\Delta_i[\sigma_1 \mid x_1, \ldots \sigma_{i-1} \mid x_{i-1}] \text{ is a type}}\right)\right)$

Thus since $H(R_t)$, $J\left(\frac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{t[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \in \Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n]}\right)$

is defined and is equal to $J(R_{\sigma_n})^* \ldots J(R_{\sigma_1})^* p(J(R_{\Lambda_\ell}), 1)^* J(R_t)$. But as $H(R_{t_i})$ it follows that $J(R_{\sigma_i}) = J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_k}), 1)^* J(R_{t_i}$

Hence $J\left(\dfrac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{t[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \in \Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n]}\right)$ =

$(J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_\ell}),1)^* J(R_{t_n}))^* \ldots (J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_\ell}),1)^*$
$J(R_{t_1}))^* p(J(R_{\Lambda_\ell}),1)^* J(R_t) = J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_\ell}),1)^* J(R_{t_n})^* \ldots J(R_{t_1})^*$
$p(J(R_{\Omega_m}),1)^* J(R_t)$, by lemma 4(iii) of §2.3.

But since $H(R_t)$, $J(R_{t_n})^* \ldots J(R_{t_1})^* p(J(R_{\Omega_m}),1)^* J(R_t) =$

$J\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right)$. Thus

$J\left(\dfrac{z_1 \in \Lambda_1, \ldots z_\ell \in \Lambda_\ell}{t[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n] \in \Delta[\sigma_1 \mid x_1, \ldots \sigma_n \mid x_n]}\right) =$

$J(R_{S_m})^* \ldots J(R_{S_1})^* p(J(R_{\Lambda_\ell}),1)^*$

$J\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t[t_1 \mid x_1, \ldots t_n \mid x_n] \in \Delta[t_1 \mid x_1, \ldots t_n \mid x_n]}\right)$, as required.

Case 3. R is a T=rule, say $R = \dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$.

We assume H(R) and we must show that $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] = \Delta'[t_1 \mid x_1, \ldots t_n \mid x_n]}\right)$

H(R) implies that $H(R_\Delta)$ and by case 1 that implies

$H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$.

Similarly we must have $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta'[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$.

$H(R_\Delta)$ also implies that $J\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n] \text{ is a type}}\right)$

$J(R_{t_n})^*\ldots J(R_{t_1})^*p(J(R_{\Omega_m}),1)^*J(R_\Delta)$. $H(R_{\Delta'})$ implies that

$$J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta'[t_1\mid x_1,\ldots t_n\mid x_n] \text{ is a type}}\right) = J(R_{t_n})^*\ldots J(R_{t_1})^*p(J(R_{\Omega_m}$$

$J(R_{\Delta'})$. But H(R) implies that $J(R_\Delta) = J(R_{\Delta'})$, hence

$$J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n] \text{ is a type}}\right) = J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta'[t_1\mid x_1,\ldots t_n\mid x_n] \text{is}}\right.$$

This completes the proof that $H\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta[t_1\mid x_1,\ldots t_n\mid x_n] = \Delta'[t_1\mid x_1,\ldots t_n\mid}\right.$

Case 4. R is an $\in$=rule. Very similar to case 3.

Lemma 12. (i). For every $n\geqslant 1$, if $1\lhd A\ldots\lhd A_n$ in $\mathbb{C}$ then (a)

$$J\left(\frac{x_1\in\overline{A_1},\ldots x_{n-1}\in\overline{A_{n-1}}(x_1,\ldots x_{n-2})}{\overline{A_n}(x_1,\ldots x_{n-1}) \text{ is a type}}\right)$$

is defined and is equal to $A_n.(b)$ for any $i, 1\leqslant i\leqslant n$,

$$J\left(\frac{x_1\in\overline{A_1},\ldots x_n\in\overline{A_n}(x_1,\ldots x_{n-1})}{x_i\in\overline{A_i}(x_1,\ldots x_{i-1})}\right)$$

is defined and is equal to '$p(A_n,A_i)$'.

(ii). For every $n\geqslant 1$, if $1\lhd A_1\ldots\lhd A_n$ in $\mathbb{C}$ and if $f:A_n\longrightarrow B$ is a non-trivial morphism of $\mathbb{C}$ then

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{(f \circ p(B))^* B}(x_1, \ldots x_n)}\right)$$

is defined and is equal to 'f'.

Proof. (i). The proof is by induction on n. If n = 1 then (a) $J(\overline{A_1}$ is a type) is by definition A_1. (b) $J\left(\frac{x_1 \in \overline{A_1}}{x_1 \in \overline{A_1}}\right)$ is by definition 'p(A_1,A_1)'. If n > 1 then (a) $J\left(\frac{x_1 \in \overline{A_1}, \ldots x_{n-1} \in \overline{A_{n-1}}(x_1, \ldots x_{n-2})}{\overline{A_n}(x_1, \ldots x_{n-1}) \text{ is a type}}\right)$

is defined to be $J(R_{n-1})^* \ldots J(R_1)^* p(J(RA_{n-1}), 1)^* A_n$, where R_i is the rule $\frac{x_1 \in \overline{A_1}, \ldots x_{n-1} \in \overline{A_{n-1}}(x_1, \ldots x_{n-2})}{x_i \in \overline{A_i}(x_1, \ldots x_{i-1})}$ and RA_{n-1} is the rule $\frac{x_1 \in \overline{A_1}, \ldots x_{n-2} \in \overline{A_{n-2}}(x_1, \ldots x_{n-2})}{\overline{A_{n-1}}(x_1, \ldots x_{n-2}) \text{ is a type}}$. By the inductive hypothesis $J(R_i)$ = 'p(A_{n-1},A_i)' and $J(RA_{n-1}) = A_{n-1}$. Thus

$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_{n-1} \in \overline{A_{n-1}}(x_1, \ldots x_{n-2})}{\overline{A_n}(x_1, \ldots x_{n-1}) \text{ is a type}}\right)$ is defined and is equal to A by lemma 3(ii) of § 2.3. (b) In view of (a),

$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{x_i \in \overline{A_i}(x_1, \ldots x_{i-1})}\right)$ is just defined to be 'p(A_n,A_i)'

(ii). By definition of J, $J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{(f \circ p(B))^* B}(x_1, \ldots x_n)}\right)$ =

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}\right) * \ldots J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}}{x_1 \in \overline{A_1}}\right.$$

$$\left.p\left(J\left(\frac{x_1 \in \overline{A_1}, \ldots x_{n-1} \in \overline{A_{n-1}}(x_1, \ldots x_{n-2})}{\overline{A_n}(x_1, \ldots x_{n-1}) \text{ is a type}}\right), 1\right)\right. \quad *'f'.$$

So by part (i this lemma, this equals $'p(A_n,A_n)'*\ldots'p(A_n,A_1)'*p(A_n,1)*$ 'f', which equals 'f' by lemma 3(ii) of § 2.3.

Corollary 13. (i) For every sort symbol $\overline{A}$ of $U(\mathbb{C})$, if $1 \lhd A_1 \ldots \lhd A_n \lhd A$ in $\mathbb{C}$ then $H\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{A}(x_1, \ldots x_n) \text{ is a type}}\right)$

(ii) For every sort symbol $\overline{f}$ of $U(\mathbb{C})$, if $1 \lhd A_1 \ldots A_n$ in $\mathbb{C}$ and $f: A_n \longrightarrow B$ in $\mathbb{C}$ then $H\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in (\overline{f \circ p(B)) * B}(x_1, \ldots x_n)}\right)$

Proof. In view of lemma 12 it only remains to show that J behaves correctly on all substitution instances of the given rules. But the definition of J ensures exactly this.

Lemma. For every derived rule R of $U(\mathbb{C})$, H(R).

Proof. By induction on the derivations in $U(\mathbb{C})$. We must check that every principle of derivation preserves the property H.

The principles L1-7 preserve property H. This can be seen at a glance. We go on to the other principles.

T1. Let R_t be a derived rule of $U(\mathbb{C})$ of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$

let R be a derived rule of $U(\mathbb{C})$ of the form $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$

Let R'_t be the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta'}$. We must show that

$H(R_t)$ and $H(R)$ implies that $H(R'_t)$.

$J\left(\dfrac{P}{t \in \Delta}\right)$ is always defined independently of Δ so since

$J(R_t)$ is defined and belongs to $Arr_{\mathbb{C}}(J(R_\Delta))$ it follows that $J(R'_t)$ is defined and belings to $Arr_{\mathbb{C}}(J(R_\Delta))$. Since $H(R)$ it follows that $J(R_\Delta) = J(R_{\Delta'})$, hence $J(R'_t)$ is defined and belongs to $Arr_{\mathbb{C}}(J(R_{\Delta'}))$. That is 2(b) holds of R'_t. 2(a) holds of R'_t because $H(R)$ implies $H(R_{\Delta'})$. 2(c) holds of R'_t because 2(c) holds of R_t and because $J\left(\dfrac{P}{t \in \Delta}\right)$ is defined independently of Δ.

CF1. Suppose that H holds of the derived rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta_{n+1}}$ is a type of $U(\mathbb{C})$. We wish to show that for all $i, 1 \leq i \leq n+1$,

$$H\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}\right).$$

The proof is by induction on i. Fix an $i, i \leq n+1$, assume that for all $j, 1 \leq j < i$, $H\left(\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_j \in \Delta_j}\right)$

we can now show that $H\left(\dfrac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}\right)$

is the case as follows:

$H(R_{\Delta_{n+1}})$ implies that for all $k, 1 \leq k \leq n+1$, $H(R_{\Delta_k})$. In particular $H(R_{\Delta_i})$. Also, $\langle x_1, \ldots x_{i-1} \rangle$ is a realisation of $\langle x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1} \rangle$ wrt $\langle x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1} \rangle$ such that for each $j, 1 \leq j \leq i-1$, $H\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_j \in \Delta_j}\right)$.

Thus by lemma 11 it follows that $H\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}\right)$

Which is to say that 2(a) holds of $\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}$

By definition of J, $J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_j \in \Delta_j}\right) =$

$'p(J(R_{\Delta_{n+1}}), J(R_{\Delta_j}))'$. Thus since $H(R_{\Delta_i})$, $J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}\right)$

$'p(J(R_{\Delta_{n+1}}), J(R_{\Delta_{i-1}}))' * \ldots 'p(J(R_{\Delta_{n+1}}), J(R_{\Delta_1}))' * p(J(R_{\Delta_{n+1}}), 1) * J(R_{\Delta_i})$

Thus, by lemma 3(ii) of §2.3, $J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}\right) =$

$p(J(R_{\Delta_{n+1}}), J(R_{\Delta_{i-1}})) * J(R_{\Delta_i})$. By definition $J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}\right)$

$= 'p(J(R_{\Delta_{n+1}}), J(R_{\Delta_i})'$. Hence

$J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}\right) \in Arr_{\mathbb{C}}\left(J\left(\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{\Delta_i \text{ is a type}}\right)\right)$.

Which is to say 2(b) holds of $\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i \in \Delta_i}$.

Now suppose that $\langle t_1,\ldots t_{n+1}\rangle$ is a realisation of $\langle x_1\in\Delta_1,\ldots x_{n+1}\in\Delta_{n+1}\rangle$ wrt $\langle y_1\in\Omega_1,\ldots y_m\in\Omega_m\rangle$ and that for each $k, 1\leqslant k\leqslant n+1$, $J(R_{t_k})$ is defined, belongs to

$$Arr_{\mathbb{C}}J\left(\frac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\Delta_k[t_1\mid x_1,\ldots t_{k-1}\mid x_{k-1}] \text{ is a type}}\right).$$

We must show that $J(R_{t_i}) = J(R_{t_{n+i}})^*\ldots J(R_{t_1})^*p(J(A_{\Omega_m}),1)^*$

$$J\left(\frac{x_1\in\Delta_1,\ldots x_{n+1}\in\Delta_{n+1}}{x_i\in\Delta_i}\right)$$

But that is immediate from lemma 4(i) of §2.3, since

$$J\left(\frac{x_1\in\Delta_1,\ldots x_{n+1}\in\Delta_{n+1}}{x_i\in\Delta_i}\right) = 'p(J(R_{\Delta_{n+1}}),J(R_{\Delta_i})'.$$

Thus 2(c) holds of $\dfrac{x_1\in\Delta_1,\ldots x_{n+1}\in\Delta_{n+1}}{x_i\in\Delta_i}$

CF2(a). Suppose that $\overline{A}$ is a sort symbol of $U(\mathbb{C})$ introduced by the rule $\dfrac{x_1\in\overline{A_1},\ldots x_n\in\overline{A_n}(x_1,\ldots x_{n-1})}{\overline{A}(x_1,\ldots x_n) \text{ is a type}}$. Suppose that for each $i, 1\leqslant i\leqslant n$, $\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{t_i\in\overline{A_i}(t_1,\ldots t_{i-1})}$ is a derived rule of $U(\mathbb{C})$ of which H holds.

We must show that $H\left(\dfrac{y_1\in\Omega_1,\ldots y_m\in\Omega_m}{\overline{A}(t_1,\ldots t_n) \text{ is a type}}\right)$

This is an immediate consequence of lemma 11 and corollary 13.

CF2(b). Similar to CF2(a).

SI1. Suppose that R is a derived rule of $U(\mathbb{C})$ of the form

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\Delta = \Delta'}$$

and that for each $i, 1 \leq i \leq n$, R_i is a derived rule of $U(\mathbb{C})$ of the form $\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{t_i = t_1' \in \Delta_i[t_1 \mid x_1, \ldots t_{i-1} \mid x_{i-1}]}$

Suppose also that $H(R)$ and that for each $i, 1 \leq i \leq n$, $H(R_i)$. We must show that

$$H\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] = \Delta'[t_1' \mid x_1, \ldots t_n' \mid x_n]}\right)$$

From $H(R)$ we deduce $H(R_\Delta)$ and from each $H(R_i)$ we deduce $H(R_{t_i})$.

By lemma 11 it follows that $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right)$.

Similarly it follows that $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta'[t_1' \mid x_1, \ldots t_n' \mid x_n] \text{ is a type}}\right)$.

Finally, $J\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] \text{ is a type}}\right) = J(R_{t_n}) * \ldots J(R_{t_1}) * p(J(R_{\Omega_n}$

$J(R_\Delta) = J(R_{t_n'}) * \ldots J(R_{t_1'}) * p(J(R_{\Omega_m}), 1) * J(R_{\Delta'}) =$

$$J\left(\frac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta'[t_1' \mid x_1, \ldots t_n' \mid x_n] \text{ is a type}}\right).$$

So $H\left(\dfrac{y_1 \in \Omega_1, \ldots y_m \in \Omega_m}{\Delta[t_1 \mid x_1, \ldots t_n \mid x_n] = \Delta'[t_1' \mid x_1, \ldots t_n' \mid x_n]}\right)$

as required.

SI2. Similar to SI1.

A1 and A2. We wish to show that these principles preserve property H

We must show that whenever $\frac{P}{\Delta = \Delta'}$ is an axiom of $U(\mathbb{C})$ such that

$H\left(\frac{P}{\Delta \text{ is a type}}\right)$ and $H\left(\frac{P}{\Delta' \text{ is a type}}\right)$ then $H\left(\frac{P}{\Delta = \Delta'}\right)$

and we must show that whenever $\frac{P}{t = t' \in \Delta}$ is an axiom of $U(\mathbb{C})$ such

that $H\left(\frac{P}{t \in \Delta}\right)$ and $H\left(\frac{P}{t' \in \Delta}\right)$ then $H\left(\frac{P}{t = t' \in \Delta}\right)$. In the fi

case this amounts to showing that $J\left(\frac{P}{\Delta \text{ is a type}}\right) = J\left(\frac{P}{\Delta' \text{ is a t}\right.$

in the second case it amounts to showing that $J\left(\frac{P}{t \in \Delta}\right) = J\left(\frac{P}{t' \in \Delta}\right.$

so we go through all of the axioms of $U(\mathbb{C})$ checking that one or the other holds as appropriate.

(i) For $n \geqslant o$, $m \geqslant 1$, $\ell \geqslant 1$, if $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1 \ldots \lhd B_m$ and $1 \lhd C_1 \ldots \lhd C_\ell$ and $f: A_n \longrightarrow B_m$, $g: B_m \longrightarrow C_\ell$ in $\mathbb{C}$ then $U(\mathbb{C})$ has the axio

$x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})$

$\overline{f \circ g}(x_1, \ldots x_n) = \overline{g}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f}(x_1, \ldots x_n))$ $(\overline{f \circ g \circ p(C_\ell)}) * C_\ell$

By lemma 12, $J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}}{\overline{f \circ g}(x_1, \ldots x_n) \in (\overline{f \circ g \circ p(C_\ell)}) * C_\ell (x_1, \ldots x_n)}\right) = \text{'}f \circ g\text{'}.$

On the other hand, by definition

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{g}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f}(x_1, \ldots x_n)) \in (\overline{f \circ g \circ p(C_\rho)) * C_\rho}(x_1, \ldots x_n)}\right)$$

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f}(x_1, \ldots x_n) \in \overline{B_m}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n))}\right)^*$$

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n) \in \overline{B_1}}\right)^* \quad p($$

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_{n-1} \in \overline{A_{n-1}}(x_1, \ldots :}{\overline{A_n}(x_1, \ldots x_n) \text{ is a type}}\right.$$

$1)^*$'g' which by lemma 12 equals 'f'* ....'$f \circ p(B_m, B_1)$'* $p(A_n, 1)$*'g'. By lemma 3(ii) of §2.3 this equals f*'g', which by lemma 4(i) of §2.3 equals 'f∘g'. Which is just whats required.

(ii). For $n \geqslant 1$, if $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{C}$ and $1 \leqslant i \leqslant n$ then $U(\mathbb{C})$ has the axiom $\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{p(A_n, A_i)}(x_1, \ldots x_n) = x_i \in \overline{A_i}(x_1, \ldots x_{i-1})}$

$J\left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}}{p(A_n, A_i)(x_1, \ldots x_n) \in \overline{A_i}(x_1, \ldots x_{i-1})}\right)$ is defined and is equal to $'p(A_n, A_i)'$ by lemma 12(ii). $J\left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{x_i \in \overline{A_i}(x_1, \ldots x_{i-1})}\right)$ is defined and eqal to $'p(A_n, A_i)'$ by virtue of lemma 12(i). Thus

$J\left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{p(A_n, A_i)}(x_1, \ldots x_n) \in \overline{A_i}(x_1, \ldots x_{i-1})}\right)$ =

$J\left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{x_i \in \overline{A_i}(x_1, \ldots x_{i-1})}\right)$ as required.

(iii). For $n \geqslant o$, $m \geqslant 1$, if $1 \lhd A_1 \ldots \lhd A_n$, $1 \lhd B_1, \ldots \lhd B_m \lhd B$ and $f: A_n \longrightarrow B_m$ in $\mathbb{C}$ then $U(\mathbb{C})$ has the axioms

$$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{f^*B}(x_1, \ldots x_n) = \overline{B}(\overline{f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f}(x_1, \ldots x_n))}$$

and $$\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}), y \in \overline{f^*B}(x_1, \ldots x_n)}{\overline{q(f,B)}(x_1, \ldots x_n, y) = y \in \overline{f^*B}(x_1, \ldots x_n)}$$

By virtue of lemma 12, $J\left(\dfrac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1})}{\overline{B(f \circ p(B_m, B_1)}(x_1, \ldots x_n), \ldots \overline{f}(x_1, \ldots x_n)) \text{ is a typ}}\right)$

is defined to be $'f'^* \ldots 'f \circ p(B_m, B_1)'^* p(A_n, 1)^* B$, which by lemma 3(i) of §2.3 is just f^*B, which is

$$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{i-1})}{f^*B(x_1, \ldots x_n) \text{ is a type}}\right) \qquad \text{as required.}$$

By definition of J, $$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}), y \in \overline{f^*B}(x_1, \ldots x_n)}{y \in \overline{f^*B}(x_1, \ldots x_n)}\right)$$

$'id_{f^*B}'$. By lemma 12, $$J\left(\frac{x_1 \in \overline{A_1}, \ldots x_n \in \overline{A_n}(x_1, \ldots x_{n-1}), y \in \overline{f^*B}(x_1, \ldots x_n)}{\overline{q(f,B)}(x_1, \ldots x_n, y) \in \overline{f^*B}(x_1, \ldots x_n)}\right)$$

$'q(f,B)'$. But by lemma 5 of §2.3, $'id_{f^*B}' = 'q(f,B)'$.

We can collect together the information about J that we really want in the following:

<u>Corollary 14</u> (i) If $\langle x_1 \in \Delta_1, \ldots x_n \in \Delta_n \rangle$ is a context of $U(\mathbb{C})$ then $1 \lhd J(R\Delta_1) \ldots \lhd J(R\Delta_n)$ in $\mathbb{C}$, where $R\Delta_i = \dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i \text{ is a type}}$.

(ii). If $\langle t_1,\dots t_m\rangle$ is a realisation of $\langle y_1\in\Omega_1,\dots y_m\in\Omega_m\rangle$ wrt $\langle x_1\in\Delta_1,\dots x_n\in\Delta_n\rangle$ in $U(\mathbb{C})$ then for each $j, 1\leq j\leq m$, $J(R_{t_j})\in \mathrm{Arr}_{\mathbb{C}}(J(R_{t_{j-1}})^*\dots J(R_{t_1})^*p(J(R_{\Delta_n}),1)^*J(R_{\Omega_j})$. Where R_{t_j} is the rule $\dfrac{x_1\in\Delta_1,\dots x_n\in\Delta_n}{t_j\in\Omega_j[\,t_1\mid y_1,\dots t_{j-1}\mid y_{j-1}]}$.

(iii). If $\langle x_1\in\Delta_1,\dots x_n\in\Delta_n\rangle \equiv \langle x_1'\in\Delta_1',\dots x_n'\in\Delta_n'\rangle$ then $J(R_{\Delta_n}) = J(R_{\Delta'_n})$.

(iv). If $\langle t_1,\dots t_m\rangle = \langle t_1',\dots t_m'\rangle$ then for each $j, 1\leq j\leq m$, $J(R_{t_j}) = J(R_{t_j'})$.

By corollary 14(ii) and by lemma 3(iii) of §2.3, whenever $\langle t_1,\dots t_m\rangle$ is a realisation of $\langle y_1\in\Omega_1,\dots y_m\in\Omega_m\rangle$ wrt $\langle x_1\in\Delta_1,\dots x_n\in\Delta_n\rangle$ in $U(\mathbb{C})$ then there exists a unique m-tuple $\langle \gamma_1,\dots \gamma_m\rangle$ of morphisms of $\mathbb{C}$ such that for each $j, 1\leq j\leq m$, $\gamma_j: J(R_{\Delta_n})\longrightarrow J(R_{\Omega_j})$ and $'\gamma_j' = J(R_{t_j})$, and such that for each $j, 1\leq j\leq m$, $\gamma_{j+1}\circ p(J(R_{\Omega_{j+1}}),J(R_{\Omega_j})) = \gamma_j$. This last condition implies that the m-tuple $\langle \gamma_1,\dots \gamma_m\rangle$ is determined by $\gamma_m: J(R_{\Delta_n})\longrightarrow J(R_{\Omega_m})$. So the statement can be reworded: for each such realisation there exists a unique $\gamma: J(R_{\Delta_n})\longrightarrow J(R_{\Omega_m})$ such that for each $j, 1\leq j\leq m$, $'\gamma\circ p(J(R_{\Omega_m}),J(R_{\Omega_j}))' = J(R_{t_j})$. Thus we can define a function $\S$ from objects and morphisms of $\mathbb{C}(U(\mathbb{C}))$ to objects and morphisms of $\mathbb{C}$ by

$$\begin{array}{ccc} [\langle x_1\in\Delta_1,\dots x_n\in\Delta_n\rangle] & \overset{\S}{\longmapsto} & J(R_{\Delta_n}) \\ \Big\downarrow\ [\langle t_1,\dots t_m\rangle] & & \Big\downarrow\ \gamma \\ [\langle y_1\in\Omega_1,\dots y_m\in\Omega_m\rangle] & & J(R_{\Omega_m}) \end{array}$$

where γ is the unique map such that for all $j, 1 \leq j \leq m$, $'\gamma_{ap}(J(R_{\Omega_m}), J(R_{\Omega_j}))' = J(R_{t_j})$.

Then γ is well defined by corollary 14(iii) and (iv).

We show that ξ is an inverse to $\eta_{\mathbb{C}}: \mathbb{C} \longrightarrow \mathbb{C}(U(\mathbb{C}))$. We need one last lemma.

Lemma 15. If $\dfrac{x_1 \in \Delta_1, \ldots x_{n-1} \in \Delta_{n-1}}{\Delta_n \text{ is a type}}$ is a derived rule of $U(\mathbb{C})$

then for all $i, 1 \leq i \leq n$, $\dfrac{x_1 \in \Delta_1, \ldots x_{i-1} \in \Delta_{i-1}}{\Delta_i = \overline{J(R_{\Delta_i})}(x_1, \ldots x_{i-1})}$

is a derived rule of $U(\mathbb{C})$.

If $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t \in \Delta}$ is a derived rule of $U(\mathbb{C})$ then

$\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t = \overline{J(R_t)}(x_1, \ldots x_n)}$ is a derived rule of $U(\mathbb{C})$.

Proof. By induction on derivations in $U(\mathbb{C})$. We show that each principle of derivation preserves the property. We have only to check those principles by which T and $\in$-rules are derived. These are T1, CF1, CF2(a) and CF2(b). If fact the checking of T1 is trivial. CF2(b) is very similar to CF2(a), so that leaves us to check CF1 and CF2(a).

CF1. Suppose that for all $j, 1 \leqslant j \leqslant n+1$,

$$\frac{x_1 \in \Delta_1, \ldots x_{j-1} \in \Delta_{j-1}}{\Delta_j = \overline{J(R_{\Delta_j})}(x_1, \ldots x_{j-1})}$$

is a derived rule of $U(\mathbb{C})$ and suppose that have some i , $1 \leqslant i \leqslant n+1$. We must show that

$$\frac{x_1 \in \Delta_1, \ldots x_{n+1} \in \Delta_{n+1}}{x_i = \overline{'p(J(R_{\Delta_{n+1}}), J(R_{\Delta_i}))'}(x_1, \ldots x_{n+1}) \in \Delta_i}$$

is a derived rule of $U(\mathbb{C})$. But this easily follows from corollary 3 of this section and the axiom

$$\frac{x_1 \in \overline{J(R_{\Delta_1})}, \ldots x_{n+1} \in \overline{J(R_{\Delta_{n+1}})}(x_1, \ldots x_n)}{\overline{p(J(R_{\Delta_{n+1}}), J(R_{\Delta_i}))}(x_1, \ldots x_{n+1}) = x_i \in \overline{J(R_{\Delta_i})}(x_1, \ldots x_{i-1})}$$

CF2(a). Suppose that $\overline{B}$ is a sort symbol of $U(\mathbb{C})$ introduced by the rule

$$\frac{y_1 \in \overline{B_1}, \ldots y_m \in \overline{B_m}(y_1, \ldots y_{m-1})}{\overline{B}(y_1, \ldots y_m) \text{ is a type}}$$

and suppose that for each $j, 1 \leqslant j \leqslant m$,

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{t_j = \overline{J(R_{t_j})}(x_1, \ldots x_n) \ \ \overline{B_j}(t_1, \ldots t_{j-1})}$$

is a derived rule of $U(\mathbb{C})$.

We must show that the rule $\dfrac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{\overline{B}(t_1, \ldots t_m) = \overline{J(R)}(x_1, \ldots x_n)}$ is a derived rule of $U(\mathbb{C})$, where R is the rule

$$\frac{x_1 \in \Delta_1, \ldots x_n \in \Delta_n}{B(t_1, \ldots t_m) \text{ is a type}}$$

Let $\gamma : J(R_{\Delta_n}) \longrightarrow B_m$ be the unique map such that for all $j, 1 \leqslant j \leqslant m$, $'\gamma \circ p(B_m, B_j)' = J(R_{t_j})$. By lemma 5(i) of this section

$$\frac{x_1 \in \overline{J(R_{\Delta_1})} \ldots x_n \in \overline{J(R_{\Delta_n})}(x_1,\ldots x_{n-1})}{\overline{\gamma^*B}(x_1,\ldots x_n) = \overline{B}(\overline{J(R_{t_1})}(x_1,\ldots x_n),\ldots \overline{J(R_{t_m})}(x_1,\ldots x_n))}$$

is a derived rule of $U(\mathbb{C})$. Thus so to is

$$\frac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\overline{\gamma^*B}(x_1,\ldots x_n) = \overline{B}(t_1,\ldots t_m)}$$

. But by lemma 3(ii) of § 2.3, $\gamma^*B = J(R_{t_m})^* \ldots J(R_{t_1})^* p(J(R_{\Delta_n}),1)^*B$, which by definition of J is just $J\left(\frac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\overline{B}(t_1,\ldots t_m) \text{ is a type}}\right)$. Thus $\frac{x_1 \in \Delta_1,\ldots x_n \in \Delta_n}{\overline{B}(t_1,\ldots t_m) = \overline{J(R)}(x_1,\ldots x_n)}$

is a derived rule of $U(\mathbb{C})$, as required.

Corollary. (i) $\xi \circ \eta_{\mathbb{C}} = id_{\mathbb{C}}$. (ii) $\eta_{\mathbb{C}} \circ \xi = id_{U(\mathbb{C}(U))}$.

Proof. (i) Follows directly from lemma 12. (ii) Follows from lemma 15 and corollary 3.

It now follows from quite general considerations that $\xi : \mathbb{C}(U(\mathbb{C})) \longrightarrow \mathbb{C}$ is a contextual functor. For example ξ preserves composition because $\xi(fog) = \xi(\eta_{\mathbb{C}}(\xi(f)) \circ \eta_{\mathbb{C}}(\xi(g)) = \xi(\eta_{\mathbb{C}}(\xi(f) \circ \xi(g))) = \xi(f) \circ \xi(g)$.

So that completes the proof that $\eta_{\mathbb{C}} : \mathbb{C} \longrightarrow (U(\mathbb{C}))$ is an isomorphism for every contextual category $\mathbb{C}$. So completing the proof that the category Con is equivalent to the category GAT.

2.5 Functorial Semantics, Universal Algebra

2.5.1 Functorial Semantics

An algebraic semantics is an equivalence between a category of theories amd a category of structures. We referred to several such in § 2.1. In all cases so far considered there is a further equivalence. In all cases the usual definition of model of a theory can be replaced by a new definition which uses only the notion of structure. Lawvere has used the term functorial semantics in describing this kind o semantics. Functorial semantics depends on an equivalence between the category of models of a theory U and the category of structure preservin morphisms from the structure $\mathbb{C}(U)$ corresponding to U to a special canonical structure (the world structure?). For example the canonical structure is taken to be the category with products Set in the case of algebraic theories (Lawvere [11]). Or in the case of classical propositional theories the canonical structure is taken to be the Boolean Algebra $\{0,1\}$.

The present situation is as well behaved as any if the canonical structure is taken to be the contextual category Fam.

If U is a generalised algebraic theory then the category of models U is equivalent to the category which has contextual functors $\mathbb{C}(U)$ to Fa as objects and natural transformations as morphisms. Thus we can assert

U-alg $\equiv$ ConFunc($\mathbb{C}(U)$, Fam).

It has turned out that the inductive construction of $\mathbb{C}(U)$ from U has enabled us to replace the usual inductive definition of model of U by the definition "a model of U is a contextual functor $M : \mathbb{C}(U) \longrightarrow \underline{Fam}$".

Every interpretation $I : U \longrightarrow U'$ induces a contextual functor $\mathbb{C}(I) : \mathbb{C}(U) \longrightarrow \mathbb{C}(U')$. Composition with $\mathbb{C}(I)$ is a functor from ConFunc $(\mathbb{C}(U'), \underline{Fam})$ to ConFunc$(\mathbb{C}(U), \underline{Fam})$. It is the functor $I\text{-alg} : U'\text{-alg} \longrightarrow U\text{-alg}$. Those functors between categories of models which are induced in this way are called generalised algebraic functors. We can show that all such functors have left adjoints. But anyhow this is equivalent to a known generalisation of Lawvere [11]'s theorem all algebraic functors have left adjoint.

2.5.2 Universal Algebra

We have been able to prove a generalisation of Birkhoffs theorem. Previously this theorem has been proved for many sorted algebraic theori see Birkhoff and Lipson [3] . Birkhoffs theorem classifies those subcategories of a category $\underline{U\text{-alg}}$ that arise as the category of models o equational extension of U. The result that we describe characterises subcategories $\underline{U'\text{-alg}}$ of $\underline{U\text{-alg}}$ in cases where U is any generalised algebraic theory and U' is any $\in$-equational extension of U.

By an $\in$-equaltional extension U' of U we mean an extension by axioms all of which are $\in$=rules. Thus U' is not permitted to have any T=axioms which are not already in U.

If we state the result then we can explain the terms afterwards.

Theorem If U is a generalised algebraic theory and if Γ is a subcategory of U-alg then equivalent are:

i. Γ is the subcategory determined by some $\in$-equational extension U' of U.

ii. Γ is a full subcategory of U closed under products, homomorphic images and subalgebras and having the property that if M is a U-algebra and if all the finitely generated subalgebras of M belong to Γ then M belongs to Γ.

The U-algebra M' is said to be a homomorphic image of the U-algebra M iff there exists a homomorphism $f : M \longrightarrow M'$ having the property that for all $1 \lhd A_1 \lhd \ldots \lhd A_n$ in $\mathbb{C}(U)$, for all $a'_1 \in M'(A_1)$, for all $a'_2 \in M'(A_2)(a'_1)$,... for all $a'_n \in M'(A_n)(a'_1,\ldots a'_1,\ldots a'_{n-1})$, there exists $a_1,\ldots a_n$ such that $a_1 \in M(a_1),\ldots a_n \in M(A_n)(a_1,\ldots a_{n-1})$ and such that $f(a_1) = a'_1,\ldots f(a_n) = a'_n$.

A U-algebra M is to be finitely generated U-algebra iff it is the homomorphic image of some finitely generalted free algebra.

Consider for a moment. Every theory U has a minimal model denoted K_U and built out of closed terms. Alternatively this minimal model is described just in terms of the structure $\mathbb{C}(U)$. For example if $1 \lhd A$ in $\mathbb{C}(U)$ then $K_U(A) = \mathrm{Hom}(1,A)$, otherwise if $1 \lhd A_1 \ldots A_n \lhd A$ in $\mathbb{C}(U)$ then if $a_1 \in K_U(A_1),\ldots$ if $a_n \in K_U(A_n)(a_1,\ldots a_{n-1})$ then $K_U(A)(a_1,\ldots a_n) = \{ a \in \mathrm{Hom}_{\mathbb{C}(U)}(1,A) \mid a \circ p(A) = a_n \}$.

Now, the free U-algebras are the algebras I-alg(K_U') for I : U $\longrightarrow$ U' an extension of U by constants alone. The finitely generated free U-algebras are these algebras where U' in an extension by just finitely many constants.

For example, take U to be the theory of categories. Take U' to be the theory of categories +

Symbol	Introductory Rule
a_1	$a_1 \in Ob$
a_2	$a_2 \in Ob$
b	$b \in Hom(a_1, a_2)$.

In this case I-alg(K_U') is the category $\cdot \longrightarrow \cdot$. It is the free category with one morphism. It is a finitely generated free category.

CHAPTER 3

ALGEBRAIC SEMANTICS OF M-L TYPE THEORY

We say what it is for a contextual category to have disjoint unions and a singleton object (comparable with category with finite products and terminal object). The category of these structures is denoted Σ -Con.

We then introduce a new notion of structure - the notion of a category with attributes. The category of these structures is denoted Attcat. We show that the category Σ -Con and the category Attcat are equivalent categories.

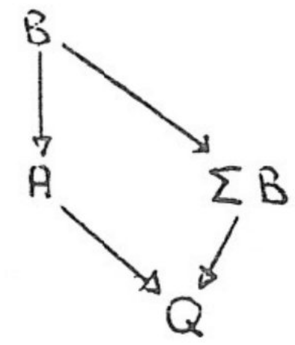
In §1.5 we alluded to a contextual category of categories, category indexed families of categories and so on. The well known fibration construction induces a structure of disjoint unions on this contextual category. We use the Attcat, Σ -Con equivalence to give a fairly brief description of this structure. This is in § 3.3.

§ 3.4 contains the definition of M-L structure intended as the model theory of a strengthened M-L type theory.

Also in §3.4 an equivalent notion of structure, based on the Attcat, Σ -Con equivalence, is put forward. Then in § 3.5 we can briefly describe a new model of M-L type theory, we refer to it as the limit space model.

3.1 Disjoint Unions and Singleton Objects

A <u>contextual category with disjoint unions</u> $\langle \mathbb{C}, \Sigma, \alpha \rangle$ consists of a contextual category $\mathbb{C}$ and for every $Q \triangleleft A \triangleleft B$ in $\mathbb{C}$, an object ΣB of $\mathbb{C}$ and a morphism $\alpha(B)$ of $\mathbb{C}$ such that $Q \triangleleft \Sigma B$ in $\mathbb{C}$, such that $\alpha(B)$ is an isomorphism in $\mathbb{C}$, $\alpha(B)\colon B \longrightarrow \Sigma B$ such that the diagram



commutes,

subject to the condition: If $f\colon Q \longrightarrow Q'$ in $\mathbb{C}$ and if $Q' \triangleleft A \triangleleft B$ then $f^* \Sigma B = \Sigma f^* B$ and $f^* \alpha(B) = \alpha(f^* B)$.

The category of contextual categories with disjoint unions has as morphisms $F\colon \langle \mathbb{C}, \Sigma, \alpha \rangle \longrightarrow \langle \mathbb{C}', \Sigma', \alpha' \rangle$ those contextual functors $F\colon \mathbb{C} \longrightarrow \mathbb{C}'$ such that for all $Q \triangleleft A \triangleleft B$ in $\mathbb{C}$, $F(\Sigma B) = \Sigma' F(B)$ and $F(\alpha(B)) = \alpha'(F(B))$.

It follows that if $\langle \mathbb{C}, \Sigma, \alpha \rangle$ is a contextual category with disjoint unions and if $f : A \longrightarrow A'$ in $\mathbb{C}$ then $\mathbb{C}_f : \langle \mathbb{C}_{A'}, \Sigma, \alpha \rangle \longrightarrow \langle \mathbb{C}_A, \Sigma, \alpha \rangle$ is a morphism of contextual categories with disjoint unions.

Actual disjoint unions of families of sets induce the structure of contextual category with disjoint unions on the contextual category <u>Fam</u>

of sets, families of sets and so on. This is as follows.

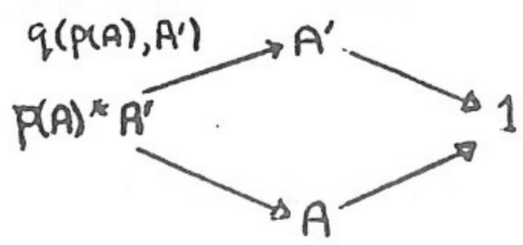
If $1 \lhd Q_1 \ldots \lhd Q_n \lhd A \lhd B$ in Fam then define ΣB, such that $Q_n \lhd \Sigma B$ in Fam, by defining $\Sigma B(r_1,\ldots r_n) = \{ \langle a,b \rangle \mid a \in A(r_1,\ldots r_n)$ and $b \in B(r_1,\ldots r_n,a) \}$ whenever $r_1 \in Q_1, \ldots r_n \in Q_n(r_1,\ldots r_{n-1})$. Define $\alpha(B) : B \longrightarrow \Sigma B$ in Fam by defining $\alpha(B) = \langle f_1,\ldots f_n,g \rangle$ where $f_i, 1 \leq i \leq n$ is defined by $f_i(r_1,\ldots r_n,a,b) = r_i$ whenever $r_1 \in Q_1, \ldots r_n \in Q_n(r_1,\ldots r_{n-1})$, $a \in A(r_1,\ldots r_n)$ and $b \in B(r_1,\ldots r_n,a)$, and where g is defined by $g(r_1,\ldots r_n,a,b) = \langle a,b \rangle$.

It is easy to see that $\alpha(B)$ has inverse $\alpha^{-1}(B)$ given by $\langle L_1,\ldots L_n,k_1,k_2 \rangle$ where $L_i, 1 \leq i \leq n$, is the operator given by $L_i(r_1,\ldots r_n,c) = r_i$ whenever $r_1 \in Q_1, \ldots r_n \in Q_n(r_1,\ldots r_{n-1})$ and $c \in \Sigma B(r_1,\ldots r_n)$, and where k_1 and k_2 are given by $k_i(r_1,\ldots r_n,c) =$ the ith component of the ordered pair c. The conditions $f^* \Sigma B = \Sigma f^* B$ and $f^*(\alpha(B)) = \alpha(f^*B)$ are automatically satisfied. Even without these conditions it is clear that the definition characterises disjoint unions in Fam up to isomorphism.

Lemma 1 If $\langle \mathbb{C}, \Sigma, \alpha \rangle$ is a contextual category with disjoint unions then Base ($\mathbb{C}$) is a category with products of pairs (and hence h finite products).

Proof Suppose that A and A' are objects of Base ($\mathbb{C}$). That is to say suppose $1 \lhd A$ and $1 \lhd A'$ in $\mathbb{C}$.

In any category, pulling back over the terminal object yields a product diagram. Since



is a pullback diagram in $\mathbb{C}$,

$$A \xleftarrow{\quad} p(A)^*A' \xrightarrow{\ q(p(A),A')\ } A'$$

is a product diagram in $\mathbb{C}$.

Since $\alpha(p(A)^*A') : p(A)^*A' \longrightarrow \sum p(A)^*A'$ is an isomorphism

$$A \xleftarrow{\ \alpha^{-1}(p(A)^*A')\circ p(p(A)^*A')\ } \sum p(A)^*A' \xrightarrow{\ \alpha^{-1}(p(A)^*A')\circ q(p(A),A')\ } A'$$

is a product diagram in $\mathbb{C}$. $1 \lhd \sum p(A)^*A'$ in $\mathbb{C}$ and so this diagram is a product diagram in Base ($\mathbb{C}$).

If this proof is interpreted in $\mathbb{C}$ = Fam then we have Base ($\mathbb{C}$) = Set, the category of sets and functions. If A and A' are sets then $p(A)^*A'$ is the constant A-indexed family with value A'. Thus AxA' in Set is given by $\sum p(A)^*A'$ in Fam which in turn is just $\{ \langle a,a' \rangle \mid a \in A \text{ ad } a' \in A' \}$. All is as it should be.

A singleton object of the contextual category $\mathbb{C}$ is defined to be an object of Base ($\mathbb{C}$) that is terminal in $\mathbb{C}$. Equivalently it is an object of Base ($\mathbb{C}$) that is isomorphic in $\mathbb{C}$ to the terminal object 1. The singleton object, if there is one, will usually be denoted $\{\cdot\}$.
The unique morphism : $1 \longrightarrow \{\cdot\}$ in $\mathbb{C}$ will be denoted e.
Thus a contextual category with singleton object $\langle \mathbb{C}, \{\cdot\}, e \rangle$ consists of a contextual category $\mathbb{C}$, an object $\{\cdot\}$ of $\mathbb{C}$ such that $1 \lhd \{\cdot\}$ in $\mathbb{C}$ and a morphism $e : 1 \longrightarrow \{\cdot\}$ such that $p(\{\cdot\})oe = id_{\{\cdot\}}$. The morphisms between contextual categories with singleton objects are taken to be those contextual functors which preserve the singleton object.

It is not difficult to see that if U is a generalised algebraic theory then the contextual category $\mathbb{C}(U)$ has a singleton object iff there

exists expressions t and Δ of U such that $x \in \Delta : t = x \in \Delta$ is a derive rule of U.

Lemma 2 If $\mathbb{C}$ is a contextual category with singleton object $\{\cdot\}$ a if $F : \mathbb{C} \longrightarrow \mathbb{D}$ is a contextual functor then $F(\{\cdot\})$ is a singleton object of $\mathbb{D}$. Proof F is a functor, hence F preserves isomorphisms. Thus $1 \lhd F(\{\cdot\})$ and $F(\{\cdot\}) \simeq 1$ in $\mathbb{D}$.

Lemma 3 If $\langle \mathbb{C}, \{\cdot\} \rangle$ is a contextual category with singleton object and if A is an object of $\mathbb{C}$ then $\langle \mathbb{C}_A, p(A,1)^* \{\cdot\} \rangle$ is a contextual category with singleton object. If $f:A \longrightarrow A'$ in $\mathbb{C}$ then $\mathbb{C}_f : \langle \mathbb{C}_{A'}, p(A',1)^* \{\cdot\} \rangle \longrightarrow \langle \mathbb{C}_A, p(A,1)^* \{\cdot\} \rangle$ is a morphism of contextual categories with singleton objects.

Proof If A is an object of $\mathbb{C}$ then $p(A,1) : A \longrightarrow 1$ in $\mathbb{C}$. Hence $\mathbb{C}_{p(A,1)} : \mathbb{C} \longrightarrow \mathbb{C}_A$. Thus by lemma 2 $\langle \mathbb{C}_A, p(A,1)^* \{\cdot\} \rangle$ is a contextual category with singleton object. Now if $f : A \longrightarrow A'$ in $\mathbb{C}$ then $\mathbb{C}_f : \mathbb{C}_{A'} \longrightarrow \mathbb{C}_A$ and $\mathbb{C}_f(p(A',1)^* \{\cdot\}) = f^* p(A',1)^* \{\cdot\} = (f \circ p(A',1))^* \{\cdot\} = p(A,1)^* \{\cdot\}$. Thus $\mathbb{C}_f : \mathbb{C}_{A'} \longrightarrow \mathbb{C}_A$ and preserves the singleton object.

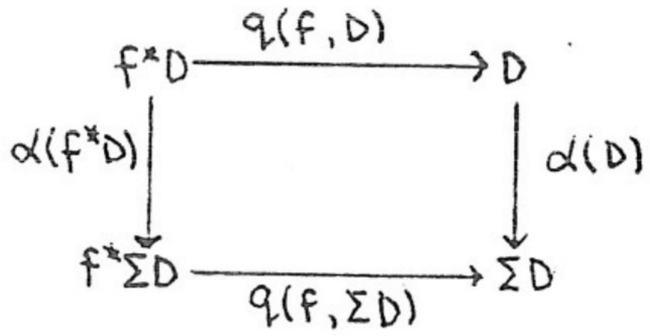
The category of contextual categories with disjoint unions and singleton objects is denoted Σ-Con.

The point is that in a contextual category with disjoint unions and singleton object there is a lot of repetition of structure. For example if in such a structure $Q \lhd A \lhd B$ then $\mathbb{C}_{\alpha(B)}$ is an isomorphism of

of $\mathbb{C}_B$ with $\mathbb{C}_{\Sigma B}$. Thus the structure of $\mathbb{C}$ above B is isomorphic to the structure of $\mathbb{C}$ above ΣB and because ΣB is at a lower level than B it turns out that the whole structure of $\mathbb{C}$ is coded up as structure at a very low level. This leads to the notion of a category with attributes. We must first introduce some new notation.

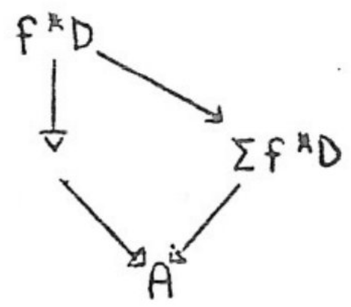
For the remainder of this section we suppose $\langle \mathbb{C}, \Sigma, \alpha, \{\cdot\} \rangle$ to be a contextual category with disjoint unions and singleton object.

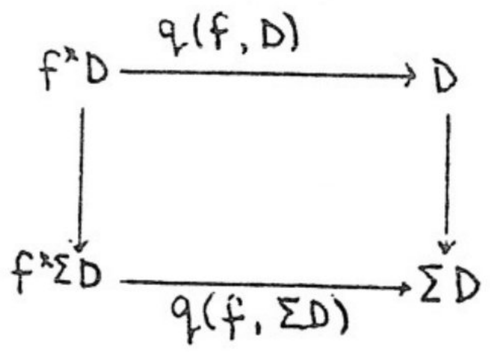
Lemma 4 If $f : A \longrightarrow A'$ in Base($\mathbb{C}$) and if $A' \lhd B \lhd D$ in $\mathbb{C}$ then the diagr



commutes.

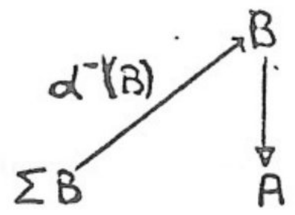
Proof. Since $f : A \longrightarrow A'$ and $A' \lhd B \lhd D$ we have by definition that $f^* \Sigma D = \Sigma f^* D$ and $f^* \alpha(D) = \alpha(f^* D)$. But $f^* \alpha(D)$ is defined to be the unique map : $f^* D \longrightarrow f^* \Sigma D$ such that the diagrams





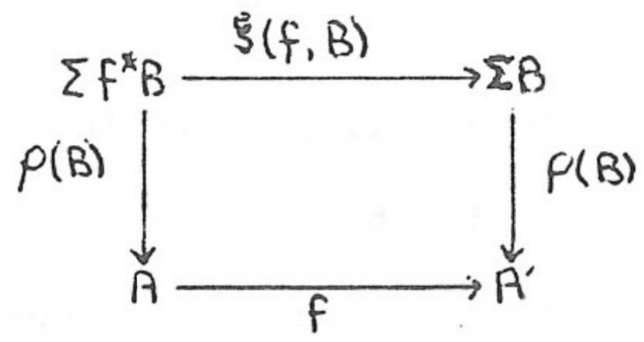
commute. Thus the statement of the lemma is a restatement of the condition $f^* \alpha(D) = \alpha(f^* D)$.

If $1 \lhd A \lhd B$ in Fam then we let $\rho(B) : \Sigma B \longrightarrow A$ be the 1st projection function from $\{\langle a,b\rangle \mid a \in A, b \in B(A)\}$ to A.
More generally if $Q \lhd A \lhd B$ in $\mathbb{C}$ then define $\rho(B) : \Sigma B \longrightarrow A$ to be $\alpha^{-1}(B) \circ \rho(B)$.



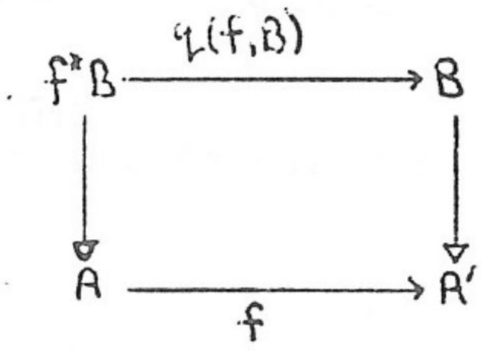
Now suppose $f : A \longrightarrow A'$ in Base ($\mathbb{C}$) and $A' \lhd B$ in $\mathbb{C}$, define $\S(f,B) : \Sigma f^*B \longrightarrow \Sigma B$ to be the morphism $\alpha^{-1}(f^*B) \circ q(f,B) \circ \alpha(B)$.

Lemma 5 If $f : A \longrightarrow A'$ in Base ($\mathbb{C}$) and if $A' \lhd B$ in $\mathbb{C}$ then



is a pullback diagram in Base ($\mathbb{C}$).

Proof Just because



is a pullback diagram in $\mathbb{C}$ and $\alpha(f^*B)$ and $\alpha(B)$ are isomorphisms.

Note that these pullback diagrams fit together in the sense that if $A \xrightarrow{f} A' \xrightarrow{f'} A''$ in $\mathbb{C}$ then $f^*f'^*B = (ff')^*B$ and

$\xi(f,f'^*B)\circ\xi(f',B) = \xi(ff',B)$. We are at this moment collapsing the structure of $\mathbb{C}$ into Base ($\mathbb{C}$).

We now introduce the $\#$, γ notation. We have to think of this as a reflection at a low level of $\mathbb{C}$, structure at higher levels.

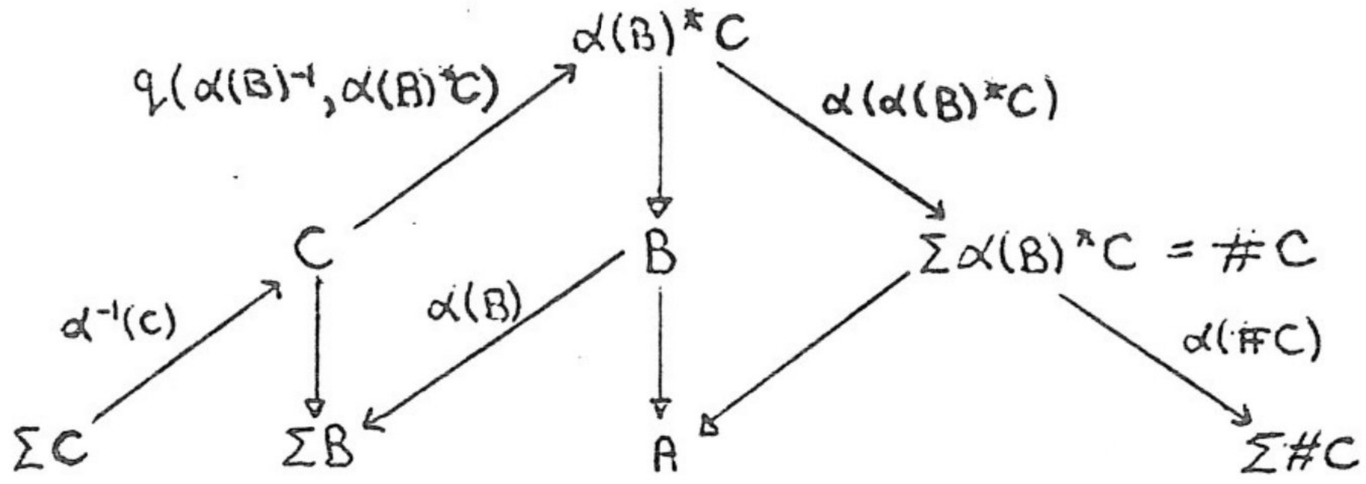
Lemma 6 If $f : A \longrightarrow A'$ is an isomorphism in $\mathbb{C}$ and If $A' < B$ then $q(f,B) : f^*B \longrightarrow B$ is an isomorphism with inverse $q(f^{-1},f^*B) : B \longrightarrow f^*B$ in $\mathbb{C}$.

Proof $q(f^{-1},f^*B) : B \longrightarrow f^*B$ in $\mathbb{C}$ because $f^{-1*}f^*B = B$.

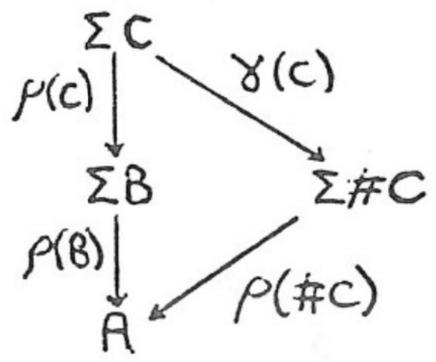
$q(f^{-1},f^*B)\ q(f,B) = q(f^{-1}f,B) = q(id_{A'},B) = id_B$.

$q(f,B)\ q(f^{-1},f^*B) = q(f,f^{-1*}f^*B)\ q(f^{-1},f^*B) = q(ff^{-1},f^*B) = id_{f^*B}$.

In future if $1 \lhd A \lhd B$ in $\mathbb{C}$ and $\Sigma B \lhd C$ then the object $\Sigma\alpha(B)^*C$ will be denoted $\#C$. Thus $A \lhd \#C$ in $\mathbb{C}$. By lemma 6 $\alpha(B)^*C \simeq C$. The composite isomorphism $C \xrightarrow{\sim} \alpha(B)^*C \xrightarrow{\sim} \Sigma\alpha(B)^*C$ will be denoted $\sigma(C)$. Thus $\sigma(C) = q(\alpha(B)^{-1},\alpha(B)^*C)\circ\alpha(\alpha(B)^*C)$. Finally we define $\gamma(C) : \Sigma C \longrightarrow \Sigma\#C$ to be the isomorphism $\alpha^{-1}(C)\circ\sigma(C)\circ\alpha(\#C)$.

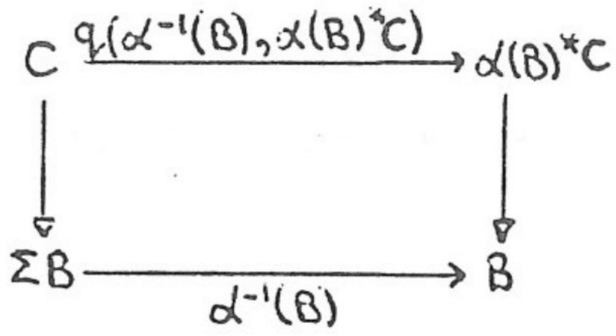


Lemma 7 If $1 \lhd A \lhd B$ and $\Sigma B \lhd C$ in $\mathbb{C}$ then the diagram.

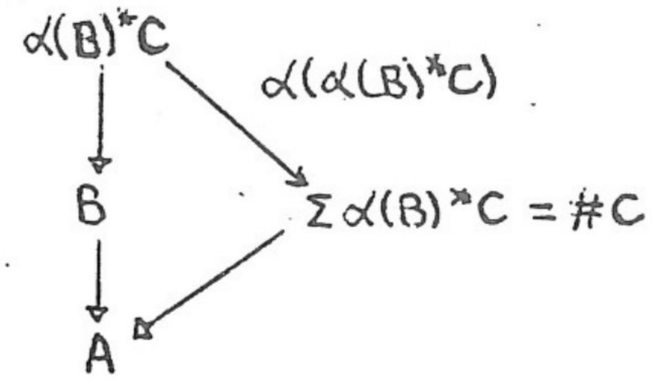


commutes in Base ($\mathbb{C}$).

Proof Use the definition of ρ and γ plus the commutivity of the diagrams

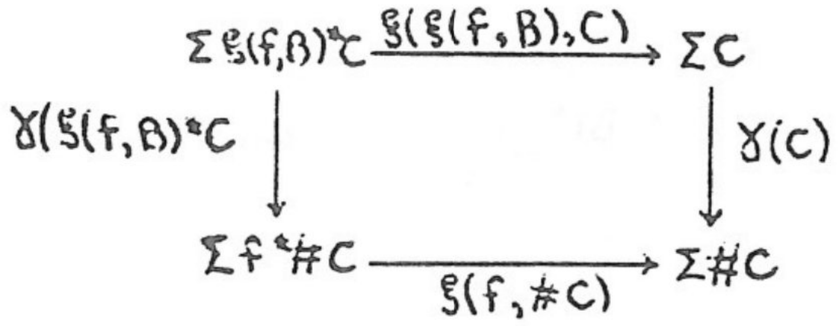


and



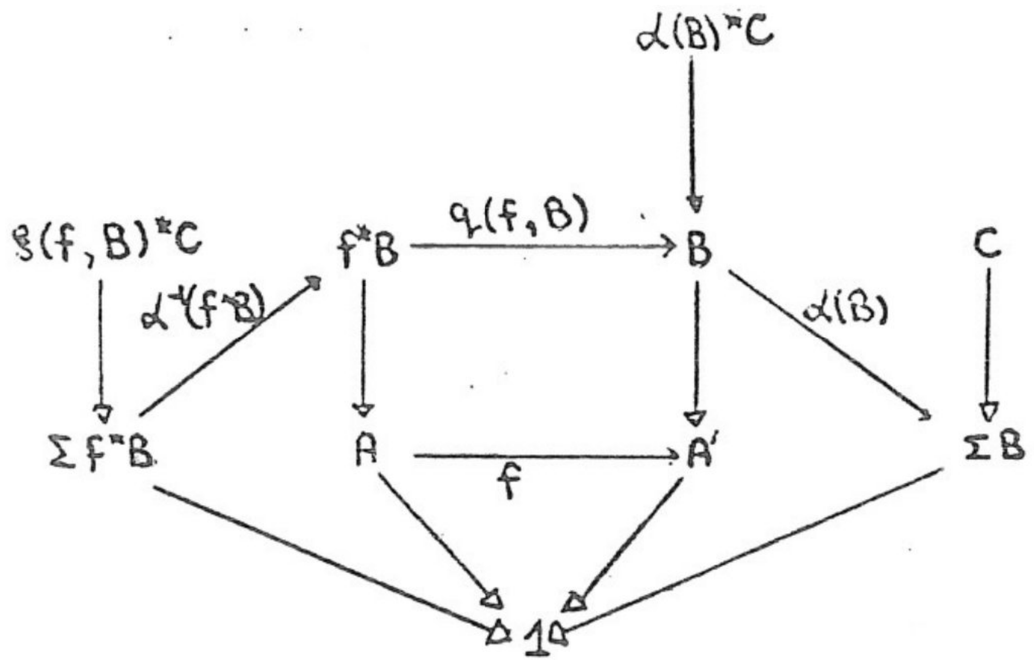
(The first diagram commutes because it is a pullback diagram, the second be definition of disjoint unions).

Lemma 8 If $f : A \longrightarrow A'$ in Base ($\mathbb{C}$) and if $A' \lhd B$, $\Sigma B \lhd C$ in $\mathbb{C}$ then $\#(\S(f,B)^*C) = f^*\#C$ and the diagram



commutes.

Proof The situation:



The identity $\#(\xi(f,B)^*C) = f^*B$ holds as follows:

$\#(\xi(f,B)^*C) = \Sigma(\alpha(f^*B)^*\xi(f,B)^*C)$ by def. of $\#$.

$= \Sigma(\alpha(f^*B)^*(\alpha^{-1}(f^*B)\circ q(f,B)\circ\alpha(B))^*C)$ by def. of ξ.

$= \Sigma(q(f,B)^*\alpha(B)^*C)$

$= \Sigma(f^*\alpha(B)^*C)$ * notation - see lemma 1 of 2

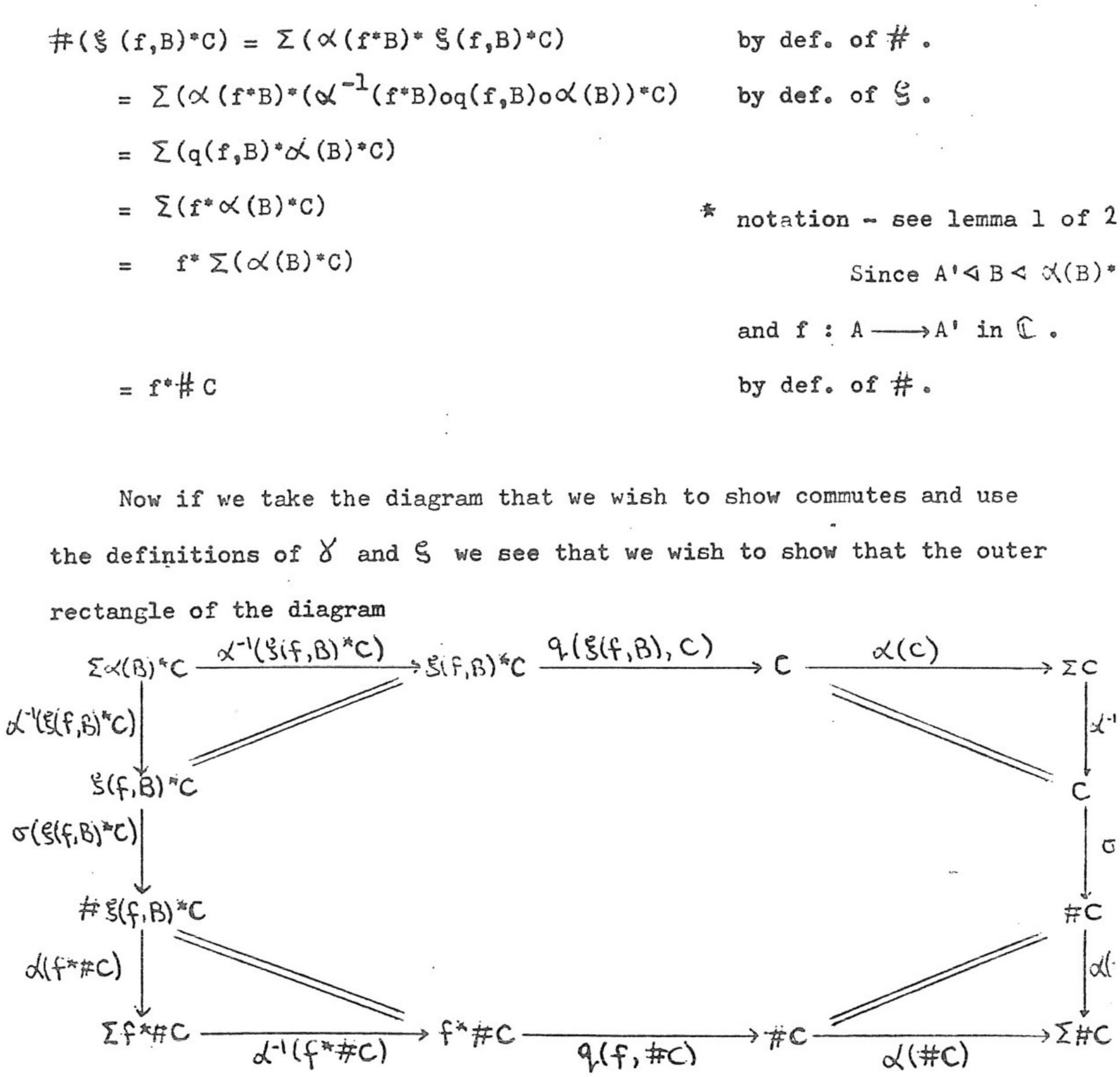
Since $A' \lhd B \prec \alpha(B)^*$

and $f : A \longrightarrow A'$ in $\mathbb{C}$.

$= f^*\Sigma(\alpha(B)^*C)$

$= f^*\#C$ by def. of $\#$.

Now if we take the diagram that we wish to show commutes and use the definitions of γ and ξ we see that we wish to show that the outer rectangle of the diagram



commutes.

Thus it suffices to show that the diagram

$$\begin{array}{ccc} \S(f,B)^*C & \xrightarrow{q(\S(f,B),C)} & C \\ {\scriptstyle \sigma(\S(f,B)^*C)}\downarrow & & \downarrow{\scriptstyle \sigma(C)} \\ f^*\#C = \#(\S(f,B)^*C) & \xrightarrow[q(f,\#C)]{} & \#C \end{array}$$

commutes.

Now $\sigma(C)$ is defined to be $q(\alpha^{-1}(B), \alpha(B)^*C) \circ \alpha(\alpha(B)^*C)$ and $\sigma(\S(f,B)^*C) = q(\alpha^{-1}(f^*B), \alpha(f^*B)^*\S(f,B)^*C) \circ \alpha(\alpha(f^*B)^*\S(f,B)^*C) = q(\alpha^{-1}(f^*B), f^*\alpha(B)^*C) \circ \alpha(f^*\alpha(B)^*C)$, by definition of $\S$. Thus we wish to show that the outer rectangle of the diagram

$$\begin{array}{ccc} \S(f,B)^*C & \xrightarrow{q(\S(f,B),C)} & C \\ {\scriptstyle q(\alpha^{-1}(f^*B), f^*\alpha(B)^*C)}\downarrow & & \downarrow{\scriptstyle q(\alpha^{-1}(B), \alpha(B)^*C)} \\ f^*\alpha(B)^*C & \xrightarrow{q(f,\alpha(B)^*C)} & \alpha(B)^*C \\ {\scriptstyle \alpha(f^*\alpha(B)^*C)}\downarrow & & \downarrow{\scriptstyle \alpha(\alpha(B)^*C)} \\ f^*\#C & \xrightarrow{q(f,\#C)} & \#C \end{array}$$

commutes.

Well the lower rectangle commutes by lemma 4 of this section. To show that the upper rectangle commutes we replace the extended $^*,q$ notation (i.e. use lemma 1 of § 2.3), use the fact that pullbacks fit together and use the definition of $\S$. This is as follows

$q(\alpha^{-1}(f^*B), f^*\alpha(B)^*C) \circ q(f, \alpha(B)^*C) =$

$q(\alpha^{-1}(f^*B), q(f,B)^*\alpha(B)^*C) \circ q(q(f,B), \alpha(B)^*C) =$

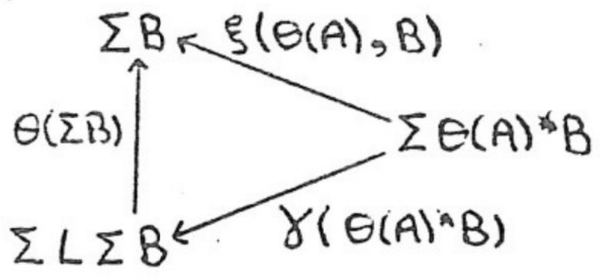
$q(\alpha^{-1}(f^*B) \circ q(f,B), \alpha(B)^*C) = q(\S(f,B) \circ \alpha^{-1}(B), \alpha(B)^*C) =$

$q(\S(f,B), \alpha^{-1}(B)^*\alpha(B)^*C) \circ q(\alpha^{-1}(B), \alpha(B)^*C) =$

$q(\S(f,B), C) \circ q(\alpha^{-1}(B), \alpha(B)^*C)$, as required.

If $A \in |\text{Base}(\mathbb{C})|$ then $p(\{\cdot\})^*A \simeq A$ because $p(\{\cdot\}) : \{\cdot\} \longrightarrow 1$ is an isomorphism. Hence $\sum p(\{\cdot\}) \simeq A$. We denote $p(\{\cdot\})^*A$ by $L(A)$ and the isomorphism of $\sum LA$ with A by $\Theta(A)$. Thus $\Theta(A) : \sum LA \longrightarrow A$ is defined by $\Theta(A) = \alpha^{-1}(p(\{\cdot\})^*A) \circ q(p(\{\cdot\}), A)$.

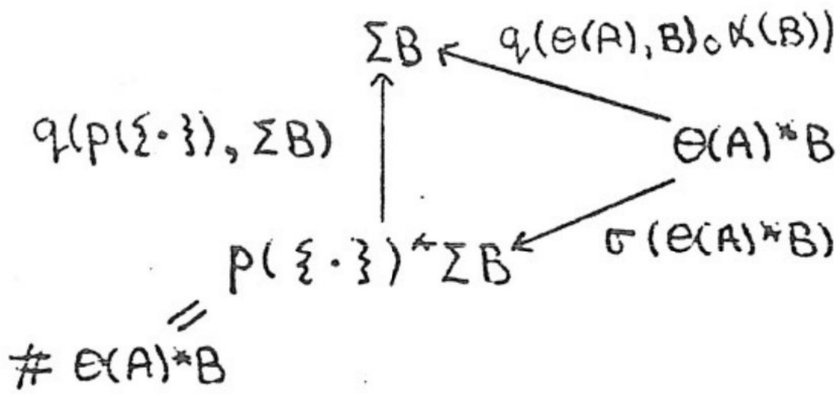
Lemma 9 If $1 \lhd A \lhd B$ in $\mathbb{C}$ then $L(\sum B) = \#(\Theta(A)^*B)$ and the diagram



commutes.

Proof $\#((\alpha^{-1}(p(\{\cdot\})^*A)^*q(p(\{\cdot\}), A)^*B = \sum q(p(\{\cdot\}, A)^*B$ by defintion of $\#$, $= \sum p(\{\cdot\})^*B$, by replacing extended *,q notation, $= p(\{\cdot\})^*\sum B$. That is $\#\Theta(A)^*B = L(\sum B)$.

As for the commuting triangle, if we cancel out the α's and α^{-1}'s, after substituting in for Θ, γ and ξ, then we see that we want to show that the diagram



commutes. Now $q(\Theta(A), B) \circ \alpha(B) = q(\alpha^{-1}(p(\{\cdot\})^*A) \circ q(p(\{\cdot\}, A), B) \circ \alpha(B)$ $q(\alpha^{-1}(p(\{\cdot\})^*A), p(\{\cdot\})^*B) \circ q(p(\{\cdot\}), B) \circ \alpha(B)$. Whereas

$\sigma(\theta(A)*B) = q(\alpha^{-1}(p(\{\cdot\})*A, p(\{\cdot\})*B) \circ \alpha(p(\{\cdot\})*B)$. Thus we must just show that the diagram

$$\begin{array}{ccc} p(\{\cdot\})*B & \xrightarrow{q(p(\{\cdot\}),B)} & B \\ \downarrow{\scriptstyle \alpha(p(\{\cdot\})*B)} & & \downarrow{\scriptstyle \alpha(B)} \\ \Sigma p(\{\cdot\})*B & \xrightarrow[q(p(\{\cdot\},\Sigma B)]{} & \Sigma B \end{array}$$

commutes. But this commutes by lemma 4.

3.2 Categories with attributes

A category with attributes $\langle \underline{C}, Att, 1_C, \Sigma, \rho, *, \xi, \#, \gamma, L, \theta \rangle$ consists of

i. A category $\underline{C}$ with terminal object 1_C. Said to be the base category.

ii. For every object A of $\underline{C}$, a set Att(A). The set of attributes of type A.

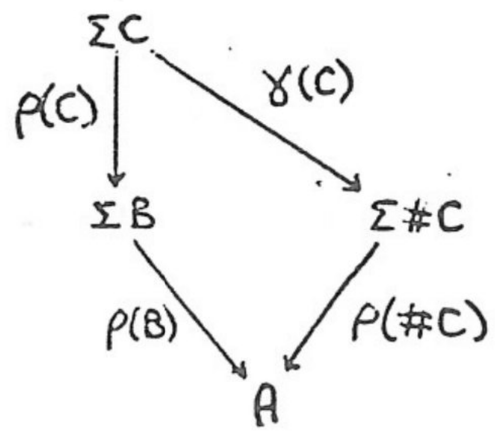
iii. For every object A of $\underline{C}$, for every $B \in Att(A)$, an object ΣB of $\underline{C}$ and a morphism $\rho(B) : \Sigma B \longrightarrow A$ in $\underline{C}$.

iv. For every morphism $f : A \longrightarrow A'$ in $\underline{C}$, for every $B \in Att(A')$, an attribute $f^*B \in Att(A)$ and a morphism $\xi(f,B) : \Sigma f^*B \longrightarrow \Sigma B$ in $\underline{C}$ such that the diagram

$$\begin{array}{ccc} \Sigma f^*B & \xrightarrow{\xi(f,B)} & \Sigma B \\ {\scriptstyle \rho(f^*B)}\downarrow & & \downarrow{\scriptstyle \rho(B)} \\ A & \xrightarrow[f]{} & A' \end{array}$$

is a pullback diagram in $\underline{C}$.

v. For every object A of $\underline{C}$, for every $B \in Att(A)$, for every $C \in Att(\Sigma B)$, an attribute $\#C \in Att(A)$ and an isomorphism $\gamma(C) : \Sigma C \longrightarrow \Sigma \#C$ in $\underline{C}$ such that the diagram



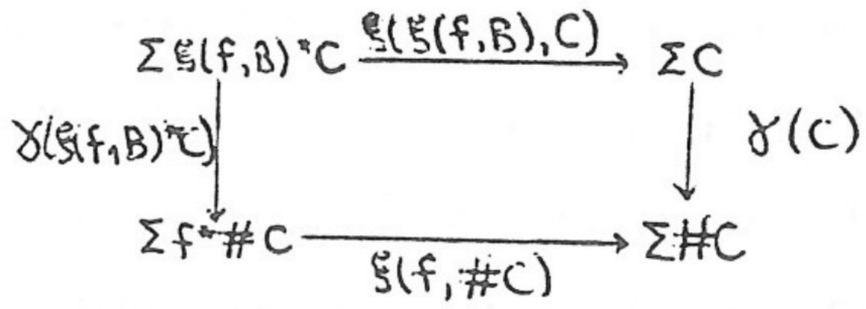
commutes.

vi. For every object A of $\underline{C}$, an attribut $L(A) \in Att(1_C)$ and an isomorphism $\Theta(A) : \Sigma LA \longrightarrow A$ in $\underline{C}$.
Such that L as a function $L : |\underline{C}| \longrightarrow Att(1_C)$ is an isomorphism of sets.
Subject to the conditions:

I. If A is an object of $\underline{C}$ and if $B \in Att(A)$ then $id_A{}^*B = B$ and $\xi(id_A, B) = id_{\Sigma B}$.

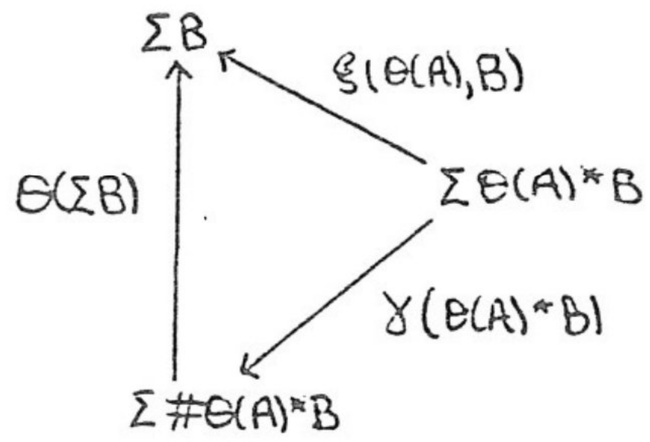
II. If $A \xrightarrow{f} A' \xrightarrow{f'} A''$ in $\underline{C}$ and if $B \in Att(A'')$ then $f^*f'^*B = B$ and $\xi(f, f'^*B) \circ \xi(f', B) = \xi(ff', B)$.

III. If $f : A \longrightarrow A'$ in $\underline{C}$ and if $B \in Att(A')$ and $C \in Att(\Sigma B)$ then $\#(\xi(f,B)^*C) = f^*\#C$ and the diagram



commutes.

IV. If A is an object of $\underline{C}$ and if $B \in Att(A)$ then $L\,\Sigma B = \# \Theta(A)^*B$ and the diagram



commutes.

The category of categories with attributes is denoted Attcat. A morphism $F : \langle \underline{C}, Att, \ldots\ldots \rangle \longrightarrow \langle \underline{C}', Att', \ldots\ldots \rangle$ consists of a functo: $F : \underline{C} \longrightarrow \underline{C}'$ preserving the terminal object and for each object A of $\underline{C}$ a function, also called F, $F : Att(A) \longrightarrow Att'(F(A))$ such that F, as a whole, preserves all the structure - $\Sigma, \rho, ^*, \xi, \#, \gamma, L$ and Θ .

In this section we prove that the category Attcat is equivalent to the category $\underline{\Sigma\text{-con}}$ of contextual categories with disjoint unions and singleton object. Part of this work has been done in §3.1 where we did, in effect, prove that every $\mathbb{D} \in |\,\underline{\Sigma\text{-con}}\,|$ induces a category with attributes now to be called $\phi(\mathbb{D})$. $\phi(\mathbb{D})$ has as base category the category Base ($\mathbb{D}$). If $A \in$ Base ($\mathbb{D}$) then $Att_{\phi(\mathbb{D})}(A)$ is taken to be the $\{B \in |\mathbb{D}| \mid A \lhd B \text{ in } \mathbb{D}\}$. $\Sigma, \rho, ^*, \xi, \#, L$ and Θ are then defined in $\phi(\mathbb{D})$ as defined in §3.1. Lemmas of that section then ensure that $\phi(\mathbb{D})$ so defined is a category with attributes.

Since any morphism $F : \mathbb{D} \longrightarrow \mathbb{D}'$ in $\underline{\Sigma\text{-con}}$ completely preserves the disjoint union and singleton object structure as well as

contextual structure and since $\varphi(\mathbb{D})$ and $\varphi(\mathbb{D}')$ are defined entirely in terms of this structure it follows that such an F induces a morphism $\varphi(F) : \varphi(\mathbb{D}) \longrightarrow \varphi(\mathbb{D}')$ in Attcat. The functoriality of $\varphi : \Sigma\text{-con} \longrightarrow$ Attcat is immediate. We wish to show that φ is an equivalence of categories. This involves the definition of a functor $\psi :$ Attcat $\longrightarrow \Sigma\text{-con}$.

Then we show that for all $\mathbb{D} \in |\ \Sigma\text{-con}|$, $\psi(\varphi(\mathbb{D}))$ is, upto isomorphism of structures, the structure $\mathbb{D}$ recovered from $\varphi(\mathbb{D})$.

If $\mathbb{E}$ is a category with attributes write Base $(\mathbb{E})$ for the base category of $\mathbb{E}$ and write $\mathrm{Att}_{\mathbb{E}}(A)$ for attributes of A when $A \in |$ Base(

Construction If $\mathbb{E}$ is a category with attributes then define a contextual category $\psi(\mathbb{E})$ as follows:

Step 1 The objects of $\psi(\mathbb{E})$ will be defined in such a way that each ob is an ordered pair $\langle n, A \rangle$ where n is a natural number and A is an attribute in $\mathbb{E}$.

The tree of objects of $\psi(\mathbb{E})$ is defined inductively. The least element of the tree is taken to be the ordered pair $\langle o, L(1_{\mathbb{E}}) \rangle$, where $1_{\mathbb{E}}$ the terminal object of Base $(\mathbb{E})$. Then if $\langle n, A \rangle$ is an object of $\psi(\mathbb{E})$ define the set of objects of $\psi(\mathbb{E})$ succeeding $\langle n, A \rangle$ to be the set $\{ \langle n+1, B \rangle \mid B \in \mathrm{Att}_{\mathbb{E}}(\Sigma A) \}$.

Thus an arbitrary path up through $\psi(\mathbb{E})$ is a path of the form $\langle o, L(1_{\mathbb{E}}) \rangle \triangleleft \langle 1, A_1 \rangle \ \ldots\ldots \triangleleft \langle n, A_n \rangle$ where $A_1 \in \mathrm{Att}_{\mathbb{E}}(\Sigma L 1_{\mathbb{E}})$ and for each $i, 2 \leq i \leq n$, $A_i \in \mathrm{Att}_{\mathbb{E}}(\Sigma A_{i-1})$.

Step 2. The morphisms of $\Psi(\mathbb{E})$. Define $\mathrm{Hom}_{\Psi(\mathbb{E})}(\langle n,A\rangle, \langle m,A'\rangle) = \{\langle n,m,f\rangle \mid f : \Sigma A \longrightarrow \Sigma A' \text{ in Base } (\mathbb{E})\}$. Define $\langle n,m,f\rangle \circ \langle m,\ell,g\rangle = \langle n,\ell,fg\rangle$.

Step 3. Projections. If $\langle n,A\rangle \lhd \langle n+1,B\rangle$ in $\Psi(\mathbb{E})$ then $B \in \mathrm{Att}_{\mathbb{E}}(\Sigma A)$ so we can define $p(\langle n+1,B\rangle) = \langle n+1,n,\rho(B)\rangle$.

Step 4. Pullbacks. If $\langle n,m,f\rangle : \langle n,A\rangle \longrightarrow \langle m,A'\rangle$ in $\Psi(\mathbb{E})$ and if $\langle m,A'\rangle \lhd \langle m+1,B\rangle$ then $B \in \mathrm{Att}_{\mathbb{E}}(\Sigma A')$ and $f : \Sigma A \longrightarrow \Sigma A'$ in $\mathbb{E}$ so define $\langle n,m,f\rangle^* \langle m+1,B\rangle = \langle n+1,f^*B\rangle$ and define $q(\langle n,m,f\rangle, \langle m+1,B\rangle) = \langle n+1,m+1,\xi(f,B)\rangle$. We must check that

$$\begin{array}{ccc} \langle n+1, f^*B\rangle & \xrightarrow{\langle n+1,\, m+1,\, \xi(f,B)\rangle} & \langle m+1, B\rangle \\ \downarrow & & \downarrow \\ \langle n, A\rangle & \xrightarrow[\langle n,m,f\rangle]{} & \langle m, A'\rangle \end{array}$$

is a pullback diagram in $\Psi(\mathbb{E})$. This is easily checked using the fact that

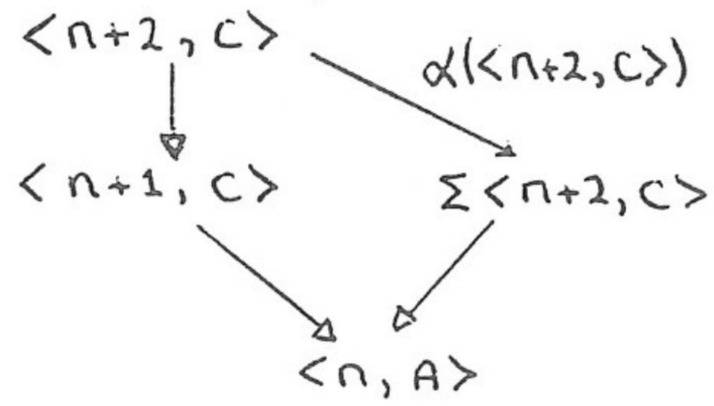
$$\begin{array}{ccc} \Sigma f^*B & \xrightarrow{\xi(f,B)} & \Sigma B \\ {\scriptstyle \rho(f^*B)}\downarrow & & \downarrow{\scriptstyle \rho(B)} \\ \Sigma A & \xrightarrow[f]{} & \Sigma A' \end{array}$$

is a pullback diagram in Base $(\mathbb{E})$.

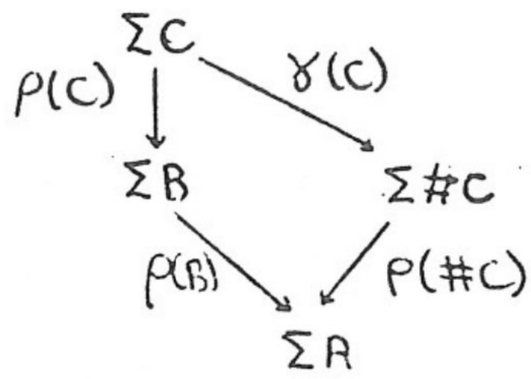
Similary the pullbacks in $\Psi(\mathbb{E})$ fit together because the corresponding pullbacks in Base $(\mathbb{E})$ fit together.

Step 4. Σ and α. Suppose that $\langle n,A\rangle \lhd \langle n+1,B\rangle \lhd \langle n+2,C\rangle$ in $\psi(\mathbb{E})$. Then $C \in Att_{\mathbb{E}}(\Sigma B)$ and $B \in Att_{\mathbb{E}}(\Sigma A)$ hence $\# C \in Att_{\mathbb{E}}(\Sigma A)$ and $\gamma(C) : \Sigma C \longrightarrow \Sigma \# C$ in Base $(\mathbb{E})$. Thus we can define $\Sigma \langle n+2,C\rangle = \langle \# C\rangle$ and $\alpha(\langle n+2,C\rangle) = \langle n+2,n+1,\gamma(C)\rangle$. Then $\alpha(\langle n+2,C\rangle)$ is an isomorphism because $\gamma(C)$ is an isomorphism.

The diagram



commutes in $\psi(\mathbb{E})$ because the diagram



commutes in Base $(\mathbb{E})$.

Finally we must show that if $\langle n,m,f\rangle : \langle n,A\rangle \longrightarrow \langle m,A'\rangle$ in $\mathbb{C}$ and if $\langle m,A'\rangle \lhd \langle m+1,B\rangle \lhd \langle m+2,C\rangle$ then $\langle n,m,f\rangle * \Sigma\langle n+2,C\rangle = \Sigma\langle n,m,f\rangle * \langle m+2,C\rangle$ and $\langle n,m,f\rangle * \alpha(\langle m+2,C\rangle$ $\alpha(\langle n,m,f\rangle * \langle m+2,C\rangle)$.

The first identity holds as follows:

$$<n,m,f>^* \Sigma <m+2,C> = <n,m,f>^* <m+1,\#C> = <n+1,f^*\#C>.$$

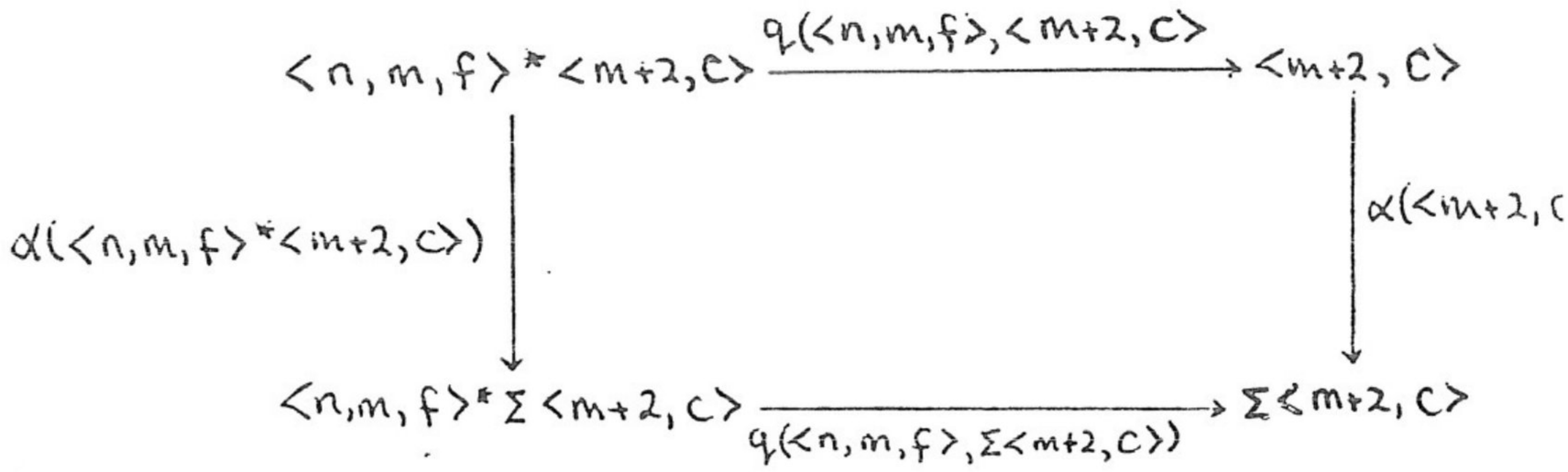
Now by condition III of the definition of Attcat, $f^*\#C = \# \S(f,B)^*C$.

Hence $<n+1,f^*\#C> = <n+1,\# \S f,B)^*C> = \Sigma <n+2, \S(f,B)^*C> =$

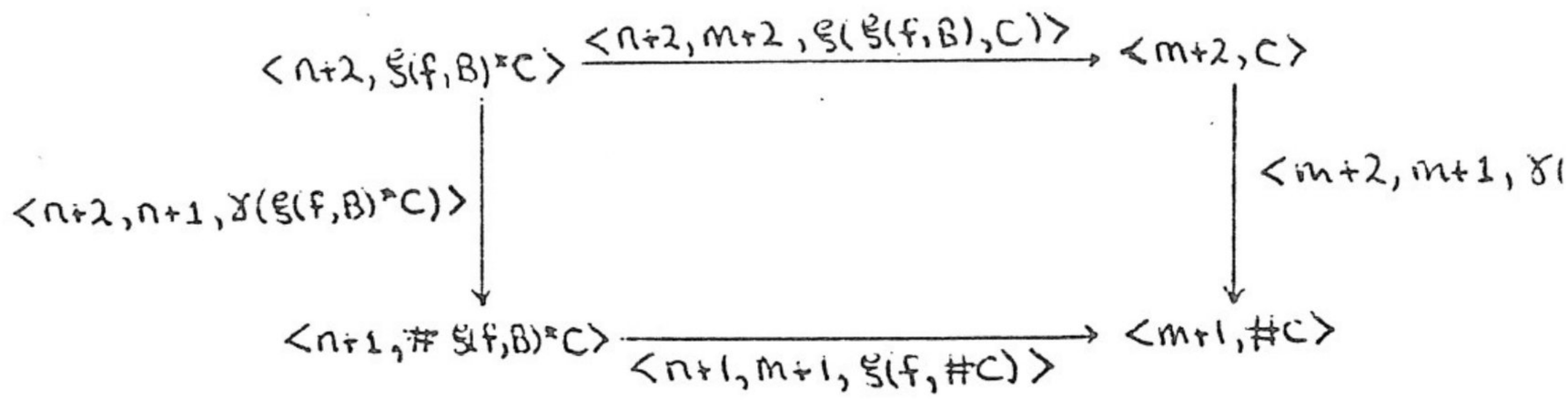
$\Sigma <n+1,m+1, \S(f,B)>^* <m+2,C> = \Sigma\, q(<n,m,f>,<m+1,B>)^* <m+2,C> =$

$\Sigma <n,m,f>^* <m+2,C>$ in the extended * notation.

By lemma 4 of § 3.1 the second identity is equivalent to the commativity of



If in this diagram the extended *,q notation is replaced, thus $q(<n,m,f>,<m+2,C>) = q(q(<n,m,f>,<m+1,B>),<m+2,C>)$ and $<n,m,f>^*<m+2,C> = q(<n,m,f>,<m+1,B>)^*<m+2,C>$. If *, q, Σ and are evaluated then the diagram required to commute is



The commutuvuty of this diagram follows from the commutivity of the corresponding diagram in Base (匡) which commutes by condition III of the definition of Attcat.

Step 5. Singleton object. Recall that in $\mathbb{E}$ we have $L(1_{\mathbb{E}}) \in Att_{\mathbb{E}}(1_{\mathbb{E}})$ and an isomorphism $\Theta(1_{\mathbb{E}}) : \Sigma L(1_{\mathbb{E}}) \longrightarrow 1_{\mathbb{E}}$ in Base ($\mathbb{E}$). It follows that $\Theta(1_{\mathbb{E}})^* L(1_{\mathbb{E}}) \in Att_{\mathbb{E}}(\Sigma L(1_{\mathbb{E}}))$ and thus th $\langle o, L(1_{\mathbb{E}})\rangle \lhd \langle 1, \Theta(1_{\mathbb{E}})^* (1_{\mathbb{E}})\rangle$ in $\psi(\mathbb{E})$. Also since $\Theta(1_{\mathbb{E}})$ is an isomorphism in Base ($\mathbb{E}$) it follows that $\Sigma\Theta(1_{\mathbb{E}})^* L(1_{\mathbb{E}})$ is isomorphic to $\Sigma L(1_{\mathbb{E}})$. Thus $\Sigma\Theta(1_{\mathbb{E}})^* L(1_{\mathbb{E}}) \simeq \Sigma L(1_{\mathbb{E}}) \simeq 1_{\mathbb{E}}$. Thus $\Sigma\Theta(1_{\mathbb{E}})^* L(1_{\mathbb{E}})$ is a terminal object of Base ($\mathbb{E}$).

It now follows that $\langle 1, \Theta(1_{\mathbb{E}})^* L(1_{\mathbb{E}})\rangle$ is a terminal object of $\psi(\mathbb{E})$. We can define the singleton object $\{\cdot\}$ of $\psi(\mathbb{E})$ to be $\langle 1, \Theta(1_{\mathbb{E}})^* L(1$

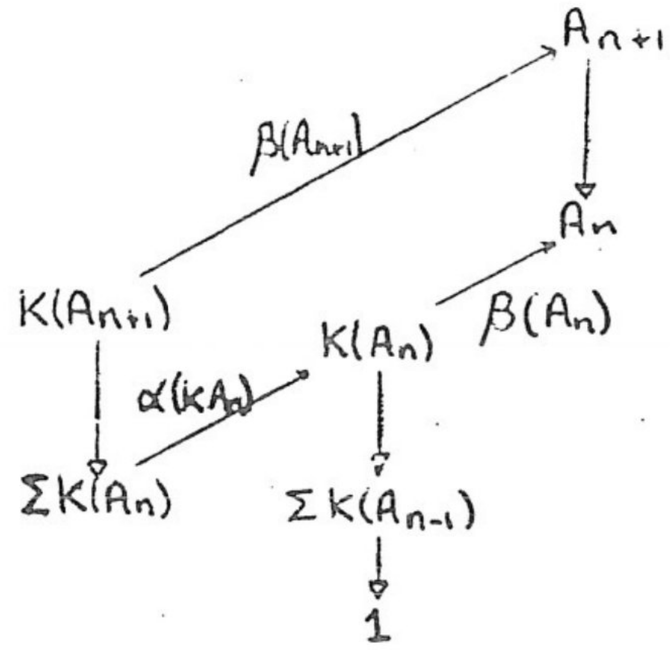
This completes the description of the contextual category with disjoint unions and singleton object induced by a category with attributes $\mathbb{E}$. Because the morphisms of the categories Attcat and Σ-con are in both cases just the structure preserving functions it follows that the above construction is just the object part of a functor ψ: Attcat $\longrightarrow$ Σ-Con.

Lemma 1. If $\mathbb{D}$ is a contextual category with disjoint unions and singleton object then $\psi(\phi(\mathbb{D})) \simeq \mathbb{D}$.

Proof. Let $S = \{ B \in |\mathbb{D}| \mid$ there is a D such that $1 \lhd D \lhd B$ in $\mathbb{D}\}$. Now define a function $K : |\mathbb{D}| \longrightarrow S$ and simultane define a morphism $\beta(A) : K(A) \longrightarrow A$ in $\mathbb{D}$, for each object A of $\mathbb{D}$.

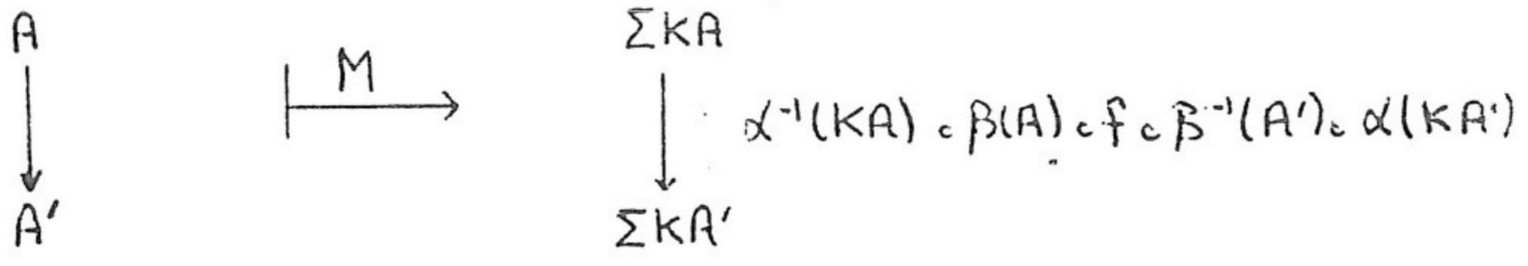
K(A) and $\beta(A)$ are defined by induction on the height of A in $\mathbb{D}$. Define $K(1) = p(\{\cdot\})^* \{\cdot\}$ and define $\beta(1) = p(K(1),1)$.

If $1 \lhd A_1 \ldots \lhd A_n \lhd A_{n+1}$ in $\mathbb{D}$ for some $n \geq 0$ and if $K(A_n)$ and $\beta(A_n)$ have already been defined then define $K(A_{n+1}) = \alpha^{-1}(K(A_n)) * \beta(A_n) * A_{n+1}$ and define $\beta(A_{n+1}) = q(\alpha^{-1}(K(A_n)) \circ \beta(A_n), A_{n+1})$. Thus $\Sigma K(A_n) \lhd K(A_{n+1})$ and $\beta(A_{n+1}) : K(A_{n+1}) \longrightarrow A_{n+1}$ in $\mathbb{D}$.



It is easy to prove by induction that for each n, $B(A_n)$ is an isomorphism (using lemma 6 of §3.1, $\beta^{-1}(A_{n+1}) = q(\beta^{-1}(A_n) \circ \alpha(K(A_n)), K(A_{n+1}))$.).

Now define a functor $M : \mathbb{D} \longrightarrow$ Base ($\mathbb{D}$) by



Finally define a would be contextual functor $\mathfrak{F} : \mathbb{D} \longrightarrow \Psi(\Phi(i\mathbb{D}))$ as follows:

If $1 \lhd A_1 \ldots \lhd A_n$ in $\mathbb{D}$ then $\mathfrak{F}(A_n) = \langle n, K(A_n) \rangle$

If $1 \lhd A_1 \ldots \lhd A_n$ and $1 \lhd B_1 \ldots \lhd B_m$ in $\mathbb{D}$ and if $f : A_n \longrightarrow B_m$ in $\mathbb{D}$ then $\mathfrak{F}(f) = \langle n, m, M(f) \rangle$.

The result follows when we show that $\mathbb{F}$ is a contextual functor, serves disjoint unions and singleton object and is 1-1 and onto from objects and morphisms of $\mathbb{D}$ to objects and morphisms of $\psi(\varphi(\mathbb{D}))$. We content ourselves with checking the bit about disjoint unions. If anything the other stuff is a bit easier to check.

We show that whenever $1 \lhd Q_1 \ldots \lhd Q_n \lhd A \lhd B \lhd C$ in $\mathbb{D}$ then $\mathbb{F}(\Sigma C) = \Sigma \mathbb{F}(C)$ and $\mathbb{F}(\alpha(C)) = \alpha(\mathbb{F}(C))$.

$\mathbb{F}(C) = \langle n+2.\ K(C)\rangle$ hence by definition of ψ,

$\Sigma\mathbb{F}(C) = \langle n+1, \#K(C)\rangle$ where $\#K(C)$ is calculated in $\varphi(\mathbb{D})$.

But by definition of $\varphi(\mathbb{D})$, $\#K(C)$ is calculated in $\mathbb{D}$, as in §3.1, and is given by $\#K(C) = \Sigma(\alpha(K(B))^*K(C))$. Thus to show that $\mathbb{F}(\Sigma C) = \Sigma\mathbb{F}(C)$ we must show just $\Sigma\alpha(K(B))^*K(C) = K(\Sigma C)$. This is as follows:

$$\begin{aligned}
\Sigma\alpha(K(B))^*K(C) &= \Sigma\,\alpha(K(B))^*(\alpha^{-1}(K(B))\circ\beta(B))^*C && \text{by def. of K}\\
&= \Sigma\beta(B)^*C\\
&= \Sigma\, q(\alpha^{-1}(K(A))\circ\beta(A), B)^*C && \text{by def. of B}\\
&= \Sigma(\alpha^{-1}(K(A))\circ\beta(A))^*C && \text{replacing extend}\\
& && \text{* notation}\\
&= (\alpha^{-1}(K(A))\circ\beta(A))^*\Sigma C && \text{Since } \alpha^{-1}(K(A))\\
& && :\Sigma KA \longrightarrow A \text{ and}\\
&= K(\Sigma C) && \text{by def. of K}
\end{aligned}$$

It remains to show that $\mathbb{F}(\alpha(C)) = \alpha(\mathbb{F}(C))$. Well $\mathbb{F}(\alpha(C))$ is just $\langle n+2, n+1, M(\alpha(C))\rangle$ and $\alpha(\mathbb{F}(C)) = \alpha(\langle n+2, K(C)\rangle$ $\alpha(\langle n+2, K(C)\rangle) = \langle n+2, n+1, \gamma(K(C))\rangle$ where $\gamma(K(C))$ is calculated in $\mathbb{D}$ as

$\alpha^{-1}(K(C))\circ\sigma(K(C))\circ\alpha(\#K(C))$. Thus we must just show that in $\mathbb{D}$, $M(\alpha(C)) = \alpha^{-1}(K(C))\circ\sigma(K(C))\circ\alpha(\#K(C))$.

Well we have $\alpha^{-1}(K(C))\circ\sigma(K(C))\circ\alpha(\#K(C)) = \alpha^{-1}(K(C))\circ q(\alpha^{-1}(K(B)$ $\alpha(K(B))*K(C))\circ\alpha(\alpha^{-1}(K(B))*K(C))\circ\alpha(\#K(C))$, (by definition of σ), $=$ $\alpha^{-1}(K(C))\circ q(\alpha^{-1}(K(B)),\beta(B)*C)\circ\alpha(\beta(B)*C)\circ\alpha(K(\Sigma C))$, since as above $\alpha(K(B)*K(C)) = \beta(B)*C$ and $\#K(C) = K(\Sigma C)$.

And $M(\alpha(C)) = \alpha^{-1}(K(C))\circ\beta(C)\circ\alpha(C)\circ\beta^{-1}(\Sigma C)\circ\alpha(K(\Sigma C))$ so we must check that $\beta(C)\circ\alpha(C)\circ\beta^{-1}(\Sigma C) = q(\alpha^{-1}(K(B)),\beta(B)*C)\circ\alpha(\beta(B)*C)$. This is as follows:

$\beta(C)\circ\alpha(C)\circ\beta^{-1}(\Sigma C) = \beta(C)\circ\alpha(C)\circ q(\beta^{-1}(A)\circ\alpha(K(A)),$ $(\alpha^{-1}(K(A))\circ\beta(A))*\Sigma C)$, by definition of β^{-1}, $= \beta(C)\circ\alpha((\beta^{-1}(A)\circ$ $\alpha(K(A)))*((\alpha^{-1}(K(A)\circ\beta(A))*C)\circ q(\beta^{-1}(A)\circ\alpha(K(A)),\Sigma(\alpha^{-1}(K(A))\circ\beta(A))*C =$ $\beta(C)\circ q(\beta^{-1}(A)\circ\alpha(K(A)),(\alpha^{-1}(K(A))\circ\beta(A))*C)\circ\alpha((\alpha^{-1}(K(A))\circ\beta(A))*C,$ by lemma 4 of §3.1, $= \beta(C)\circ q(\beta^{-1}(A)\circ\alpha(K(A)),q(\alpha^{-1}(K(A))\circ\beta(A),B)*C)\circ$ $\alpha(q(\alpha^{-1}(KA)\circ\beta(A),B)*C) = \beta(C)\circ q(\beta^{-1}(A)\circ\alpha(K(A)),\beta(B)*C)\circ\alpha(\beta(B)*C) =$ $\beta(C)\circ q(q(\beta^{-1}(A)\circ\alpha(K(A)),K(B)),\beta(B)*C) = \beta(C)\circ q(\beta^{-1}(B),\beta(B)*C)\circ\alpha(\beta(B)*C)$ $q(\alpha^{-1}(K(B))\circ\beta(B),C)\circ q(\beta^{-1}(B),\beta(B)*C)\circ\alpha(\beta(B)*C) = q(\alpha^{-1}(K(B)),\beta(B)*C)\circ\alpha($ $B)*C)$ as required.

Lemma 2. If $\mathcal{Q}: \mathbb{D} \longrightarrow \mathbb{D}'$ is a morphism in Σ-Con then the diagram

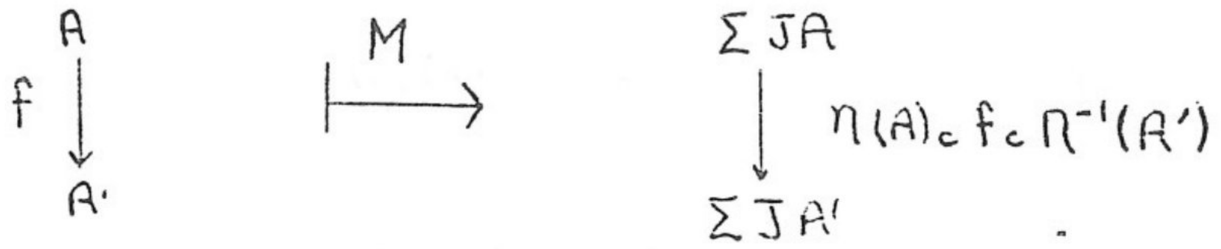
$$\begin{array}{ccc} \mathbb{D} & \xrightarrow{\sim} & \psi(\varphi(\mathbb{D})) \\ \mathcal{Q}\downarrow & & \downarrow\psi(\varphi(\mathcal{Q})) \\ \mathbb{D}' & \xrightarrow{\sim} & \psi(\varphi(\mathbb{D}')) \end{array}$$

commutes in Σ-Con.

Proof Reduces to showing that if $A \in |\mathbb{D}|$ then $\mathcal{G}(K(A)) = K(\mathcal{G}(A))$ and th $f : A \longrightarrow A'$ in $\mathbb{D}$ then $\mathcal{G}(M(f)) = M(\mathcal{G}(f))$. But $\mathcal{G}$ preserves the structur that K and M are defined in terms of .

This completes the demonstation that $\Phi \circ \Psi \simeq id_{\Sigma\text{-Con}}$. For the remainder of this section we aim at proving that $\Psi \circ \Phi \simeq id_{\underline{Attca}}$ and thus that Σ-Con and Attcat are equivalent categories. Throughout we suppose that $\mathbb{E}$ is a category with attributes. Eventually we show that $\Phi(\Psi(\mathbb{E})) \simeq \mathbb{E}$.

If $A \in |\text{Base}(\mathbb{E}))|$ then $\Theta(1_{\mathbb{E}})^*L(A) \in \text{Att}_{\mathbb{E}}(\Sigma L(1_{\mathbb{E}}))$ and $\xi(\Theta(1_{\mathbb{E}}), L(A)) \circ \Theta(A) : \Sigma\, \Theta(1_{\mathbb{E}})^*L(A) \longrightarrow A$ is an isomorphism. We define $J(A) = \Theta(1_{\mathbb{E}})^*L(A)$ and $\eta(A) = \xi(\Theta(1_{\mathbb{E}}), L(A)) \circ \Theta(A)$ then for each $A \in |\text{Base}(\mathbb{E})|$, $J(A) \in \text{Att}_{\mathbb{E}}(\Sigma L(1_{\mathbb{E}}))$ and $\eta(A) : \Sigma JA \longrightarrow A$ is an isomorphism in Base ($\mathbb{E}$). We can then define a functor M : Base ($\longrightarrow$Base ($\mathbb{E}$) by



We shall eventually define an isorphism $\mathcal{J} : \mathbb{E} \longrightarrow \Phi(\Psi(\mathbb{E}))$ in terms of M and η but we will come back to that later. Lemmas 3-9 just function to show that $\mathcal{J}$ once defined, is a morphism.

Lemma 3. If A is an object of Base ($\mathbb{E}$) and if $B \in Att_{\mathbb{E}}(A)$ then

$\#(\xi(\theta(1_{\mathbb{E}}),LA)^* \theta(A)^*B) = \theta(1_{\mathbb{E}})^* \#(\theta(A)^*B)$ and

$\gamma(\eta(A)^*B) \circ \xi(\theta(1_{\mathbb{E}}),L\Sigma B) = \xi(\xi(\theta(1_{\mathbb{E}}),LA), \theta(A)^*B) \circ \gamma(\theta(A)^*B)$ i

Base ($\mathbb{E}$).

Proof This is just condition III of the defintion of Attcat where $A \xrightarrow{f} A'$, B and C are taken to be $\Sigma L(1_{\mathbb{E}}) \xrightarrow{\theta(1_{\mathbb{E}})} 1_{\mathbb{E}}$, LA and $\theta(A)$

Lemma 4. If A is an object of Base ($\mathbb{E}$) and if $B \in Att_{\mathbb{E}}(A)$ then

$\gamma(\eta(A)^*B) \circ \eta(\Sigma B) = \xi(\eta(A),B)$.

Proof

$\gamma(\eta(A)^*B) \circ \eta(\Sigma B) = \gamma(\eta(A)^*B) \circ \xi(\theta(1_{\mathbb{E}}),L\Sigma B) \circ \theta(\Sigma B)$ by def of η

$= \xi(\xi(\theta(1_{\mathbb{E}}),LA), \theta(A)^*B) \circ \gamma(\theta(A)^*B) \circ \theta(\Sigma B)$ by lemma 1.

$= \xi(\xi(\theta(1_{\mathbb{E}}),LA), \theta(A)^*B) \circ \xi(\theta(A),B)$ by condition of def of At

$= \xi(\xi(\theta(1_{\mathbb{E}}),LA) \circ \theta(A),B)$

$= \xi(\theta(A),B)$ by def of η

Lemma 5. If A is an object of Base ($\mathbb{E}$) and $B \in Att_{\mathbb{E}}(A)$ then

$J(\Sigma B) = \#(\eta(A)^*B)$

and $\eta(\Sigma B) \circ \rho(B) \circ \eta^{-1}(A) = \gamma^{-1}(\eta(A)^*B) \circ \rho(\eta(A)^*B)$.

Proof

$J(\Sigma B) = \theta(1_{\mathbb{E}})^* L(\Sigma B)$ by def of J

$= \theta(1_{\mathbb{E}})^* \# \theta(A)^*B$ by condition

$= \#(\xi(\theta(1_{\mathbb{E}}),LA)^* \theta(A)^*B)$ by lemma 1

$= \#(\eta(A)^*B)$ by definition

Which is the first identity. As for the second:

$$\eta(\Sigma B)\circ\rho(B)\circ\eta^{-1}(A) = \gamma^{-1}(\eta(A)^*B)\circ \xi(\eta(A),B)\circ\rho(B)\circ\eta^{-1}(A) \quad \text{by lemma 2}$$

$$= \gamma^{-1}(\eta(A)^*B)\circ\rho(\eta(A)^*B)\circ\eta(A)\circ\eta^{-1}(A) \quad \text{because of commutivit the pullba diagram fo B along } \eta($$

$$= \gamma^{-1}(\eta(A)^*B)\circ\rho(\eta(A)^*B) \quad \text{as require}$$

Lemma 6. If $f : A \longrightarrow A'$ in Base ($\mathbb{E}$) and $B \in Att_{\mathbb{E}}(A')$ then $\eta(\Sigma f^*B)\circ \xi(f,B)\circ\eta^{-1}(\Sigma B) = \gamma^{-1}(\eta(A)^*f^*B)\circ \xi(\eta(A)\circ f\circ\eta^{-1}(A'),\eta(A')^*B)\circ \gamma(\eta(A)^*B)$.

Proof. $\eta(\Sigma f^*B)\circ \xi(f,B)\circ\eta^{-1}(\Sigma B) = \xi(\Theta(1_{\mathbb{E}}),L(\Sigma f^*B))\circ\Theta(f^*B)\circ \xi(f\circ\Theta^{-1}(\Sigma B)\circ \xi(\Theta^{-1}(1_{\mathbb{E}}), \Theta(1_{\mathbb{E}})^*L(\Sigma B))$, by def of η.

$= \xi(\Theta(1_{\mathbb{E}}),L(\Sigma f^*B))\circ \gamma^{-1}(\Theta(A)^*f^*B)\circ \xi(\Theta(A),f^*B)\circ \xi(f,B)\circ \xi(\Theta^{-1}(A'), \Theta(A')^*B)\circ \gamma(\Theta(A')^* \quad \xi(\Theta^{-1}(1_{\mathbb{E}}), \Theta(1_{\mathbb{E}})^*L(\Sigma B))$, using condition IV of attcat.

$= \gamma^{-1}(\eta(A)^*f^*B)\circ \xi(\xi(\Theta(1_{\mathbb{E}}),LA), \Theta(A)^*f^*B)\circ \xi(\Theta(A),f^*B)\circ \xi(f,B)\circ \xi(\Theta^{-1}(A'), \Theta(A')^*B)\circ \xi(\xi(\Theta^{-1}(1_{\mathbb{E}}), \Theta(1_{\mathbb{E}})^*A'),\eta(A')^*B)\circ \gamma^{-1}(\eta(A')^*B,$ by lemma 1.

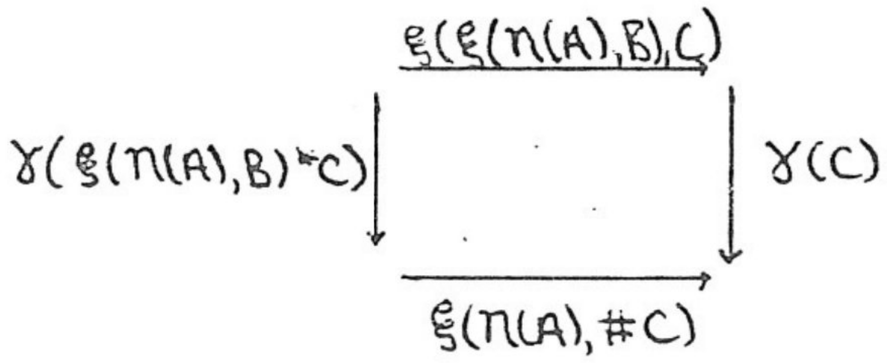
$= \gamma^{-1}(\eta(A)^*f^*B)\circ \xi(\eta(A)\circ f\circ\eta(A'),\eta(A')^*B)\circ \gamma(\eta(A')^*B)$, by fitting the pullbacks together.

Lemma 7. If A is an object of Base ($\mathbb{E}$) and if $B \in Att_{\mathbb{E}}(A)$ and $C \in Att_{\mathbb{E}}(\Sigma B)$ then $\eta(A)^*\#C = \#\gamma(\eta(A)^*B)^*\eta(\Sigma B)^*C$ and

$\eta(\Sigma C)\circ\gamma(C)\circ\eta^{-1}(\Sigma\#C) = \gamma^{-1}(\eta(\Sigma B)^*C)\circ \xi(\gamma(\eta(A)^*B),\gamma(\eta(A)^*B)^*\eta(\Sigma B)^*$

$\circ\gamma(\gamma(\eta(A)^*)^*\eta(\Sigma B)^*C)\circ\gamma(\#(\gamma(\eta(A)^*B)^*\eta(\Sigma B)^*C)$.

Proof Use condition III of attcat at $\eta(A) : \Sigma JA \longrightarrow A$ and $B \in Att_{\mathbb{E}}(A), C \in Att_{\mathbb{E}}(\Sigma B)$. We get $\eta(A) * \# C = \# \xi(\eta(A), B) * C)$ and



commutes.

The first identity follows from lemma 2 which says that $\gamma(\eta(A) * B) \circ \eta(\Sigma B) = \xi(\eta(A), B)$.

The R.H.S. of the second identity is simplified by the same, lemma It becomes $\gamma^{-1}(\eta(\Sigma B) * C) \circ \xi(\gamma^{-1}(\eta(A) * B), \xi(\eta(A), B) * C) \circ \gamma(\xi(\eta(A), B) * C$ $\gamma(\# \xi(\eta(A), B) * C)$. Using both the above identities from condition III the R.H.S. becomes $\gamma^{-1}(\eta(\Sigma B) * C) \circ \xi(\gamma^{-1}(\eta(A) * B) \circ \xi(\eta(A), B)C) \circ$ $\gamma(C) \circ \xi(\eta^{-1}(A), \eta(A), \eta(A) * \# C) \circ \gamma(\eta(A) * \# C)$.

Use lemma 2 twice and we get the L.H.S.

Lemma 8. If A is an object of Base ($\mathbb{E}$) then $\eta(\Sigma LA) \circ \Theta(A) \circ \eta^{-1}(A) = \gamma^{-1}((\rho(\Theta(1_{\mathbb{E}}) * L(1_{\mathbb{E}})) * L(A)) \circ \xi(\rho(\Theta(1_{\mathbb{E}}) *$ $L(1_{\mathbb{E}})), \Theta(1_{\mathbb{E}}) * LA)$.

Proof Both $\rho(\Theta(1_{\mathbb{E}}) * L(1_{\mathbb{E}}))$ and $\xi(\Theta(1_{\mathbb{E}}), L(1_{\mathbb{E}})) :$ $\Sigma J(1_{\mathbb{E}}) \longrightarrow \Sigma L(1_{\mathbb{E}})$ in Base ($\mathbb{E}$). Since $\Sigma L(1_{\mathbb{E}})$ is terminal, $\rho(\Theta(1_{\mathbb{E}}) * L(1_{\mathbb{E}})) = \xi(\Theta(1_{\mathbb{E}}), L(1_{\mathbb{E}}))$. Now the R.H.S. of required identity becomes $\gamma^{-1}(\eta(1_{\mathbb{E}}) * L(A)) \circ \xi(\xi(\Theta(1_{\mathbb{E}}), L(1_{\mathbb{E}})), \Theta(1_{\mathbb{E}}) * L(A))$.

Using lemma 1 on LHS we get $\xi(\theta(1_{\mathbb{E}}), L\Sigma LA)\circ\theta(\Sigma LA)\circ\theta(A)\circ\theta^{-1}(A)\circ\xi(\theta^{-1}(1_{\mathbb{E}}), \theta(1_{\mathbb{E}})^{*}LA)$.

Using the same lemma 1 on R.H.S. we get

$\xi(\theta(1_{\mathbb{E}}), L\Sigma LA)\circ\gamma^{-1}(\theta(1_{\mathbb{E}})^{*}LA)$.

Now LHS = RHS follows from the case of condition IV :

$\gamma^{-1}(\theta(1_{\mathbb{E}})^{*}LA) = \theta(\Sigma LA)\circ\xi(\theta^{-1}(1_{\mathbb{E}}), \theta(1_{\mathbb{E}})^{*}LA)$.

Before we can define $\mathcal{J} : \mathbb{E} \longrightarrow \varphi(\psi(\mathbb{E}))$ we must know what $\varphi(\psi(\mathbb{E}))$ looks like. We know what the $^{*}, q, \Sigma, \alpha$ structure of $\psi(\mathbb{E})$ is 1 The $\xi, \#, \gamma, L, \theta$ structure is defined in terms of this and induces the category with attributes $\varphi(\psi(\mathbb{E}))$. The next lemma calculates the effects of the definition of $\xi, \#, \gamma, L, \theta$ in $\psi(\mathbb{E})$.

Lemma 9. If $\mathbb{E}$ is a category with attributes then

i. If $\langle 1,1,f\rangle : \langle 1,A\rangle \longrightarrow \langle 1,A'\rangle$ and $\langle 1,A'\rangle \triangleleft \langle 2,B\rangle$ in $\psi(\mathbb{E})$ then $\xi(\langle 1,1,f\rangle, \langle 2,B\rangle) = \langle 1,1, \gamma^{-1}(f^{*}B)\circ\xi(f,B)\circ\gamma(B)\rangle$.

ii. If $1 \triangleleft \langle 1,A\rangle \triangleleft \langle 2,B\rangle$ and $\Sigma\langle 2,B\rangle \triangleleft \langle 2,C\rangle$ in $\psi(\mathbb{E})$ then

$\#\langle 2,C\rangle = \langle 2, \#\gamma(B)^{*}C\rangle$ and

$\gamma(\langle 2,C\rangle) = \langle 1,1, \gamma^{-1}(C)\circ\xi(\gamma^{-1}(B), \gamma(B)^{*}C)\circ\gamma(\gamma(B)^{*}C)\circ\gamma(\#\gamma(B$

iii. If $1 \triangleleft \langle 1,A\rangle$ in $\psi(\mathbb{E})$ then $L(\langle 1,A\rangle) = \langle 2, \rho(\theta^{-1}(1_{\mathbb{E}})^{*}L(1_{\mathbb{E}}))^{*}A\rangle$ and $\theta(\langle 1,A\rangle) = \langle 1,1, \gamma^{-1}(\rho(\theta(1_{\mathbb{E}})^{*}L(1_{\mathbb{E}}))^{*}A)\circ\xi(\rho(\theta(1_{\mathbb{E}})), A)$.

Proof i. $\xi(\langle 1,1,f\rangle, \langle 2,B\rangle) = \alpha^{-1}(\langle 1,1,f\rangle^{*}\langle 2,B\rangle)\circ q(\langle 1,1,f\rangle, \langle 2,B\rangle)\circ\alpha(\langle 2,B\rangle)$, by definition of ϕ.

$= \alpha^{-1}(\langle 2,f^{*}B\rangle)\circ\langle 2,2, \xi(f,B)\rangle\circ\alpha(\langle 2,B\rangle)$ by def of ψ.

$= \langle 1,2, \gamma^{-1}(f^{*}B)\rangle\circ\langle 2,2, \xi(f,B)\rangle\circ\langle 2,1, \gamma(B)\rangle$, by def. of ψ .

$= \langle 1,1, \gamma^{-1}(f^{*}B)\circ\xi(f,B)\circ\gamma(B)\rangle$.

Cases ii. and iii. are just as straightforward.

Given the definition of φ we can now give a complete description of $\varphi(\psi(\mathbb{E}))$ as follows:

$|\text{Base}(\varphi(\psi(\mathbb{E}))| = \{ \langle 1,A\rangle \mid a \in \text{Att}_{\mathbb{E}}(\Sigma L(1_{\mathbb{E}})) \}$

$\text{Hom}_{\varphi(\psi(\mathbb{E}))}(\langle 1,A\rangle, \langle 1,A'\rangle) = \{ \langle 1,1,f\rangle \mid f : \Sigma A \longrightarrow \Sigma A' \text{ in Base } (\mathbb{E}$

$\text{Att}_{\varphi(\psi(\mathbb{E}))}(\langle 1,A\rangle) = \{\langle 2,B\rangle \mid B \in \text{Att}_{\mathbb{E}}(\Sigma A)\}$

$1_{\varphi(\psi(\mathbb{E}))} = \langle 1, \Theta^{-1}(1_{\mathbb{E}})^* L(1_{\mathbb{E}})\rangle$

If $\langle 2,B\rangle \in \text{Att}_{\varphi(\psi(\mathbb{E}))}(\langle 1,A\rangle)$ then $\Sigma\langle 2,B\rangle = \langle 1,\#B\rangle$ and $\rho(\langle 2,B\rangle) = \langle 1,1, \gamma^{-1}(B)\circ\rho(B)\rangle$.

If $\langle 2,B\rangle \in \text{Att}_{\varphi(\psi(\mathbb{E}))}(\langle 1,A\rangle)$ and if $\langle 2,C\rangle \in \text{Att}_{\varphi(\psi(\mathbb{E}))}(\Sigma\langle 2,B$

then $\#\langle 2,C\rangle = \langle 2,\#\gamma(B)^*C\rangle$ and $\gamma(\langle 2,C\rangle) = \langle 1,1,$ $\gamma^{-1}(C)\circ\xi(\gamma^{-1}(B), \gamma(B)^*C)\circ\gamma(\gamma(B)^*C)\circ\gamma(\#\gamma(B)^*C)\rangle$.

If $\langle 1,1,f\rangle : \langle 1,A\rangle \longrightarrow \langle 1,A'\rangle$ in Base $(\mathbb{E})$ and if $\langle 2,B\rangle$ $\in \text{Att}_{\varphi(\psi(\mathbb{E}))}(\langle 1,A'\rangle)$ then $\langle 1,1,f\rangle^* \langle 2,B\rangle = \langle 2,f^*B\rangle$ and $\xi(\langle 1,1,f\rangle, \langle 2,B\rangle) = \langle 1,1, \delta^{-1}(f^*B)\circ\xi(f,B)\circ\gamma(B)\rangle$.

If $\langle 1,A\rangle \in |\text{Base}(\varphi(\psi(\mathbb{E})))|$ then $L(\langle 1,A\rangle) = \langle 2,\rho(\Theta(1_{\mathbb{E}})^*$ $L(1_{\mathbb{E}}))^*A\rangle$ and $\Theta(\langle 1,A\rangle) = \langle 1,2,\gamma^{-1}(\rho(\Theta(1_{\mathbb{E}})^*L(1_{\mathbb{E}}))^*A)\circ$ $\xi(\rho(\Theta(1_{\mathbb{E}})^*L(1_{\mathbb{E}})),A)\rangle$.

Define $\mathfrak{J} : \mathbb{E} \longrightarrow \varphi(\psi(\mathbb{E}))$ by i. if A is an object of Base $(\mathbb{E})$ then $\mathfrak{J}(A) = \langle 1, \Theta(1_{\mathbb{E}})^*L(A)\rangle$. ii. If $f : A \longrightarrow A'$ in Base $(\mathbb{E})$ then $(f) = \langle 1,1,\eta(A)\circ f\circ\eta^{-1}(A)\rangle$. iii. If $B \in \text{Att}_{\mathbb{E}}(A)$ then $\mathfrak{J}(B) = \langle 2,\eta(A)^*B\rangle$.

$\mathfrak{J}$ is easily seen to be 1-1 and onto from objects, morphisms and attributes of $\mathbb{E}$ to objects, morphisms, respectively attributes of $\Phi(\Psi(\mathbb{E}))$. $\mathfrak{J}$ preserves Σ and ρ by virtue of lemma 3. * is easily seen to be preserved. $\mathfrak{S}$ is preserved by $\mathfrak{J}$, use lemma 4.

$\#$ and γ are preserved, use lemma 5.

Terminal object easily seen to be preserved. L preserved because $1_{\mathbb{E}}$ terminal implies $(\Theta(1_{\mathbb{E}})^{*}L(1_{\mathbb{E}}))\circ\theta(1_{\mathbb{E}}) = \eta(1_{\mathbb{E}})$.

Θ preserved - use lemma 6.

Finally, the collection of isomorphism $\{\mathfrak{J}_{\mathbb{E}} \mid \mathbb{E} \in |\underline{\text{Attcat}}|\}$ cannot help but be natural in $\mathbb{E}$. Thus we have a natural isomorphism $\Phi\circ\Psi \simeq \text{id}_{\underline{\text{Attcat}}}$, completing the proof that $\underline{\Sigma\text{-Con}} \cong \underline{\text{Attcat}}$.

3.3 Categories and Fibrations

In §1.5 we alluded to the contextual category of categories, category indexed families of categories, category indexed families of category indexed families of categories and so on. In fact this contextual category has disjoint unions and singleton object. By far the easiest way of describing the structure is by describing the corresponding category with attributes which is to be called Fib The significance of this structure is twofold. In the first place the most attractive notion of a category indexed family of categories does lead to a contextual category and thus generalised algebraic theories can be interpreted in this manner.

Further more this notion is not reducible to any morphism with codomain notion and so there is no parellel interpretation of essentially algebraic theories.

This was discussed in §1.5. On the otherhand disjoint unions in this structure are calculated by taking fibrations. Thus we have a new way of looking at the fibration construction. This compares with the interpretation of Gray [7] .

The category with attributes Fib of categories indexing categories is described as follows:

i. Base (Fib) = Cat, the category of all (small) categories.

ii. If A is a category then an attribute of type A is an A-indexed family of categories, ie is a functor $B : A \longrightarrow Cat$

iii. If $B \in Att_{\underline{Fib}}(\underline{A})$, ie if $B : \underline{A} \longrightarrow \underline{Cat}$ is a functor, then $\underline{\Sigma B}$ is the category defined as follows:

$$|\underline{\Sigma B}| = \{ \langle a,b \rangle \mid a \in |\underline{A}| \text{ and } b \in |\underline{B(a)}| \}.$$

$$\mathrm{Hom}_{\underline{\Sigma B}}(\langle a,b \rangle, \langle a',b' \rangle) = \{ \langle f,g \rangle \mid f : a \longrightarrow a' \text{ in } \underline{A} \text{ and } g : B(f)(b) \longrightarrow b' \text{ in } \underline{B(a')} \}.$$

If $\langle a,b \rangle \in |\underline{\Sigma B}|$ then $\mathrm{id}_{\langle a,b \rangle} = \langle \mathrm{id}_a, \mathrm{id}_b \rangle$.

If $\langle a,b \rangle \xrightarrow{\langle f,g \rangle} \langle a',b' \rangle \xrightarrow{\langle f',g' \rangle} \langle a'',b'' \rangle$ in $\underline{\Sigma B}$ then

$$\langle f,g \rangle \circ \langle f',g' \rangle = \langle ff', B(f')(g) \circ g' \rangle.$$

$\rho(B) : \underline{\Sigma B} \longrightarrow \underline{A}$ is defined to be the 1st projection functor.

iv. If $F : \underline{A} \longrightarrow \underline{A'}$ in $\underline{Cat}$ and if $B \in Att_{\underline{Fib}}(A')$ then $F^*B \in Att_{\underline{Fib}}(A)$ is defined to be the functor $FoB : \underline{A} \longrightarrow \underline{Cat}$. $\xi(F,B) : \underline{\Sigma F^*B} \longrightarrow \underline{\Sigma}$ is defined to be the functor "Apply F to the 1st component leave other component unchanged". For example, if $\langle a,b \rangle \in |\underline{\Sigma F^*B}|$ then $a \in |\underline{A}|$ and $b \in |\underline{B(F(a))}|$, hence $\langle Fa,b \rangle \in |\underline{\Sigma B}|$. Similarly on morphisms of $\underline{\Sigma F^*B}$.

v. If $B : \underline{A} \longrightarrow \underline{Cat}$ and $C : \underline{\Sigma B} \longrightarrow \underline{Cat}$ then define $\#C : \underline{A} \longrightarrow \underline{Cat}$ as follows:

If $a \in |\underline{A}|$ then $\underline{\#C(a)}$ is the category such that

$$|\underline{\#C(a)}| = \{ \langle b,c \rangle \mid b \in |\underline{B(a)}| \text{ and } c \in |\underline{C(\langle a,b \rangle)}| \}$$

$$\mathrm{Hom}_{\underline{\#C(a)}}(\langle b,c \rangle, \langle b',c' \rangle) = \{ \langle g,h \rangle \mid g : b \longrightarrow b' \text{ in } \underline{B(a)} \text{ and } h : C(\langle a,g \rangle)(c) \longrightarrow c' \text{ in } \underline{C(\langle a,b' \rangle)} \}.$$

If $f : a \longrightarrow a'$ in $\underline{A}$ then $\#C(f)$ is the functor

$\underline{\#C(a)} \longrightarrow \underline{\#C(a')}$ given by

$$\begin{array}{ccc} \langle b, c\rangle & & \langle B(f)(b),\ C(\langle f, id_{B(f)(b)}\rangle)(c)\rangle \\ \Big\downarrow \langle g,h\rangle & \overset{\#C(f)}{\longmapsto} & \Big\downarrow \langle B(f)(g),\ C(\langle f, id_{B(f)(b')}\rangle)(h)\rangle \\ \langle b', c'\rangle & & \langle B(f)(b'),\ C(\langle f, id_{B(f)(b')}\rangle)(c')\rangle \end{array}$$

Check that ΣC and $\Sigma \#C$ are isomorphic. In fact the isomorphism $\gamma(C) : \Sigma C \longrightarrow \Sigma \# C$ is given by

$$\begin{array}{ccc} \langle\langle a, b\rangle, c\rangle & & \langle a, \langle b, c\rangle\rangle \\ \Big\downarrow \langle\langle f,g\rangle, h\rangle & \overset{\gamma(C)}{\longmapsto} & \Big\downarrow \langle f, \langle g.h\rangle\rangle \\ \langle\langle a', b'\rangle, c'\rangle & & \langle a', \langle b', c'\rangle\rangle \end{array}$$

vi. Chose $1_{\underline{Fib}}$ to be any terminal object of $\underline{Cat}$. If $\underline{A}$ is a category then $L(\underline{A}) \in Att_{\underline{Fib}}(\underline{1})$ is taken to be the functor : $\underline{1} \longrightarrow \underline{Cat}$ whose value at the unique object of $\underline{1}$ is $\underline{A}$. Then

$\Sigma\ L(\underline{A}) = \underline{1} \times \underline{A}$.

Define $\Theta(A) : \underline{1} \times \underline{A} \longrightarrow \underline{A}$ to be the projection functor.

That completes the description of $\underline{Fib}$. Checking that $\underline{Fib}$ is a category with attributes is very straightforward and rather tedious.

3.4 Martin-Lof Type Theory

Martin-Lof type theory is a generalisation of the typed λ-calculus. It is more general in that in the syntax there are variable types - the structure of substitution is of the general form. The algebraic semantics of Martin-Lof type theory is thus provided for by a extension of the theory of contextual categories. This extension we call the theory of weak M-L structures. Every weak M-L structure is a model of Martin-Lof type theory. In fact the definition of weak M-L structure is a set theoretic definition of the notion 'model of Martin-Lof type theory'. It is the most general possible such definition.

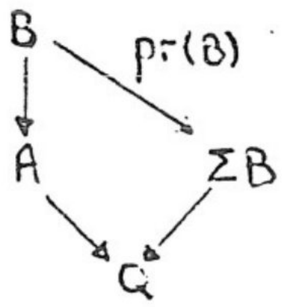
We need to explain that Martin-Lof type theory generalises a weak version of the typed λ-calculus. The η-rule of λ-calculus, $S = \lambda x.\ Ap(x,S)$, corresponding to the uniqueness of the term $\lambda x.t$ subject to the condition $Ap(x, \lambda x.t) = t$, is not assumed. Neither is the rule $pr(p_1(z),p_2(z)) = z$ assumed where pr is the pairing function and p_1 and p_2 are the projections. On the other hand cartesian closed categories correspond to a strong version of the typed λ-calculus which includes these two rules. The effect of all this is that whereas cartesian closedn is defnable in terms of universal arrows, weak M-L strucure is defined in terms of weak universal arrows (the definition of weak universal arrow i like the definition of universal arrow except that the uniqueness condition is dropped). If we strengthen the definition of weak M-L structure by replacing weak universality by universality, then we can drop the adjective weak and call the structures strong M-L structures. The strong M-L structures provide us with the algebraic semantics of a strengthened version of M-L type theory, stronger just by the inclusion of

a rule expressing the uniqueness of $\lambda x.t$ subject to the condition $Ap(x, \lambda x.t) = t$ and similar uniqueness conditions, one for each logical scheme.

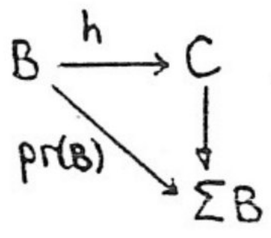
In the following defintion of strong M-L structure we use the same notation as Martin-Lof uses except in the case of notation for the type with precisely one element - Martin-Lof uses the notation N_1 for this type but we use the notation $\{\cdot\}$. Our use of Σ and $\{\cdot\}$ notation in this section will be consistent with our use of the same notation for disjoint unions and singleton object in previous sections.

A strong <u>M-L structure</u> $\langle \mathbb{C}, \Sigma, pr, \Pi, Ap, Id, r, +, i, j, \{\cdot\}, e, N, o, S \rangle$ consists of a contextual category $\mathbb{C}$ and the following additional structure:

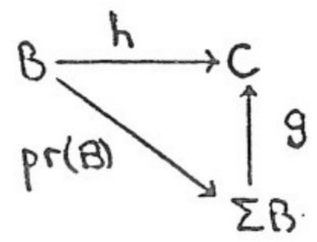
i. Whenever $Q \lhd A \lhd B$ in $\mathbb{C}$, an object ΣB of $\mathbb{C}$ and a morphism $pr(B)$ of $\mathbb{C}$ such that $Q \lhd \Sigma B$ in $\mathbb{C}$ and $pr(B) : B \longrightarrow \Sigma B$ such that the diagram



commutes and such that ΣB and $pr(B)$ have the following property: for every object C of $\mathbb{C}$ such that $\Sigma B \lhd C$, for every morphism $h : B \longrightarrow C$ such that the diagram

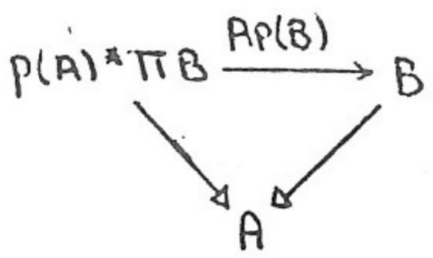


commutes, there exists a unique morphism $g \in \mathrm{Arr}_{\mathbb{C}}(C)$* such that the diagram

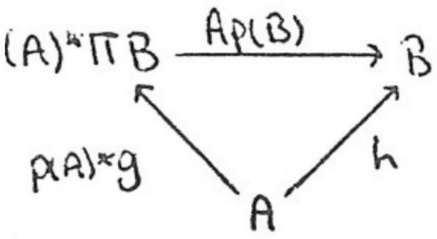


commutes.

ii. Whenever $Q \lhd A \lhd B$ in $\mathbb{C}$, an object ΠB of $\mathbb{C}$ such that $Q \lhd \Pi B$ in $\mathbb{C}$ and a morphism $Ap(B) : p(A)^*\Pi B \longrightarrow B$ in $\mathbb{C}$ such that the diagram



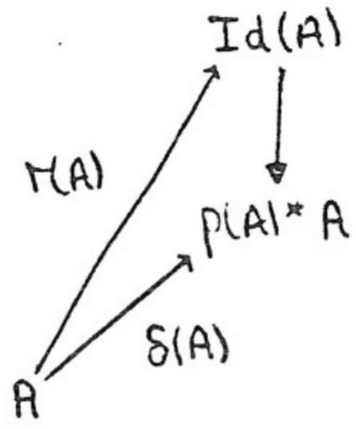
commutes and such that ΠB and $Ap(B)$ have the following property: for every morphism $h \in \mathrm{Arr}_{\mathbb{C}}(B)$, there exists a unique morphism $f \in \mathrm{Arr}_{\mathbb{C}}(\Pi B)$ such that the diagram



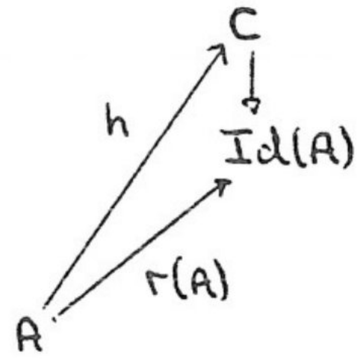
commutes. In future the unique morphism corresponding to $h : A \longrightarrow B$ will be denoted λh. Thus $\lambda h \in \mathrm{Arr}_{\mathbb{C}}(\Pi B)$ when $h : A \longrightarrow B$.

iii. Whenever $Q \lhd A$ in $\mathbb{C}$, an object IdA of $\mathbb{C}$ such that $p(A)^*A \lhd \mathrm{Id}A$ and a morphism $r(A) : A \longrightarrow \mathrm{Id}A$ in $\mathbb{C}$ such that the diagram

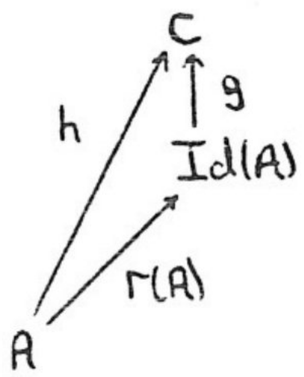
* The definition of $\mathrm{Arr}_{\mathbb{C}}$ is on page 2.19.



commutes and such that IdA and r(A) have the following property: for every object A of $\mathbb{C}$ such that IdA $\lhd$ C, for every morphism h: A $\longrightarrow$ C such that the diagram



commutes, there exists a unique morphism $g \in Arr_{\mathbb{C}}(C)$ such that the diagram



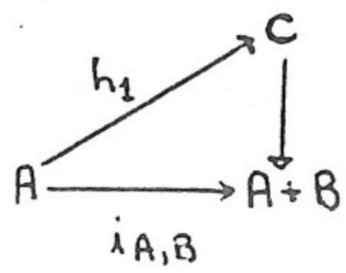
commutes.

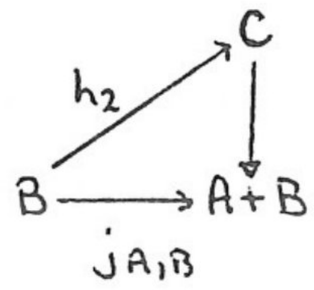
iv. Whenever $Q \lhd A$ and $Q \lhd B$ in $\mathbb{C}$, an object A+B of $\mathbb{C}$ such that $Q \lhd A+B$ and morphisms $i_{A,B}: A \longrightarrow A+B$ and $j_{A,B}: B \longrightarrow A+B$ such that the diagrams



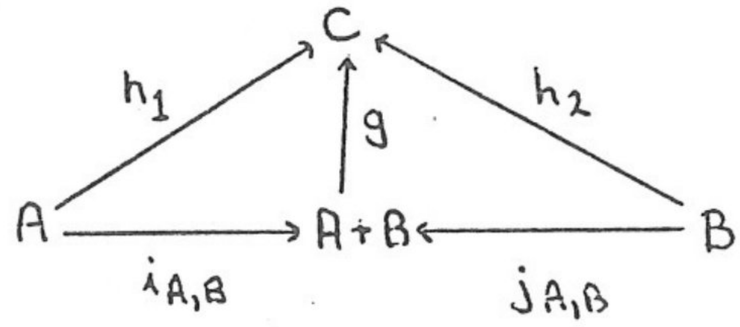
commute and such that A+B, $i_{A,B}$, $j_{A,B}$ have the following property:

for every object C of $\mathbb{C}$ such that $A+B \triangleleft C$, for all pairs of morphisms $h_1 : A \longrightarrow C$ and $h_2 : B \longrightarrow C$ such that the diagrams



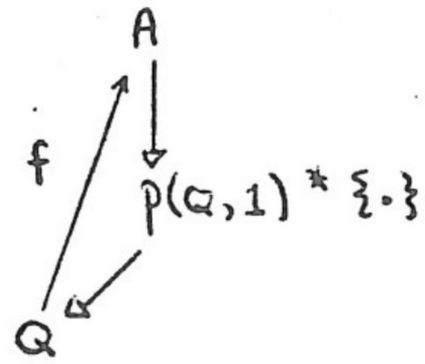


commute, there exists a unique $g \in \mathrm{Arr}_{\mathbb{C}}(C)$ such that in the diagram

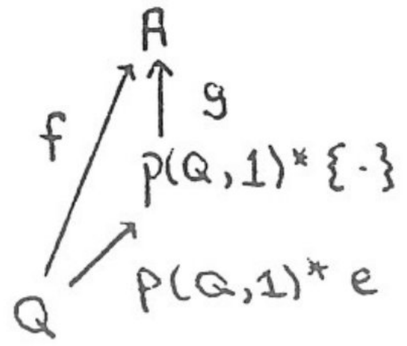


both triangles commute.

v. An object $\{\cdot\}$ of $\mathbb{C}$ such that $1 \triangleleft \{\cdot\}$ and a morphism $e : 1 \longrightarrow \{\cdot\}$ in $\mathbb{C}$ having the following property: for all objects Q of $\mathbb{C}$, for all objects A of $\mathbb{C}$ such that $p(Q,1)^* \{\cdot\} \triangleleft A$ in $\mathbb{C}$, for all morphisms $f:Q \longrightarrow A$ in $\mathbb{C}$ such that the diagram

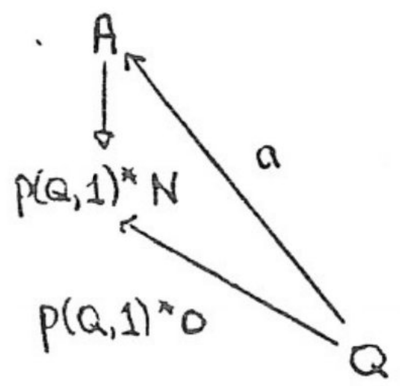


commutes, there exists a unique $g \in \mathrm{Arr}_{\mathbb{C}}(A)$ such that the diagram

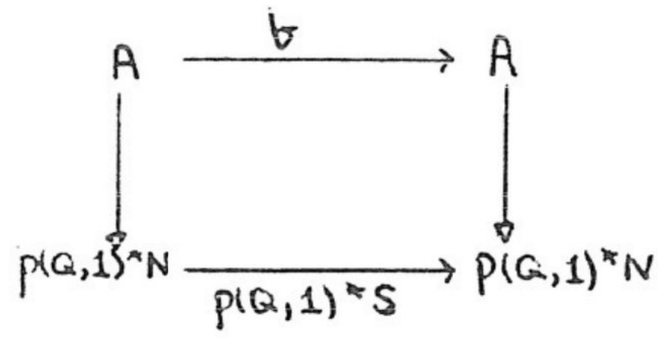


commutes.

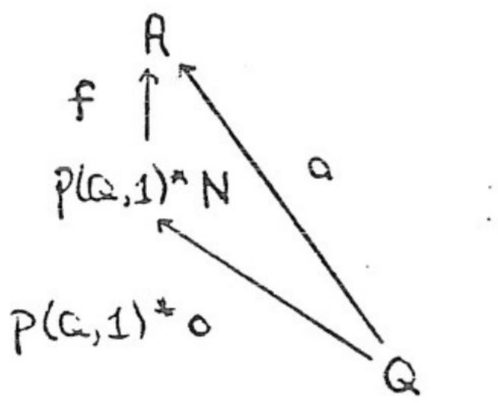
vi. An object N of $\mathbb{C}$ such that $1 \lhd N$ and morphisms $o : 1 \longrightarrow N$ and $S : N \longrightarrow N$ in $\mathbb{C}$ having the following property: for all objects Q of $\mathbb{C}$, for all objects A of $\mathbb{C}$ such that $p(A,1)^*N \lhd A$, for all morphisms $a : Q \longrightarrow A$ and $b : A \longrightarrow A$ such that the diagrams

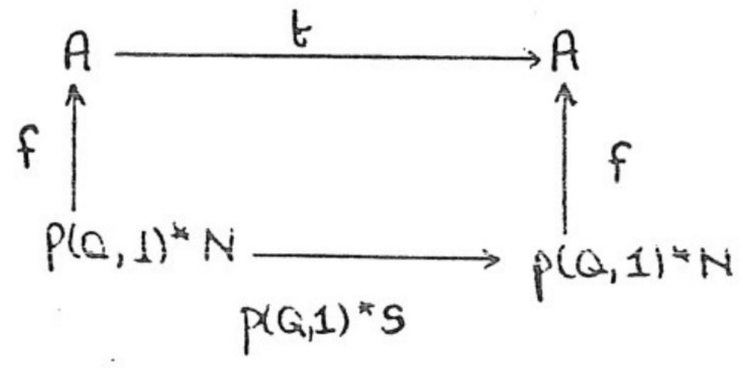


and



commute, there exists a unique $f \in Arr_{\mathbb{C}}(A)$ such that the diagrams





commute.

Subject to the following conditions. If $f : Q \longrightarrow Q'$ is a morphism in $\mathbb{C}$ then

if $Q' \lhd A \lhd B$ then $f^*\Sigma B = \Sigma f^*B$ and $f^*pr(B) = pr(f^*B)$,

if $Q' \lhd A \lhd B$ then $f^*\Pi B = \Pi f^*B$ and $f^*Ap(B) = Ap(f^*B)$,

if $Q' \lhd A$ then $f^*IdA = Idf^*A$ and $f^*r(A) = r(f^*A)$,

if $Q' \lhd A$ and $Q' \lhd B$ then $f^*(A+B) = f^*A+f^*B$, $f^*i_{A,B} = i_{f^*A,f^*B}$ and $f^*j_{A,B} = j_{f^*A,f^*B}$.

Clauses i. ... vi. in the definition of strong M-L structure correspond to the schemes for Σ, Π, Id, +, N_1 and N in Martin-Lof type theory. If the word unique is dropped from any of these clauses then what remains is an exact rewriting of the corresponding Martin-Lof scheme within the language of contextual categories.

The final condition says that f^* always preserves Σ, pr, , Ap etc.. Thus it says that the contextual functor $\mathbb{C}_f : \mathbb{C}_{Q'} \longrightarrow \mathbb{C}_Q$ is a structure preserving morphism whenever $f : Q \longrightarrow Q'$ in $\mathbb{C}$. It is, then, the 'substitution is a homomorphism between algebras of terms' condition. This is an implicit property of syntactical substitution which must always be stated explicitly in any algebraic semantics. That this is a property of syntactical substitution is because if we substitute into any expression that is the result of binding simpler expressions by a logical symbol then the result is identical with the result of first substituting into the simpler expressions and then binding by the logical symbol.

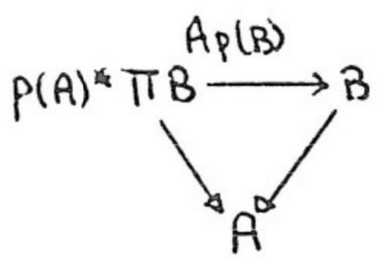
This is work which syntax always does for us and which algebra has got to be made to do. Hence the condition.

Not surprisingly the contextual category Fam of sets, families of sets and so on is in a natural way a strong M-L structure. Σ is calculated by taking actual disjoint unions of families of sets, Π calculated by taking cartesian products of families of sets. Id(A), when A a set, is the characteristic family of the identity predicate on A. + is interpreted by coproducts and N is taken to be the set of natural numbers.

The definition of strong M-L structure can be simplified considerably. In the first place those parts of the definition that are about Σ, pr, $\{\cdot\}$ and e are equivalent to $\langle \mathbb{C}, \Sigma, pr, \{\cdot\}, e\rangle$ being a contextual category with disjoint unions and singleton object. This leads to simplifications of the other clauses. Clause ii. can be used to simplify clause vi. It turns out, then, that the definition that we have given is equivalent to the following definition.

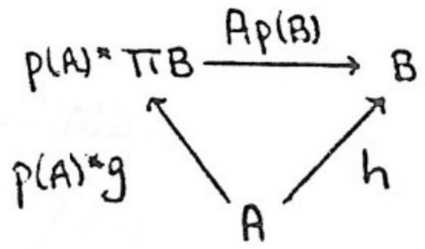
A strong M-L structure $\langle \mathbb{C}, \Sigma, pr, \Pi, Ap, Id, r, +, i, j, \{\cdot\}, e, N, o, S\rangle$ consists of a contextual category with disjoint unions and singleton object $\langle \mathbb{C}, \Sigma, pr, \{\cdot\}, e\rangle$ and the following additional structure.

ii'. For all $1 \lhd Q \lhd A \lhd B$ in $\mathbb{C}$, an object $\prod B$ of $\mathbb{C}$ such that $Q \lhd \Pi B$ and a morphism $Ap(B) : p(A)^* \Pi B \longrightarrow B$ such that the diagram



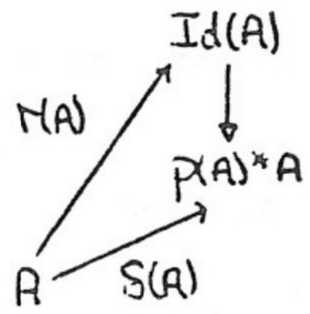
commutes and having the property: for all $h \in Arr_{\mathbb{C}}(B)$, there

exists a unique $g \in Arr_{\mathbb{C}}(\pi B)$ such that the diagram

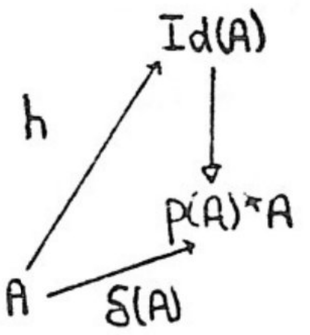


commutes.

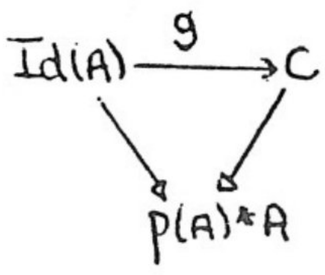
iii'. For all $1 \lhd Q \lhd A$ in $\mathbb{C}$, an object Id(A) of $\mathbb{C}$ such that $p(A)^*A \lhd IdA$ and a morphism $r(A) : A \longrightarrow IdA$ in $\mathbb{C}$ such that the diagram



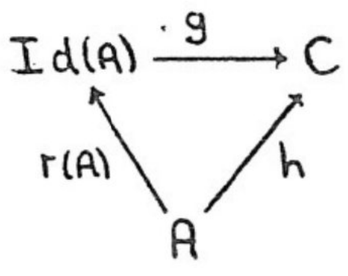
commutes and having the property: for every object C of $\mathbb{C}$ such that $p(A)^*A \lhd C$, for every morphism $h : A \longrightarrow C$ such that the diagram



commutes, there exists a unique $g : IdA \longrightarrow C$ such that



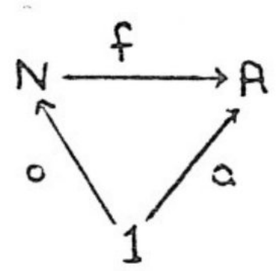
and

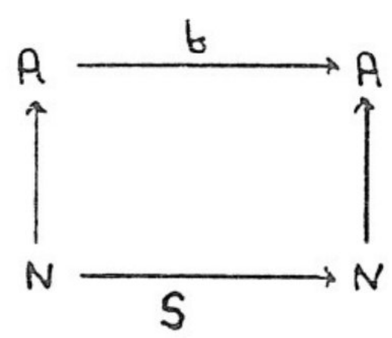


commute.

iv'. For all $1 \lhd Q \lhd A$ and $1 \lhd Q \lhd B$ in $\mathbb{C}$, an object $A+B$ of $\mathbb{C}$ such that $Q \lhd A+B$ and morphisms $i_{A,B}$ and $j_{A,B}$ such that the diagram $A \xrightarrow{i_{A,B}} A+B \xleftarrow{j_{A,B}} B$ is a coproduct diagram in Base $(\mathbb{C}_Q)$.

vi'. An object N of $\mathbb{C}$ such that $1 \lhd N$ and morphisms $o : 1 \longrightarrow N$, $S : N \longrightarrow N$ such that $< N\ o, S>$ is a natural number object in AugBase $(\mathbb{C})$. ($<N,o,S>$ is a natural number object in the category $\underline{C}$ with terminal object 1 iff $o : 1 \longrightarrow N$ and $s : N \longrightarrow N$ and for all objects A of $\underline{C}$, for all morphisms $a : 1 \longrightarrow A$ and $b : N \longrightarrow N$, there exists a unique morphism $f : N \longrightarrow A$ such that





commute.)

Subject to the condition : If $f : Q \longrightarrow Q'$ in $\mathbb{C}$, where $1 \lhd Q$ and $1 \lhd Q'$ in $\mathbb{C}$, then

if $Q' \lhd A \lhd B$ then $f^*\Pi B = \Pi f^*B$ and $f^*Ap(B) = Ap(f^*B)$,

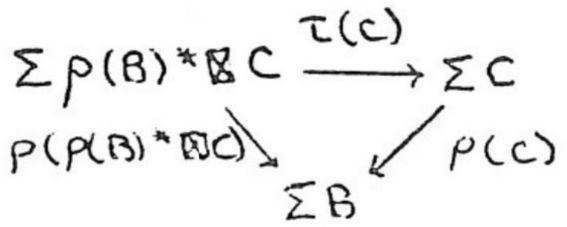
if $Q' \lhd A$ then $f^*IdA = Idf^*A$ and $f^*r(A) = r(f^*A)$,

if $Q' \lhd A$ and $Q' \lhd B$ then $f^*(A+B) = f^*A+f^*B$, $f^*i_{A,B} = i_{f^*A,f^*B}$ and $f^*j_{A,B} = j_{f^*A,f^*B}$.

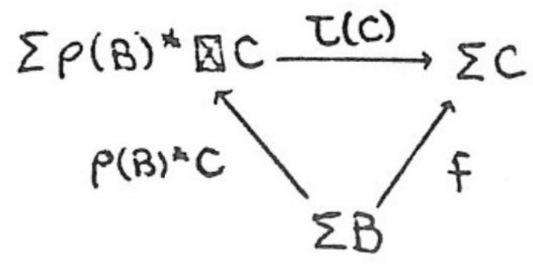
Finally, the definition of strong M-L structure can be rewritten in terms of categories with attributes :

An <u>M-L Hyperdoctrine</u> $\langle \mathbb{E}, \boxtimes, \tau, Id, r, +, i, j, N, o, S \rangle$ consists of a category with attributes $\mathbb{E}$ with the following additional structure:

i. For every object A of Base($\mathbb{E}$), for every $B \in Att_{\mathbb{E}}(A)$, for every $C \in Att_{\mathbb{E}}(\Sigma B)$, an attribute $\boxtimes C \in Att_{\mathbb{E}}(A)$ and a morphism $\tau(C)$: $\Sigma \rho(B)^* \boxtimes C \longrightarrow \Sigma C$ in Base $\mathbb{E}$ such that the diagram

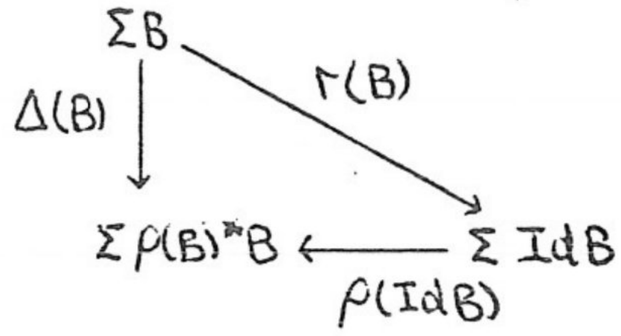


commutes and having the property that for all morphisms $f : \Sigma B \longrightarrow \Sigma C$ in Base($\mathbb{E}$) such that $f o \rho(C) = id_{\Sigma B}$, there exists a unique $g : A \longrightarrow \Sigma \boxtimes C$ in Base $\mathbb{E}$ such that $g o \rho(\boxtimes C) = id_A$ and such that the diagram

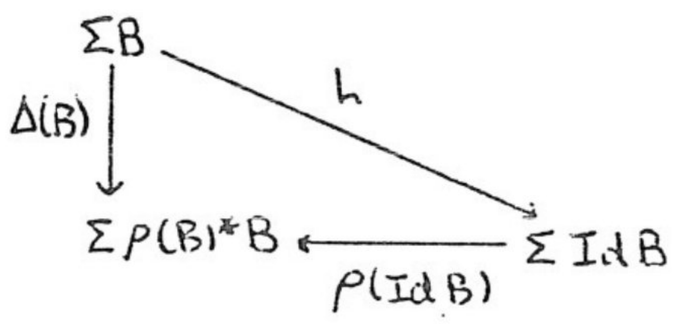


commutes. (Where $\rho(B)^*g$ is defined to be the unique morphism $h : \Sigma B \longrightarrow \Sigma \rho(B)^* \boxtimes C$ such that $h \circ \rho(\rho(B)^* \boxtimes C) = id_{\Sigma B}$ and $h \circ \xi(\rho(B), \boxtimes C) = \rho(B) \circ g$.)

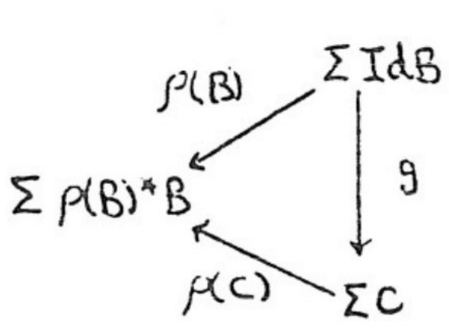
ii. For all objects A of Base($\mathbb{E}$), for all $B \in Att_{\mathbb{E}}(A)$, an attribute $IdB \in Att_{\mathbb{E}}(\Sigma \rho(B)^*B)$ and a morphism $r(B) : \Sigma B \longrightarrow \Sigma IdB$ in Base $\mathbb{E}$ such that the diagram



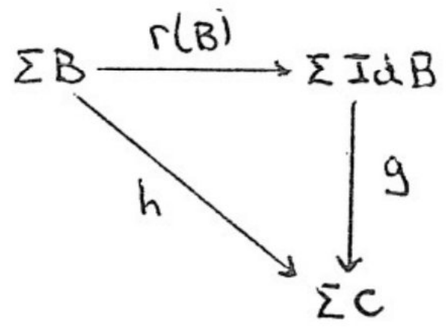
commutes (where $\Delta(B)$ is the unique morphism such that $\Delta(B) \circ \rho(\rho(B)^*B) = id$ and $\Delta(B) \circ \xi(\rho(B), B) = id$) and having the property that for all attributes $C \in Att_{\mathbb{E}}(\Sigma \rho(B)^*B)$ and for all morphisms $h : \Sigma B \longrightarrow \Sigma C$ such that the diagram



commutes, there exists a unique morphism $g : \Sigma IdB \longrightarrow \Sigma C$ such that the diagrams

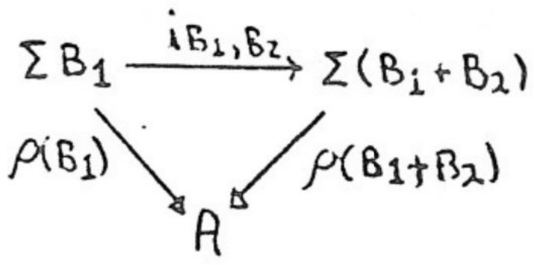


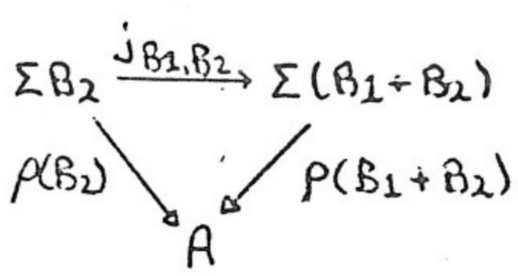
and



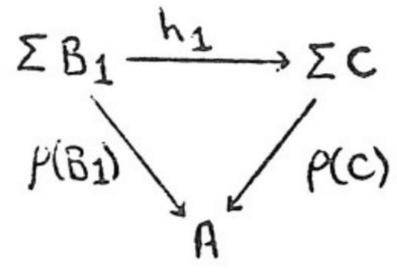
commute.

iii. For all objects A of Base ($\mathbb{E}$), for all $B_1, B_2 \in Att(A)$, an object $B_1 + B_2 \in Att(A)$ and morphisms $i_{B_1,B_2} : \Sigma B_1 \longrightarrow \Sigma(B_1+B_2)$ and $j_{B_1,B_2} : \Sigma B_2 \longrightarrow \Sigma(B_1+B_2)$ such that the diagrams

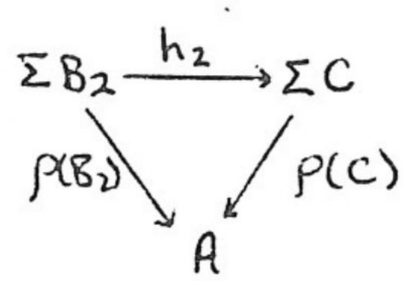




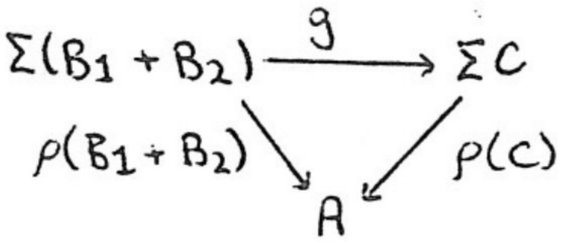
commute and having the property that for all $C \in Att_{\mathbb{E}}(A)$, for all pairs of morphisms $h_1 : \Sigma B_1 \longrightarrow \Sigma C$ and $h_2 : \Sigma B_2 \longrightarrow \Sigma C$ such that the diagrams



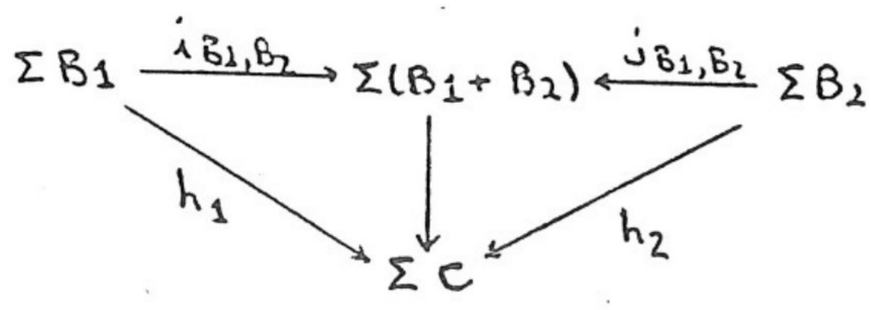
and



commute, there exists a unique morphism $g : \Sigma(B_1+B_2) \longrightarrow \Sigma C$ such that



commutes and such that in the diagram



both triangles commute.

iv. An object N of Base ($\mathbb{E}$), morphisms o : 1 $\longrightarrow$ N and S : N $\longrightarrow$ N such that $\langle N, o, S \rangle$ is a natural number object in Base ($\mathbb{E}$).

Subject to the conditions:

v. If f : A $\longrightarrow$ A' in Base $\mathbb{E}$, if B $\in$ Att$_{\mathbb{E}}$(A') and C $\in$ Att$_{\mathbb{E}}$(ΣB) then $f^{*}\boxtimes C = \boxtimes \xi(f,B)^{*}C$ and the diagram

$$\begin{array}{ccc} \Sigma\, \rho(f^{*}B)^{*} f^{*} \boxtimes C & \xrightarrow{\xi(\xi(f,B),\, \rho(B)^{*}\boxtimes C)} & \Sigma\, \rho(B)^{*}\boxtimes C \\ \Big\downarrow \tau(\xi(f,B)^{*}C) & & \Big\downarrow \tau(C) \\ \Sigma\, \xi(f,B)^{*}C & \xrightarrow[\xi(\xi(f,B),C)]{} & \Sigma C \end{array}$$

commutes.

vi. If f : A $\longrightarrow$ A' in Base ($\mathbb{E}$) and if B $\in$ Att$_{\mathbb{E}}$(A') then $\xi(\xi(f,B), \rho(B)^{*}B)^{*}\mathrm{Id}B = \mathrm{Id}f^{*}B$ and the diagram

$$\begin{array}{ccc} \Sigma\, \mathrm{Id}f^{*}B & \xrightarrow{\xi(\xi(\xi(f,B),\, \rho(B)^{*}B),\, \mathrm{Id}B)} & \Sigma\, \mathrm{Id}B \\ \Big\uparrow \nu(f^{*}B) & & \Big\uparrow \nu(B) \\ \Sigma f^{*}B & \xrightarrow[\xi(f,B)]{} & \Sigma B \end{array}$$

commutes.

vii. If $f : A \longrightarrow A'$ in Base($\mathbb{E}$) and $B_1, B_2 \in Att_{\mathbb{E}}(A')$ then $f^*(B_1+B_2) = f^*B_1 + f^*B_2$ and the diagrams

$$\begin{array}{ccc} \Sigma f^*B_1 & \xrightarrow{\xi(f,B_1)} & \Sigma B_1 \\ {\scriptstyle i_{f^*B_1, f^*B_2}}\downarrow & & \downarrow{\scriptstyle i_{B_1,B_2}} \\ \Sigma f^*(B_1+B_2) & \xrightarrow[\xi(f,B_1+B_2)]{} & \Sigma(B_1+B_2) \end{array} \quad \text{and} \quad \begin{array}{ccc} \Sigma f^*B_2 & \xrightarrow{\xi(f,B_2)} & \Sigma B_2 \\ {\scriptstyle j_{f^*B_1, f^*B_2}}\downarrow & & \downarrow{\scriptstyle j_{B_1,B_2}} \\ \Sigma f^*(B_1+B_2) & \xrightarrow[\xi(f,B_1+B_2)]{} & \Sigma(B_1+B_2) \end{array}$$

commute.

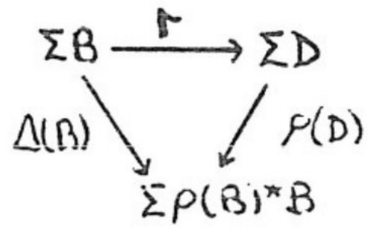
We are claiming, then, that the category of M-L structures and the category of M-L hyperdoctrines are equivalent. The proof of this result is an extension of the proof that $\underline{\Sigma\text{-con}}$ and $\underline{\text{Attcat}}$ are equivalent categories.

3.5. Limit Spaces - a model of Martin-Lof type theory.

We wish to describe a model of Martin-Lof type theory in which types are interpreted as limit spaces and in which families of types indexed by a type are, roughly speaking, interpreted as 'morphisms with codomain' in the category of limit spaces. The model is described as an M-L hyperdoctrine with base category the category of limit spaces. If we first describe the M-L hyperdoctrine of sets and families of sets then the M-L hyperdoctrine of limit spaces can be described without much trouble.

But first, two very useful and trivial lemmas:

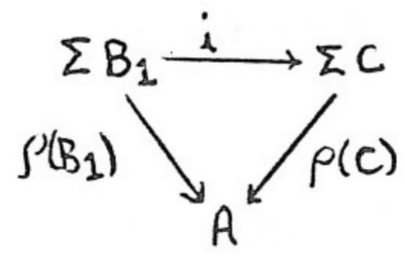
Lemma 1. If $\mathbb{E}$ is a category with attributes, if A is an object of Base ($\mathbb{E}$), If $B \in Att_{\mathbb{E}}(A)$, if $D \in Att(\Sigma\, \rho(B)^*B)$ and If $r : \Sigma\, B \xrightarrow{\sim} \Sigma\, D$ is an isomorphism in Base ($\mathbb{E}$) such that the diagram



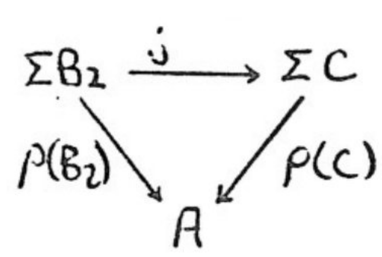
commutes then $\langle D, r\rangle$ satisfies clause ii. of the definition of M-L hyperdoctrine. That is clause ii. is satisfied if Id(B) is taken to D and if r(B) is taken to be r.

Lemma 2. If $\mathbb{E}$ is a category with attributes, if A is an object of $\mathbb{E}$, If $B_1, B_2 \in Att_{\mathbb{E}}(A)$, If $C \in Att_{\mathbb{E}}(A)$ and if $i : \Sigma\, B_1 \longrightarrow \Sigma\, C$, $j : \Sigma B_2 \longrightarrow \Sigma\, C$

in Base ($\mathbb{E}$) such that the diagrams



and



commute and if $\Sigma B_1 \xrightarrow{i} \Sigma C \xleftarrow{j} \Sigma B_2$ is a coproduct diagram in Base ($\mathbb{E}$) then $\langle C,i,j\rangle$ satisfy clause iii. of the definition of M-L hyperdoctrine. That is if B_1+B_2 is taken to be C and if i_{B_1,B_2} is taken to be i and similarly if j_{B_1,B_2} is taken to be j then clause iii. is satisfied.

The M-L hyperdoctrine Fam of sets and families of sets is as follows:

1. Base (Fam) = Set.
2. If $A \in |\text{Base}(\underline{\text{Fam}})|$, ie if A is a set, then $\text{Att}_{\underline{\text{Fam}}}(A) = \{ \text{A-indexed families of sets} \}$.
3. If $A \in |\text{Base}(\underline{\text{Fam}})|$ and if $B \in \text{Att}(A)$ then $\Sigma B = \dot{\bigcup}_{a \in A} B(a) = \{\langle a,b\rangle \mid a \in A \text{ and } b \in B(a)\}$.
4. If $f : A \longrightarrow A'$ in Base (Fam) if $B \in \text{Att}(A')$ then $f^*B = \lambda a \in A.\ B(f(a))$. $\xi(f,B) : f^*B \longrightarrow B$ is given by $\xi(f,B)(\langle a,b\rangle) = \langle f(a),b\rangle$.
5. If $A \in |\text{Base}(\underline{\text{Fam}})|$, if $B \in \text{Att}(A)$, if $C \in \text{Att}(\Sigma B)$ then $\#C \in \text{Att}(A)$ is given by $\#C(a) = \{ \langle b,c\rangle \mid b \in B(a) \text{ and } c \in C(\langle a,b\rangle$
$\gamma(C) : \Sigma \#C \longrightarrow \Sigma C$ is given by $\gamma(C)(\langle a,\langle b,c\rangle\rangle) = \langle\langle a,b\rangle ,c$

6. If $A \in |\text{Base}(\underline{\text{Fam}})|$ then $L(A) = \lambda x \in \{\cdot\} . A$. Then $\Theta(A) : \Sigma LA \quad A$ is given by $\Theta(A)(<\cdot, a>) = a$.

7. If $A \in |\text{Base}(\underline{\text{Fam}})|$, If $B \in \text{Att}(A)$, if $C \in \text{Att}(\Sigma B)$ then define $\boxtimes C \in \text{Att}(A)$ by

$$\boxtimes C(a) = \prod x \in B(a) . C(<a,x>).$$

If $\boxtimes C$ is so defined then $\Sigma \rho(B)^* \boxtimes C =$

$$\{ <<a,b>, g> \mid a \in A, b \in B(a) \text{ and } g \in \prod x \in B(a). \ C(<a,x>) \} .$$

Define $Ap(C) : \Sigma \rho(B)^* \boxtimes C \longrightarrow \Sigma C$ by defining

$$Ap(C)(<<a,b>, g>) = <<a,b>, Ap(b,g)> .$$

8. If $A \in |\text{Base}(\underline{\text{Fam}})|$ and if $B \in \text{Att}(A)$ then $\Sigma \rho(B)^* B = \{ <<a,b>, b'> \mid a \in A \text{ and } b, b' \in B(a) \}$. Define $Id(B)$, $r(B)$ by

$Id(B)(<<a,b>, b'>) = \{\cdot\}$, if $b = b'$, $= \emptyset$ otherwise,

$r(B)(<a,b>) = <<<a,b>, b>, \cdot>$.

9. If $A \in |\text{Base}(\underline{\text{Fam}})|$, If $B_1, B_2 \in \text{Att}(A)$ then define $B_1 + B_2$ by $(B_1 + B_2)(a) = B_1(a) + B_2(a)$. Then $\Sigma(B_1 + B_2) = \Sigma B_1 + \Sigma B_2$. Define i_{B_1, B_2} and j_{B_1, B_2} accordingly.

10. As is well known the set of natural numbers is a natural number object in the category $\underline{\text{Set}}$.

Completing the description of $\underline{\text{Fam}}$ as an M-L hyperdoctrine.

The set of all filters on a set A ordered by inclusion is a lattice. The meets in this lattice are given by intersection of filters. If φ and ψ are filters on A then $\varphi \vee \psi = \{ u \cap v \mid u \in \varphi \text{ and } v \in \psi \}$.

If A is a set then the ordered set of all filters on A will be denoted $\mathcal{F}(A)$.

If $f : A \longrightarrow A'$ is a function then define $\mathcal{F}(f) : \mathcal{F}(A) \longrightarrow \mathcal{F}(A')$ by $\mathcal{F}(f)(\Theta) = \{ w' \subseteq A' \mid \exists w \in \Theta \text{ s.t. } w' \supseteq f(w) \}$. Then $\mathcal{F}(f)$ is an order preserving map and it is easy to

see that $\mathcal{F}$ is functorial:

$\mathcal{F}(fof') = \mathcal{F}(f')$ and $\mathcal{F}(idA) = id_{\mathcal{F}(A)}$. $\mathcal{F}$ can be considered to be a functor from Set to the category of ordered sets.

A limit space $\langle A, C\rangle$ consists of a set A and for each $a \in A$, a set $C(a) \subseteq \mathcal{F}(A)$, said to be the set of filters converging to a, and subject to the conditions:

i. If $a \in A$ then $\langle a\rangle \in C(a)$, where $\langle a\rangle$ is the principle filter generated by a.

ii. If $a \in A$ and $\varphi, \psi \in C(a)$ then $\varphi \wedge \psi \in C(a)$.

iii. If $a \in A$ and $\varphi \in C(a)$ and if $\varphi \leq \psi$ in $\mathcal{F}(A)$ then $\psi \in C(a)$.

We usually say that φ conv a in preference to $\varphi \in C(a)$.

A morphism of limit spaces $f : \langle A, C\rangle \longrightarrow \langle A', C'\rangle$ is a function $f : A \longrightarrow A'$ such that whenever φ conv a in $\langle A, C\rangle$ then $\mathcal{F}(f)(\varphi)$ conv $f(a)$ in $\langle A', C'\rangle$.

The category of limit spaces will be denoted

Limsp .

We shall make use of the fact that for every $f : A \longrightarrow A'$ in Set, the morphism $\mathcal{F}(f) : \mathcal{F}(A) \longrightarrow \mathcal{F}(A')$ has a left adjoint. We define

an order preserving morphism $\check{\mathcal{T}}(f) : \mathcal{T}(A') \longrightarrow \mathcal{T}(A)$ by

$\check{\mathcal{T}}(f)(\varphi) = \{ w \subseteq A \mid \exists w' \in \varphi \text{ s.t. } w \supseteq f^{-1}(w') \}$.

That $\check{\mathcal{T}}(f)$ adj $\mathcal{T}(f)$ is expressed by Lemma 3 If $\Theta \in \mathcal{T}(A)$ and if $\varphi \in \mathcal{T}(A'$ and if $f : A \longrightarrow A'$ then $\check{\mathcal{T}}(f)(\varphi) \leq \Theta$ iff $\varphi \leq \mathcal{T}(f)(\Theta)$.

Proof 1. $\check{\mathcal{T}}(f)(\varphi) \leq \Theta \Rightarrow \varphi \leq \mathcal{T}(f)(\Theta)$. Assume that $\check{\mathcal{T}}(f)(\varphi) \leq \Theta$. Then $u \in \varphi \Rightarrow f^{-1}(u) \in \check{\mathcal{T}}(f)(\varphi) \Rightarrow f^{-1}(u) \in \Theta \Rightarrow f(f^{-1}(u)) \in \mathcal{T}(f)(\Theta)$. Hence $u \in \varphi \Rightarrow u \in \mathcal{T}(f)(\Theta)$ because $f(f^{-1}(u)) \subseteq u$. Thus $\check{\mathcal{T}}(f)(\varphi) \leq \Theta \Rightarrow \varphi \leq \mathcal{T}(f)(\Theta)$.

2. The converse. Suppose $\varphi \leq \mathcal{T}(f)(\Theta)$. Suppose $v \in \check{\mathcal{T}}(f)(\varphi)$. Then $\exists u \in \varphi$ such that $f^{-1}(u) \subseteq v$. $u \in \varphi \Rightarrow u \in \mathcal{T}(f)(\Theta)$. Hence $u \supseteq f(w)$, for some $w \in \Theta$. But $u \supseteq f(w)$ iff $f^{-1}(u) \supseteq w$. Hence $v \in \Theta$. As required, $\varphi \leq \mathcal{T}(f)(\Theta) \Rightarrow \check{\mathcal{T}}(f)(\varphi) \leq \Theta$.

Corollary 4 (a) $\check{\mathcal{T}} : \underline{\mathrm{Set}}^{\mathrm{op}} \longrightarrow$ the category of ordered sets, is a functor.

(b) For each $f : A \longrightarrow A'$, $\mathcal{T}(f) : \mathcal{T}(A) \longrightarrow \mathcal{T}(A')$ preserves meets.

The base category of the M-L hyperdoctrine of limit spaces is taken to be the category Limsp. Then if $\mathcal{A} \in |\,\underline{\mathrm{Limsp}}\,|$ and if $\mathcal{A}$ has underlying set A then define $\mathrm{Att}(\mathcal{A}) = \{ \langle B, C \rangle \mid$ B is an A-indexed family of sets and C is such that $\langle \Sigma B, C \rangle$ is a limit space such that $\rho(B) : \langle \Sigma B, C \rangle \longrightarrow \mathcal{A}$ is continuous $\}$.

If $\mathcal{B} \in \mathrm{Att}(A)$ then define $\Sigma \mathcal{B} = \langle \Sigma B, C \rangle$, where B is the underlying family of sets associated with $\mathcal{B}$. Define $\rho(\mathcal{B}) = \rho(B)$. Thus every $\mathcal{B} \in \mathrm{Att}(\mathcal{A})$ is determined by an A-indexed family of sets B and a limit space

$\Sigma\mathcal{B}$ which has underlying set ΣB and which is such that $\rho(B)$: $\Sigma\mathcal{B} \longrightarrow \mathcal{A}$ is continuous.

Pullbacks If $f : \mathcal{A} \longrightarrow \mathcal{A}'$ and if $\mathcal{B} \in Att(\mathcal{A})$ then the underlying family of sets of $f^*\mathcal{B}$ is taken to be f^*B. Then $\Sigma f^*\mathcal{B}$ is taken to be the unique limit space with underlying set Σf^*B such that

$$\begin{array}{ccc} \Sigma f^*\mathcal{B} & \xrightarrow{\xi(f,B)} & \Sigma\mathcal{B} \\ \rho(f^*\mathcal{B}) \downarrow & & \downarrow \rho(\mathcal{B}) \\ \mathcal{A} & \xrightarrow[f]{} & \mathcal{A}' \end{array}$$

is a pullback diagram in Limsp.

Convergence in $f^*\mathcal{B}$ is given by Θ conv x iff $\mathcal{F}(\rho(f^*B))(\Theta)$ conv $\rho(f^*B)(x)$ and $\mathcal{F}(\xi(f,B))(\Theta)$ conv $\xi(f,B)(x)$. Then that the diagram is a pullback diagram follows immediately from the fact that the underlying diagram in Set is a pullback diagram. Uniqueness of this convergence relation subject to the diagram being a pullback diagram is the case because if X is a new limit space with underlying set Σf^*B and such that

$$\begin{array}{ccc} X & \xrightarrow{\xi(f,B)} & \Sigma\mathcal{B} \\ \rho(f^*\mathcal{B}) \downarrow & & \downarrow \rho(\mathcal{B}) \\ \mathcal{A} & \xrightarrow[f]{} & \mathcal{A}' \end{array}$$

is a pullback diagram then there exists an isomorphism

$g : \Sigma f^*\mathcal{B} \longrightarrow X$ such that $g \circ \rho(f^*B) = \rho(f^*B)$ and $g \circ \xi(f,B) = \xi(f,B)$.

But then since the corresponding diagrams in Set are pullback diagrams so $g = id_{\Sigma f^*B}$. So $id_{\Sigma f^*B}$ is a morphism of limit spaces $id_{\Sigma f^*B} : \Sigma f^*\mathcal{B} \longrightarrow X$ and also $id_{\Sigma f^*B} : X \longrightarrow \Sigma f^*\mathcal{B}$. Hence φ conv x in $\Sigma f^*\mathcal{B}$ implies $\mathcal{T}(id_{\Sigma f^*B})(\varphi)$ conv x in X ie φ conv x in X. And vice versa. Thus $\Sigma f^*\mathcal{B}$ is unique.

That pullbacks fit together follows from the fact that they fit together at the level of underlying sets and from the uniqueness of the pullbacks once underlying sets are decided.

<u># and γ.</u> If $\mathcal{A} \in |$ <u>Limsp</u> $|$, If $\mathcal{B} \in Att(\mathcal{A})$ and If $\mathcal{C} \in Att(\Sigma\mathcal{B})$ then the underlying family of $\#\mathcal{C}$ is taken to be $\#C$. The limit space $\Sigma\#\mathcal{C}$ is then uniquely determined by the requirement that $\gamma(c) : \Sigma\mathcal{C} \longrightarrow \Sigma\#\mathcal{C}$ be an isomorphism of limit spaces.

<u>L and Θ.</u> <u>Limsp</u> has a terminal object 1. If $\mathcal{A} \in$ $|$ <u>Limsp</u> $|$ then $L(\mathcal{A})$ is taken to have underlying family of sets L(A) and $\Sigma L\mathcal{A}$ is then determined by the requirement that $\Theta(A) : \Sigma L(\mathcal{A}) \longrightarrow \mathcal{A}$ be an isomorphism of limit spaces.

<u>Id, + and N</u> are all trivial to define. For Id and + lemmas 1 and 2 of this section can be used.

<u>$\boxtimes$ and Ap</u> The slightly non trivial bit. If $\mathcal{A} \in |$ <u>LimSp</u> $|$, If $\mathcal{B} \in Att(\mathcal{A})$ and if $\mathcal{C} \in Att(\Sigma\mathcal{B})$ then define $\boxtimes\mathcal{C}(a) = \{ f \in \Pi x \in B(a). C(\langle a,x \rangle) \mid$ the corresponding $f : a^*B \longrightarrow \xi(a,B)^*C$ is continuous $\}$. Ap is taken to be application. $\Sigma\boxtimes\mathcal{C}$ is topologised by Φ conv y in $\Sigma\boxtimes\mathcal{C}$ iff $\mathcal{T}(\rho(\boxtimes\mathcal{C}))(\Phi)$ conv $\rho(\boxtimes\mathcal{C})(y)$ and $\forall x \in \xi(\rho(B), \boxtimes\mathcal{C})^{-1}(\{y\})$, $\forall \Psi \in \mathcal{T}(\Sigma B)$ s.t. Ψ conv $\rho(\rho(B)^*\boxtimes\mathcal{C})(x)$, $\mathcal{T}(Ap)(\check{\mathcal{T}}(\xi(\rho(B), \boxtimes\mathcal{C}))(\Phi) \vee \check{\mathcal{T}}(\rho(\rho(B)^*\boxtimes\mathcal{C}))(\Psi)$ convAp(x).

REFERENCES

(1) Barendregt, H. P. — Some extensional models for combinatory logics and λ-calculi, thesis, Utrecht, 1971

(2) Benabou, J. — Introduction to bicategories, Springer LNM 47, 1971

(3) Birkhoff, G. and Lipson, J. D. — Heterogeneous algebras, J. Combinatory Theory 8, 1970

(4) Fourman, M. P. — Higher order intuitionistic logic and topoi, thesis, Oxford 1974

(5) Freyd, P. — Aspects of topoi, Bull. Austral. Math. Soc. 7, 1972

(6) Hyland, M. — Recursion theory on the countable functionals, thesis, Oxford, 1975

(7) Gray, J. W. — The categorical comprehension scheme, Springer LNM 99, 1969

(8) Gray, J. W. — Formal category theory: adjointness for 2-categories, Springer LNM 391, 1974

(9) Kelly, G. M. — (Editor) Category seminar, Springer LNM 420, 1974

(10) Kelly, G. M. and Street, R. — Review of the elements of 2-categories, in (7)

(11) Lawvere, F. W. — Functorial Semantics of algebraic theories, Proc. Nat. Acad. Sci. 50, 1963

(12) Lawvere, F. W. — Functorial Semantics of elementary theories, abstract, J. of symbolic logic 31, 1966

(13) Lawvere, F. W. Theories as categories and the completeness theorem, abstract, J. of symbolic logic 32, 1967

(14) Lawvere, F. W. Some algebraic problems in the context of functorial semantics of algebraic theories, Springer LNM 99, 1968

(15) Lawvere, F. W. Adjointness in foundations, Dialectica 23, 1968

(16) Lawvere, F. W. Equality in hyperdoctrines and comprehension schema as an adjoint functor. Proc. AMS Symposia Pure Math. 17, 1970

(17) Lawvere, F. W. Introduction, Model theory and topoi, Springer LNM 445, 1975

(18) Lambek, J. Deductive systems and categories I, Math. Systems Theory 2, 1968

(19) Lambek, J. Deductive systems and categories II, Springer LNM 86, 1969

(20) Lambek, J. Deductive systems and categories III, Springer LNM 274, 1972

(21) Linton, F. E. J. Some aspects of equational theories, Proc. conf. categorical algebra, La Jolla (1965), Springer 1966

(22) Linton, F. E. J. An outline of functorial semantics Springer LNM 80, 1969

(23) MacLane, S. Categories for the working mathematician, Springer-Verlag, 1972

(24) MacLane, S. (Editor) Coherence in categories, Springer LNM 281, 1972.

(25) Martin-Löf, P. An intuitionistic theory of types, Logic Colloquium '73, North Holland

(26) Myers, A. Cartesian closed categories and λ-calculus, dissertation, Bristol 1974

(27) Pareigis, B. Categories and functors, Acad. Press, 1970

(28) Reyes, G. Sheaves and concepts, to appear

(29) Szabo, M. E. Algebra of proofs, North Holland, 1976