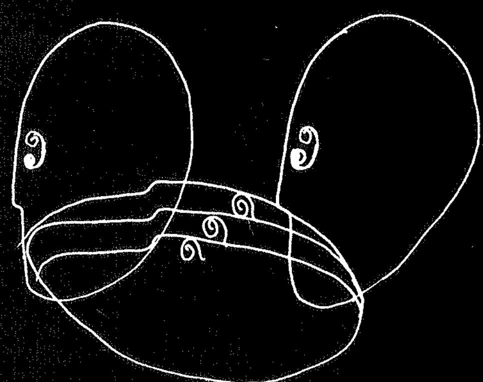
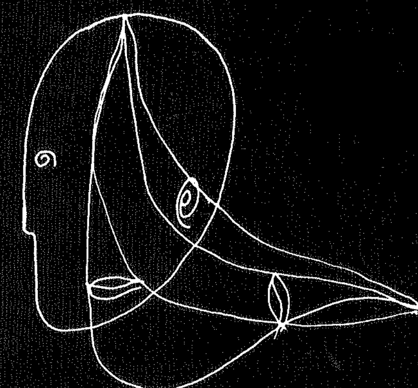


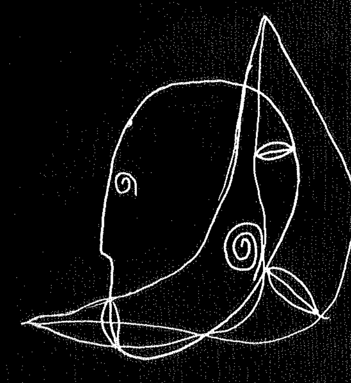
PREDICATES AND FIBRATIONS

Duško Pavlović

PREDICATES AND FIBRATIONS

Duško Pavlović





PREDICATES AND FIBRATIONS

FROM TYPE THEORETICAL TO CATEGORY THEORETICAL PRESENTATION OF CONSTRUCTIVE LOGIC

(MET EEN SAMENVATTING IN HET NEDERLANDS)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE RIJKSUNIVERSITEIT TE UTRECHT OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF. DR. J. A. VAN GINKEL, INGEVOLGE HET BESLUIT VAN HET COLLEGE VAN DEKANEN IN HET OPENBAAR TE VERDEDIGEN OP MAANDAG 26 NOVEMBER 1990 DES NAMIDDAGS TE 12.45 UUR

DOOR

DUŠKO PAVLOVIĆ

GEBOREN OP 18 OKTOBER 1957 TE AZANJA, JOEGOSLAVIË

PROMOTOR	PROF. DR. D. VAN DALEN
COPROMOTOR	DR. I. MOERDIJK

FACULTEIT DER WISKUNDE EN INFORMATICA,
RIJKSUNIVERSITEIT UTRECHT

IN PART PREPARED AT	DEPARTMENT OF PURE MATHEMATICS AND MATHEMATICAL STATISTICS, UNIVERSITY OF CAMBRIDGE
WITH SUPPORT FROM	NETHERLANDS ORGANISATION FOR SCIENTIFIC RESEARCH (NWO)

Preface

Nuances, nuances sans couleurs
Paul Valéry

Abstract.

The commonplace simplification of constructive logic as logic without the excluded middle conceals a more authentic idea of constructivism - that proofs are constructive functions, while quantifiers should be sums and products. An exact realisation of this idea requires a very special mathematical setting. Expressed categorically, the idea is that there is a small category (of "propositions" and "proofs"), with all small sums and products (as "quantifiers"). In ordinary category theory - or in a Grothendieck topos - such a category must be a preorder (and there can be at most one "proof" from one "proposition" to another; so these "proofs" do not really look like functions). In Hyland's effective topos, however, a nondegenerate small complete category has been discovered recently. It can be regarded as the first mathematical model of logic with constructive proofs. On the other hand, a significant impulse to the formal development of such a logic has been given by computer science (especially in the work of Coquand, Huet and their collaborators).

In this thesis, we consider two mathematical formulations of constructive logic: a type theoretical, and a category theoretical. In the end, the former is completely interpreted in the latter. The purpose of such a connection is to yield a characterisation of type theoretical *structures* by categorical *properties*.

In chapter I, we define the *theory of predicates*, a type theoretical generalisation of higher order predicate logic with a type of truth values and the comprehension scheme. Although presented rather differently, it is closely related to the theory of constructions with Σ-operations (as in Hyland-Pitts 1987). It deviates from the theory of constructions at two points: on one hand, a severe, but intuitively justified restriction is imposed on the contexts in it; on the other hand, a new operation of *extent* is

introduced, similar to the set theoretical operation $\varphi(X) \mapsto \{X | \varphi(X)\}$. In a sense, this operation compensates for the imposed restriction. We show that the theory of predicates is somewhat more general than the theory of constructions (which is, so far, the strongest type theory known to be logically consistent). With an additional rule, which forces some types to become isomorphic, the theory of predicates has exactly the same expressive power as the theory of constructions (modulo a translation, of course).

The general background for our categorical interpretation of constructive logic is the picture of a fibration as a variable category of predicates: its base is a category of "sets" and "functions", while the objects and arrows of a fibre are regarded as "predicates" over a set, with "proofs" between them. Chapter II surveys the concepts of propositional and predicate logic in this setting. The most important ideas are lifted from the context of indexed categories, where Lawvere introduced them some twenty years ago. Although many results in this chapter can be considered as basic, for very few of them a reference can be found. Some of them, however, surely belong to the folklore.

Chapter III is concerned with some fundamental notions of set theory in the setting of fibred categories. Having reviewed some basic ideas - mostly due to Bénabou this time, we begin in section 2 a categorical analysis of the comprehension principle. The property of fibrations, which is proposed as an interpretation of this principle, the structures induced by it, and the resulting representation of a fibration in its base are studied in detail in the next two sections. The induced structures include Lawvere's comprehension scheme as a special case - despite the apparent conceptual differences. Another special case are D-categories, used by Ehrhard in his interpretation of the theory of constructions, though in no connection with the concept of comprehension.

Putting together all the described categorical notions, in section 1 of chapter IV we define *categories of predicates*, small fibrations with small products and coproducts, and some fibrewise structure. There are some well known special cases again. The most prominent are, of course, elementary toposes: they can be presented as fibred *preorders* of predicates, with equality. The categorical structure, introduced by Hyland and Pitts in their study of the theory of constructions, can be regarded as a category of predicates generated by 1 (in appropriate sense).

Two pictures of constructive logic, built in the preceding chapters, are superimposed in chapter IV. In section 1, the interpretation of a theory of predicates in a category of

predicates is defined; in section 2, a category of predicates is built up from a given theory of predicates. This gives a relatively simple way to systematically obtain artificial examples of small categories with small sums and products. Completeness of this semantical construction is then proved, so that the two pictures conceptually coincide. Technically, however, they are complementary: the theory of predicates is a "programming language"; a category of predicates is its "computer". Each of them seems too complex to be developed alone.

An unusual phenomenon occurs in categories of predicates: the *weak*, i.e. *nonunique factorisations* play a structural role. We discuss two examples: weakly cocartesian liftings, and weak equalisers. The former are the "existential quantifiers" induced by comprehension on "sets". (To express the Beck-Chevalley property in a form appropriate for this situation, we characterized it in terms of inverse images only, in section 3a of chapter II.) These "quantifiers" are weak because logic is not extensional: a predicate may contain more than its extent. The weak equalisers, on the other hand, arise as extents of equality predicates. Multiple proofs of an equality predicate exactly correspond to the multiple factorisations through the weak equaliser belonging to it.

The final section is mostly devoted to various aspects of the equality predicates. A topos allows only one; but there can be a lot of them in a category of predicates. Despite their wekness, all equality predicates support much of the usual set theoretical approach to functions as graphs. (A connection between comprehension and the Cauchy completeness seems plausible.)

At the end, we use an arbitrary equality predicate to formulate some internal category theory (based on weak equalisers) under a category of predicates; then we construct another category of predicates over the same base - as a category of "internal presheaves" in the given one. Starting from any of the known mathematical models for the theory of constructions, this construction yields plenty of proper categories of predicates. (It thus multiplies the known examples of nondegenerate small categories with small products and coproducts.)

Acknowledgements.

The ideas and the friendship of the following people helped me to produce this thesis:

- my promotor prof. D. van Dalen, who accepted me as his student, supported me when I needed, and taught me, unnoticeably;
- my copromotor dr. I. Moerdijk, whose encouragement and scepticism stimulated me so often; his pertinent remarks led to numerous improvements in the text;
- dr. J.M.E. Hyland, who invited me in Cambridge although everything told him not to; before that - it was his ideas that led me to consider constructive logic categorically;
- prof. A.S. Troelstra, whose vivid criticism clarified some of my ideas to me, and made me understand how much more needs to be clarified;
- the members of Peripatetic Seminar for Sheaves and Logic, who not only showed me what category theory is and how to give a talk, but also - by enduring through my own talks - how not to;
- the members of prof. H.P. Barendregt's Intercity Seminar λ-Calculus, who endured a number of my talks too, and taught me valuable lessons in theory of types; in particular, B. Jacobs has discovered a mistake in an early version of chapter I, and always showed intensive interest in my ideas; he even wrote a paper on comprehension in type theory;
- dr. T. Streicher, who supplied many useful questions and advices;
- prof. C. Smoryński, who was my neighbour for two years; he told me that mathematics should be friendly and relaxed;
- prof. V. Perić and prof. M. Vuković, the first ones who believed that I could become a mathematician.

Finally, let me confess that I always found it a bit silly when people address their wives in mathematical books. But looking in this book, I recognize between every two words a little white thought on my wife Koštana. If you pull them all out, you get a big white thought, and all the words get mixed up.

To the reader.

Although it contains, I believe, no propositions which are not reused in the end, this thesis grew uncomfortably long. But for a reader, I am afraid, it might be uncomfortably short at some places. While trying to give a better oversight of a proof, I frequently introduce notations, where a sentence or two would do as well. This is perhaps wrong, perhaps a matter of taste. I like pictures rather than explanations.

Proofs, or outlines of proofs, are usually enclosed between thick points: •...•. Routine arguments are often omitted. The reader is assumed to understand category theory sufficiently to be able to look up, say, Johnstone's *Topos Theory* (1977) without difficulties. For the first and the last chapters, some acquaintance with type theory and its semantics is probably necessary (e.g. Martin-Löf 1984, and Seely 1984). A reader who wants to supply the inductive arguments omitted in section I.2, will perhaps need a bit more than that.

The list of references contains only those papers and books which are actually refered to somewhere in the text. I do not see the theme of this thesis as ripe for an exhaustive list of relevant literature.

Chapters are divided in sections, sections in subsections. A subsection is usually organized by bold subtitles; when necessary, it is subdivided by decimal numbers. For instance, paragraphs 3.12 and 3.111 are both in subsection 3.1, and 3.111 comes before 3.12. "II.3.12" denotes paragraph 2 in subsection 1, section 3 of chapter II; within chapter II, this paragraph is called "3.12"; within section 3, it is just "12".

Contents

Preface

Abstract iii

Acknowledgements viii

To the reader ix

Introduction 1

I. Theory of predicates 7

1. Type theories

1. Examples 9
2. Basic concepts 12
3. Sums and products. 18
4. Unit 22
5. Predicates. 22
6. Example 27
7. Constructions 29
8. Isomorphisms. 29

2. Translations

1. Plan 33
2. Definitions. 35
3. Comparing theories 36
4. Comparing calculi 40

II. Variable categories 45

1. Fibrations. A brief introduction

1. A motive. 47
2. Cartesianness 47
3. Fibrations. 50
4. The Grothendieck construction 54
5. Examples 56

6. Discrete fibrations....57
7. Lifting homotopies....60
2. Vertical structure
1. Basic notions, fibrewise....67
2. Fibrewise vs. global limits....71
3. Fibrewise fibrations....75
4. A lemma about adjointness....77
3. Horizontal structure
1. Hyperdoctrines....79
2. Co-, bi-, trifibrations....80
3. Beck-Chevalley condition/property....84
4. A characterisation of the Beck-Chevalley property....89
5. The Beck-Chevalley condition over all commutative squares....90
6. Hyperfibrations....92
7. Closure properties for co-, bi-, tri-, and hyperfibrations....94
8. Fibrewise co-, bi-, tri-, hyperfibrations....98
3a. The Beck-Chevalley condition without direct images
1. Interpolation condition....105
2. Uniform interpolation....109
3. Lemmas....111
4. Families of arrows
1. Examples....117
2. Relativisation....118
3. Intrinsic structures and closure conditions for families....119
4. Arrow cofibrations and factorisations....121
5. Right arrow bifibrations....127
6. rccc....128
7. Fibrewise rccc....131

III. Notions of size....133
1. Yoneda
1. Representable fibrations....135
2. Small fibrations....137
2. Fibred classes and predicates
1. Examples....145
2. Locally and globally small....147
3. Definability....153
4. Constructive comprehension....155
5. Associated arrow fibrations....160
3. Aspects of comprehension
1. Homogenous descriptions....161
2. Descriptions with colimits....162
3. Comprehension structures....164
4. Ehrhard's comprehension....165
5. Lawvere's comprehension....169
4. About extent fibrations
1. What do they inherit?....171
2. Weak coproducts....173
3. Which comprehensive fibrations are arrow fibrations?....178

IV. Semantics....181
1. Interpretation
1. Conceptual elements of the interpretation....183
2. Categories of predicates....188
3. Simplifying contexts....191
4. Definition of the interpretation....194
5. Internal language of a category of predicates....199
2. Term models
1. From a theory to a category of predicates....201
2. Semantical completeness....207
3. First steps
1. Generating small complete categories....211
2. Equality predicates....212
3. Describing functions....215
4. Procreation of models....223

References....231

Index....I

Curriculum vitae....V

Samenvatting (Summary in Dutch)....VII

Introduction

Question: What is a proof?

Answer 1. Formal logic. Begin with obviously true propositions and derive the conclusion by inference rules which preserve the truth, in the sense that the truth value of the premises always remains less or equal than the truth value of the conclusion.

Truth values are the elements of a poset (Lindebaum algebra).

Answer 2. Type theory. A *constructive* proof is some kind of a function (according to Brouwer, Heyting and Kolmogorov):

f proves $\forall x \in A.\varphi(x)$ <u>means</u> $f{:}A \longrightarrow \text{Proofs}$
<u>where</u> for every $a \in A$ $f(a)$ proves $\varphi(a)$;
f proves $\exists x \in A.\varphi(x)$ <u>means</u> $f = \langle a, f_a \rangle$
<u>where</u> f_a proves $\varphi(a)$,

etc.

The questions remain: *What kind of a function is a proof?* and: *What kind of a set is a proposition?* To realizability and beyond, various answers have been proposed. In general, they were some formal systems in which the terms had been recognized the dignity of constructive functions. By considering proofs-as-terms and formulas-as-types, the practice of *constructive-logic-as-type-theory* has been developed. - Not without a gain of generality and a loss of intuition: $\forall x \in A.\varphi(x)$ has become $\prod_{x \in A} \varphi(x)$, and there can be many different proofs for the same formula now.

Truth values are now collected in a category.

Answer 3. Semantics. A proof can be given by showing a model. "When the Eleatics argued that the movement did not exist, Diogenes stood up and walked around." Another example, characteristic for the New Age is: "Cogito, ergo sum": "The existence of this thought (or of this sentence, if you like) is a model of my existence". Methodically, the answer is being produced from the material given in the question itself; by its actions, the subject constitutes itself as an object.

In mathematics, any abstract group is concretized by its actions on its own underlying set (Cayley); any category is concretized by its representable functors (Yoneda)[1].

Upon this basic idea a sophisticated superstructure of spectral theories has been developed. It started from the Dedekind-style completisation of a poset by embedding it in the set of its own lower sets, and evolved through Stone's representation theorems to the theory of Grothendieck toposes. In it, the *dialectic* of certain

theories	- categorically presented in geometric logic,
their completions	- the classifying toposes, and
their models	- the points of these toposes

has been disclosed as the central issue.

Theme. We shall be concerned with the constructive proofs - in the sense of answer 2. We shall try to approach them semantically - in the sense of answer 3.

Constructive logic has at least three levels:

- *intuitionism* - i.e. Brouwer's ideology;
- *formally constructive logic* - i.e. formal logic without excluded middle;
- *logic with constructive proofs* - i.e. logic-as-type-theory.

[1]The importance of the Yoneda lemma for category theory can hardly be overestimated. It tells what it's all about. For instance, about the objects: In any category $\mathcal{C}$ the objects are determined (up to an iso) by the arrows to (or from) them, i.e. by the representable functors:

$$\mathcal{C}(A, B) \simeq \underline{Nat}(\nabla A, \nabla B)$$

(where $\nabla A: \mathcal{C}^{o} \to \underline{Set} : X \mapsto \mathcal{C}(X, A)$ is the functor represented by A, and $\underline{Nat}$ the natural transformations) - just as the sets are uniquely determined by their elements:

$$A = B \iff \forall x (x \in A \leftrightarrow x \in B).$$

A universe of sets with formally constructive logic is investigated in topos theory. Although this theory probably hasn't reached its maturity yet, computer science has propelled the question of a universe of sets with constructive logic in the strongest sense - i.e. with constructive proofs. This is our theme.

In chapter I a type theoretical generalisation of higher order predicate logic with a type of truth values and a constructive extent operation is introduced: the *theory of predicates*. In chapters II and III the corresponding categorical structures are considered: those corresponding to logical operations in chapter II, those characteristic for a universe of sets in chapter III. Putting them together, we define in chapter IV *categories of predicates*. The correspondence between the introduced categorical and type theoretical concepts is then spelled out. Some ways to produce categories of predicates as models for the theory of predicates are studied.

Method. The central part of this thesis is a *categorical* interpretation of logic with constructive proofs. To help us approach it, a *type theoretical* interpretation has been introduced. An effort has been made to keep the latter simple; the syntactical machinery has not been spelt out in detail. This is, of course, a subjective decision, and there is no doubt that a different approach, concentrated on syntactical aspects, would be at least as appropriate.

Conceived ninety years ago on the soil of creative subject, constructivism now surfaces in computers and in some mathematical structures. One often cannot help to feel that *there is not enough intuition for constructive logic* any more. Equality, structure, complexity of proofs do not seem to be a part of our everyday logical experience. Proof theory - direct syntactical study of formal systems - is probably one of the branches of mathematics with the highest price per result.

In recent years, a development of a semantical approach to type theory has started. In a very straightforward way, formal systems are interpreted in the metalanguage of category theory, and then some models are constructed, as categories with appropriate structure. The word "model" is in fact a bit stretched here: not the "meaning" - as in model theory - but the structure of a system is being modelled. The formal type theoretical expressions are actually just rewritten in terms of categorical operations. - So what is the gain? A tactical gain is that one more easily finds examples of a given structure, and perhaps "understands" it better (whatever that might mean!), since category theory is some kind of a natural language, rooted in mathematical practice. The strategic gain from interpreting a type theory categorically is that given *structures are characterized as properties*. Namely, while a type theoretical operation is decreed by syntactical rules, a category theoretical operation (e.g. $\varinjlim$) in principle originates from a property (cocompleteness). Connecting one with another creates a movement in both directions: category theoretical considerations sometimes lead to meaningful syntactical rules for constructive logic, when intuition does not help (e.g. the rule $\eta\Sigma$ in I.1.2); and type theory equips some complex categories (e.g. toposes) with an intuitive *internal language*: approaching them without this language can be like programming a computer using the machine code.

A similar interplay is going on between type theory and computer science. A system of logic with constructive proofs might serve as a "natural programming language", in which programs could closely follow the given specifications. Conversely, the complexity of such a system makes a computational approach to proof-checking in it indispensable.

Presently, the most efficient approach to logic with constructive proofs seems to be the combination of type theoretical and category theoretical formalisms and intuitions: type theory gives a picture with sharp lines, while category theory adds a third dimension to it. However, the inexorable difference between category theory and type theory - that one is about properties, the other about structures - which makes their contact fruitful, also makes a formal fusion impossible: cf. remark IV.1.1.

Context. The analogies shown in the rows of the following table (or two tables, glued in the middle) might offer some readers a rough orientation. - Of course, the alignments like this must be taken with a grain of salt!

concrete	abstract	generalized	
algebra	ring	Abelian category	
geometry	space	Grothendieck topos	
propositional logic	Heyting algebra	cartesian closed category	typed λ-calculus
higher order predicate logic with truth object	tripos	PL-category	le système Fω
... and with comprehension	elementary topos	category of predicates	theory of predicates
logic	Lindebaum algebra	categorical structure	type theory

Conceptually, categories of predicates should generalize elementary toposes! - This certainly doesn't have to mean that their theory will be as rich. In the worst case, they may turn out to be just another symptom of a "generalize!"-disease. I can only say that they did not arise from a pretension to generalize: I was only trying to understand the conception of constructive proofs at the confluence of three sources:

- *theory of constructions* (Coquand-Huet 1986, 1988, Hyland-Pitts 1987);
- *hyperdoctrines* (Lawvere 1970, Hyland-Johnstone-Pitts 1980, Seely 1987);
- *fibred categories* (Grothendieck 1959, Gray 1966, Bénabou 1975b, 1983, 1985).

The theory of predicates arose from the theory of constructions, and an observation how the multiplicity of constructive proofs spoils the *comprehension* principle. (Cf. I.1.54.) A lead towards a solution was found in Lawvere (1970), together with the complete conceptual equipment for categorical interpretation of logic with constructive proofs. The passage from Lawvere's indexed categories to fibrations is essential only as much as it is the step from structure to properties.

In the introduction to his thesis (1988), T. Ehrhard claimed that already the theory of constructions and a corresponding categorical structure - which he christened *dictos* -

appropriately generalize the notion of topos. (In a subsequent article (Ehrhard 1989) he went on to suggest that even Grothendieck toposes and geometric morphisms should be generalized on the same drift!) My work evolved independently from Ehrhard's, but it did start off from the conceptual basis of theory of constructions - which ultimately remains a special case of the theory of predicates, namely that in which truth values can be reduced to some sets. The relation between the two theories will be discussed in section 2 of chapter I. We shall see in chapter IV that a category of predicates is a model for the theory of constructions exactly when its category of propositions is fully generated by 1 (in the sense of III.4).

I. Theory of predicates

We are concerned with a type theoretical generalisation of (1) higher order (2) predicate logic with (3) constructive proofs and (4) the comprehension principle. The presentation which we are about to give reveals that these four components are just echoes of the same basic structure between two *sorts of types*, Ω and Θ, which we'll have all reasons to call *propositions* and *sets* respectively. This echoing may seem amazing, amusing, dubious or disappointing - it certainly makes the presentation shorter.

The idea for this presentation comes from H.P. Barendregt; it is simply to extend the typing relation (_:_) by one more level. This allows for a many sorted type theory: an expression $p{:}S{:}\Delta$ tells that p is a *term* of *type* S, and that type S is of *sort* Δ. The typing relation can then receive different meanings for different sorts. For instance, while $p{:}S{:}\Omega$ means that p is a proof of a proposition S, $p{:}S{:}\Theta$ tells that p is an element of a set S.

The same homonymy is then extended on the operations: $\prod$ represents a product on one side, a quantifier on the other. This is so because the rules, which define this product and this quantifier, also appear to coincide.

In section 1 we present the *theory of predicates* and the *theory of constructions*. The latter has recently been developed in computer science. It is the strongest type theory known to be consistent. Following a "set theoretical" intuition, we drop a "quarter" of the theory of constructions, and introduce a new, very simple operation - to define the theory of predicates. In section 2 these two theories are compared. The theory of constructions is not stronger: it can be translated as a special case of the theory of predicates - namely the one in which *logic is extensional* in the sense that propositions can be identified with a special class of sets.

I. Theory of predicates

Warning: Formal systems do their best here to look simple and natural. Some subtle structural questions, which, for instance, an implementation would have to answer, remain hidden behind the natural deduction notation.

1. Type theories

1. Examples.

Conceptually, type theory stems from logic. Formally, it can be regarded as generalized algebra. Algebra is about operations and equations on a set. Type theory is about partial operations and equations on indexed families of sets (which are now called types; their elements - terms). This point of view has been explained in detail by Cartmell (1986). We just give some examples of variable types and operations on them.

A category $\mathbb{A}$ consists of:

a constant type: $Ob_{\mathbb{A}}$

a variable type:

$$\frac{X:Ob_{\mathbb{A}} \quad Y:Ob_{\mathbb{A}}}{Hom_{\mathbb{A}}(X,Y)}$$

terms:

$$\frac{X:Ob_{\mathbb{A}}}{id(X):Hom_{\mathbb{A}}(X,X)}$$

$$\frac{\dfrac{X:Ob_{\mathbb{A}} \quad Y:Ob_{\mathbb{A}}}{f:Hom_{\mathbb{A}}(X,Y)} \qquad \dfrac{Y:Ob_{\mathbb{A}} \quad Z:Ob_{\mathbb{A}}}{g:Hom_{\mathbb{A}}(Y,Z)}}{\circ(f,g):Hom_{\mathbb{A}}(X,Z)}$$

satisfying

$$\frac{\dfrac{X:Ob_{\mathbb{A}} \quad Y:Ob_{\mathbb{A}}}{f:Hom_{\mathbb{A}}(X,Y)} \qquad \dfrac{Y:Ob_{\mathbb{A}}}{id:Hom_{\mathbb{A}}(Y,Y)}}{\circ(f,id) = f}$$

and two more equations.

A particular category $\mathbb{A}$ can be given by specifying its objects P,Q... as constant terms of type $Ob_{\mathbb{A}}$, and its arrows as constant terms of types $Hom_{\mathbb{A}}(P,Q)$. Alternatively, a particular category can be regarded as a *model* of this type theory. Indeed, by the usual categorical interpretation of indexed families as arrows (cf. Cartmell 1986, or Seely 1984), every model of the theory above in a (finitely complete) category $\mathcal{C}$ will be an internal category in it.

To present large categories, we should introduce two *sorts* of types: Sets and Classes, such that

P : Sets <u>implies</u> P : Classes

Then $Hom_{\mathbb{A}}(X,Y)$: Sets, while $Ob_{\mathbb{A}}$: Classes. Enriched categories could be defined in a similar fashion.

To perform in this type theory the categorical constructions involving the commutativity conditions, we must express the equality of arrows as a type. Therefore a sort of Propositions is needed, and an operation I:

$$\frac{f,g: P: \text{Sets}}{I(f,g) : \text{Propositions}}$$

where some additional rules give a term r:I(f,g) <u>iff</u> f=g.

If, furthermore, a category $\mathbb{B}$ and a functor $F : \mathbb{A} \longrightarrow \mathbb{B}$ are given:

$$\frac{X : Ob_{\mathbb{A}} : \text{Classes}}{F_0(X) : Ob_{\mathbb{B}} : \text{Classes}}$$

$$\frac{\dfrac{X,Y : Ob_{\mathbb{A}} \quad : \text{Classes}}{f : Hom_{\mathbb{A}}(X,Y) \quad : \text{Sets}}}{F_1(f) : Hom_{\mathbb{B}}(F_0(X),F_0(Y)) : \text{Sets}}$$

(plus equations saying that F_1 preserves id and $\circ$), then the operation I allows us, for instance, to express the fact that F is faithful:

$$\frac{\dfrac{\dfrac{X,Y : Ob_{\mathbb{A}} \quad : \text{Classes}}{f,g : Hom_{\mathbb{A}}(X,Y) \quad : \text{Sets}}}{x : I(F_1(f),F_1(g)) : \text{Propositions}}}{p : I(f,g) \quad : \text{Propositions}}$$

The *comprehension principle* plays a fundamental role in the type theory of Propositions and Sets. It can be expressed by the operation

$$\frac{\dfrac{[X:P \quad : \text{Sets}]}{\varphi(X) \quad : \text{Propositions}}}{\{X \in P \mid \varphi(X)\} : \text{Sets}}$$

which *binds* its variable X - i.e. $\{X \in P \mid \varphi(X)\}$ does not vary over X:P any more. This binding is denoted by [_] around X:P.

In a type theory with the operation Σ representing disjoint union, it is better to use the nonbinding operation ι of *extent* which formalizes the notion "such that". The intended meaning of

$$\frac{\dfrac{X:P : \text{Sets}}{\varphi(X) : \text{Propositions}}}{\iota(\varphi) : \text{Sets}}$$

is that $\iota(\varphi(X)) \neq \emptyset$ <u>iff</u> $\varphi(X)$ is true. Instead of $\{X \in P \mid \varphi(X)\}$ we now use $\Sigma X{:}P.\iota(\varphi)$.

Using the extent operation, we define, for instance, the slice category $\mathbb{A}/P$ for $P{:}Ob_{\mathbb{A}}$:

$Ob_{\mathbb{A}/P} := \Sigma X{:}Ob_{\mathbb{A}}.\ Hom_{\mathbb{A}}(X,P)$: Classes,

$Hom_{\mathbb{A}/P}(t,u) := \Sigma f{:}Hom_{\mathbb{A}}(\pi_0 t,\pi_0 u).\ \iota\big(I(\pi_1 t,\ \pi_1 u \circ f)\big)$: Sets,

where $t{:}Ob_{\mathbb{A}/P}$ gives $\pi_0 t : Ob_{\mathbb{A}}$, and $\pi_1 t : Hom_{\mathbb{A}}(\pi_0 t,P)$.

2. Basic concepts.

Now we shall sketch a picture of the formal type theory, touching up those nuances of this variegated field which are important in the sequel, or which might cause some confusion. By the way, the relevant keywords will be just mentioned; it is assumed that the reader already has some idea of their meaning. A discussion of standard notions can be found at the beginning of Martin-Löf 1984, and in Troelstra-van Dalen 1988, ch. 11.

Many-sorted algebra. To understand what kind of a formal system is a type theory, let us first consider the simplest fragment. There are two levels of articulation:

- *expressions* (or *words*) are strings of *operation* symbols; there is a set of distinguished letters Δ, Δ'..., which denote *sorts*;
- *statements* (i.e. *formulas*): besides *equations* P=Q, many sorted algebra admits the *sorting statements* P:Δ.

Each operation Φ is introduced by a *formation rule*

$$\frac{P_0{:}\Delta_0 \quad P_1{:}\Delta_1 \quad \ldots \quad P_{n-1}{:}\Delta_n}{\Phi P_0 \ldots P_n : \Delta'}$$

The set of *premises* $\Gamma = (P_i{:}\Delta_i)_{i \in n+1}$ describes the *arity* of Φ. The statement $\Phi P_0 \ldots P_n{:}\Delta'$ is the *conclusion* of Γ in the above rule. The operations with empty arity are *constants*. *Generators* (of an algebra) can be regarded as constants, and vice versa. Starting from either of them, by iterated application of formation rules, the *well-formed expressions* are obtained. Note that the class Λ of all the well-formed expressions of a many-sorted algebra L comes equipped with the relation of *derivability*

$$(\vdash) \subseteq \Lambda^* \times \Lambda,$$

the transitive reflexive closure of all the instances of formation rules given for L, where $\Lambda^* := \bigcup_{i \in \omega} \Lambda^i$. Assuming that there are no constants given (only generators), $\Gamma \vdash \gamma$ means that Γ is a bar in the parsing tree of γ.

Of course, there is also *equality*

$$(=) \subseteq \Lambda \times \Lambda,$$

the equivalence relation generated by all the instances of equations imposed on L. An *algebraic* study of L is concerned with the equality. A *grammatical* study is concerned with the derivability, parsing, and the structural recursion by which Λ is generated.

A *model for a many-sorted algebra* is a system of sets ⟦Δ⟧, ⟦Δ'⟧... to represent sorts, with some structure on them, to represent operations. If U is the union of all these sets ⟦Δ⟧..., the *model assignment* will be a mapping

$$[\![\_]\!] : \Lambda \longrightarrow U$$

such that P:Δ <u>implies</u> ⟦P⟧∈⟦Δ⟧, and P=Q <u>implies</u> ⟦P⟧=⟦Q⟧. The relation (⊢) plays no role here.

Type algebra is many-sorted algebra extended by one more level: besides sorted *types* P:Δ, there are typed *terms* p:P. Thus,

- expressions can be *terms* p, q..., *types* P, Q..., or sorts Δ, Δ'...;
- each kind of statements can appear terms as well as for types: there are equations p=q, and P=Q, typing statements p:P, and sorting statements P:Δ.

(These four kinds of statements correspond to Martin-Löf's *judgements*.) The well-formed terms are obtained by iterated application of *typing rules*, in the form:

$$\frac{p_0{:}P_0{:}\Delta_0 \quad p_1{:}P_1{:}\Delta_1 \quad \ldots \quad p_m{:}P_n{:}\Delta_n \quad P_{n+1}{:}\Delta_{n+1} \quad \ldots \quad P_{n+m}{:}\Delta_{n+m}}{\varphi p_0 \ldots p_m : \Phi P_0 \ldots P_{n+m} : \Delta'}$$

Here p:P:Δ abbreviates p:P <u>and</u> P:Δ. Each typing rule must reduce to a formation rule, when all the typing statements are omitted - i.e. when all the terms, outlined above, are removed from it. Inductively, one easily shows that every well-formed term must have a well-formed type, just as every well-formed type must be sorted. A well-formed term/type always occurs in a typing/sorting statement. To recover the intuitive difference between the subject "term f with type Q" and the statement "term f has type Q", lost in the language of type theory, we shall sometimes write f^Q in place of f:Q.

A *model for a type algebra* is again a system of sets ⟦Δ⟧, ⟦Δ'⟧...; their elements, representing types, are to be some sets again, containing representants for terms. Of course, it is now required that p:P <u>implies</u> ⟦p⟧∈⟦P⟧.

Contexts. Type theory is built up around the relations (⊢) and (=) in a similar fashion as type algebra. In fact, type algebras are missing only one dimension of type theory, though probably the most important one: *indexing*, as exemplified above (in part 1). It is represented formally by the device of *contexts*. The context of a type or term T is the set Γ of the *declared variables*, on which T may depend. We write $\Gamma \Rightarrow T$. (Variables, of course, constitute a distinguished class of terms.)

Each atomic type or term - i.e. a generator - must be given with a context. And the derivation rules now interfere with the contexts: they impose conditions not only on the expressions in premises, but also on the contexts; and they yield a conclusion with a context. The contexts are derived together with the well-formed expressions; the relations ($\vdash$) and ($\Rightarrow$) are defined by simultaneous recursion. In this way, *each well-formed type and term comes with a unique context.* The relation ($\Rightarrow$) can thus be regarded as a mapping DV which assigns to each type or term T a finite set of declared variables DV(T). T is said to be *closed* when its context DV(T) is empty.

Each type or term thus presents itself by an expression - its name - and a context. The name can be arbitrary - e.g. any letter will do for an atom - but the context an intrinsic structure of a type or term. Generating type theory is a dynamic process, because an atom may contain complex derived types in its context - and can be used in derivations only when all these types have been formed. (Cf. "Derivations" below.) Moreover, there are operations which act on the context of a type or term, without leaving any trace on the expression which denotes this type/term. Therefore, the derivations in type theory are not just the parsing trees of well-formed expressions. They are more like logical derivations. In fact, every logician will recognize the simultaneous recursion of ($\vdash$) and ($\Rightarrow$) as a *sequent calculus.*

The logical aspect of type theory is reflected in ($\vdash$), the algebraic aspect - in (=). The former was historically far more important: type theory was developed as logic with constructive proofs, the algebraic side being just a study of the equivalence of proofs. The equations imposed in a type theory are therefore usually called *conversion rules*, while the relation (=) is called *convertibility*[2]. Generators are called *atoms*, and constants – 0-ary operations - are *axioms*.

[2]Troelstra-van Dalen (1988, 9.4.17) call *conversion* the relation which consists of all the instances of conversion rules. For its transitive, reflexive and symmetric closure we use the term *convertibility* following Barendregt (1981, 3.1.5.).

To model type theory, discrete sets are not sufficient: they must be structured by some arrows, to interprete the indexing. In chapter IV we shall see how categories can be used for this.

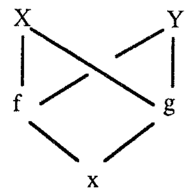
Declared variables vs. free variables. A paradigmatic picture of contexts can be acquired by considering the indexed families of eminently constructive sets: *data types*; their elements are the *functional programs*. In this setting, a variable represents an input gate; a context is a list of the *declarations of input data.* All the data used in a program must be declared. Some declared data may not be used in the program. And yet, even if the value of the output does not depend on the values of some of the declared input, the *existence* of the output always depends on the *existence* of the input: if some of the declared data do not exist, the program can never become executable. In other words, the context of a type/term T may contain a *dummy* variable x^P, i.e. one which is not used in the calculation of T; nevertheless, T depends on x^P, in the sense that it exists only if the type P is *inhabited*, i.e. if there is a closed term t^P, to be substituted for x^P.

Hence the difference between the context DV(T) and the set FV(T) of *free* variables of T, namely those variables which determine the value of T. Clearly, there is an inclusion $FV(T) \subseteq DV(T)$, and it can be proper. We use the common convention in accounting for the (relevant) elements of FV(T) in parentheses behind (the expression) T.

Structure of contexts. (Cf. Hyland-Pitts 1987, 1.3-4.) The type of a variable y can be indexed by another variable z: in order to know where to choose a value for y, we must be given a value of z. This provides a notion of natural partial order for each context:

$$y^Q \leq z^R :\Leftrightarrow z^R \in DV(y^Q).$$

(Note that $DV(x^P) := DV(P) \cup \{x^P\}$, so that $z^R \in DV(y^Q) \Leftrightarrow DV(z^R) \subseteq DV(y^Q)$.) This partial order can be seen as the relation of "being above" in the trees of variables in our examples. For instance, the context of the term p:I(f,g) is



We shall denote by MIN(T) the set of minima of DV(T). E.g., MIN(x)={f,g}.

Although this may be syntactically nontrivial, contexts will often be collapsed, to allow the sequent notation, e.g.

X,Y:Ob, f,g:Hom(X,Y), x:I(F(f),F(g)) $\Rightarrow$ p:I(f,g);

they will be truncated, or even omitted when no confusion seems likely.

Natural deduction. In concrete derivations, we shall expand contexts as trees of variables, like in the examples above. This means that *"being above" will denote both* ($\Rightarrow$) *and* ($\vdash$). This is the basic idea of the *natural deduction*. ("Being above" is conventionally denoted by a separating line: premises are written on a horizontal line above the conclusion, variables belonging to the context of a type or term are displayed on a horizontal line above its name. A double line represents several steps in a derivation.) A practical advantage of this notation is that the contexts need not be rewritten in derivations: a variable which was above a premise is above the conclusion too; the context of the conclusion can be made from the contexts of premises. Moreover, the partial order of a context presents itself in this notation in such a way that the permutations under which a context should be invariant are completely obvious, while the structural rules governing the manipulation with contexts, come as "natural". - This convenience does cause certain formal disadvantages, but they seem less important for our purposes.

Structural rules. When building derivations in natural deduction, one should certainly keep in mind the difference between ($\Rightarrow$) and ($\vdash$) - i.e. between *open assumptions* under which a formula is valid, and *premises* from which it can be derived. In this *type theoretical* natural deduction, the role of open assumptions is played by variables. This refers to the "coincidence" that the *free variables in a predicate obey the same structural rules as the open assumptions in a derivation*:

they can be used in any order (provided they are independent of each other), several times, or not at all.

(These structural rules will be derivable in the type theories which we are about to introduce. Cf. "Substitution" below.) In fact, variables are just *labels* for assumptions. If an ordinary assumption is in the form "Suppose that P is provable", assuming a variable X:P in a context can be understood as saying "Suppose that X is a proof of P".

By nature, variables satisfy the requirement that there are always sufficiently many of them: for every derivable type there must always exist a *fresh* variable. This can be expressed by the following rule, which will be assumed in all our systems.

$$\frac{P:\Delta}{X:P:\Delta}$$

In other words, before we assume that P has a proof, we must know that it is well-formed.

Derivations. We start a derivation from atoms, and build it by iterated application of derivation rules. Whenever a type has been formed, its variables can be assumed. A type/term can be introduced in a derivation only below its context, i.e. after all the types which occur there have been formed. When introduced, it is again available as a premise for a rule.

In these interactions between the rules and the introduced atoms, a *derivation tree* is built. Both ($\vdash$) and ($\Rightarrow$) are displayed in it. The context of every type or term must be contained in each of its derivations. (The bottom is considered as a part of the derivation too. A variable always occurs in its own context.)

Variations. This notion is specific for sorted type theories. Let Δ' and Δ" be sorts of types. If a variable X:P:Δ' occurs in the context of q:Q:Δ" (or of Q:Δ" alone), then the term q (resp. the type Q) is said to have *variation* Δ'Δ". A theory has variation Δ'Δ" if the types and terms in it are allowed to have this variation. If a variation is not allowed, we assume that neither types *nor terms* may have it.

A type with ΔΔ as its only variation is called *dependent*.

In type theories below, the rules will be parametrized by sorts: a Δ'Δ"-rule acts in principle on the variation Δ'Δ", i.e. on Δ'-variables in the context of Δ"-types/terms.

3. Sums and products.

The essence of logic-as-type-theory is that the fundamental type theoretical operations Π and Σ satisfy similar *introduction* and *elimination* rules as $\forall$ and $\exists$. Namely, by removing the terms - which we ouline for better visibility - from the typing rules for Π and Σ, the usual logical rules for $\forall$ and $\exists$ are obtained. (Recall: *Every typing rule reduces to a formation rule when stripped of terms.*)

Where $\forall$ and $\exists$ discharge an assumption in a derivation, Π and Σ *bind* a variable. As usually, binding is denoted by [_].

(The "coincidence" of the *declared* variables and the *open* assumptions is thus extended to the *bound* variables and the *closed* assumptions, used in derivations.)

Typing rules.

IΠΔ'Δ"

$$\dfrac{\dfrac{[X:P:\Delta']}{q:Q:\Delta''}}{\lambda X.q : \Pi X:P.Q : \Delta''}$$

condition: $X \in MIN(q:Q)$

EΠΔ'Δ"

$$\dfrac{p:P:\Delta' \qquad r:\Pi X:P.Q(X):\Delta''}{rp : Q[X:=p]:\Delta''}$$

IΣΔ'Δ"

$$\dfrac{p:P:\Delta' \qquad \dfrac{[X:P:\Delta']}{Q:\Delta''} \qquad q:Q(p):\Delta''}{\langle p,q\rangle : \Sigma X:P.Q : \Delta''}$$

condition: $X \in MIN(Q)$

EΣΔ'Δ"

$$\dfrac{r:\Sigma X:P.Q:\Delta'' \qquad \dfrac{\dfrac{[X:P:\Delta']}{[Y:Q(X):\Delta'']}}{s(X,Y):S(\langle X,Y\rangle):\Delta}}{v(r,(X,Y).s) : S(r) : \Delta}$$

conditions: $Y \in MIN(s:S)$; $\Delta \in \{\Delta',\Delta''\}$

Conversion rules.

$\beta\Pi\Delta'\Delta''$	$(\lambda X.\, q)p \;= q[X{:=}p]$	
$\eta\Pi\Delta'\Delta''$	$\lambda X.\, (tX) \;= t$	condition: $X \notin DV(t)$
$\beta\Sigma\Delta'\Delta''$	$\nu(\langle X,Y\rangle, (X,Y).s) \;= s$	
$\eta\Sigma\Delta'\Delta''$	$\nu(r, (X,Y).t(\langle X,Y\rangle)) = t(r)$	condition: $X,Y \notin DV(t)$

Equality. It is, of course, intended that the equals can replace each other. Hence the *equality rules*, which tell for each operation that it is well defined with respect to the relation of convertibility. We won't write them down.

Although no conversion rules have been given for types, the nontrivial convertibility is induced on them by the equality rule (in a condensed notation)

$$\frac{p = q}{R(p) = R(q)}$$

Remark. The type theory of Π and Σ is *an algebra of constructive proofs*. The typing rules are easily recognized as well known logical rules, enriched with terms to encode proofs[3]. The conversion rules then define an equivalence of proofs. But in formal logic any two proofs of a formula are equivalent. So the logical experience doesn't help us to choose the conversion rules. They are actually determined in the categorical interpretation of type theory. In chapter IV we shall see that the conversion rules given above just say that Σ and Π are respectively left and right adjoint to substitution. (Cf. also II.3.1.)

Substitution. Given $X{:}P \Rightarrow q{:}Q$, and $p{:}P$, we can *define* $Q[X{:=}p]$ and $q[X{:=}p]$ as the type and term obtained by first applying $I\Pi$, and then $E\Pi$. (In the rule $\beta\Pi$, the left side then defines the notation on the right side.) Otherwise, we could give a separate rule for

[3]Only $E\Sigma\Delta'\Delta''$ deviates from the usual form of the $\exists$-elimination by the presence of $Z{:}\Sigma X{:}P.Q$ in the "formula" S. Lemmas 3 explain that this rule in fact presents the disjoint union, rather than existential quantifier. Compare also proposition 52 and remark IV.1.4.

substitution, which would consist in replacing a minimal variable $X{:}P$ in the context of $q{:}Q$, by a term $p{:}P$ (under its context). ($\beta\Pi$ could then read: $(\lambda X.\, q)X = q$.)

In our systems, substitution includes the structural operations on contexts: *weakening* (i.e. adding a dummy) can be represented as a substitution (of two variables for one) along a *projection* π_i (see below), while *contraction* (i.e. reusing a variable) can be viexed as a substitution along a *diagonal* ρ (idem).

To provide a simplified notation and some bookkeeping of substitutions, we shall sometimes refer to some declared (and not just to free) variables in parentheses after an expression. When Z is clear from the context, we shall write $T(r)$ for $T[Z{:=}r]$; and $T[X{:=}p][Y{:=}q][Z{:=}r]$ will be abbreviated by $T(p,q,r)$.

Notation. As usually, if $X \notin DV(Q)$, then

$\Pi X{:}P.Q$ will be written as $P \to Q$, and

$\Sigma X{:}P.Q$ as $P \times Q$.

Furthermore, we abbreviate

$id_P := \lambda X^P.\, X^P$

$f \circ g := \lambda X.f(gX)$ where $f{:}Q \to R,\ g{:}P \to Q$

$\pi_0 := \lambda Z.\nu(Z,(X,Y).X)$

$\pi_1 := \lambda Z.\nu(Z,(X,Y).Y)$

$\rho := \lambda X.\langle X,X\rangle$.

Lemmas. 31. $\pi_i\langle X_0, X_1\rangle = X_i,\ i \in 2.$

• This is just $\beta\Sigma$.•

32. $\langle \pi_0 Z, \pi_1 Z\rangle = Z$

• $\langle \pi_0 Z, \pi_1 Z\rangle \overset{\eta}{=} \nu(Z, (X_0,X_1).\langle \pi_0\langle X_0,X_1\rangle, \pi_1\langle X_0,X_1\rangle\rangle) = \nu(Z, (X_0,X_1).\langle X_0,X_1\rangle) \overset{\eta}{=} Z$.•

33. $s(\pi_0 Z, \pi_1 Z) = \nu(Z, (X_0,X_1).s)$

34. $s(\langle X,Y\rangle) = t(\langle X,Y\rangle)$ then $s = t$

• $s(Z) \overset{\eta}{=} \nu(Z, (X,Y).s(\langle X,Y\rangle)) = \nu(Z, (X,Y).t(\langle X,Y\rangle)) \overset{\eta}{=} t(Z)$.•

4. Unit.

We shall assume that every sort in any type theory contains the *unit type* 1. It represents the empty context. A *closed* type - containing no free variables - can be viewed as varying over 1.

Rules.

$$1\Delta \qquad\qquad \emptyset{:}1{:}\Delta$$

$$\emptyset\Delta \qquad\qquad \frac{p{:}1{:}\Delta}{p = \emptyset}$$

Notation. $\emptyset_P{:}1_P$ will denote $\emptyset(X^P){:}1(X^P)$ (i.e. Ø:1 with a dummy X:P.)

5. Predicates.

In our theory of sets and propositions, the following notational convention will be respected whenever possible

Sorts	Types	Terms
Ω	propositions (truth values): α, β...; ξ. (variable)	proofs: a, b ,f...; x... (variable)
Θ	sets: H, K, M...; Ω	elements (functions): h, k..., u, v; X... (variable)
both sorts:		
Δ, Δ', Δ''	P, Q, R, S	p, q, r, s; X... (variable)

The axiom

$$\Omega{:}\Theta$$

("the collection of propositions is a set") entails that every proposition - a type in Ω - is also a term in Θ. Hence propositional variables ξ,ζ... : Ω among the variable propositions[4]. In principle, there are two sorts of variables: for elements and for proofs. ξ is just a notation for X^Ω. If T is a type or term we denote by EV(T) the set of the element variables in its context, while PV(T) is the set of its proof variables.

The symbols Ø:1 will be reserved for *singleton* Ø:1:Θ; *truth*, the unit in Ω, will be denoted by $\phi{:}\top{:}\Omega$.

Four variations (cf. part 1) are now possible and they correspond to the logical components which we invoked at the beginning:

(1) ΘΘ - higher order,
(2) ΘΩ - predicate logic,
(3) ΩΩ - propositional logic with constructive proofs,
(4) ΩΘ - comprehension principle.

51. (Ad 1) *Orders* are the sets generated by × and → from Ω alone. *Polymorphic types* are the propositions generated only from order variables.

ΠΘΘ and ΣΘΘ give products and sums of sets indexed by sets. Higher order is: being able to quantify over them, over exponents, over orders in particular.

52. (Ad 2) A *predicate* is a proposition indexed *only* over sets - an assignment of truth values to their elements.

In order to represent the quantifiers, ΠΘΩ and ΣΘΩ must be restricted to predicates. Furthermore, the rule EΣΘΩ must be restricted to Δ = Ω: we cannot obtain an element

[4]In general, note the difference between the element variables X^K and variable elements - i.e. functions $k^K(Z^M)$; and the difference between the proof variables x^β and variable proofs $f^\beta(X,y)$.

k:K from the truth value $\Sigma X{:}K.\alpha(X)$. (Otherwise a paradox can be derived: cf. Pavlović 1989.) This Σ-elimination rule can be further weakened: just as in the $\exists$-elimination in logic, forbid the dependency of the predicate S on the active variables X and Y. However, this last restriction is inessential: the resulting operation isn't any weaker.

Definition. *Quantifiers* $\forall$ and $\exists$ are respectively the operations $\Pi\Theta\Omega$ and $\Sigma\Theta\Omega$ defined with the following additional conditions:

on $\mathrm{I}\Pi\Theta\Omega$ and $\mathrm{I}\Sigma\Theta\Omega$	$PV(q{:}Q) = \emptyset$
on $\mathrm{E}\Sigma\Theta\Omega$	$X,Y \notin DV(S),\ \Delta := \Omega.$

The groups $\Pi\Theta\Omega$ and $\Sigma\Theta\Omega$, restricted like this, will be denoted by $\forall$ and $\exists$ respectively; the rules in these groups are $\mathrm{I}\forall$, $\eta\forall$, $\mathrm{E}\exists$ etc.

Proposition. The rule $\mathrm{E}\Sigma\Theta\Omega$, restricted to $\Delta=\Omega$, is derivable by means of $\exists$ and $\Sigma\Omega\Omega$. (The operation v', introduced in $\mathrm{E}\Sigma\Theta\Omega$, is defined in terms of v and $\langle\_,\_\rangle$ belonging to $\exists$ and $\Sigma\Omega\Omega$; and the conversion rules are satisfied.)

•

$$\dfrac{r : \exists X{:}K.\alpha : \Omega \qquad \dfrac{\dfrac{\dfrac{[X{:}K{:}\Theta]}{[y{:}\alpha{:}\Omega]}}{s(X,y) : \sigma(\langle X,y\rangle) : \Omega}}{\langle\langle X,y\rangle,\ s(X,y)\rangle{:}\Sigma z{:}(\exists X{:}K.\alpha).\sigma(z) : \Omega}}{v(r,\ (X,y).\langle\langle X,y\rangle,\ s(X,y)\rangle) : \Sigma z{:}(\exists X{:}K.\alpha).\sigma(z) : \Omega}$$

Denote the last term by n(r,s). The fact that $\pi_0 n(r,s) = r$ follows from lemma 34 (because $\pi_0 n(\langle X,y\rangle,s) = \langle X,y\rangle$). It is routine to check that

$$v'(r,\ (X,y).s) := \pi_1 n(r,s) : \sigma(r) : \Omega$$

satisfies the conversion rules.•

53. (Ad 3) The variation $\Omega\Omega$ - propositions indexed by proofs of other propositions - must be understood from the notion of constructive proof. If a constructive proof - an inference of one proposition from another - is a function, then a proposition must be the

set of proofs that it is true. If the constructions are (postulated to be) proofs, then the sets (of constructively given elements) and the propositions (with constructive proofs) boil down to the same thing. In that case, *there is a higher order predicate logic already within propositional logic*: the type $x{:}\alpha \Rightarrow \beta(x)$ is a predicate, if α is viewed as a set and every $\beta(x)$ as a truth value.

Built upon this idea, the Martin-Löf type theory opens an almost royal road to constructive logic. The constructive quantifiers are the products and sums, as intended from the beginning. The matter seems closed.

Fortunately, we can still ask for more: Martin-Löf type theory doesn't allow a type of truth values. Since all the types are truth values, it should have to be the type of all types, and Girard has shown that this causes a paradox. (See Troelstra-van Dalen 1988, 11.7.4.)

Therefore, the sets and propositions must remain in two separate sorts, each with its own sums and products, and with a calculus of predicates between them.

54. (Ad 4) The variation $\Omega\Theta$ is needed to pass from a proposition describing a set to the set itself. For instance - going back to examples 1 - if we define the equaliser

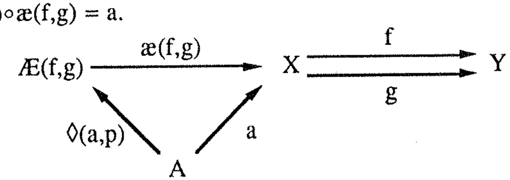
$$X,Y{:}Ob{:}\Theta,\ f,g{:}Hom(X,Y){:}\Theta \Rightarrow æ(f,g){:}Hom(Æ(f,g),X){:}\Theta$$

the factorisation through it will be given by

$$a{:}Hom(A,X){:}\Theta,\ p{:}I(f\circ a,\ g\circ a){:}\Omega \Rightarrow \Diamond(a,p){:}Hom(A,Æ){:}\Theta$$

satisfying

$$\Diamond(a,p)\circ æ(f,g) = a.$$



However, at a closer look this is a bit "too constructive": the factorisation $\lozenge(a,p)$ must exist when the proof p exists, but it shouldn't really depend on the choice of p. This dependency can be suppressed by a conversion rule.[5]

Whenever a set must be described *uniformly* by a proposition, we could first use $\Omega\Theta$, and then suppress it. But we shall rather dump $\Omega\Theta$ completely, and let the comprehension take care of itself.

Definition. The *extent* operation ι is given by the following rules:

$$\text{I}\iota \qquad \frac{a:\alpha:\Omega}{\delta a:\iota\alpha:\Theta}$$

condition: $PV(a:\alpha) = \emptyset$

$$\text{E}\iota \qquad \frac{k:\iota\alpha:\Theta}{\tau k:\alpha:\Omega}$$

$\beta\iota \qquad \tau(\delta a) = a$

$\eta\iota \qquad \delta(\tau k) = k$

$1\iota \qquad \iota\top = 1$

55. Definition. A *theory of predicates* (TOP) is a type theory with
- sorts Θ and Ω;
- variations $\Theta\Theta$, $\Omega\Omega$, and $\Theta\Omega$;
- operations $\prod\Theta\Theta$, $\sum\Theta\Theta$, $\prod\Omega\Omega$, $\sum\Omega\Omega$, $\forall$, $\exists$, ι.

The fragment without $\sum$ and $\exists$ is the *calculus of predicates* (COP).

A *strong* theory of predicates (STOP) is a theory of predicates with an additional operation, defined by:

[5]In fact, the premis $I(f\circ a, g\circ a)$ can even be completely avoided in this case: cf. Lambek-Scott 1986, 0.5.4.

$$\delta ab \qquad \frac{\langle a,b\rangle : \ \sum x{:}\alpha.\ \beta(x)}{\langle \delta a,b\rangle : \exists X{:}\iota\alpha.\beta(\tau X)}$$

$\beta\delta \qquad \nu(\langle\delta a,b\rangle, (X,y).s(\tau X,y)) = s(a,b)$

$\eta\delta \qquad \langle\delta(\tau k),b\rangle = \langle k,b\rangle.$

Comment. The pairing notation for $\langle\delta a,b\rangle$ is a mnemotechnic device which allows us to "derive" the conversion rules from those for $\sum$ and ι. In fact, $\eta\delta$ tells that the term $\langle\delta(\tau k),b\rangle$ obtained by δab for $a{:=}\tau k$ is equal to the "honest" pair $\langle\delta(\tau k),b\rangle$, obtained by $\text{I}\sum$. But if the term a contains a proof variable, the term δa cannot be formed and $\langle\delta a,b\rangle$ is really not a pair.

6. Example.

In chapters II and III the means will be developed to assign a semantics to the theory of predicates as a set theory with constructive proofs. Just for orientation, let us take a quick look at a degenerate model: one for set theory with formally constructive logic[6].

Let $\mathcal{S}$ be an elementary topos, $\mathbf{\Omega} := \left(\leq \rightrightarrows \Omega\right)$ the Heyting algebra of truth values in it. (Every $\mathcal{S}(K,\Omega)$ with the induced pointwise partial order $\overset{K}{\leq}$ is a Heyting algebra.) Clearly, $\mathcal{S}$ will give the sets and functions, $\mathbf{\Omega}$ the propositions and proofs. The main semantical framework is the *category of predicates* of $\mathcal{S}$:

$$|\mathcal{S}/\mathbf{\Omega}| \quad := |\mathcal{S}/\Omega|,$$

$$\mathcal{S}/\mathbf{\Omega}\,(\kappa,\mu) := \{u\in \mathcal{S}(K,M)\colon \kappa \overset{K}{\leq} \mu\circ u\},$$

[6]Cf. the introduction for the meaning of this expression.

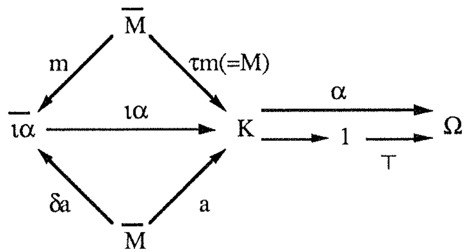
where $\kappa \in \mathcal{S}(K,\Omega)$, $\mu \in \mathcal{S}(M,\Omega)$. This category comes equipped with the obvious projection functor $\nabla\Omega : \mathcal{S}/\Omega \longrightarrow \mathcal{S}$. Denote by $\wp K$ the full subcategory of $\mathcal{S}/\Omega$ spanned by $\mathcal{S}(K,\Omega)$ (i.e. $\wp K := (\nabla\Omega)^{-1}K$.)

The interpretation is roughly as follows:

- The sets and functions (ΘΘ) are the objects and arrows in $\mathcal{S}$. The sets depending on a set K are the objects of $\mathcal{S}/K$... (Standard interpretation of a Martin-Löf type theory: Seely 1984.)
- The propositions and proofs (ΩΩ) are the objects and arrows in $\mathcal{S}/\Omega$. The predicates over K are in $\wp K$. (Since this is a Heyting algebra, there is at most one proof from α to β.) The propositions depending on $\alpha \in |\wp K|$ are the objects of $\wp K/\alpha$... (By the standard interpretation, a dependent proposition $x{:}\alpha \Rightarrow \beta(x)$ must be an "arrow" $\beta \leq \alpha$; $\prod x{:}\alpha.\beta(x)$ is then $\alpha \to \beta$, while $\sum x{:}\alpha.\beta(x)$ is just β.)
- The quantifiers (ΘΩ) are interpreted by the quantifiers from the topos.
- The extent (ι) of $\alpha \in |\wp K|$ is interpreted by

$$\iota\alpha := \{X \in K \colon \alpha(X)\} := æ(\alpha, \top_K),$$

where $\top_K := \top \circ \emptyset_K \colon K \longrightarrow 1 \longrightarrow \Omega$, while æ(f,g) denotes an equaliser of f and g.



Given $m{:}M \longrightarrow \iota\alpha$ in $\mathcal{S}/K$, $\tau m{:}\top_M \longrightarrow \alpha$ is $\iota\alpha \circ m$ (=M) in $\mathcal{S}/\Omega$. Given $a{:}\top_M \longrightarrow \alpha$ (i.e. $\alpha \circ a = \top_M$) in $\mathcal{S}/\Omega$, $\delta a{:}M \longrightarrow \iota\alpha$ (in $\mathcal{S}/K$) is its unique factorisation through $\iota\alpha$.

In a trivial way, this interpretation supports a strong theory of predicates: note that $\sum X{:}\iota\alpha.\iota\beta$ has the same meaning as $\iota(\sum x{:}\alpha.\beta)$.

7. Constructions.

In every topos there is a one-to-one correspondence between the predicates and their extents (provided that some representants are chosen for the subobjects). These can be identified: propositions can be regarded as a distinguished class of sets (namely, the subobjects of 1). With such an *extensional* logic there is no reason to avoid ΩΘ any more, hence no need for ι: the sets varying over propositions can be regarded as indexed over extents. Toposes contain such a logic.

Definition. A *theory of constructions* (TOC) is a type theory with

- sorts Θ and Ω;
- variations ΘΘ, ΩΩ, ΘΩ, ΩΘ;
- operations $\prod$ and $\sum$ for each of these variations; $E\sum\Theta\Omega$ is restricted to $\Delta{:=}\Omega$.

The $\prod$-fragment is called the *calculus of constructions* (COC).

Remark. The calculus of constructions was defined by Coquand and Huet (1986, 1988). The theory of constructions is due to Hyland and Pitts (1987). (The presentations were different.) Both attracted much attention, from computer science as well as logic. They appear to be the strongest consistent calculus/theory of types presently available. Allowing all the complicated contexts they look much stronger than the calculus/theory of predicates for one. In section 2 we shall see that this impression is not quite true.

8. Isomorphisms.

But before we start relating theories, we must relate operations in each of them. Are all these $\prod$s and $\sum$s really independent from each other?

Definition. Types P and Q are *isomorphic* ($P \simeq Q$) if there are terms (isos) $f^P(Y^Q)$ and $g^Q(X^P)$, such that $f(g(X)) = X$ and $g(f(Y)) = Y$. (Using λ, there are $f'\colon Q \to P$ and $g'\colon P \to Q$ such that $f' \circ g' = id$ and $g' \circ f' = id$.)

Lemmas. The first two of the following lemmas are about the theory of constructions, the remaining seven about the theory of predicates. However, to prove each particular statement, only the rules for the operations mentioned in it are needed.

81. $P \times 1 \simeq P$.

82. The statement:

$$\underline{\text{if}}\ P \simeq P'\ \underline{\text{and}}\ X{:}P \Rightarrow Q(X) \simeq Q'(X)\ \underline{\text{then}}\ \Box X{:}P.\,Q(X) \simeq \Box X{:}P'.\,Q'(f(X))$$
$$\underline{\text{where}}\ X{:}P' \Rightarrow f(X){:}P \text{ realizes } P \simeq P'.$$

holds for all the combinations of sorts for P, P', Q, Q', and for $\Box \in \{\Sigma, \Pi\}$, *with one exception*:

$$M \simeq \mu\ \underline{\text{does not imply}}\ \Sigma X{:}K.M \simeq \Sigma X{:}K.\mu.$$

• The exception will become clear in the semantics. The positive part is very easy for $P, P'\colon \Delta'$ and $Q, Q'\colon \Delta''$. With terms $k^M(x^\mu)$ and $a^\mu(X^M)$ realizing $M \simeq \mu$, the following seven cases remain:

$$\Pi X{:}K.M \simeq \Pi X{:}K.\mu$$
$$\Box x{:}\alpha.\ M \simeq \Box x{:}\alpha.\ \mu$$
$$\Box x{:}\mu.\ K(x) \simeq \Box X{:}M.\ K(a(X))$$
$$\Box x{:}\mu.\ \alpha(x) \simeq \Box X{:}M.\ \alpha(a(X)).$$

We prove only $\Sigma x{:}\mu.\ \alpha(x) \simeq \Sigma X{:}M.\ \alpha(a(X))$. The terms

$$z : \Sigma x{:}\mu.\alpha(x) \Rightarrow \langle k(\pi_0 z), \pi_1 z\rangle : \Sigma X{:}M.\ \alpha(a(X)) \text{ and}$$
$$Z\colon \Sigma X{:}M.\ \alpha(a(X)) \Rightarrow \nu(Z, (X,y).\langle a(X), y\rangle) : \Sigma x{:}\mu.\alpha(x)$$

should realize the isomorphism.

1) $\nu(\langle k(\pi_0 z), \pi_1 z\rangle, (X,y).\langle a(X), y\rangle) \overset{\beta}{=} \langle a(k(\pi_0 z)), \pi_1 z\rangle = \langle \pi_0 z, \pi_1 z\rangle \overset{\eta}{=} z$

2) $\langle k\big(\pi_0 \nu(Z, (X,y).\langle a(X),y\rangle)\big), \pi_1 \nu(Z, (X,y).\langle a(X),y\rangle)\rangle \overset{\eta}{=}$
$\nu\Big(Z,(U,v).\langle k\big(\pi_0 \nu(\langle U,v\rangle,(X,y).\langle a(X),y\rangle)\big),\pi_1 \nu(\langle U,v\rangle,(X,y).\langle a(X),y\rangle)\rangle\Big) \overset{\beta}{=}$
$\nu\big(Z,(U,v).\langle k(\pi_0\ \langle a(U),\ v\rangle),\ \ \pi_1\ \langle a(U),\ v\rangle\ \rangle\big) =$
$\nu\ (Z,\ (U,v).\ \langle k(a(U)),\ v\rangle) = \nu\ (Z,\ (U,v).\ \langle U,\ v\rangle) \overset{\eta}{=} Z$•

83. $\Sigma X{:}\iota\alpha.\iota\big(\beta(\tau X)\big) \simeq \iota(\Sigma x{:}\alpha.\beta)$

84. $\Pi X{:}\ K.\iota\beta \simeq \iota(\forall X{:}K.\beta)$

85. $\iota\alpha \simeq \iota(\iota\alpha \times \top)$

• The iso is realized by

$$X{:}\iota\alpha \Rightarrow \delta\langle X,\emptyset\rangle : \iota(\iota\alpha \times \top)$$
$$W{:}\iota(\iota\alpha \times \top) \Rightarrow \delta\nu(\tau W, (X,\emptyset).\tau X) : \iota\alpha$$

As usually, one identity is trivial, and the other requires $\eta\Sigma$:

$$\delta\langle\delta\nu(\tau W, (X,\emptyset).\tau X),\ \emptyset\rangle \overset{\eta}{=} \delta\nu\big(\tau W,\ (X,\emptyset).\langle\delta\nu(\langle X,\emptyset\rangle,\ (X,\emptyset).\tau X),\ \emptyset\rangle\big) \overset{\beta}{=}$$
$$\delta\nu\big(\tau W,\ (X,\emptyset).\langle\delta\tau X,\ \emptyset\rangle\big) = \delta\tau W = W\text{•}$$

86. $\Pi X{:}\iota\alpha.\iota\big(\beta(\langle X,\emptyset\rangle)\big) \simeq \iota(\Pi x{:}(\iota\alpha\times\top).\beta(x))$

• We just give the isos.

$$Z{:}\Pi X{:}\iota\alpha.\iota\big(\beta(\langle X,\emptyset\rangle)\big) \Rightarrow \delta\lambda x.\nu\big(x,\ (X,\emptyset).\tau(ZX)\big){:}\ \iota(\Pi x{:}(\iota\alpha\times\top).\beta(x))$$
$$W : \iota(\Pi x{:}(\iota\alpha\times\top).\beta(x)) \Rightarrow \lambda X.\delta\big((\tau W)\langle X,\emptyset\rangle\big) : \Pi X{:}\iota\alpha.\iota\big(\beta(\langle X,\emptyset\rangle)\big)\text{•}$$

87. Assuming δab:

$$\alpha \simeq \iota\alpha \times \top$$

• This is realized by

$$x : \alpha \Rightarrow \langle\delta x,\ \emptyset\rangle : \iota\alpha \times \top \text{ and}$$
$$z\colon \iota\alpha \times \top \Rightarrow \nu\big(z,\ (Z,\emptyset).\tau Z\big) :\alpha.$$

Here is one of identities:

$$\langle\ \delta\nu\big(z,\ (Z,\emptyset).\tau Z\big),\ \emptyset\rangle \overset{\eta}{=}$$
$$\nu\big(z,\ (Z,\emptyset).\langle\delta\nu(\langle Z,\emptyset\rangle,\ (Z',\emptyset).\tau Z'),\ \emptyset\rangle\big) \overset{\beta}{=}$$
$$\nu\big(z,\ (Z,\emptyset).\langle\delta(\tau Z),\ \emptyset\rangle\big) = \nu\big(z,\ (Z,\emptyset).\langle Z,\ \emptyset\rangle\big) = z\text{•}$$

88. Assuming δab:

$$\Sigma x{:}\alpha.\beta(x) \simeq \exists X{:}\iota\alpha.\beta(\tau X).$$

• Consider

$$z : \Sigma x{:}\alpha.\beta(x) \;\Rightarrow\; \langle \delta\pi_0 z,\, \pi_1 z\rangle : \exists X{:}\iota\alpha.\ \beta(\tau X) \text{ and}$$

$$w: \exists X{:}\iota\alpha.\ \beta(\tau X) \;\Rightarrow\; \nu\big(w,\ (X,y).\langle \tau X, y\rangle\big) : \Sigma x{:}\alpha.\beta(x).$$

The identity on $\Sigma x{:}\alpha.\beta(x)$ is easy; the one on $\exists X{:}\iota\alpha.\ \beta(\tau X)$ is obtained similarly as the one shown in 88.•

89. Combining the above results, we conclude that with δab

$$\iota(\Box x{:}\alpha.\beta) \;\simeq\; \iota(\Box X{:}\iota\alpha.\beta) \;\simeq\; \Box X{:}\iota\alpha.\iota\beta$$

$$\Box x{:}\alpha.\beta \;\simeq\; \Box X{:}\iota\alpha.\beta \;\simeq\; (\Box X{:}\iota\alpha.\iota\beta)\times\top$$

holds for $\Box \in \{\Sigma, \Pi\}$.

Remark. The rule δab is derivable in a theory of predicates <u>iff</u> $\Sigma x{:}\alpha.\beta \simeq \exists X{:}\iota\alpha.\beta$. The rule δaø, obtained by restricting δab to $\beta{=}\top$, is derivable <u>iff</u> $\alpha \simeq \iota\alpha\times\top$.

The following chain of isomorphisms shows that δab is derivable from δaø:

$$\Sigma x{:}\alpha.\beta \overset{\delta a\emptyset}{\simeq} \iota(\Sigma x{:}\alpha.\beta)\times\top \overset{83}{\simeq} (\Sigma X{:}\iota\alpha.\iota\beta)\times\top \overset{\#}{\simeq} \exists X{:}\iota\alpha.(\iota\beta\times\top) \overset{\delta a\emptyset}{\simeq}$$

$$\simeq \exists X{:}\iota\alpha.\beta.$$

The step (#) is based on: $\exists Z{:}(\Sigma X{:}K.L).\varphi \simeq \exists X{:}K.\exists Y{:}L.\varphi$.

2. Translations

1. Plan.

In this section the theories of predicates and of constructions will be compared. It will turn out that the strong theory of predicates is equivalent with the theory of constructions, and that the calculus of predicates is stronger than calculus of constructions (in a sense defined below). This will mainly result from some considerations about subtheories without variations over propositions.

Terminology. A *system* A is a triple

$$A = \langle \text{Sorts}_A, \text{Variations}_A, \text{Operations}_A\rangle$$

used to define a class of type theories. For instance, TOP, COC etc. are systems. A particular A-theory $\Lambda = A(\Xi)$ is generated by the Operations_A (defined by a set of rules) from a given class Ξ of atomic types and terms, which only have Sorts_A and Variations_A.

A *subsystem* $B \subseteq A$ has $X_B \subseteq X_A$ for $X \in \{\text{Sorts, Variations, Operations}\}$. (More restrictions can be imposed on the typing rules in B. But B is assumed to contain the whole group of rules by which any of its operations is defined in A; in particular all the conversion rules.)

If Λ is an A-theory, a class of types and terms $M \subseteq \Lambda$ is its B-*subtheory* if it is a B-theory, for $B \subseteq A$. (So M has Sorts_B and Variations_B, and it is closed in Λ under Operations_B.)

Systems. For every system A introduced in section 1, we define a system $A_\Theta = A{\upharpoonright}\Theta$ in which only the variations and operations over sets are allowed. Since $\Omega\Omega$ is not allowed, there are no proof variables in these systems.[7] There are only sets and predicates.

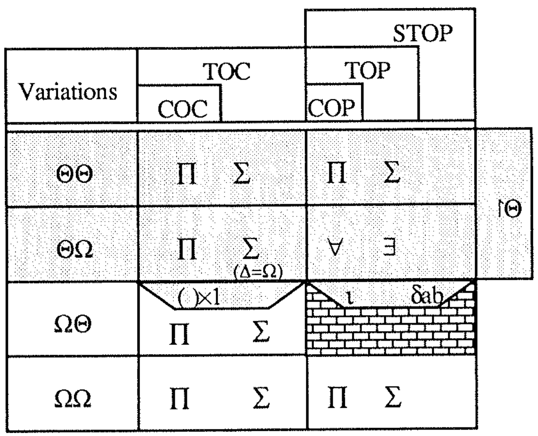
In addition to the operations for ΘΘ and ΘΩ, we stipulate that TOP_Θ and COP_Θ contain ι, while TOC_Θ and COC_Θ have the ΩΘ-operation $(\_)\times 1$:

$$\frac{a:\alpha:\Omega}{\langle a,\emptyset\rangle:\alpha\times 1:\Theta}$$

$$\frac{k:\alpha\times 1:\Theta}{\pi_0 k:\alpha:\Omega}$$

Note that $COC_\Theta \not\subsetneq COC$; all other A_Θ are proper restrictions, i.e. subsystems of A.

So we have the following systems:



[7] Logically, this can be understood as scholastic rigidity: "Nothing can be predicated about predicates."

2. Definitions.

21. Let Λ and M be two arbitrary type theories. A *translation* $F:\Lambda \longrightarrow M$ is an algorithm which prescribes how to transform every type from Λ into a type from M and every term from Λ into a term from M, so that

variables go to variables, and

the relations (_:_), (=) and (⊢) are preserved.

Let A and B be two systems. A *translation* $F:A \longrightarrow B$ is an algorithm which assigns to every A-theory Λ a B-theory FΛ and a translation $F_\Lambda:\Lambda \longrightarrow F\Lambda$.

22. A (B-)subtheory M of (an A-theory) Λ is a *retract* of Λ if there is a translation $F:\Lambda \longrightarrow M$ such that for every type P from Λ

$$F(P) \simeq P.$$

A system A is *conservative* over a subsystem B if there is a translation $F:A \longrightarrow B$ such that every FΛ is a retract of Λ by F_Λ.

23. The systems A and B are *equivalent* if there are translations $F:A \longrightarrow B$ and $G:B \longrightarrow A$, such that for every A-theory Λ and B-theory M

$$GF\Lambda \subseteq \Lambda \ \underline{\text{and}}\ FGM \subseteq M,$$

and for all types P from Λ and Q from M

$$GF(P) \simeq P \ \underline{\text{and}}\ FG(Q) \simeq Q$$

(with obvious subscripts).

Comments. The idea is that systems should be equivalent if and only if they have the same class of models. For instance, group theory using $\langle \cdot, (\_)^{-1}, 1\rangle$ is equivalent to the one with $\langle -, 0\rangle$ (where $-$ is the subtraction).

The theory of Boolean algebras with $\langle \vee,\wedge,\rightarrow,\neg,0,1\rangle$ is conservative over the one using only $\langle \wedge,\rightarrow,0\rangle$; the sequent calculus with cut rule is conservative over the one without.

Remark. To eliminate superfluous operations means to remove some synonyms from a language. In principle, the language becomes harder to speak: e.g. the cut-elimination yields somewhat unnatural proofs. But it becomes easier to understand: less ambiguous, closer to semantics.

As far as type theories are concerned, we want to consider synonymous exactly those isomorphic types that will be identified semantically. We shall be looking for isomorphisms that will be interpreted as identities in the models.

3. Comparing theories.

Propositions.

31. TOC is conservative over TOC_Θ.

32. STOP is conservative over $STOP_\Theta$.

33. STOP and TOC are equivalent.

Proof of 31. Given a TOC-theory Λ, we define a translation $E:\Lambda \longrightarrow \Lambda$, such that

$$E(P) \simeq P$$

for every type P, while the image of E is a TOC_Θ-subtheory $\Lambda_\Theta \subseteq \Lambda$.

The idea is that E should translate every variation $\Delta'\Delta''$ - where $\Delta',\Delta'' \in \{\Omega,\Theta\}$ - into variation $\Theta\Delta''$ and all the $\Delta'\Delta''$-operations into the corresponding $\Theta\Delta''$-operations. Roughly speaking, E just replaces every proposition α in a context by the set $\alpha\times 1$. The variables are substituted into an E-image along the terms d to keep the image isomorphic with the original.

For an arbitrary type or term T we define

$$E\big(\ldots X^Q \ldots \Rightarrow T(\ldots X^Q \ldots)\big) := \ldots X^{D(Q)} \ldots \Rightarrow \llbracket T \rrbracket\big(\ldots d_Q X^{D(Q)} \ldots\big)$$

$$D\big(\ldots X^Q \ldots \Rightarrow T(\ldots X^Q \ldots)\big) := \ldots X^{D(Q)} \ldots \Rightarrow [T]\ \big(\ldots d_Q X^{D(Q)} \ldots\big)$$

<u>where</u>

$$\llbracket A \rrbracket := A \qquad \text{\underline{where} A is a letter}$$

$$\llbracket \Box X{:}P.Q \rrbracket := \Box X{:}\ [P].\ \llbracket Q \rrbracket$$

$$\llbracket \lambda X.q \rrbracket := \lambda X.\ \llbracket q \rrbracket$$

$$\llbracket pq \rrbracket := \llbracket p \rrbracket\ [q]$$

$$\llbracket \langle p,q \rangle \rrbracket := \langle [p], \llbracket q \rrbracket \rangle$$

$$\llbracket \nu(r,\ (X,Y).s) \rrbracket := \nu(\llbracket r \rrbracket,\ (X,Y).\llbracket s \rrbracket)$$

$$[\alpha] := \llbracket \alpha \rrbracket \times 1 \qquad [K] := \llbracket K \rrbracket$$

$$[a] := \langle \llbracket a \rrbracket, \emptyset \rangle \qquad [k] := \llbracket k \rrbracket$$

$d_Q : D(Q) \rightarrow Q$

$$d_\alpha := e_\alpha \circ \pi_0 \qquad \tilde{d}_\alpha := \lambda x.\langle \tilde{e}_\alpha x, \emptyset \rangle$$

$$d_K := e_K \qquad \tilde{d}_K := \tilde{e}_K$$

$e_Q : E(Q) \rightarrow Q$

$$e_A := id_A \qquad \tilde{e}_A := id_A \qquad \text{\underline{where} A is a letter}$$

$$e_{\Box X:P.Q} := v_\Box \circ w \qquad \tilde{e}_{\Box X:P.Q} := \tilde{w} \circ \tilde{v}_\Box$$

$v_\Box : \big(\Box X{:}E(P).E(Q)\big) \rightarrow \big(\Box X{:}P.Q\big)$

$$v_\Pi := \lambda Z.\ e_Q \circ Z \circ \tilde{e}_P \qquad \tilde{v}_\Pi := \lambda Z.\ \tilde{e}_Q \circ Z \circ e_P$$

$$v_\Sigma := \nu(Z,\ (X,Y).\langle e_P X,\ e_Q Y \rangle) \qquad \tilde{v}_\Sigma := \nu(Z,\ (X,Y).\langle \tilde{e}_P X,\ \tilde{e}_Q Y \rangle)$$

$w : \big(\Box X{:}D(P).E(Q)\big) \rightarrow \big(\Box X{:}E(P).E(Q)\big)$ is an iso from lemma 1.82.

The substitution operation is translated

$$E(T[X:=f]) := E(T)[X:=D(f)]$$

$$D(T[X:=f]) := D(T)[X:=D(f)].$$

An inspection of the definition of E now shows that it preserves (_:_) and (=). An inductive argument which follows the recursive definition of the iso $e_Q{:}E(Q) \rightarrow Q$ shows that $\Gamma \vdash p$ <u>implies</u> $E(\Gamma) \vdash E(p)$. So E is a translation. Its image Λ_Θ is obviously a TOC_Θ-subtheory (i.e. closed under the TOC_Θ-operations).

So TOC is conservative over TOC_Θ.•

Proof of 32. The approach is completely the same: Given a STOP-theory M, we define a translation $H:M \longrightarrow M$, with $H(P) \simeq P$ for every P, and such that $M_\Theta := im(H)$ is a $STOP_\Theta$-subtheory. H will translate the variation $\Omega\Omega$ into $\Theta\Omega$ by replacing every α in a context by $\iota\alpha$.

H is defined using algorithms $[\![\_]\!]=[\![\_]\!]_H$ and $[\_]=[\_]_H$ for names, and some terms j for the substitution of variables

$$H(\ldots X^Q \ldots \Rightarrow T(\ldots X^Q \ldots)) := \ldots X^{J(Q)} \ldots \Rightarrow [\![T]\!](\ldots j_Q X^{J(Q)} \ldots)$$
$$J(\ldots X^Q \ldots \Rightarrow T(\ldots X^Q \ldots)) := \ldots X^{J(Q)} \ldots \Rightarrow [T](\ldots j_Q X^{J(Q)} \ldots)$$

for an arbitrary type or term $T(\ldots X^Q \ldots)$.

The definition of $[\![\_]\!]_H$ is the same as that of $[\![\_]\!]_E$, with

$$[\![\iota\alpha]\!] := \iota[\![\alpha]\!]$$
$$[\![\delta a]\!] := \delta[\![a]\!]$$
$$[\![\tau k]\!] := \tau[\![k]\!]$$

as additional items. But $[\_]_H$ is

$$[\alpha] := \iota[\![\alpha]\!] \qquad [K] := [\![K]\!]$$
$$[a] := \delta[\![a]\!] \qquad [k] := [\![k]\!]$$

The difference with the translation of the context is that there is no terms from propositions to sets in STOP, hence no iso from $H(\alpha)$ to $J(\alpha)$. However, lemma 1.87 tells that in a strong theory of predicates α can still be recovered from $\iota\alpha$ - up to an iso. The basic idea for the translations in this proposition is to extend $\iota\alpha\times\top \simeq \alpha$ to $J(\alpha)\times\top \simeq H(\alpha)$ $(\simeq\alpha)$. This way, H(P) and P can be kept isomorphic despite the fact that their contexts cannot always be connected with each other by isos.

$j_Q : J(Q) \rightarrow Q$

$$j_\alpha := h_\alpha \circ \tau$$
$$j_K := h_K$$

$h_Q : H(Q) \rightarrow Q$

$$h_A := id_A \qquad \tilde{h}_A := id_A \qquad \underline{\text{where}}\text{ A is a letter}$$
$$h_{\iota\alpha} := \delta \circ h_\alpha \circ \tau \qquad \tilde{h}_{\iota\alpha} := \delta \circ \tilde{h}_\alpha \circ \tau$$
$$h_{\Box X:P.Q} := v_\Box \circ w \qquad \tilde{h}_{\Box X:P.Q} := \tilde{w} \circ \tilde{v}_\Box$$

$v_\Box : (\Box X{:}H(P).H(Q)) \rightarrow (\Box X{:}P.Q)$ is defined exactly as above, everywhere with h in place of e.

$w : (\Box X{:}J(P).H(Q)) \rightarrow (\Box X{:}H(P).H(Q))$ is an iso from lemma 1.89. (We need δab when $\Box = \Sigma$ and $P,Q{:}\Omega$.)

The substitution is:

$$H(T[X:=f]) := H(T)[X:=J(f)]$$
$$J(T[X:=f]) := J(T)[X:=J(f)].$$

The preservation properties of H are checked in the same way as above, for E. And just as above, the image M_Θ of H is clearly closed under $STOP_\Theta$-operations in M - i.e. it is a $STOP_\Theta$-subtheory. Hence the result.•

Proof of 33. (We use the notation from the preceding proofs.)

The translations $F_\Theta : TOC_\Theta \longrightarrow STOP_\Theta$ and $G_\Theta : STOP_\Theta \longrightarrow TOC_\Theta$ are easy to guess. F_Θ just rewrites the expressions from a TOC_Θ-theory Λ_Θ and replaces:

$$\alpha\times 1 \mapsto \iota\alpha$$
$$\langle a,\emptyset\rangle \mapsto \delta a$$
$$\pi_0 k \mapsto \tau k,$$

while G_Θ goes in the opposite direction with the expressions from a $STOP_\Theta$-theory M_Θ. The preservation properties are immediate: the rules which define each pair of corresponding operations are completely analogous. (The restrictions which distinguish $\forall$, $\exists$ and ι from respectively $\prod\Theta\Omega$, $\sum\Theta\Omega$ and $(\_)\times 1$ are superfluous in the absence of $\Omega\Omega$.) We can say that TOC_Θ and $STOP_\Theta$ are isomorphic.

Given a TOC-theory Λ, define $F\Lambda$ to be the smallest STOP-theory containing the $STOP_\Theta$-theory $F_\Theta\Lambda_\Theta$ (i.e. its closure under $\Omega\Omega$-operations). Given a STOP-theory M, let GM be the smallest TOC-theory which contains $G_\Theta M_\Theta$ (i.e. its closure under $\Omega\Omega$-, and $\Omega\Theta$-operations). Obviously, $GF\Lambda \subseteq \Lambda$ <u>and</u> $FGM \subseteq M$.

Further define for every Λ and M the translations $F = F_\Lambda : \Lambda \longrightarrow F\Lambda$ and $G = G_M : M \longrightarrow GM$ by

$$F := F_\Theta \circ E, \text{ and}$$

$G := G_\Theta \circ H$

(with E and H as in the proofs of the preceding propositions). Obviously, arbitrary types P from Λ and Q from M are translated

$G \circ F(P) = E(P)$ <u>and</u> $F \circ G(Q) = H(Q)$.

The required isos

$G \circ F(P) \simeq P$ <u>and</u> $F \circ G(Q) \simeq Q$

are thus e_P and h_Q constructed above.

So F and G realize an equivalence of TOC and STOP.•

Remark. The danger of working modulo isos is that whole groups (of automorphisms) can be swept away: reduced to an identity. This doesn't happen if the types (i.e. the type schemes) are identified only along *unique canonical* isos. "Canonical" here means: uniformly defined for all types, natural, "meta-polymorphic". The isos used in the previous propositions are obviously canonical. A curious reader will perhaps want to check that they are unique. (The assertions are: For every canonical iso $f_Q:E(Q)\to Q$, $D(f_Q)=id_{D(Q)}$ <u>implies</u> $f_Q=e_Q$; for every canonical iso $g_Q:H(Q)\to Q$, $J(g_Q)=id_{J(Q)}$ <u>implies</u> $g_Q=h_Q$.) - Strictly speaking, the unique canonical isos should have been demanded already by definitions 2. We refrained from this for the sake of simplicity.

4. Comparing calculi.

The naive idea behind our manipulations with systems is: "Reduce everything to sets". However, if you simply "hit" every atomic proposition in TOC by $(\_)\times 1$ and transform all the $\prod$s and $\sum$s to $\prod\Theta\Theta$ and $\sum\Theta\Theta$ respectively, some types will become isomorphic which previously weren't - due to the exception in lemma 1.82. It took the simultaneous recursion of D and E in 31 above, using not only $\prod\Theta\Theta$ and $\sum\Theta\Theta$, but $\prod\Theta\Omega$ and $\sum\Theta\Omega$ too, to circumvent this - as to obtain $D(P)\simeq D(Q)$ <u>iff</u> $P\simeq Q$. A similar story can be told with STOP, ι, J and H instead of TOC, $(\_)\times 1$, D and E. - The translations D and J

"reduce everything to sets" in such a way that the images are isomorphic exactly when the originals are.

However, propositions 3 actually required a bit more: an iso between each type and its image (by a translation). It was convenient (even necessary in 32) that translations preserve sorts. We therefore used E and H rather than D and J. Besides, E and H helped us to define Λ_Θ and M_Θ neatly from Λ and M. *But* the fact is that $E:\Lambda \longrightarrow \Lambda_\Theta$ and $H:M \longrightarrow M_\Theta$ are rather poor as morphisms: there are no terms $x:E(\alpha)\Rightarrow f:E(\beta)$, even if $\alpha\simeq\beta$. - The translations E and H are not "functorial", while D and J are.

Definition. Let Λ and M be two arbitrary type theories. A *translation* $F:\Lambda \longrightarrow M$ is *sound* if it also preserves the relation (⇒) (i.e. the contexts) and the substitution.

A sound translation F is *full* if for every term $F(\Gamma)\Rightarrow r:F(P)$ in M

there is $\Gamma\Rightarrow p:P$ in Λ such that $r=F(p)$,

where $F(\Gamma)$ is obtained from Γ by translating each element. It is *faithful* if for every pair $\Gamma\Rightarrow p,q:R$

$F(p)=F(q)$ <u>implies</u> $p=q$.

If a sound translation is both full and faithful, we say that it is *complete*.

Let A and B be two systems. A translation $F:A \longrightarrow B$ is *sound/full/faithfull/complete* if all its components $F_\Lambda:\Lambda \longrightarrow F\Lambda$ are.

We say that a system B is *stronger* than A if there is a complete translation $F:A \longrightarrow B$.[8]

Comments. A stronger system is meant to have a greater expressive power: every A-theory must be completely interpreted in some B-theory if B is stronger. Ring theory is stronger than group theory; predicate logic is stronger than propositional logic.

A complete translation F establishes for every type $\Gamma\Rightarrow P$ a bijection between collections of terms

[8]Note that the relation "stronger" is reflexive; it should perhaps be called "at least as strong as". But as far as the following proposition is concerned, COP is strictly stronger than COC, i.e. COC is not stronger than COP.

$F_{\Gamma,P} : \{\Gamma \Rightarrow p{:}P\} \longrightarrow \{F(\Gamma) \Rightarrow r{:}F(P)\}$.

From the formulas-as-types point of view, this correspondence between the proofs in the original and those in the image seems to generalize the usual notion of a complete interpretation: "Exactly the images of provable formulas are provable".

Remark. An inclusion $B \subseteq A$ (with components $M \hookrightarrow M^A$, where M^A is the closure of M under the A-operations) is always a faithful translation, but it needn't be full. The inclusions $TOC_\Theta \subseteq TOC$ and $STOP_\Theta \subseteq STOP$ are obviously not complete. On the other hand, $D{:}TOC \longrightarrow TOC_\Theta$ and $J{:}STOP \longrightarrow STOP_\Theta$ are complete translations. The semantics will later tell us that complete translations $TOP \longrightarrow TOP_\Theta$ and $COP \longrightarrow COP_\Theta$ do not exist: the operations in theory of predicates are too independent - TOP and COP are *essentially* richer than their respective restrictions.

Proposition. COP is stronger than COC.

Proof. The idea is that $D{:}TOC \longrightarrow TOC_\Theta$ restricts to a complete translation

$C{:}COC \longrightarrow COC_\Theta$.

Given a COC-theory Λ, define Λ_1 to be its closure under the $\Omega\Theta$-operation $(\_)\times 1$. $C\Lambda$ is then the image by E of Λ_1 in itself. The translation $C_\Lambda : \Lambda \longrightarrow C\Lambda$ is just a restriction of the algorithm for D.

By the definiton (of D), C preserves the contexts and the substitution. The other preservation properties follow similarly as for E.

For every type $\Gamma \Rightarrow P$ in Λ there is a mapping

$C_{\Gamma,P} : \{\Gamma \Rightarrow p{:}P\} \longrightarrow \{C(\Gamma) \Rightarrow r{:}C(P)\} : p \longmapsto C(p)$.

The difference between $\Gamma \Rightarrow p{:}P$ and $C(\Gamma) \Rightarrow C(p){:}C(P)$ is that the latter eventually contains $(\_)\times 1$, $\langle \_,\emptyset \rangle$ and π_0, which C introduced. The algorithm $\tilde{C}_{\Gamma,P}$:= "remove $(\_)\times 1$, $\langle \_,\emptyset \rangle$ and π_0 from $C(\Gamma) \Rightarrow r{:}C(P)$" is easily seen to define a mapping

$\tilde{C}_{\Gamma,P} : \{C(\Gamma) \Rightarrow r{:}C(P)\} \longrightarrow \{\Gamma \Rightarrow p{:}P\}$,

- inverse to $C_{\Gamma,P}$. - So C is a complete translation.

Since F_Θ and G_Θ restrict to isomorphisms of COC_Θ and COP_Θ, all this remains true for ι, δ, τ in place of respectively $(\_)\times 1$, $\langle \_,\emptyset \rangle$ and π_0. Hence there is a complete translation

$K{:}COC \longrightarrow COP$,

with $K\Lambda$ defined to be the smallest COP-theory containing the COP_Θ-theory $F_\Theta C\Lambda$ and

$K_\Lambda := F_\Theta \circ C_\Lambda$.•

II. Variable categories

This chapter is about *fibrations* as variable categories. The theory of fibrations, or *fibred categories*, is just category theory relative to a base category. Ordinary categories can be viewed as fibred over 1.

Fibred categories have been defined by Grothendieck (1959) for the purposes of algebraic geometry. Accidentally or not, no appropriate introductory text on them is available yet. The only general references (known to me) are: Grothendieck 1971, Gray 1966 and Bénabou 1983. The first two have been written more than 25 years ago, the third one is unfinished and unpublished. So we have to start by working our own way through, to the facts which are in part probably well known to some people, or used to be well known some time ago.

The main definitions and basic facts about fibrations are surveyed in section 1. We couldn't afford to give complete proofs, but an effort has been made to arrange this folklore material in such a way that a dilligent reader could supply them using elementary category theory. The fibrewise versions of some common categorical notions are examined in section 2. Section 3, on the other hand, is devoted to some concepts having no ancestors in the ordinary category theory: the left and right direct images, and the Beck-Chevalley property. In subsection 3a a characterisation of the Beck-Chevalley property is given in which the direct images are not mentioned. (It becomes possible to extend this property from bifibrations to fibrations in general.) Section 4 finally lists some facts about *arrow fibrations*, which are particularly important for interpretation of type theories.

It should be stressed that this chapter is not intended as an introducion into fibred category theory. It introduces *only* those aspects of fibrations which are really needed for our interpretation of the theory of predicates in chapter IV. Nothing has been included here that could be left out - without causing even more work. A reader with no patience for categorical abstraction should perhaps just skim through this chapter, and come back later, when he needs to.

1. Fibrations. A brief introduction

1. A motive.

If a family of sets indexed over a set B is a functor from (the discrete category) B to Set, then a family of categories indexed over a category $\mathbb{B}$ can be viewed as a (pseudo)functor from $\mathbb{B}$ to Cat. For every set B there is a trivial one-to-one correspondence of the B-indexed sets and the functions to B:

$$\int_B : \underline{Set}^B \longrightarrow \underline{Set}/B : \{\gamma_x \mid x \in B\} \mapsto \left(\sum_{x\in B} \gamma_x \longrightarrow B \right).$$

A similar correspondence $\int_{\mathbb{B}} : \underline{Cat}^{\mathbb{B}} \longrightarrow \underline{Cat}/\mathbb{B}$ exists for indexed categories, but it is not trivial, and not surjective. In the first approximation, fibrations over $\mathbb{B}$ are the functors to $\mathbb{B}$ which lie in the image of $\int_{\mathbb{B}}$, i.e. those which correspond to some indexed categories. (Cf. 4 below.) In fact, the nuance neglected in this approximation contains much of the conceptual power of fibrations. (Cf. Bénabou 1983, 1985.)

2. Cartesianness.

21. Conventions, notations. Categories will be denoted by script letters $\mathcal{A}$, $\mathcal{B}$, $\mathcal{E}$...; small categories by $\mathbb{A}$, $\mathbb{B}$, $\mathbb{E}$... Small categories are objects of the (two-)category Cat. We shall also use metacategories such as CAT, which contains Set, Cat, $\mathcal{A}$, $\mathcal{B}$, $\mathcal{E}$,... - but mostly for commodity and better view[1].

[1]This "commodity" relies upon the effectiveness: the functors on metacategories are defined as *procedures*, i.e. receipes how to transform an input an output. One can interprete the quantifiers in this way too, and define certain universal constructions in

If $F:\mathcal{A}\to\mathcal{B}$ and $G:\mathcal{C}\to\mathcal{B}$ are functors, we denote their comma category by F/G. The category of arrows $\mathcal{B}/\mathcal{B}$ is comma category of two identity functors on $\mathcal{B}$. The slice category $\mathcal{B}/I$ (over an object $I\in|\mathcal{B}|$) is the comma category of the identity on $\mathcal{B}$ and the constant functor $\ulcorner I\urcorner:1\to\mathcal{B}$. The functors

$$\text{Cod}: \mathcal{B}/\mathcal{B}\to\mathcal{B} : (v:J\to I) \mapsto I \text{ and}$$

$$\text{Dom}: \mathcal{B}/I\to\mathcal{B} : (v:J\to I) \mapsto J$$

will be denoted by $\nabla\mathcal{B}$ and ∇I respectively.

Some data will tacitly be carried over from statement to statement. In particular, we shall mostly be concerned with a category $\mathcal{E}$, fibred over a base $\mathcal{B}$ by a functor $E:\mathcal{E}\to\mathcal{B}$. The following convention on letters will be respected whenever possible:

Categories	Objects	Arrows
fibred: $\mathcal{E},\mathcal{E}',\mathcal{F},\mathbb{E}$	X, Y, Z, A, B, C	f, g, s, t cartesian: ϑ; vertical: a, b, c
base: $\mathcal{B},\mathcal{S},\mathbb{B}$	H, I, J, K, M	h, k, m, u, v

Now we proceed to define what is a fibred category, cartesian arrows, vertical arrows.

22. Terminology. Let $E:\mathcal{E}\to\mathcal{B}$ be a functor, $I\in|\mathcal{B}|$. The *fibre* of E over I is the category $\mathcal{E}_I$:

$$|\mathcal{E}_I| := E^{-1}(I)$$

$$\mathcal{E}_I(X,Y) := E^{-1}(id_I)\cap\mathcal{E}(X,Y).$$

Furthermore, for every $u\in\mathcal{B}(I,J)$, $X\in\in|\mathcal{E}_I|$ and $Z\in|\mathcal{E}_J|$, we denote

$$\mathcal{E}_u(X,Z) := E^{-1}(u)\cap\mathcal{E}(X,Z),$$

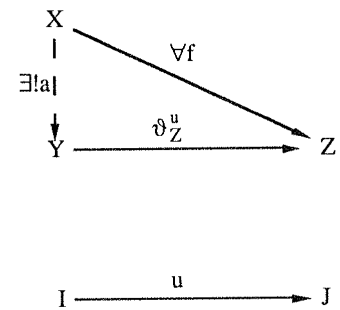
i.e., the set of all arrows $X\to Z$ over u. Hence functors

$$\mathcal{E}_u(\_,Z) : \mathcal{E}_I{}^o \to \underline{\text{Set}}.: X\mapsto\mathcal{E}_u(X,Z)$$

23. Proposition. The following statements are equivalent:

a) Every functor $\mathcal{E}_u(\_,Z)$ is representable.

b) Every arrow $u\in\mathcal{B}(I,J)$ has a terminal lifting at every $Z\in|\mathcal{E}_J|$. In other words, there is an arrow ϑ^u_Z over u (i.e. $E\vartheta^u_Z=u$) through which every other arrow f over u factorizes by a unique arrow a over id.



• If Y represents $\mathcal{E}_u(\_,Z)$, then $\vartheta^u_Z\in\mathcal{E}_u(Y,Z)$ corresponds to $id\in\mathcal{E}_I(Y,Y)$ by the representation isomorphism $\mathcal{E}_u(X,Z)\simeq\mathcal{E}_I(X,Y)$ natural in X.•

24. Definitions. An arrow $\vartheta\in\mathcal{E}(Y,Z)$ is called (E-)*cartesian* if it is a terminal lifting of $E\vartheta$ at Z. Y is then an *inverse image* of Z along $E\vartheta$.

An arrow $a\in\mathcal{E}$ is called (E-)*vertical* if Ea=id.

$F:\mathcal{E}'\to\mathcal{E}$ is a *cartesian functor* from $E':\mathcal{E}'\to\mathcal{B}$ to $E:\mathcal{E}\to\mathcal{B}$ if EF=E' and F takes the E'-cartesian arrows to the E-cartesian ones. A *natural transformation* $\varphi:F'\to F$ between $F,F':\mathcal{E}'\to\mathcal{E}$, EF=EF', is (E-)*cartesian* if all its components are E-vertical. An *adjointness* $\langle F:\mathcal{E}'\to\mathcal{E}, G:\mathcal{E}\to\mathcal{E}', \eta: id\to GF, \varepsilon: FG\to id\rangle$ is *cartesian* with respect to $E':\mathcal{E}'\to\mathcal{B}$ and $E:\mathcal{E}\to\mathcal{B}$ if all its components are cartesian.

25. Comment. The habit of calling all these notions cartesian stems presumably from the fact that the Cod-cartesian arrows are the cartesian (i.e. pullback) squares.

26. Facts. Under the (equivalent) conditions from proposition 23, E is

full <u>iff</u> every hom-set of every fibre is inhabited (nonempty),

some metacategories. (Comma categories, for instance.) - Metacategories are considered as far as this can be done *locally*, and *effectively*.

faithful <u>iff</u> every hom-set of every fibre has at most one element (i.e. the fibres are preorders).

An adjointness $\langle F, G, \eta, \varepsilon\rangle$ is cartesian with respect to E' and E (all data as in 24) <u>iff</u> F and G are cartesian <u>and</u> for all arrows $f \in \mathcal{E}(FA,C)$ holds $Ef = E'f'$, where $f' \in \mathcal{E}'(A,GC)$ is the right transpose of f.

3. Fibrations.

31. Proposition. The statements below are related as follows:

$$\begin{array}{ccc} (a)\Leftrightarrow(b)\Leftrightarrow(c) \\ \Uparrow \qquad\qquad \Downarrow AC \\ (d)\Leftrightarrow(e)\Leftrightarrow(f) \end{array}$$

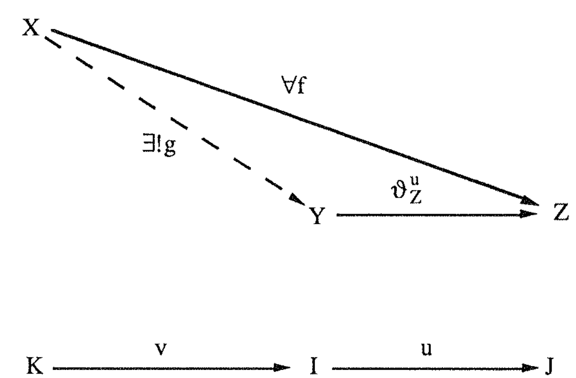
("AC" means that the axiom of choice is needed for this implication.)

a) For every $u \in \mathcal{B}(I,J)$, $Z \in |\mathcal{E}_J|$ there is $Y \in |\mathcal{E}_I|$, such that

$$\mathcal{E}_{u \circ v}(\_,Z) \simeq \mathcal{E}_v(\_,Y)$$

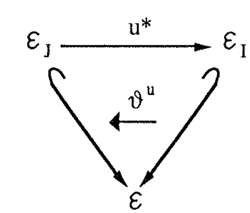
is realized by composition with an arrow $\vartheta^u_Z \in \mathcal{E}_u(Y,Z)$.

b) Every arrow $u \in \mathcal{B}(I,J)$ has at every $Z \in |\mathcal{E}_J|$ a lifting ϑ^u_Z such that for every $v \in \mathcal{B}(K,I)$ and f over uv there is a unique g over v, $f = \vartheta^u_Z \circ g$.



c) Every arrow $u \in \mathcal{B}(I,J)$ has a cartesian lifting at every $Z \in |\mathcal{E}_J|$ and cartesian arrows are closed under composition.

d) For every $u \in \mathcal{B}(I,J)$ there is an inverse image functor u^* and a natural transformation ϑ^u over u (i.e. $E(\vartheta^u_Z) = u$ for all $Z \in \mathcal{E}_J$):



Composing with ϑ^u_Z gives $\mathcal{E}_u(\_,Z) \simeq \nabla(u^*Z)$.

Furthermore, the natural transformations

$$c^{uv} : v^*u^* \longrightarrow (uv)^*,$$

induced as unique factorisations by ϑ, are isomorphisms.

e) For every $X \in |\mathcal{E}|$ the functor

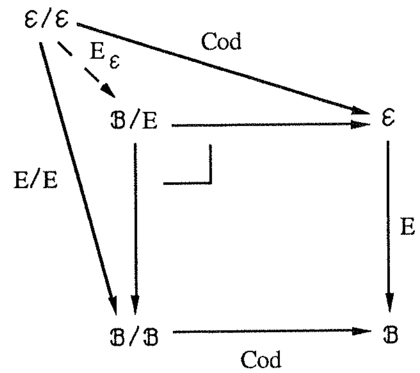
$$E_X : \mathcal{E}/X \longrightarrow \mathcal{B}/EX : f \mapsto Ef$$

has a right adjoint right inverse Θ_X.

f) The functor

$$E_{\mathcal{E}} : \mathcal{E}/\mathcal{E} \longrightarrow \mathcal{B}/E \;\; : f \mapsto \langle Ef, Cod(f)\rangle$$

has a right adjoint right inverse $\Theta_{\mathcal{E}}$.



• (c)⇒(a): For arbitrary u,v and Z, (c) gives Y, W and cartesian liftings $\vartheta_Z^u \in \mathcal{E}_u(Y,Z)$ and $\vartheta_Y^v \in \mathcal{E}_v(W,Y)$. The fact that $\vartheta_Z^u \circ \vartheta_Y^v$ is cartesian means that it induces $\mathcal{E}_{u\circ v}(\_,Z) \simeq \nabla W$ (by composition). But ϑ_Y^v induces $\mathcal{E}_v(\_,Y) \simeq \nabla W$.

(d)⇔(e): $\vartheta_Z^u = \Theta_Z(u)$. The unit of the adjunction, $\eta_f: f \to \Theta_Z E_Z(f)$, is the factorisation of f through ϑ_Z^{Ef}. (e)⇔(f): $\Theta_Z(u) = \Theta_{\mathcal{E}}(u,Z)$.•

32. Definition. A *fibration* is a functor $E : \mathcal{E} \to \mathcal{B}$ which satisfies (any of) the conditions (a-c) from the preceding proposition; the category $\mathcal{E}$ is *fibred* over $\mathcal{B}$. If E satisfies conditions (d-f) too, it is a *cloven* fibration: the triple $\langle (\_)^*, \vartheta, c\rangle$ is its *cleavage*. A cleavage (and the fibration E to which it belongs) is *normal* if all id_I^* are identity functors, $I \in |\mathcal{B}|$; it is *split* if $u^*v^* = (vu)^*$ (i.e. c^{vu}=id) for every composable $u,v \in \mathcal{B}$. (Without loss of generality, it can usually be assumed that cleavages are normal. But see example 51.)

The morphisms between fibrations are *cartesian functors*. The morphisms of cloven fibrations must preserve cleavages. The category of cartesian functors $E' \to E$ is denoted by $\underline{CART}_{\mathcal{B}}$ (E',E) or $\underline{FIB}/\mathcal{B}$(E',E) (- if it exists); the category of cleavage preserving functors will be $\underline{CLEAV}_{\mathcal{B}}$ (E',E). $\underline{Fib}/\mathbb{B}$ is the (two-)category of small fibred categories over $\mathbb{B}$, with cartesian functors (and cartesian natural transformations). Cleav/$\mathbb{B}$ is the analogous (two-)category of small cloven fibred categories with cleavage preserving functors.

33. Potential structure, subfibrations. Being a fibration is a property of a functor (just as having products is a property of a category). By the axiom of choice, this property can be fixed as a structure: a cleavage can be chosen. In practice, one works with this *potential structure* of fibred categories, i.e. *as if* it were given. If a fibration is not cloven, u*X denotes in it the domain of an *arbitrary* cartesian lifting ϑ_X^u (just as X×Y denotes an arbitrary product when no particular choice of products is given). Locally, we can work as if u* were a functor: given u*X, u*Y and $f \in \mathcal{E}(X,Y)$, the inverse image $u^*(f) \in \mathcal{E}(u^*X, u^*Y)$ is defined as the unique factorisation of $f \circ \vartheta_X^u$ through ϑ_Y^u. Then u*(id) = id and $u^*(f) \circ u^*(g) = u^*(f \circ g)$ hold.

A subalgebra must be closed under all the operations from its signature. By analogy, a *subfibration* $E':\mathcal{E}' \to \mathcal{B}$ of $E:\mathcal{E} \to \mathcal{B}$ (given with a cartesian inclusion $\mathcal{E}' \hookrightarrow \mathcal{E}$) should be closed under all the *potential* operations: with every object X, $\mathcal{E}'$ must contain *all* the E-cartesian arrows $\vartheta_X \in \mathcal{E}(Y,X)$. (In other words, every $\mathcal{E}'_I$ must be closed under the isomorphisms in $\mathcal{E}_I$. Without this requirement, the intersection $\mathcal{E}' \cap \mathcal{E}''$ of fibred categories $E':\mathcal{E}' \to \mathcal{B}$ and $E'':\mathcal{E}'' \to \mathcal{B}$ with cartesian inclusions $\mathcal{E}' \hookrightarrow \mathcal{E}$ and $\mathcal{E}'' \hookrightarrow \mathcal{E}$ - may fail to be a fibred category. Cf. Bénabou 1983, 1.4.)

34. Closure properties. For every functor E the class of E-cartesian arrows is closed under left division (i.e. if f and f∘g are cartesian <u>then</u> g is cartesian). When E is a fibration, this class is also closed (in $\mathcal{E}$) under composition and stable under pullbacks along vertical arrows (and vertical arrows are stable under pullbacks along cartesian arrows), which allways exist.[2]

The class of fibrations is stable under all pullbacks, and closed under the composition. Every fibration $F:\mathcal{E}' \to \mathcal{E}$ is a cartesian functor to the fibration $E:\mathcal{E} \to \mathcal{B}$ from E':=EF. The converse does not hold (i.e. not every cartesian functor is a fibration) and the class of fibrations is not closed under left division. However, if E'=EF and F are fibrations,

[2]Given a partial binary operation φ on the arrows of $\mathcal{B}$, we say that a class of arrows $\mathfrak{a} \subseteq \mathcal{B}$ is *closed* under φ if $a,a' \in \mathfrak{a}$ <u>and</u> φ(a,a') exists <u>imply</u> $\varphi(a,a') \in \mathfrak{a}$; and we say that $\mathfrak{a}$ is *stable* under φ if for every $u \in \mathcal{B}$, $a \in \mathfrak{a}$ <u>and</u> φ(u,a) exists <u>imply</u> $\varphi(u,a) \in \mathfrak{a}$. When φ=pulling back, φ(u,a)=u*a is the arrow obtained by pulling back a along u.

and if all the fibres of F are inhabited, <u>then</u> E must be a fibration too. (• An F-image of an E'-cartesian lifting of $u \in \mathcal{B}(I,J)$ at an object $W \in |\mathcal{E}'|$ such that $FW=Z$ is an E-cartesian lifting of u at Z.•)

Since small fibred categories are stable under pullbacks, the functor Cod: $\underline{\mathrm{Fib}} \longrightarrow \underline{\mathrm{Cat}}$ is a fibration, with fibres $\underline{\mathrm{Fib}}/\mathbb{B}$, and with cartesian liftings defined by pulling back.

4. The Grothendieck construction.

41. The correspondence $\int_B : \underline{\mathrm{Set}}^B \longrightarrow \underline{\mathrm{Set}}/B$ comes down to the fact that every set C with a function $c: C \longrightarrow B$ can be recovered (up to a bijection) from the indexed set

$$\gamma : B \longrightarrow \underline{\mathrm{Set}}: x \longmapsto c^{-1}(x)$$

of its fibres. Given a category $\mathbb{E}$ with a functor $E: \mathbb{E} \longrightarrow \mathbb{B}$, in order to define the arrow part of the corresponding indexed category

$$\Gamma : \mathbb{B}^o \longrightarrow \underline{\mathrm{Cat}} : I \longmapsto \mathbb{E}_I$$

in the first place, we need a representant u^*Z for each of the functors

$$\mathbb{E}_u(\_,Z) : \mathbb{E}_I{}^o \longrightarrow \underline{\mathrm{Set}}.$$

But each function

$$u^* : |\mathbb{E}_J| \longrightarrow |\mathbb{E}_I| \ : \ Z \longmapsto u^*Z,$$

obtained in this way, can be uniquely extended to a functor (• $f \in \mathbb{E}_J(Z,W)$ induces $\varphi : \mathbb{E}_u(\_,Z) \longrightarrow \mathbb{E}_u(\_,W)$ by composition, and φ induces $u^*(f)$ by the Yoneda lemma•), and we can define

$$\Gamma(u) := u^* : \mathbb{E}_J \longrightarrow \mathbb{E}_I.$$

This is why we want the functor $E: \mathbb{E} \longrightarrow \mathbb{B}$ to have cartesian liftings: by proposition 23, all functors $\mathbb{E}_u(\_,Z)$ will then be representable, and E will correspond to an indexed category Γ. We see from proposition 31 that $\Gamma(u \circ v) \simeq \Gamma(v) \circ \Gamma(u)$ and $\Gamma(id) \simeq id$ hold if these cartesian liftings are closed under the composition. Hence the notion of a fibration. And if $E: \mathbb{E} \longrightarrow \mathbb{B}$ is a fibration, then it can be recovered from the corresponding indexed category - i.e. pseudofunctor.

42. Proposition. Let $\underline{\mathrm{Cat}}^{\mathbb{B}^o}$ be the category of pseudofunctors $\mathbb{B}^o \longrightarrow \underline{\mathrm{Cat}}$. There is an isomorphism of categories

$$\int_{\mathbb{B}} : \underline{\mathrm{Cat}}^{\mathbb{B}^o} \longrightarrow \underline{\mathrm{Cleav}/\mathbb{B}} : \Gamma \longmapsto \big(E : \mathbb{E} \longrightarrow \mathbb{B}\big),$$

where

$$|\mathbb{E}| := \sum_{I \in |\mathbb{B}|} \Gamma I,$$

$$\mathbb{E}(\langle I,X\rangle, \langle J,Z\rangle) := \sum_{u \in \mathbb{B}(I,J)} \Gamma I(X, \Gamma u(Z)),$$

$$\langle u,a\rangle \circ \langle v,b\rangle := \langle u \circ v, c^{uv} \circ \Gamma v(a) \circ b\rangle.$$

(The canonical natural isomorphisms $c^{uv} : \Gamma v \circ \Gamma u \longrightarrow \Gamma(u \circ v)$ are given with Γ.) Split fibrations correspond to strong functors from $\underline{\mathrm{Cat}}^{\mathbb{B}^o}$.

• Cf. Grothendieck 1971, 8., Gray 1966, 1.5., Gray 1974, I,3.5, or Bénabou 1983, 1.2.•

43. Some advantages of fibrations over pseudofunctors are:

i) by keeping the cleavage implicite (i.e. as a "possible structure"), considerable complications with canonical isomorphisms are avoided;

ii) considering fibrations over a large category and/or with large fibres does not involve the metacategory <u>CAT</u>;

iii) from every pseudofunctor Γ a fibration $\int_{\mathbb{B}} \Gamma$ can *always* be obtained.

(Conversely, as we saw in proposition 31, every fibration can be cloven - and expressed as a pseudofunctor - only if the axiom of choice is assumed.)

Bénabou (1985) is a programmatic discussion on these matters.

5. Examples.

51. Let $\mathbb{P}$ and $\mathbb{M}$ be categories with one object, i.e. monoids. A morphism P: $\mathbb{P} \longrightarrow \mathbb{M}$ is a fibration if $\mathbb{P}$ decomposes on isomorphic right Ker(P)-cosets: for every $m \in \mathbb{M}$ there is a ϑ^m such that $P^{-1}(m) = \vartheta^m \mathrm{Ker}(P)$ and for all $a,b \in \mathrm{Ker}(P)$ the equality $\vartheta^m a = \vartheta^m b$ implies a=b. If $\mathbb{P}$ is a group, this is the case whenever P is an epi.

If the axiom of choice is assumed, then we can choose a cleavage for P. When $\mathbb{P}$ is a group, every element of $P^{-1}(m)$ will do as ϑ^m. In that case, thus, cleavages are just the splittings of P as a function, i.e the functions ϑ: $\mathbb{M} \longrightarrow \mathbb{P}$, such that $P\vartheta$=id. A cleavage ϑ will be normal if $\vartheta^1=1$; it will be split if $\vartheta^{mn}=\vartheta^m\vartheta^n$. The former can always be achieved (for every fibration); the latter not: e.g.

$$P: \mathbb{Z} \longrightarrow \mathbb{Z}_n : x \longmapsto x (\mathrm{mod}\ n)$$

is a fibration which cannot be split. In fact, when $\mathbb{P}$ is an abelian group, then P: $\mathbb{P} \longrightarrow \mathbb{M}$ is a split fibration iff it is the projection from direct product $\mathbb{P} \simeq \mathrm{Ker}(P) \otimes \mathbb{M}$. For groups in general, P: $\mathbb{P} \longrightarrow \mathbb{M}$ is a split fibration iff it is the projection from semidirect product $\mathbb{P} \simeq \mathrm{Ker}(P) \times_P \mathbb{M}$. - Namely, when restricted to groups, the Grothendieck construction produces the (right) semidirect product.

52. Let $\mathbb{U}$, $\mathbb{H}$ be posets. A monotone map U : $\mathbb{U} \longrightarrow \mathbb{H}$ is a fibration iff for every $k \leq i \leq j$ in $\mathbb{H}$ and every $x \leq z$ in U, such that U(x)=k and U(z)=j, the set

$$\{y \in U^{-1}(i) : x \leq y \leq z\}$$

has a supremum. It is easy to see that this supremum does not depend on k and x; it is an inverse image of z above i. Since $\leq$ is antisymmetrical, this inverse image is unique. Hence, every fibration U from a poset $\mathbb{U}$ is canonically cloven, and also normal and split.

53. Let a right action α of a monoid $\mathbb{M}$ on a set A be given, i.e. a function

$\alpha : A \times \mathbb{M} \longrightarrow A$ such that

1) $\alpha(x, 1) = x$ and

2) $\alpha(\alpha(x,m), n) = \alpha(x, mn)$

hold for every $x \in A$, $m,n \in \mathbb{M}$. If we define the hom-sets by

$$A(x,y) := \{\langle m,x,y\rangle : m \in \mathbb{M}, \alpha(y,m)=x\},$$

A becomes a category fibred over $\mathbb{M}$ by the obvious projection. All the fibres of this fibration are discrete (as categories, i.e. they are sets). By the Grothendieck construction it corresponds to a *presheaf*, an object of $\underline{\mathrm{Set}}^{\mathbb{M}^\circ}$, i.e. a variable *set* over $\mathbb{M}$.

Conversely, every presheaf over $\mathbb{M}$ can be presented as a right action of $\mathbb{M}$ in a unique way.

54. Let $\mathbb{H}$ be a complete Heyting algebra. Regarded as a category, it has at most one arrow per hom-set (i.e. $\mathbb{H}(p,q) \neq \emptyset :\Leftrightarrow p \leq q$). The arrow part of a functor $G \in \underline{\mathrm{Set}}^{\mathbb{H}^\circ}$ - an $\mathbb{H}$-presheaf - can be thought of as an operation of *restriction*, i.e.

$$G(p \leq q) : Gq \longrightarrow Gp : x \longmapsto x{\upharpoonright}q,$$

and G can be represented as an action of $\mathbb{H}$ on $A := \sum_{p \in \mathbb{H}} G(p)$ by:

$$\upharpoonright : A \times \mathbb{H} \longrightarrow A : \big\langle\langle x,p\rangle,q\big\rangle \longmapsto \langle x{\upharpoonright}p\wedge q,\ p\wedge q\rangle.$$

For every $\xi \in A$, $p,q \in \mathbb{H}$, this action satisfies:

1) $\xi{\upharpoonright}E\xi = \xi$,

2) $(\xi{\upharpoonright}p){\upharpoonright}q = \xi{\upharpoonright}(p\wedge q)$,

3) $E(\xi{\upharpoonright}p) = E\xi\wedge p$,

where

$$E : A \longrightarrow \mathbb{H} : \langle x,p\rangle \longmapsto p.$$

The Grothendieck construction now suggests a partial order on A:

$$x \leq y :\Leftrightarrow Ex \leq Ey \text{ and } x = y{\upharpoonright}Ex,$$

which extends E to a fibration. The operation $\upharpoonright$ assigns to every pair $\langle x,p\rangle$ the inverse image of x along the arrow $Ex\wedge p \leq Ex$ by this fibration. (3) says that $x{\upharpoonright}p \leq x$ is a lifting of this arrow. It is cartesian by the definition of $\leq$ in A. (2) says that these cartesian liftings are closed under composition, (1) that the identities are lifted to identities.

6. Discrete fibrations.

61. Proposition. For every functor $E:\mathcal{E} \longrightarrow \mathcal{B}$ the following statements are equivalent:

a) E is a fibration and all its fibres are discrete.

b) Every $u \in \mathcal{B}(I,J)$ has at every $Z \in |\mathcal{E}_J|$ a unique lifting. (Every arrow in $\mathcal{E}$ is thus cartesian.)

c) For every $X \in |\mathcal{E}|$ the functor

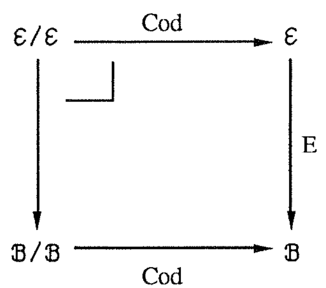
$E_X : \mathcal{E}/X \longrightarrow \mathcal{B}/EX : f \mapsto Ef$

is an isomorphism.

d) The functor

$E_{\mathcal{E}} : \mathcal{E}/\mathcal{E} \longrightarrow \mathcal{B}/E \;\; : f \mapsto \langle Ef, Cod(f) \rangle$

is an isomorphism, i.e. there is pullback



62. Definition. A fibration which satisfies (any of) the conditions from the preceding proposition is called *discrete*.

63. Given $\mathcal{B}$, any discrete fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ is uniquely determined by its object part,

$|E|:|\mathcal{E}| \longrightarrow |\mathcal{B}|$, and by

$|Dom|:|\mathcal{E}/\mathcal{E}| \longrightarrow |\mathcal{E}|$.

In view of the fact that $|\mathcal{E}/\mathcal{E}| \simeq |\mathcal{E}| \underset{|\mathcal{B}|}{\times} |\mathcal{B}/\mathcal{B}|$ (i.e. the diagram under (d) above remains a pullback if just the object parts of functors are considered), |Dom| is in fact an action

$\restriction \, : |\mathcal{E}| \underset{|\mathcal{B}|}{\times} |\mathcal{B}/\mathcal{B}| \longrightarrow |\mathcal{E}|$.

In this way the discrete fibrations generalize examples 53 and 54. (See e.g. Johnstone 1987, 2.14-15.)

64. The category <u>Dfib</u>/$\mathbb{B}$ of discrete fibrations over $\mathbb{B}$ is isomorphic with the topos $\underline{Set}^{\mathbb{B}^{\circ}}$. The category of internal categories in $\underline{Set}^{\mathbb{B}^{\circ}}$ (or in <u>Dfib</u>/$\mathbb{B}$) is isomorphic with the category of functors $\mathbb{B}^{o} \longrightarrow \underline{Cat}$. By the Grothendieck construction, this last category is isomorphic with the category Sfib/$\mathbb{B}$ of split fibrations over $\mathbb{B}$. So we have:

$$\frac{Sfib/\mathbb{B}}{Cat} = \frac{Dfib/\mathbb{B}}{Set} = \frac{Set^{\mathbb{B}^{\circ}}}{Set}$$

65. The class of discrete fibrations is not only stable under all pullbacks and closed under composition; it is closed under left division too. I.e.: For any discrete fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$, a functor $F:\mathcal{E}' \longrightarrow \mathcal{E}$ (a fortiori cartesian) is a fibration <u>iff</u> E':=EF is. (• An F-cartesian lifting of $f \in \mathcal{E}$ can be obtained as an E'-cartesian lifting of Ef.•) Fibrations E and F are discrete <u>iff</u> EF is. Hence the isomorphisms

$FIB/\mathcal{E} \cong (FIB/\mathcal{B})/E$, and

$DFIB/\mathcal{E} \cong (DFIB/\mathcal{B})/E$

for every *discrete* E.

66. The notion of a fibration is not closed under equivalence of categories! Take two groups $\mathbb{P}$ and $\mathbb{M}$ and a fibration $P: \mathbb{P} \longrightarrow \mathbb{M}$. Let $\mathbb{N}$ be a groupoid (i.e. a category where all the arrows are isomorphisms) consisting of several copies of $\mathbb{M}$, each pair connected by one isomorphism. Every inclusion $U : \mathbb{M} \hookrightarrow \mathbb{N}$ is then an equivalence of categories (i.e. full and faithful essentially surjective functor). But UP is not a fibration. Considering $U \in \underline{Cat}/\mathbb{N}(U, id_{\mathbb{N}})$, we see that Fib/$\mathbb{N}$ is not closed under the equivalences in <u>Cat</u>/$\mathbb{N}$ either, since $id_{\mathbb{N}}$ is a fibration and U is not.

However, for functors $E:\mathcal{E} \longrightarrow \mathcal{B}$ and $F:\mathcal{E}' \longrightarrow \mathcal{E}$, if F is full and faithful and each $F_I:\mathcal{E}'_I \longrightarrow \mathcal{E}_I$ is essentially surjective, then E is a fibration <u>iff</u> E'=EF is. (Here is $\mathcal{E}'_I:=(E')^{-1}(I)$ and $F_I:=F\restriction\mathcal{E}'_I$ for $I \in |\mathcal{B}|$.) Functors like F will be called *fibrewise equivalences.*

7. Lifting homotopies.

We finish this section by considering the topological origin of the notion of fibration. A "topological" characterisation of cloven fibrations is given, with natural transformations playing the role of homotopies. (This is a slight simplification of the story told in §2 of Gray (1966).)

71. Definition. (Spanier 1966, chapter 2) Let E, B, J be topological spaces, $p:E \longrightarrow B$, $c:J \longrightarrow E$, $\chi:J \times [0,1] \longrightarrow B$ continuous maps. χ is called a *homotopy*. A *lifting* of χ along p at c is a homotopy $\vartheta:J \times [0,1] \longrightarrow E$, such that

$p \circ \vartheta \quad = \chi$, and

$\vartheta(x, 1) = c(x)$, for all $x \in J$.

p is a *Hurewicz fibration* if every homotopy χ has a lifting along p at every c.

72. Every topological space X gives rise to a category πX in a natural way: the objects of πX are the points of X, while the arrows are the homotopy classes of paths, i.e. continuous functions $g:[0,1] \longrightarrow X$ taken modulo equivalence relation $\simeq$, defined

$$g_0 \simeq g_1 :\Leftrightarrow \exists \gamma:[0,1] \times [0,1] \longrightarrow X.\ \gamma(\_,0) = g_0 \wedge \gamma(\_,1) = g_1.$$

Clearly, all these arrows are isos. πX is the *fundamental groupoid* X. By the same idea, the whole category Esp of topological spaces and continuous maps is groupoid-enriched. The arrows from c_0 to c_1 in Esp(X,Y) are the homotopies $\varphi: X \times [0,1] \longrightarrow Y$, $\varphi(\_,0) = c_0$, $\varphi(\_,1) = c_1$, taken modulo $\simeq$ again:

$$\varphi \simeq \psi :\Leftrightarrow \forall x \in X.\varphi(x) \simeq \psi(x).$$

Noticing that $\pi X = \text{Esp}(1,X)$, we define a two-functor

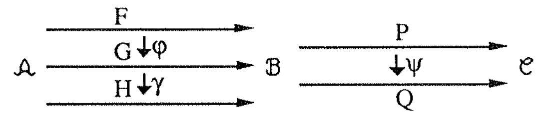
$$\pi := \text{Esp}(1,\_): \text{Esp} \longrightarrow \text{Cat}.$$

On fundamental groupoids the topological and categorical notions of fibration tend to coincide:

73. Proposition. If $p:E \longrightarrow B$ is a Hurewicz fibration, then πp is a cloven fibration.

• This is a special case of proposition 77, in view of the fact that the functor π makes every homotopy into a natural transformation.•

74. Notation. In a two-category, we reserve the symbol $\circ$ for the composition within a hom-category, while the "horizontal" composition is denoted either by $*$ or by juxtaposition. In the two-category of categories, we have thus



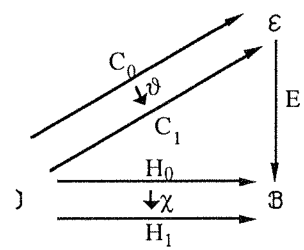
$$(\gamma \circ \varphi)_A := \gamma_A \circ \varphi_A : FA \longrightarrow GA \longrightarrow HA$$

$$(\psi\varphi)_A := (\psi * \varphi)_A := \psi_{GA} \circ P(\varphi_A) : PFA \longrightarrow PGA \longrightarrow QGA$$

$$= Q(\varphi_A) \circ \psi_{FA} : PFA \longrightarrow QFA \longrightarrow QGA.$$

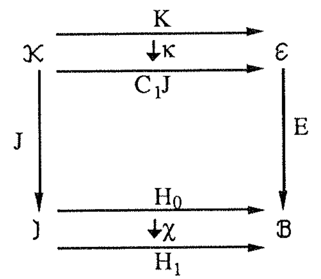
The one-cells (functors) can be identified with the identity two-cells (natural transformations); we shall rather write PF than P*F. (A standard reference for two-categories is Kelly-Street 1974.)

75. Definition. Consider the diagram

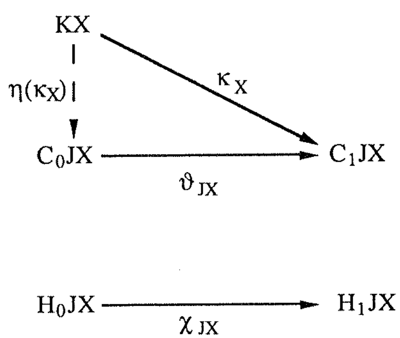


A *lifting* along a functor E of a natural transformation χ at a functor C_1 such that $EC_1 = H_1$ consists of a functor C_0 and a natural transformation ϑ, such that $\chi = E * \vartheta$ (which, of course, implies $EC_0 = H_0$).

A lifting ϑ is called *cartesian* if for every κ



such that $E*\kappa = \chi * J$, there is an arrow $\eta : \kappa \longrightarrow \vartheta * J$, consisting of unique vertical arrows.



(Officially, η is an arrow between natural transformations: a modification. In fact, it can be seen as an honest natural transformation $\eta : «\kappa» \longrightarrow «\vartheta * J»$ where the functors $«\kappa», «\vartheta * J»: \mathcal{K} \longrightarrow \mathcal{E}/\mathcal{E}$ are obtained by)

76. The couniversal property of comma categories. To every comma category F/G belongs a pair of projections P_0 and P_1 and a natural transformation α:

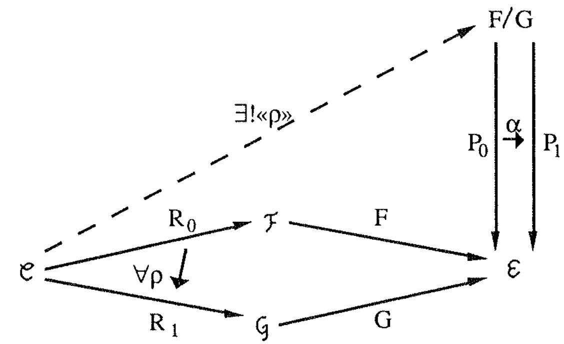
$P_0\langle X, u{:}FX \longrightarrow GY, Y\rangle := FX,$

$P_1\langle X, u{:}FX \longrightarrow GY, Y\rangle := GY$

$\alpha_{\langle X, u:FX \longrightarrow GY, Y\rangle} := u.$

For every ($\mathcal{C}$, R_0, R_1 and) ρ as below, there is a unique functor $«\rho»$, such that

$\rho = \alpha * «\rho»$.

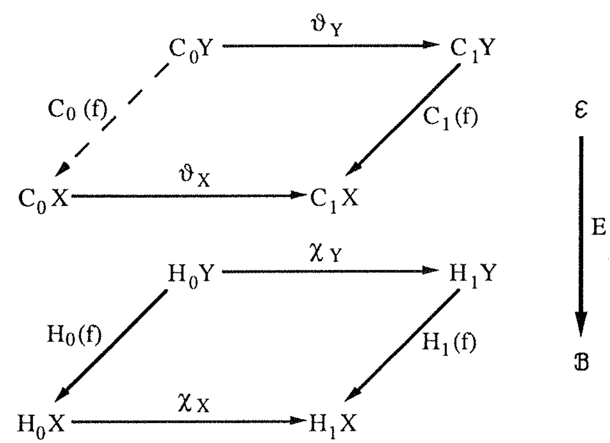


$«\rho»(X) := \langle R_0X, \rho_X, R_1X\rangle,$

$«\rho»(f) \;\; := \langle R_0f, R_1f\rangle.$

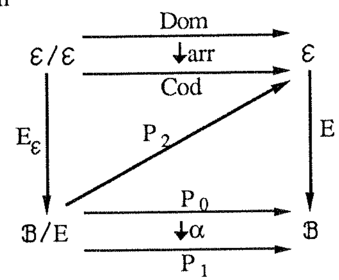
77. Proposition. E is a cloven fibration <u>iff</u> every χ has a cartesian lifting at every C_1 such that $H_1 = EC_1$.

Proof. <u>Then:</u> The components of the lifting ϑ of χ are the cartesian liftings of the components of χ.



Clearly, the unique factorisations constitute η, which is required for ϑ to be a cartesian lifting of χ.

<u>If</u>: Consider the diagram



where $P_2\langle u{:}I \longrightarrow EX, X\rangle := X$. (The other arrows are as previously defined. The natural transformation $arr_f = f$ is the α belonging with $\mathcal{E}/\mathcal{E}$ as the comma category of two identity functors. Dom and Cod are of course P_0 and P_1 of this comma.) Clearly, $EP_2 = P_1$ and $P_2E_{\mathcal{E}} = \text{Cod}$.

Every lifting ϑ of α at P_2 induces a functor «ϑ» : $\mathcal{B}/E \longrightarrow \mathcal{E}/\mathcal{E}$. Since

$$\alpha * E_{\mathcal{E}} * «\vartheta» = E * arr * «\vartheta» = E * \vartheta = \alpha,$$

by (the uniqueness part of) the couniversal property of α, $E_{\mathcal{E}} * «\vartheta» = id$ holds.

If the lifting ϑ is also cartesian, there is η: «arr» $\longrightarrow$ «ϑ»$*E_{\mathcal{E}}$. Note that «arr» $= id_{\mathcal{E}/\mathcal{E}}$. Now

$E_{\mathcal{E}} * \eta = id$, since the components of η are vertical; while

$\eta * «\vartheta» = id$ follows from the uniqueness of $\eta_{«\vartheta»(u,X)}$ as the arrow

$$\vartheta_{\langle u,X\rangle} = «arr» * «\vartheta»(u,X) \longrightarrow «\vartheta» * E_{\mathcal{E}} * «\vartheta»(u,X) = \vartheta_{\langle u,X\rangle}.$$

These two equalities mean that η is the unit of the adjunction $E_{\mathcal{E}} \dashv$ «ϑ», while the counit is identity. Hence $E_{\mathcal{E}}$ has a right adjoint right inverse functor, i.e. E is a fibration by proposition 31.•

78. Corollary. Given a cloven $E : \mathcal{E} \longrightarrow \mathcal{B}$, functors $H : \mathcal{C} \longrightarrow \mathcal{B}$ and $S : \mathcal{C} \longrightarrow \mathcal{E}$ and a natural transformation $\chi : H \longrightarrow ES$, there is a functor $T : \mathcal{C} \longrightarrow \mathcal{E}$, such that $H = ET$ (and a natural transformation $\vartheta : T \longrightarrow S$, such that $\chi = E * \vartheta$).

For instance, if a cloven fibration E has a left adjoint S, by lifting the unit of this adjunction we get a left inverse T of E. If the left adjoint S is full and faithful, then T is a left adjoint too.

79. The last proposition can be used to abstractly define cloven fibrations in an arbitrary bicategory with the comma construction.

To avoid cleavage means to lift only paths (single arrows), and not homotopies (natural families of arrows); i.e., in an abstract bicategory, to lift only the two-cells $1 \overset{\longrightarrow}{\underset{\longrightarrow}{\chi\Downarrow}} B$.

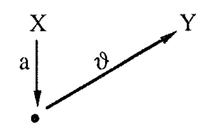
2. Vertical structure

1. Basic notions, fibrewise.

Fibrewise structure/property is one which all the fibres possess, and all the inverse images preserve.

Fibrewise versions of some standard categorical notions can be obtained in a completely straightforward way. The only question in these cases remains to relate the so obtained fibrewise notions with the old global ones, in the style: "Fibred category $E:\mathcal{E}\longrightarrow\mathcal{B}$ has fibrewise property P <u>iff</u> categories $\mathcal{E}$ and $\mathcal{B}$, and functor E satisfy condition Q". On the other hand, some constructions from ordinary category theory are not easily lifted in fibred categories. We begin by such an example.

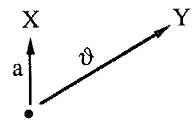
The functor op. (Bénabou 1983, 4.4.) The functor $(\_)^o$: <u>CAT</u> $\longrightarrow$ <u>CAT</u> formally changes the direction of all the arrows in a category. The corresponding fibred construction $(\_)^{op}$: <u>FIB</u>/$\mathcal{B}\longrightarrow$ <u>FIB</u>/$\mathcal{B}$[3] should change the direction of all the *vertical* arrows in E: $\mathcal{E}\longrightarrow\mathcal{B}$ and leave the other arrows somehow unchanged. While this is quite hopeless with functors in general, a fibration E induces for every $f\in\mathcal{E}$ a vertical-cartesian decomposition $f=\vartheta a$, unique up to a *vertical* isomorphism. Well, since the arrows in $\mathcal{E}(X,Y)$ can thus be regarded as the equivalence classes (modulo vertical isos) of the diagrams in the form



one is tempted to try to take the arrows in $\mathcal{E}^{op}(X,Y)$ to be the equivalence classes of the diagrams

---

[3]We must denote it differently, because $(\_)^o$ also acts on every small fibration as an arrow in <u>Cat</u>.



by the equivalence relation

$\langle a_1,\vartheta_1\rangle \sim \langle a_2,\vartheta_2\rangle :\Leftrightarrow \exists$ vertical (iso) b. $a_1=a_2\circ b \wedge \vartheta_1=\vartheta_2\circ b$.

And this works! If we use the fraction notation:

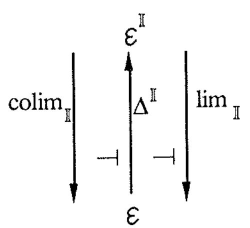
$\vartheta_0/a_0 := \{\langle a,\vartheta\rangle \mid \langle a,\vartheta\rangle \sim \langle a_0,\vartheta_0\rangle\}$,

the composition can be defined by:

$\vartheta^u/a \circ \vartheta^v/b := \vartheta^u\circ\vartheta^v \,/\, b\circ v^*(a)$.

If E is cloven, a canonical representant of each of these classes is given. Clearly, we get a cloven E^{op}.

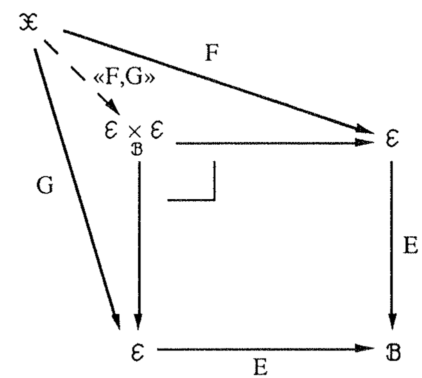
Limits and colimits. Given a small category $\mathbb{I}$, the diagrams of type $\mathbb{I}$ in a category $\mathcal{E}$ are the objects of the functor category $\mathcal{E}^{\mathbb{I}}$. There is an obvious faithful functor $\Delta^{\mathbb{I}} : \mathcal{E} \longrightarrow \mathcal{E}^{\mathbb{I}}$ which sends every object of $\mathcal{E}$ to the corresponding constant diagram. Some representants of the limits and colimits of the diagrams of type $\mathbb{I}$ are given by a right resp. left adjoint functor of $\Delta^{\mathbb{I}}$:



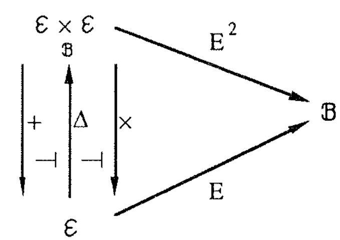
This conception of limits and colimits can directly be generalized from the constant categories to the variable ones - from Cat to Fib/$\mathbb{B}$ - by simply putting the fibred adjointness in place of the constant one, i.e. by requiring the functors and natural transformations to be cartesian over $\mathbb{B}$. Of course, the question: What is a small category $\mathbb{I}$ in Fib/$\mathbb{B}$? - must first be answered. Leaving it aside for a while, let us look at the simplest special cases.

Remark. The axiom of choice is built in into the concept of limits and colimits as functors. This can be avoided (by using universal constructions), but this leads in fibred categories to much longer definitions.

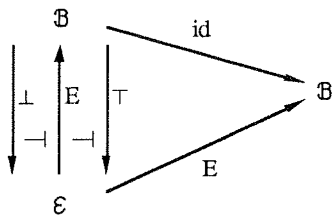
Products and coproducts. Define the fibration $E^2 := \mathrm{pb}(E,E)$: $\mathcal{E} \underset{\mathcal{B}}{\times} \mathcal{E} \longrightarrow \mathcal{B}$, and let the cartesian diagonal functor be the factorisation $\Delta := «id_{\mathcal{E}},id_{\mathcal{E}}» : \mathcal{E} \longrightarrow \mathcal{E} \underset{\mathcal{B}}{\times} \mathcal{E}$.



The binary fibrewise product $\times$ and the coproduct $+$ are then respectively the right and the left cartesian adjoint of Δ.



Terminal and initial objects. The fibration id:$\mathcal{B} \longrightarrow \mathcal{B}$ is terminal in FIB/$\mathcal{B}$. The cartesian right adjoint $\top$ of $E \in$ FIB/$\mathcal{B}$(E, id) chooses a terminal object in each E-fibre; the left adjoint $\perp$ chooses fibrewise initial objects.

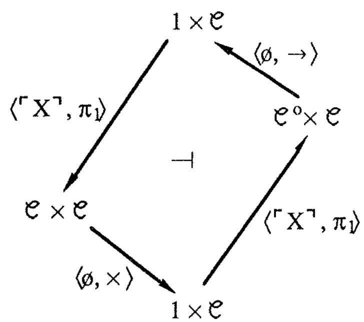


Since all the triangles must commute, we have $E\top = E\bot = \mathrm{id}$; i.e. $\top$ and $\bot$ must be *cartesian sections* (i.e. right inverses) of E. However, as soon as a full and faithful cartesian right (resp. left) adjoint $S:\mathfrak{B} \longrightarrow \mathcal{E}$ of E is given, so that $ES \simeq \mathrm{id}$, a cartesian section $\top$ (resp. $\bot$) can be obtained by corollary 1.78.

Exponents. In a category $\mathcal{C}$ the exponents by $X \in |\mathcal{C}|$ are given by the functor $X \to \_ : \mathcal{C} \longrightarrow \mathcal{C}$ which is right adjoint to $X \times \_ : \mathcal{C} \longrightarrow \mathcal{C}$. In other words, there is a functor

$$\to : \mathcal{C}^o \times \mathcal{C} \longrightarrow \mathcal{C}$$

and an adjunction



for every $X \in |\mathcal{C}|$, i.e. for every constant functor $\ulcorner X \urcorner : 1 \longrightarrow \mathcal{C}$ - which is also $\ulcorner X \urcorner : 1 \longrightarrow \mathcal{C}^o$.

For a fibred category $\mathcal{E}$, the objects $X \in |\mathcal{E}_I|$ can be represented by "constant functors"

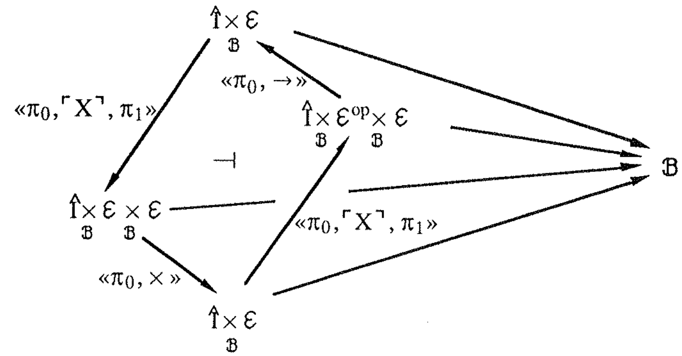
$$\ulcorner X \urcorner := \mathrm{Dom}{*}\Theta_X : \mathfrak{B}/I \longrightarrow \mathcal{E}/X \longrightarrow \mathcal{E},$$

with Θ_X as in 1.31(e). $\ulcorner X \urcorner$ is a cartesian functor $\nabla I \longrightarrow E$; and every cartesian functor $\nabla I \longrightarrow E$ can be obtained in the form $\ulcorner X \urcorner$ for some $X \in |\mathcal{E}_I|$. (This correspondence is in

fact the Yoneda lemma, III.1.1.) The fibrewise exponents in a fibration $E:\mathcal{E} \longrightarrow \mathfrak{B}$ can now be described as given by a cartesian functor

$$\to : \mathcal{E}^{op} \underset{\mathfrak{B}}{\times} \mathcal{E} \longrightarrow \mathcal{E}$$

(where $\mathcal{E}^{op} \underset{\mathfrak{B}}{\times} \mathcal{E} := \mathrm{pb}(E^{op}, E)$) with an adjunction



for every "constant functor" $\ulcorner X \urcorner : \hat{I} \longrightarrow \mathcal{E}$, where we write $\hat{I}$ for $\mathfrak{B}/I$. It is not hard to see that $\ulcorner X \urcorner$ is also $\ulcorner X \urcorner : \hat{I} \longrightarrow \mathcal{E}^{op}$ (• since $\ulcorner X \urcorner$ is cartesian, all the arrows of $\hat{I}$ are cartesian, and the cartesian arrows of $\mathcal{E}$ and of $\mathcal{E}^{op}$ coincide•).

fccc. Putting the finite products and the exponents together, we have the notion of the *fibrewise cartesian closed structure*. This is what we shall really need in chapter IV.

2. Fibrewise vs. global limits.

Now we shall inquire into the connections between fibrewise limits in $E:\mathcal{E} \longrightarrow \mathfrak{B}$ and (ordinary) global limits in $\mathcal{E}$. We first consider the terminal objects, then the pullbacks, and finally limits in general. (These propositions will be used many times later, and a corollary reducing the fibrewise products to some ordinary pullbacks will be useful in semantics, which is our main goal.)

Fact. $E:\mathcal{E} \longrightarrow \mathcal{B}$ is a cloven fibration <u>iff</u> every functor $E_X:\mathcal{E}/X \longrightarrow \mathcal{B}/EX$ (from proposition 1.3(e)) is a cloven fibration, $X \in |\mathcal{E}|$. When this is the case, all E_X have fibrewise terminal objects.

• <u>Then</u> follows from the isomorphism $(\mathcal{E}/X)/y \cong \mathcal{E}/Y$ for all $y:Y \longrightarrow X$. <u>If</u> follows from the fact that every E_X has a fibrewise terminal object whenever it is a cloven fibration.•

Propositions.

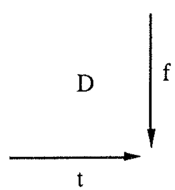
21. E has fibrewise terminal objects and $\mathcal{B}$ has a global one <u>iff</u> $\mathcal{E}$ has a global terminal object and E preserves it.

• <u>Then</u>: $\top_{\mathcal{E}} := \top(\top_{\mathcal{B}})$ is a terminal object of $\mathcal{E}$, where $\top_{\mathcal{B}}$ is a terminal object of $\mathcal{B}$, and $\top:\mathcal{B} \longrightarrow \mathcal{E}$ a fibrewise terminal object of E. <u>If</u>: $\top I := \emptyset_I{}^*(\top_{\mathcal{E}})$, where $\emptyset_I \in \mathcal{B}(I, \top_{\mathcal{B}})$. (If we want $\top$ to be a functor, we need the axiom of choice here, to choose one inverse image. But *any* inverse image $\emptyset_I{}^*(\top_{\mathcal{E}})$ is terminal in $\mathcal{E}_I$.)•

22. Let a commutative square Q in a fibred category $\mathcal{E}$ have two parallel sides cartesian. Then Q is a pullback <u>iff</u> EQ is.

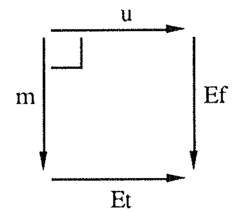
• Chase diagrams.•

23. In a fibred category $\mathcal{E}$, the diagram D:

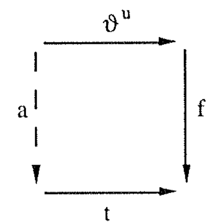


with t cartesian, has a limit which E preserves <u>iff</u> ED has a limit in $\mathcal{B}$. When this is the case, D is completed to a pullback square in which the arrow parallel to t is also cartesian.

• <u>If</u> the square



is a pullback in $\mathcal{B}$, take



in $\mathcal{E}$ (where ϑ^u is a cartesian lifting and a is a unique factorisation over m of $f \circ \vartheta^u$ through t). By the preceding proposition, this must be a pullback square.•

24. Let D be a connected diagram in a fibre of $\mathcal{E}$. Every fibrewise limit of D is also its global limit. (A diagram is connected if it is connected as a nonoriented graph.)

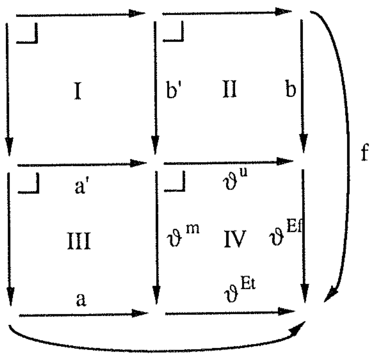
• Since D is a connected diagram of vertical arrows, all the components of an arbitrary cone $\alpha : A \longrightarrow D$ in $\mathcal{E}$ must lie over the same arrow, say, $u \in \mathcal{B}$. α must factorize uniquely through a cone of vertical arrows $\alpha' : A \longrightarrow u^*D$. But the fibrewise limits are by definition[4] preserved by the inverse images, so that α' must factorize uniquely through $u^*(\lambda)$: $u^*L \longrightarrow u^*D$, if $\lambda : L \longrightarrow D$ is a limit cone in the fibre of D. Therefore α factorizes uniquely through λ.•

25. A fibration E has fibrewise pullbacks and $\mathcal{B}$ has global pullbacks of the arrows in the image of E <u>iff</u> $\mathcal{E}$ has global pullbacks and E preserves them.

---

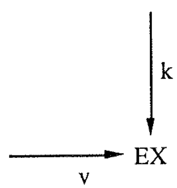
[4]in the first sentence of this section

• Then: Let $t = \vartheta^{Et} \circ a$ and $f = \vartheta^{Ef} \circ b$ be some vertical-cartesian factorisations.



Since $\mathcal{B}$ has pullbacks, $\mathcal{E}$ has pullbacks II, III and IV by proposition 23. (If ϑ^u is cartesian because of this proposition, ϑ^m is cartesian because the class of the cartesian arrows is closed under the composition and left division). Clearly, a' and b' are vertical because a and b are. By proposition 24, the pullback I of a' and b' in their fibre is their pullback in $\mathcal{E}$.

If: Given



in $\mathcal{B}$, make in $\mathcal{E}$ a pullback of ϑ_X^v and ϑ_X^k. Its E-image is a pullback because E preserves them. This is, furthermore, the reason why every (global) pullback of vertical arrows in $\mathcal{E}$ must remain within its fibre - i.e. why the global pullbacks of vertical arrows in $\mathcal{E}$ are fibrewise.•

26. A fibration E has fibrewise limits and $\mathcal{B}$ has global ones iff $\mathcal{E}$ has and E preserves the global ones.

• Just put propositions 21 and 25 together, using the fact that all limits can be constructed from (possibly infinite) pullbacks and terminal objects. (Note: When $\mathcal{E}$ has a terminal object which is preserved by functor E, then E must be surjective.)•

Corollary. A fibration E has fibrewise products and $\mathcal{B}$ has global ones iff $\mathcal{E}$ has and E preserves global products, and pullbacks of the vertical arrows to the terminal object in each fibre. (The products in this statement include the empty ones, i.e. the terminal objects.)

Remark. By proposition 25, *a fibrewise product*, i.e. a pullback over the terminal object in a fibre, *is a global pullback*: for $X, Y \in |\mathcal{E}_I|$, $X \times_I Y \simeq pb(\eta_X, \eta_Y)$ in $\mathcal{E}$, where $\eta : id \longrightarrow \top E$ is the unit of the adjunction $E \dashv \top$.

3. Fibrewise fibrations

The formula

fibred structure := structure in fibres + preserved by inverse images

can also be applied on a *potential* structure, such as the notion of fibration itself (cf. 1.33). And again, the fibrewise notion can be characterized in global terms.

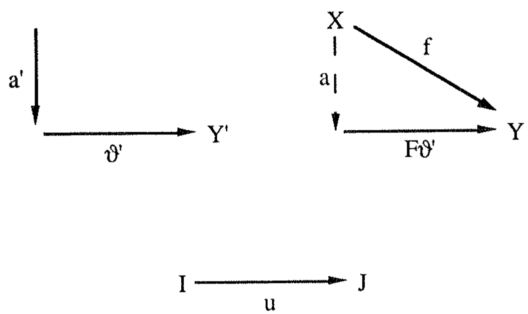
Definition. Let $E : \mathcal{E} \longrightarrow \mathcal{B}$ and $E' : \mathcal{E}' \longrightarrow \mathcal{B}$ be fibrations, $F : \mathcal{E}' \longrightarrow \mathcal{E}$ a cartesian functor and $F_I : \mathcal{E}'_I \longrightarrow \mathcal{E}_I$, $I \in |\mathcal{B}|$, the restrictions of F. We say that F is a *fibrewise fibration over* E if all the F_I are fibrations and if E'-inverse images preserve the F_I-cartesian arrows. (I.e. for every $u \in \mathcal{B}(I,J)$ and every F_J-cartesian arrow f, every inverse image $u^*(f)$ must be F_I-cartesian.) Given fibrewise fibrations $F \in \underline{FIB}/\mathcal{B}(E',E)$ and $G \in \underline{FIB}/\mathcal{B}(E'',E)$, a *fibrewise cartesian functor* $H : F \longrightarrow G$ is $H \in \underline{FIB}/\mathcal{B}(E',E'')$, such that F=GH, and every $H_I \in \underline{FIB}/\mathcal{E}_I(F_I,G_I)$. Denote by $\underline{FIB}_{\mathcal{B}}/E$ the category of fibrewise fibrations and fibrewise cartesian functors over E.

Proposition. $\underline{FIB}_{\mathcal{B}}/E = \underline{FIB}/\mathcal{E}$.

31. We first show that

$F\in \underline{FIB}/\mathcal{B}(E',E)$ is a fibrewise fibration iff $F\in |\underline{FIB}/\mathcal{E}|$.

• Then: Take $f\in \mathcal{E}(X,Y)$, $Ef = u$, and $Y'\in|\mathcal{E}'|$ such that $FY' = Y$.



If ϑ' is an E'-cartesian lifting of u, then $F\vartheta'$ is an E-cartesian lifting of u. Let $a\in|\mathcal{E}_I|$ be the unique factorisation of f through $F\vartheta'$, and a' its F_I-cartesian lifting. $\vartheta'\circ a'$ is then an F-cartesian lifting of f at Y'. Moreover, every F-cartesian arrow must be in this form: Factorize it through an E'-cartesian and a vertical arrow; the vertical one must be F_I-cartesian.

To show that the composition of two arrows in the form $\vartheta'\circ a'$ is still in this form use the fact that the E'-inverse images preserve F_J-cartesian arrows.

If: Every F_I is a fibration because it is obtained by pulling back F along $\mathcal{E}_I \hookrightarrow \mathcal{E}$. Every E'-cartesian arrow ϑ^u_X as well as every F_J-cartesian $b\in\mathcal{E}'_J(X,Y)$ is also F-cartesian. The arrow $\vartheta^u_Y\circ u^*(b) = b\circ\vartheta^u_X$ is therefore F-cartesian; thus $u^*(b)$ is. But then it must be F_I-cartesian.•

32. $H\in\underline{FIB}/E(F,G)$ iff $H\in\underline{FIB}/\mathcal{E}(F,G)$.

• Note that all the E'-cartesian arrows, as well as the F_I-cartesian arrows are F-cartesian too; and that, conversely, every F-cartesian arrow decomposes on an F_I-cartesian and an E'-cartesian.•

4. A lemma about adjointness.

It is easy to see that functors $F\in\underline{FIB}/\mathcal{B}(E',E)$ and $G\in\underline{FIB}/\mathcal{B}(E,E')$ are cartesian adjoint (cf.1.24) iff their fibrewise parts $F_I: \mathcal{E}'_I \to \mathcal{E}_I$ and $G: \mathcal{E}_I\to\mathcal{E}'_I$ are adjoint for all $I\in|\mathcal{B}|$. But we shall need a slightly stronger statement.

Lemma. Let $E:\mathcal{E}\to\mathcal{B}$, $E':\mathcal{E}'\to\mathcal{B}$ be fibrations, and $F:\mathcal{E}'\to\mathcal{E}$, $G:\mathcal{E}\to\mathcal{E}'$ functors such that $E' = EF$, $E = E'G$. Then

$F\dashv G$ and η, ε cartesian $\Leftrightarrow \forall I\in|\mathcal{B}|.\ F_I\dashv G_I$ and G cartesian.

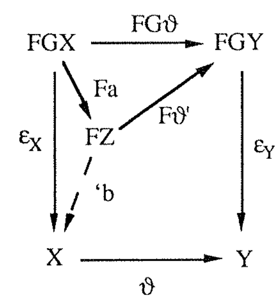
• $\Leftarrow$: $\mathcal{E}(FX, Y) = \bigsqcup_{u\in\mathcal{B}(E'X,EY)} \mathcal{E}_u(FX,Y) \simeq \bigsqcup_{u\in\mathcal{B}(E'X,EY)} \mathcal{E}'_u(X,GY) =$
$= \mathcal{E}'(X, GY)$,

since

$$\mathcal{E}_u(FX,Y) \simeq \mathcal{E}_I(FX, u^*Y) \simeq \mathcal{E}'_I(X, Gu^*Y) \simeq \mathcal{E}'_I(X, u^*GY) \simeq \mathcal{E}'_u(X,GY).$$

$\Rightarrow$: The nontrivial part is that G is a cartesian functor.

Take a cartesian arrow ϑ in $\mathcal{E}$ and consider the vertical-cartesian decomposition $G\vartheta = \vartheta'\circ a$. Back in $\mathcal{E}$, there is



where 'b is the unique vertical arrow by which $\varepsilon_Y\circ F\vartheta'$ factorizes through ϑ. Let b be the right transpose of 'b. Since $G\vartheta\circ b$ is the right transpose of $\vartheta\circ$'b, while ϑ' is the right transpose of $\varepsilon_Y \circ F\vartheta'$, from $\vartheta \circ {}^{\backprime}b = \varepsilon_Y \circ F\vartheta'$ follows $G\vartheta \circ b = \vartheta'$. But from $\vartheta' \circ a \circ b = G\vartheta \circ b = \vartheta'$ follows $a \circ b = id$.

On the diagram we see that $\vartheta \circ {}^{\backprime}b \circ Fa = \vartheta \circ \varepsilon_X$. Since ε_X, ${}^{\backprime}b$, Fa are vertical, ${}^{\backprime}b \circ Fa = \varepsilon_X$ must hold. Transposing both sides of this equality, we obtain $b \circ a = id$.

So a is an iso and $G\vartheta$ is cartesian.•

Corollary. Consider a functor $F \in \underline{FIB}/\mathcal{B}(E',E)$. A functor $G \in \underline{CAT}/\mathcal{B}(E,E')$ is its cartesian adjoint <u>iff</u> it is cartesian and adjoint to it fibrewise.

3. Horizontal structure

1. Hyperdoctrines.

The conceptual basis for the categorical interpretation of higher order logic is the notion of *hyperdoctrine*, introduced by Lawvere (1969, 1970). It is basically a (pseudo)functor

$$\wp : \mathcal{S}^o \longrightarrow \underline{Cat}\ ,$$

where

- the category $\mathcal{S}$ is cartesian closed, as well as every fibre $\wp K$, $K \in |\mathcal{S}|$;
- for every $u \in \mathcal{S}(J,K)$, there are functors $u_! \dashv \wp u \dashv u_* : \wp J \longrightarrow \wp K$.

We shall be concerned with this structure, as translated (by the Grothendieck construction) into fibred categories. The horizontal structure, which is dealt with in this section, corresponds to the functors $u_! \dashv \wp u \dashv u_*$.

Logical motivation. $\mathcal{S}$ is meant to be a "category of sets and functions". The objects of $\wp K$ represent predicates $\varphi(y^K)$ over a set K (i.e. families $\{\varphi(y) \mid y \in K\}$ of truth values). An arrow $f \in \wp K\big(\varphi(y^K), \psi(y^K)\big)$ can be understood as a proof

$$\varphi(y^K) \overset{f}{\vdash} \psi(y^K).$$

The functor $\wp u$, usually written u^*, represents substitution along the function $u \in \mathcal{S}(J,K)$, i.e.

$$u^* : \wp K \longrightarrow \wp J : \varphi(y^K) \longmapsto \varphi\big(u(x^J)\big).$$

Note that substitution along a projection $\pi \in \mathcal{B}(K \times L, K)$ means adding a dummy variable. Let us write $\varphi(y^K, \not{z}^L)$ for $\pi^*\big(\varphi(y^K)\big)$.

By adding dummies, it can be achieved that all the elements of a given finite set of predicates have the same set of variables. This is tacitly supposed when the operations of propositional logic are performed on predicates: a predicate $\gamma(x) \wedge \varphi(y)$, depending on both x and y, is in fact $(\gamma \wedge \varphi)(x, y) = \gamma(x, \not{y}) \wedge \varphi(\not{x}, y)$. In this sense, *propositional logic is done fibrewise*: one first brings all the predicates in one common fibre by taking their inverse images along projections. The *fibrewise product* and *coproduct* thus correspond to the *conjunction* and *disjunction* respectively, the *fibrewise exponent* to the *implication*.

The crucial observation made by Lawvere is that *the quantifiers are adjoint to the substitution*. Namely, they are characterized (or defined) by

$$\varphi(y) \vdash \forall z.\psi(z, y) \quad \Leftrightarrow \quad \varphi(\not{z}, y) \vdash \psi(z, y),$$
$$\exists z.\psi(z, y) \vdash \varphi(y) \quad \Leftrightarrow \quad \psi(z, y) \vdash \varphi(\not{z}, y).$$

Moreover, in logic with equality $\equiv$ holds

$$\varphi(y) \vdash \forall x.u(x) \equiv y \rightarrow \gamma(x) \quad \Leftrightarrow \quad \varphi(u(x)) \vdash \gamma(x),$$
$$\exists x.u(x) \equiv y \wedge \gamma(x) \vdash \varphi(y) \quad \Leftrightarrow \quad \gamma(x) \vdash \varphi(u(x)).$$

The logical picture of adjoints is thus:

$$u_!(\gamma(x)) = \exists x.u(x) \equiv y \wedge \gamma(x), \text{ and}$$
$$u_*(\gamma(x)) = \forall x.u(x) \equiv y \rightarrow \gamma(x),$$

i.e. they just slightly generalize the quantifiers.

In short, the logical meanings are:

$\times$	is	$\wedge$,
$+$	is	$\vee$,
π_*	is	$\forall$, and
$\pi_!$	is	$\exists$.

2. Co-, bi-, trifibrations.

Definitions. We say that a functor $E:\mathcal{E} \to \mathcal{B}$ is a *cofibration* if $E^o:\mathcal{E}^o \to \mathcal{B}^o$ is a fibration. A functor $F:E' \to E$ is *cocartesian* if $F^o:E'^o \to E^o$ is cartesian. $\underline{COFIB}/\mathcal{B}$ is the category of cofibrations over $\mathcal{B}$, with cocartesian functors.

E is a *bifibration* if E and E^o are fibrations - i.e. if E is a fibration and cofibration. $\underline{BIFIB}/\mathcal{B}$ denotes the category of bifibrations over $\mathcal{B}$, with cartesian and cocartesian functors.

E is a *trifibration* if E, E^o, and $(E^{op})^o$ are fibrations - i.e. if E and E^{op} are bifibrations. A functor $F \in \underline{FIB}/\mathcal{B}(E',E)$ is *opcartesian* if its op-image $op(F) \in \underline{FIB}/\mathcal{B}(E'^{op},E^{op})$ is cocartesian. $\underline{TRIFIB}/\mathcal{B}$ is the category of trifibrations over $\mathcal{B}$ with cartesian, cocartesian and opcartesian functors.

Remark. By the Grothendieck construction, the cofibrations correspond to the *co*variant (pseudo)functors to $\underline{Cat}$, i.e. there is

$$\int_{\mathbb{B}} : \underline{Cat}^{\mathbb{B}} \to \underline{Cofib/\mathbb{B}}.$$

Terminology, notation. Let $f^o \in \mathcal{E}^o(Y,X)$ denote the arrow $f \in \mathcal{E}(X,Y)$ as seen in the opposite category.

We say that $\sigma \in \mathcal{E}(Y,Z)$ is E-*cocartesian* if σ^o is E^o-cartesian. σ is thus an initial lifting of $u=E\sigma$ at Y. $u_!Y=Z$ is a *(left) direct image* of Y along u.

The E^{op}-cocartesian arrows are called E-*opcartesian.* If $\psi \in \mathcal{E}^{op}(Y,W)$ is an E-opcartesian lifting of u, then $u_*Y=W$ is a *right direct image* of Y along u.

We say that a co-/bi-/trifibration is cloven if all the fibrations involved in it are.

If $(u^o)^*:\mathcal{E}^o{}_I \to \mathcal{E}^o{}_J$ is an E^o-inverse image functor for $u \in \mathcal{B}(I,J)$, the *(left) direct image* functor is

$$u_! := \left((u^o)^*\right)^o : \mathcal{E}_I \to \mathcal{E}_J.$$

This functor is related to the inverse image functor by the adjointness:

$$u_! \dashv u^*$$

because $\mathcal{E}_J(u_!X, Z) \simeq \mathcal{E}_u(X, Z) \simeq \mathcal{E}_I(X, u^*Z)$ holds naturally in X and Z.

E^{op} is, of course, a cloven fibration whenever E is. Clearly, its inverse image functors $(u^{op})^*$ will be

$$(u^{op})^* : (\mathcal{E}^{op})_J \to (\mathcal{E}^{op})_I \cong (u^*)^o : (\mathcal{E}_J)^o \to (\mathcal{E}_I)^o.$$

Given the left direct image functors $(u^{op})_!:(\mathcal{E}_I)^o \to (\mathcal{E}_J)^o$ of E^{op}, we define the *right direct image* functors of E

$$u_* := \left((u^{op})_!\right)^o : \mathcal{E}_I \to \mathcal{E}_J.$$

The adjunction is now:

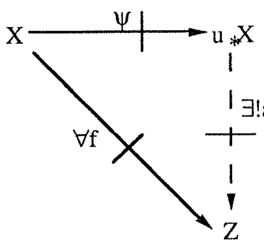
$u^* \dashv u_*$.

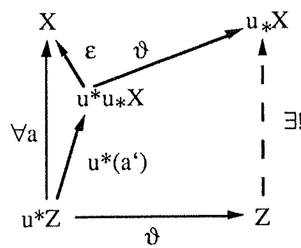
A *cocleavage* of E is a triple $\langle(\_)_!,\sigma,c'\rangle$ where $(\_)_!=(\_)^{o*o}$, $\sigma=\vartheta^o$ and $c'=c^o$ for some cleavage $\langle(\_)^*,\vartheta,c\rangle$ of E^o. An *opcleavage* of E is a triple $\langle(\_)_*,\psi,c\rangle$, $(\_)_*=(\_)_!^{\,o}$, $\psi=\sigma$ for some cocleavage $\langle(\_)_!,\sigma,c\rangle$ of E^{op}.

In brief, we systematically replace the prefix "co-" by "op-" when considering a cofibration E^{op}. We sometimes say that E is a *right bifibration* when E^{op} is a bifibration.

Opcartesian arrows. Translated from $\mathcal{E}^{op}$ to $\mathcal{E}$, the statement that $\psi=\vartheta/\varepsilon$ is E-opcartesian reads:

$$\forall a\in\mathcal{E}_I(u^*Z,X)\ \exists! a'\in\mathcal{E}_J(Z,u_*X).\ a=\varepsilon\circ u^*(a')$$





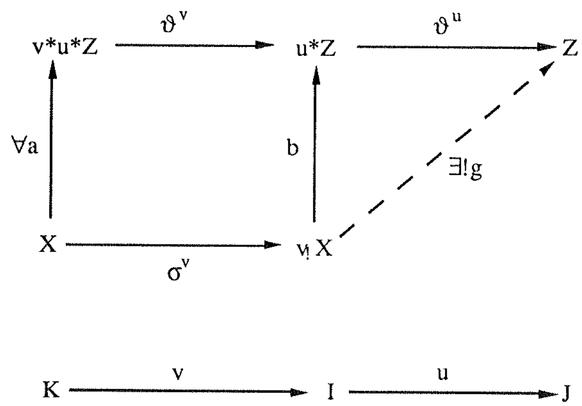
("$\forall f\in\mathcal{E}^{op}{}_u(X,Z)$" is translated to "$\forall a\in\mathcal{E}_I(u^*Z,X)$", where $f=\vartheta/a$; and "$a'\circ\psi=f$" in $\mathcal{E}^{op}$ comes down in $\mathcal{E}$ to "$a=\varepsilon\circ u^*(a')$".)

Note that the class of arrows $\psi\in\mathcal{E}^{op}$ satisfying this condition must automatically be closed under the composition. This follows from the first proposition below. The second one answers the question: When can we put an arrow $\hat\psi\in\mathcal{E}_u(X,u_*X)$ in place of an opcartesian arrow $\psi\in\mathcal{E}^{op}(X,u_*X)$, so that "the diagrams remain commutative"? (The answer is: When u_* is a full functor.)

Propositions.

21. Let E be a fibration. If every arrow $u\in\mathcal{B}(I,J)$ has a cocartesian lifting at every object $X\in|\mathcal{E}_I|$, then E is a bifibration.

• We show that E satisfies the dual of condition 1.31(b). Given an arrow $f\in\mathcal{E}_{uv}(X,Z)$, there is unique vertical arrow a, such that $f=\vartheta^u\circ\vartheta^v\circ a$.



Since σ^v is cocartesian, there is unique vertical arrow b, with $b\circ\sigma^v=\vartheta^v\circ a$. Taking

$$g:=\vartheta^u\circ b,$$

we have $f=g\circ\sigma^u$, and $Eg=u$, as required by the dual of 1.31(b).

To show the uniqueness of g, suppose that g' satisfies the same pair of conditions. From $Eg'=u$ follows that there is a vertical arrow b', such that $g'=\vartheta^u\circ b'$. From $f=g'\circ\sigma^u$, we conclude that $\vartheta^u\circ b\circ\sigma^v=\vartheta^u\circ b'\circ\sigma^v$. Since both $b\circ\sigma^v$ and $b'\circ\sigma^v$ are over v, and ϑ^u satisfies condition 1.31(b), $b\circ\sigma^v=b'\circ\sigma^v$ must hold. But σ^v is cocartesian, and therefore b=b' must be true. Hence g=g'.•

22. Let E be a right bifibration. Choose for $u\in\mathcal{B}(I,J)$ and $X\in|\mathcal{E}_I|$ an E-opcartesian lifting $\psi=\theta/\varepsilon\in\mathcal{E}^{op}(X,u_*X)$. Thus, for every E-cartesian $\vartheta^u\in\mathcal{E}(u^*Z,Z)$ a bijection

$$(\_)':\mathcal{E}_I(u^*Z,X)\longrightarrow\mathcal{E}_J(Z,u_*X)$$

is given. Then holds

$$(\alpha)\Leftrightarrow(\beta)$$

for

α) $\exists\hat\psi\in\mathcal{E}_u(X,u_*X)\ \forall f=\vartheta^u/a\in\mathcal{E}^{op}(X,Z).\ \hat\psi\circ a=a'\circ\vartheta^u$ and

β) $\forall Y\in|\mathcal{E}_I|\ \ \forall b\in\mathcal{E}_J(u_*X,u_*Y)\ \ \exists a\in\mathcal{E}_I(X,Y).\ b=u_*(a)$.

• It is easy to see that

$(\alpha) \Leftrightarrow \exists \hat{\psi} \in \mathcal{E}_u(X, u_* X).\ \hat{\psi} \circ \varepsilon = \theta.$

On the other hand,

$(\beta) \Leftrightarrow \exists e.\ e \circ \varepsilon = id$ (i.e. ε is a split mono).

Namely, the functions

$\mathcal{E}_I(X,Y) \longrightarrow \mathcal{E}_I(u^* u_* X, Y): a \longmapsto a \circ \varepsilon$

are surjective for every Y <u>iff</u> $\exists e.\ e \circ \varepsilon = id$. Extending these functions along

$\mathcal{E}_I(u^* u_* X, Y) \simeq \mathcal{E}_J(u_* X, u_* Y)$

we get

$\mathcal{E}_I(X,Y) \longrightarrow \mathcal{E}_J(u_* X, u_* Y): a \longmapsto u_*(a).$

So we need to prove

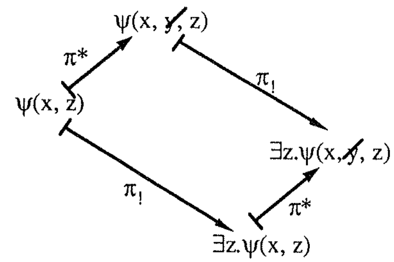
$\exists \hat{\psi}.\ \hat{\psi} \circ \varepsilon = \theta \Leftrightarrow \exists e.\ e \circ \varepsilon = id.$

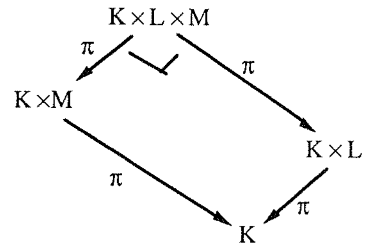
$\Rightarrow$: The splitting e is the vertical factorisation of $\hat{\psi}$ through θ.

$\Leftarrow$: $\hat{\psi} := \theta \circ e$.•

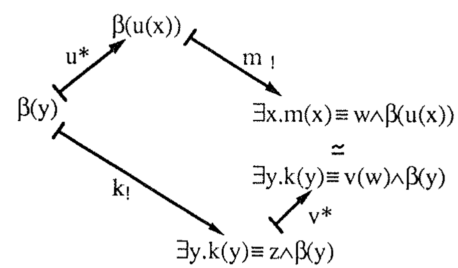
3. Beck-Chevalley condition/property.

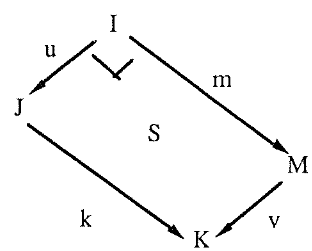
Motivation. If direct images are to represent quantifiers, they must be stable under substitution. A logical picture of this is that *the variables must be independent*, in the sense that y must be invariant under $\exists z$. One way to express this is to require that quantifying over z commutes with adding a dummy y.





For the quantifiers generalized by means of an equality predicate ≡, the independence of variables can be expressed over an arbitrary square S.





A derivation $\exists x.m(x) \equiv w \wedge \beta(u(x)) \vdash \exists y.k(y) \equiv v(w) \wedge \beta(y)$ can be given if the commutativity of S is provable, i.e.

$\vdash k(u(x)) \equiv v(m(x))$.

On the other hand, a proof $\exists y.k(y) \equiv v(w) \wedge \beta(y) \vdash \exists x.m(x) \equiv w \wedge \beta(u(x))$ follows from

$k(y) \equiv v(z) \vdash \exists x.u(x) \equiv y \wedge m(x) \equiv z$,

which tells that S is a (weak) pullback. In this way, logic suggests the requirement that

$\exists x.m(x) \equiv w \wedge \beta(u(x)) \simeq \exists y.k(y) \equiv v(w) \wedge \beta(y)$

holds over the pullback squares S.

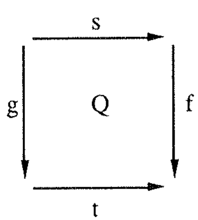
Definitions. A cloven bifibration $E : \mathcal{E} \longrightarrow \mathcal{B}$ is said to satisfy the *Beck condition* over the square S (as above) in $\mathcal{B}$ if there is a cartesian natural isomorphism

$m_! \circ u^* \simeq v^* \circ k_!$.

We say that E has the *Beck property* if it satisfies the Beck condition over all the pullback squares.

A cloven trifibration has the Beck property if both E and E^{op} do.

The *Chevalley condition* on a commutative square



in a bifibred category $\mathcal{E}$ is:

C) if s and t are cartesian and f is cocartesian, then g is cocartesian.

A bifibration E is said to satisfy the Chevalley condition over a square S if every commutative square Q over S satisfies this condition. E has the *Chevalley property* if it satisfies the Chevalley condition over all the pullback squares.

A trifibration has the Chevalley property if both E and E^{op} do.

Lemma. A bifibration E satisfies the Chevalley condition over S iff every Q over S satisfies:

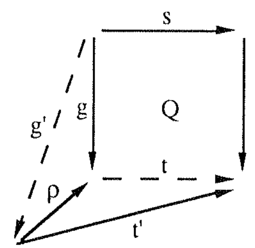
C°) if f and g are cocartesian and s is cartesian, then t is cartesian,

or equivalently

C$^{\underline{o}}$) if s is cartesian and f cocartesian, then

t is cartesian iff g is cocartesian.

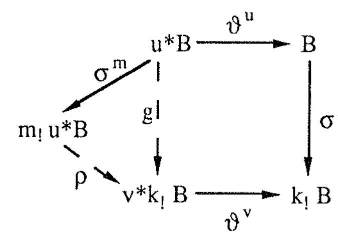
• (C)⇒(C°): Take a cartesian t' over v.



By (C), the unique factorisation g' of fs through t' is cocartesian. Both g and g' are cocartesian liftings of m; hence there is a vertical isomorphism ρ, such that g'=ρg. From tg=(fs=t'g'=)t'ρg follows t=t'ρ by the cocartesianness of g. Thus t is cartesian.•

Remark. The Chevalley condition *is* the Beck condition expressed without cleavage.

• If the unique arrow g over m by which $\sigma^k \circ \vartheta^u$ factorizes through ϑ^v is cocartesian, then the unique vertical arrow ρ induced by g



must be an iso. (Alternatively, ρ can be induced by $t : m_!u^*B \longrightarrow k_!B$, the unique arrow over v by which $\sigma^k \circ \vartheta^u$ factorize through σ^m.) Given *any* vertical iso $m_!u^*B \simeq v^*k_!B$, g must be cocartesian, and the *canonical* iso ρ is obtained.• Since we can, on the other hand, interprete the Beck condition as a statement about "all the *possible* inverse and direct images" along the given arrows (i.e. cleavage-free), it makes sense to talk about a *Beck-Chevalley condition/property* - as everybody already does. We shall sometimes abbreviate it to "BC-property". It is a standard notion in topos theory. However, neither Beck nor Chevalley ever published anything on their condition(s). (Early references are: Bénabou-Roubaud 1970, Lawvere 1970. A recent one: Hyland-Moerdijk 1990.)

Fact. A trifibration E has the Beck-Chevalley property <u>iff</u> either E or E^{op} has it as a bifibration.

• $\mathcal{E}_J(B,k^*v_*D) \simeq \mathcal{E}_K(k_!B,v_*D) \simeq \mathcal{E}_M(v^*k_!B,D) \simeq \mathcal{E}_M(m_!u^*B,D) \simeq \mathcal{E}_K(u^*B,m^*D)$
$\simeq \mathcal{E}_J(B,u_*m^*D)$•

A simple non-example. Let <u>Pos</u> be the category of posets and monotone maps and

$|\underline{Pos}/\Omega| := \{\langle \mathbb{A}, I\rangle: \mathbb{A} \text{ a poset}, I=\uparrow I\subseteq \mathbb{A}\}$,

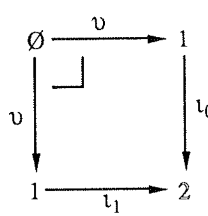
$\underline{Pos}/\Omega(\langle \mathbb{A}, I\rangle, \langle \mathbb{B}, J\rangle) := \{f\in \underline{Pos}(\mathbb{A}, \mathbb{B}): f(I)\subseteq J\}$,

where $\uparrow I:=\{y\in \mathbb{A}: \exists x\in I. x\leq y\}$. <u>Pos</u>/Ω is bifibred over <u>Pos</u> by the obvious projection, and

$f^*\langle \mathbb{B},J\rangle = \langle \mathbb{A}, \uparrow f^{-1}(J)\rangle$,

$f_!\,\langle \mathbb{A},I\rangle = \langle \mathbb{B}, \uparrow f(I)\rangle$.

Consider the pullback square



in <u>Pos</u>, where $\mathbb{2} = \{0<1\}$, $1 = \{\emptyset\}$, $\iota_k(\emptyset) = k$. Then

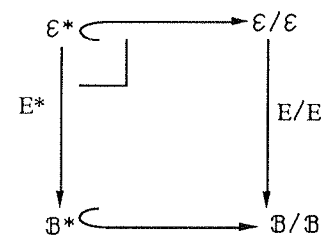
$\upsilon_!\circ\upsilon^*\langle 1,1\rangle = \langle 1, \emptyset\rangle \not\simeq \langle 1, 1\rangle = \iota_1{}^*\circ\iota_{0!}\langle 1, 1\rangle$.

With the discrete "poset" $2 = \{0,1\}$ instead of $\mathbb{2}$, the Beck-Chevalley condition would be satisfied.

4. A characterisation of the Beck-Chevalley property.

Propositions 2.22 and 2.23 told us that in every fibred category $\mathcal{E}$ the class of the cartesian arrows is stable under all pullbacks preserved by E; and that a square with two cartesian sides is a pullback if it is over a pullback. If the square S in the definition of the Chevalley condition is a pullback, then so is the square Q. In fact, the Chevalley *property* can be reformulated by saying that the class of the cocartesian arrows in $\mathcal{E}$ is stable under all the E-preserved pullbacks along cartesian arrows. These facts point towards an abstract characterisation of the Beck-Chevalley property.

Lemma. Given a functor $E:\mathcal{E} \longrightarrow \mathcal{B}$, define E* by the pullback



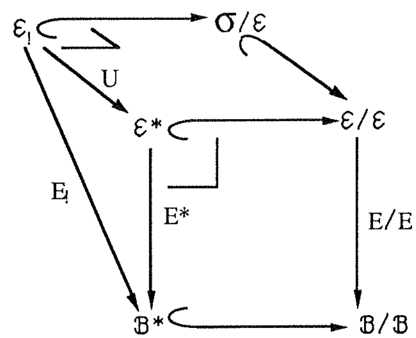
where $\mathcal{B}^* \hookrightarrow \mathcal{B}/\mathcal{B}$ is the category of arrows of $\mathcal{B}$, with pullback squares as the arrows between them; and

$E/E : \mathcal{E}/\mathcal{E} \longrightarrow \mathcal{B}/\mathcal{B} : f \mapsto Ef.$

Then E is a fibration <u>iff</u> E/E is <u>iff</u> E* is.

• An E/E-cartesian lifting is a pair of E-cartesian liftings. An E-cartesian lifting of $u\in \mathcal{B}(I,J)$ at X is obtained from an E*-cartesian lifting of $\langle u,u\rangle\in \mathcal{B}^*(id_I, id_J)$ at id_X.•

Proposition. If E is a bifibration, let $\sigma/\mathcal{E} \hookrightarrow \mathcal{E}/\mathcal{E}$ be the full subcategory spanned by the cocartesian arrows. Define U and $E_!$ by the following commutative diagram:



Then E has the Beck-Chevalley property <u>iff</u> $E_!$ is a fibration and U a cartesian functor.

• $\mathcal{E}_!$-arrows are the squares in $\mathcal{E}$ which lie over pullbacks in $\mathfrak{B}$ and have two sides cocartesian. That U is cartesian, means that the $E_!$-cartesian arrows are pairs of E-cartesian arrows. The Chevalley property tells us that every object of $\mathcal{E}_!$ has an inverse image in $\mathcal{E}_!$ along every arrow from $\mathfrak{B}^*$.•

5. The Beck-Chevalley condition over all commutative squares.

Proposition. The statements below are related as follows:

(a)⇔(b)⇒(c);

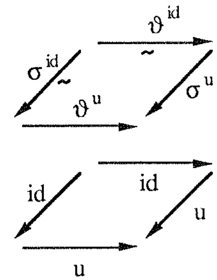
(c)⇒(a) - if $\mathfrak{B}$ has pullbacks.

a) E satisfies the Chevalley condition over every commutative square.

b) An arrow in $\mathcal{E}$ is cartesian <u>iff</u> it is cocartesian.

c) E has the Chevalley property; <u>and</u>: if an arrow in E is cartesian <u>then</u> it is cocartesian.

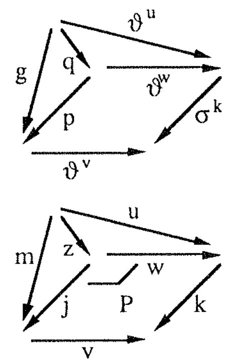
• (a)⇒(b):



(b)⇒(a): Take an arbitrary commutative square S, and let g be the unique factorisation over m of $\sigma^k \circ \vartheta^u$ through ϑ^v, as in remark 3. Using (b) and the closure properties of cartesian arrows (1.34) we have:

σ^k cocartesian <u>then</u> σ^k cartesian <u>then</u> $\sigma^k \circ \vartheta^u = \vartheta^v \circ g$ cartesian <u>then</u> g cartesian <u>then</u> g cocartesian.

(c)⇒(a): Since $\mathfrak{B}$ has pullbacks, the square S factorizes through a pullback square P, and then we lift:



The factorisation q over z of ϑ^u through ϑ^w is cartesian. By (c), it is then cocartesian. The factorisation p over j of $\sigma^k \circ \vartheta^w$ through ϑ^v is cocartesian by the Chevalley property. Hence the factorisation $g = p \circ q$ over m of $\sigma^k \circ \vartheta^u$ through ϑ^v must be cocartesian.•

Corollary. For a cloven fibration E, the conditions below are related:

(a)$\Leftrightarrow$(b)$\Rightarrow$(c);
(c)$\Rightarrow$(a) - if $\mathcal{B}$ has pullbacks.

a) E satisfies the Beck condition over every commutative square.

b) All the inverse and direct image functors are equivalences, i.e. for every $u \in \mathcal{B}$

$u_! \circ u^* \simeq \mathrm{id}$ and $u^* \circ u_! \simeq \mathrm{id}$.

c) E has the Beck property, <u>and</u> every inverse image functor is full and faithful, i.e.

$u_! \circ u^* \simeq \mathrm{id}$.

• (c)$\Rightarrow$(a) is now simply: $l_! \circ u^* \simeq p_! \circ z_! \circ z^* \circ w^* \simeq p_! \circ w^* \simeq v^* \circ k_!$.•

Remark. The diagram "(a)$\Rightarrow$(b)" above shows that in every bifibration E with the BC-property a lifting of a monic is cartesian <u>iff</u> it is cocartesian; i.e., every direct and every inverse image functor along a monic is an equivalence of categories in such E.

6. Hyperfibrations.

Terminology. In a bifibration with the Beck-Chevalley property, the direct images are called *coproducts*. The right direct images in a right bifibration with the BC-property are called *products*. (I.e., Products and coproducts are direct images which are stable under the inverse images.)

Definitions. A fibration E is said to *have small coproducts* if it is a bifibration with the Beck-Chevalley property. The full subcategory of <u>BIFIB</u>/$\mathcal{B}$ spanned by the fibrations with small coproducts is denoted by <u>FIB$_!$</u>/$\mathcal{B}$.

We say that E *has small products* if E^{op} has small coproducts. The corresponding subcategory is <u>FIB$_*$</u>/$\mathcal{B} \subseteq$ <u>BIFIB</u>/$\mathcal{B}$.

A fibration with small products and coproducts is called *hyperfibration.* (I.e., it is a trifibration with Beck-Chevalley property.) <u>HYP</u>/$\mathcal{B}$ is the full subcategory of <u>TRIFIB</u>/$\mathcal{B}$ which consists of hyperfibrations.

Example. Every (ordinary) category $\mathcal{C}$ gives rise to a split *family fibration*

$\nabla\mathcal{C}$: <u>Set</u>/$\mathcal{C} \longrightarrow$ <u>Set</u>,

in which the fibres

$(\underline{\mathrm{Set}}/\mathcal{C})_I := \mathcal{C}^I$

consist of the functors $I \longrightarrow \mathcal{C}$, i.e. the I-indexed families $(C_x \mid x \in I)$ of objects of $\mathcal{C}$, with I-indexed families of $\mathcal{C}$-arrows. The inverse image functor over $u: I \longrightarrow J$ is

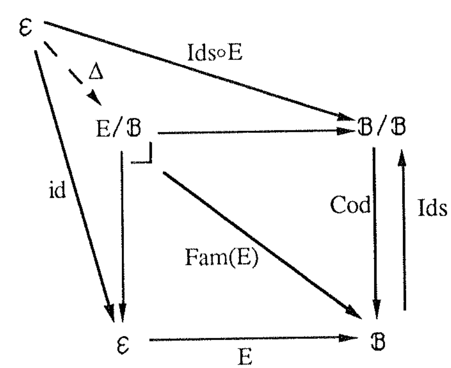
$u^*(B_y \mid y \in J) := (B_{u(x)} \mid x \in I)$.

The fibration $\nabla\mathcal{C}$ has small (co)products <u>iff</u> the category $\mathcal{C}$ has small (i.e. set indexed) (co)products in the usual sense. They are

$$u_! \,(A_x \mid x \in I) := \Big(\sum_{u(x)=y} A_x \mid y \in J\Big),$$

$$u_* \,(A_x \mid x \in I) := \Big(\prod_{u(x)=y} A_x \mid y \in J\Big).$$

Proposition. (Bénabou 1975b) (AC) Let $\mathcal{B}$ be a category with pullbacks. A fibration $E: \mathcal{E} \longrightarrow \mathcal{B}$ has small coproducts <u>iff</u> the cartesian functor $\Delta: E \longrightarrow \mathrm{Fam}(E)$



has a cartesian left adjoint. (The functor Ids takes every object to its identity arrow.)

• The BC-property boils down to the fact that the functor $\Sigma \dashv \Delta$,

$$\Sigma(X,u:EX \to J) := u_! X,$$

is cartesian.•

7. Closure properties for co-, bi-, tri-, and hyperfibrations.

The class of cocartesian arrows is closed under the right division and stable under the pushouts preserved by E - *dually* to the class of cartesian arrows (cf. 1.34 and 2.23).

The class of cofibrations has, on the other hand, obviously the *same* closure properties as the class of fibrations: it is stable under all pullbacks, and closed under composition. Idem for the class of bifibrations, of course. As for trifibrations, it is routine to see that they are stable under all pullbacks (• for every functor $C:\mathcal{C} \to \mathcal{B}$ holds $(C^*E)^{op} \cong C^*(E^{op})$, where $C^*E \in |\underline{FIB}/\mathcal{C}|$ is a pullback of E along C; hence if E^{op} is a bifibration <u>then</u> $(C^*E)^{op}$ is one•); but they don't seem to be closed under the composition.

We shall now show that hyperfibrations are closed under the composition (74), and stable under pullbacks along pullback-preserving functors (71).

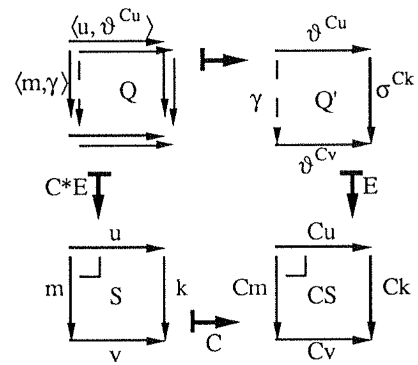
Terminology. A bifibration $\tilde{E}:\mathcal{F} \to \mathcal{E}$ over a fibred category $\mathcal{E}$ has the *vertical(-cartesian) Beck-Chevalley property* (abbreviated vBC and vcBC) if it satisfies the Chevalley condition over all the pullback squares in $\mathcal{E}$ in which two opposite sides are vertical (while the other two are cartesian).

Propositions. Consider a functor $C:\mathcal{C} \to \mathcal{B}$ and fibrations $E:\mathcal{E} \to \mathcal{B}$, $\tilde{E}:\mathcal{F} \to \mathcal{E}$, $C^*E : \mathcal{C} \times_{\mathcal{B}} \mathcal{E} \to \mathcal{C}$ and $F := E \circ \tilde{E} : \mathcal{F} \to \mathcal{B}$. (In the proofs, ϑ will denote the F-cartesian arrows in $\mathcal{F}$ and the E-cartesian arrows in $\mathcal{E}$; generic notation for $\tilde{E}$-cartesian arrows in $\mathcal{F}$ will be $\tilde{\vartheta}$. Same for the cocartesian arrows. $\mathcal{F}_I := F^{-1}(I)$.)

71. If C preserves pullbacks <u>and</u> E has small (co)products, <u>then</u> C*E has small (co)products.

• We prove that C*E has the BC-property <u>if</u> E has it.

A C*E-cartesian lifting of $p \in \mathcal{C}(U,V)$ is in the form $\langle p, \vartheta^{Cp} \rangle$; a C*E-cocartesian lifting of p is in the form $\langle p, \sigma^{Cp} \rangle$. A square Q in $\mathcal{C} \times_{\mathcal{B}} \mathcal{E}$, over a pullback square S in $\mathcal{C}$, and such that s and t are cartesian and f is cocartesian, is projected as follows:



Since CS is a pullback square, and E has the BC-property, the arrow γ is cocartesian in $\mathcal{E}$. The arrow $g = \langle m, \gamma \rangle$ is therefore cocartesian in $\mathcal{C} \times_{\mathcal{B}} \mathcal{E}$.•

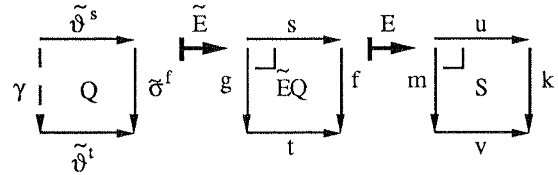
72. If E <u>and</u> $\tilde{E}$ have small coproducts, <u>then</u> F has them.

• Again we know that F is a bifibration, and just prove its BC-property.

The *formula* is:

F-(co)cartesian arrows are $\tilde{E}$-(co)cartesian liftings of E-(co)cartesian arrows.

(Namely, if h is an E-cartesian lifting of u at $\tilde{E}Z$, $\tilde{\vartheta}^h_Z$ is an F-cartesian lifting of u at Z. Every F-cartesian lifting t of u at Z is $t \simeq \tilde{\vartheta}^h_Z$ for $h := \tilde{E}t$.) The two F-cartesian sides of a square Q in $\mathcal{F}$ are thus $\tilde{E}$-cartesian too; its F-cocartesian side is $\tilde{E}$-cocartesian.



Since bifibration $\tilde{E}$ is a cartesian and cocartesian functor, s and t are E-cartesian, f is E-cocartesian; because of the BC-property of E, g is E-cocartesian. By 2.22, $\tilde{E}Q$ is a pullback square. Because of the BC-property of $\tilde{E}$, γ is $\tilde{E}$-cocartesian. By the *formula*, γ must be F-cocartesian.•

73. If E^{op} and $\tilde{E}^{op}$ are bifibrations, the latter with the vcBC-property, then F^{op} is a bifibration. If $\tilde{E}$ is a hyperfibration and E a trifibration, then is F a trifibration.

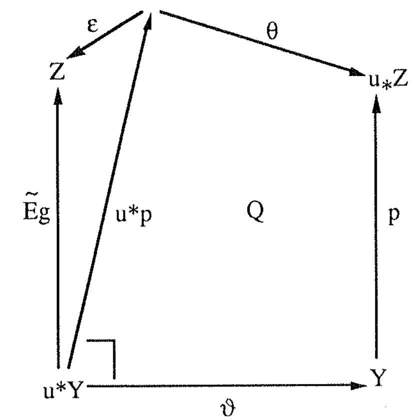
• F^{op} is certainly a fibration, because F is. We show that F^{op} is a cofibration: given $u\in\mathcal{B}(I,J)$ and $C\in|\mathcal{F}_I|$, we define

an F-opcartesian lifting $\mathring{\psi} = \mathring{\theta}/\mathring{\varepsilon}$ of u at C, using
an E-opcartesian lifting $\psi = \theta/\varepsilon$ of u at $\tilde{E}C$, and
an $\tilde{E}$-opcartesian lifting $\tilde{\psi} = \tilde{\theta}/\tilde{\varepsilon}$ of θ at ε^*C (an $\tilde{E}$-inverse image)

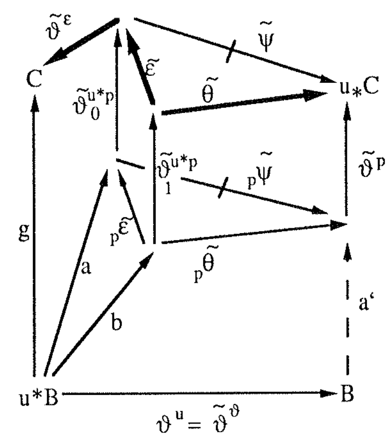
as follows:

$\mathring{\varepsilon} := \tilde{\vartheta}^{\varepsilon}_C \circ \tilde{\varepsilon}$
$\mathring{\theta} := \tilde{\theta}.$

We show that $\mathring{\psi}$ is indeed opcartesian. For an arbitrary $g\in\mathcal{F}_I(u^*B,C)$, consider $\tilde{E}g\in\mathcal{E}_I(u^*Y,Z)$. Since ψ is opcartesian, there is a unique $p\in\mathcal{E}_J(Y,u_*Z)$ such that $\tilde{E}g$ factorizes through ε by u^*p.



The square Q consists of two cartesian and two vertical arrows in $\mathcal{E}$. By the hypothesis, $\tilde{E}^{op}$ has the vcBC-property. The unique factorisation ${}_p\tilde{\psi} := {}_p\tilde{\theta}/{}_p\tilde{\varepsilon}$ over ϑ of $\tilde{\psi}\circ\tilde{\vartheta}^{u^*p}_0$ through $\tilde{\vartheta}^p$ must therefore be $\tilde{E}$-opcartesian.



Let a be the unique $\tilde{E}$-vertical factorisation of g through $\tilde{\vartheta}^{\varepsilon}\circ\tilde{\vartheta}^{u^*p}$. The cocartesianness of ${}_p\tilde{\psi}$ means that there is a unique $\tilde{E}$-vertical a‘ such that $a = {}_p\tilde{\varepsilon}\circ b$, where b is an $\tilde{E}$-inverse image of a‘ along $\vartheta^u\in\mathcal{E}(u^*Y,Y)$. Now let

$g' := \tilde{\vartheta}^p\circ a'.$

Since p was E-vertical, g' is F-vertical. Since $\tilde{\vartheta}$ is (by the *f*ormula) an F-cartesian lifting of u, the F-inverse image is $u^*(g') = \tilde{\vartheta}u_1^*p \circ b$. Hence

$$g = \mathring{\varepsilon} \circ u^*(g').$$

The second assertion follows by just adding to this reasoning the fact that F is a bifibration if E and $\tilde{E}$ are.•

74. Hyperfibrations are closed under composition.

• If E and $\tilde{E}$ are hyperfibrations, then F is a trifibration by 73. By 72, the bifibration F has the BC-property. From fact 3 follows that F^{op} has this property too. (Otherwise, check directly that the BC-property holds for the arrow $\mathring{\psi}$ defined in 73.)•

8. Fibrewise co-, bi-, tri-, hyperfibrations.

A category of predicates, structure on which we shall focus in chapter IV, will be a hyperfibration over a hyperfibration, i.e., a hyperfibration in which every fibre is again a hyperfibration, and the inverse images preserve this fibrewise structure. This may sound complicated, but we shall show that it boils down to a bit less than two honest hyperfibrations. So we first give some lengthy definitions in the by now common "fibrewise-structure-preserved-by-inverse-images" style, and then characterize the defined notions globally.

Definitions. F∈ FIB/$\mathfrak{B}$(E',E) is a *(fibrewise) cofibration in* FIB/$\mathfrak{B}$ if all its restrictions $F_I:\mathcal{E}'_I \to \mathcal{E}_I$ are cofibrations and if E'-inverse images preserve the F_I-cocartesian arrows.

F∈ BIFIB/$\mathfrak{B}$(E',E) is a *(fibrewise) bifibration in* BIFIB/$\mathfrak{B}$ if all $F_I:\mathcal{E}'_I \to \mathcal{E}_I$ are bifibrations, E'-inverse images preserve F_I-cartesian as well as F_I-cocartesian arrows, while E'-direct images preserve the F_I-cocartesian arrows.

F∈ TRIFIB/$\mathfrak{B}$(E',E) is a *(fibrewise) trifibration in* TRIFIB/$\mathfrak{B}$ if all $F_I:\mathcal{E}'_I \to \mathcal{E}_I$ are trifibrations, E'-inverse images preserve F_I-cartesian, F_I-cocartesian, and F_I-opcartesian arrows, and E'-direct images preserve F_I-cocartesian arrows.

Facts. Denote by op(F)∈ FIB/$\mathfrak{B}$(E'^{op},E^{op}) the image of F∈ FIB/$\mathfrak{B}$(E',E) by the arrow part of the functor op.

81. F∈ FIB/$\mathfrak{B}$(E',E) is a fibrewise cofibration iff op(F)∈ FIB/$\mathfrak{B}$(E'^{op},E^{op}) is a fibrewise fibration.

82. F∈ BIFIB/$\mathfrak{B}$(E',E) is a fibrewise bifibration iff F, op(F) and F^o are fibrewise fibrations.

83. F∈ TRIFIB/$\mathfrak{B}$(E',E) is a fibrewise trifibration iff F, op(F), F^o and op(F^{op}) are fibrewise fibrations. (• The domain of op(F^{op}) is obtained by first making $F^{op}:\mathcal{E}'^{op} \to \mathcal{E}$, and then op($F^{op}$): $(EF^{op})^{op} \to E^{op}$. The fibration $(EF^{op})^{op}$ has the same cartesian arrows as EF^{op}, thus the same as E'=EF.•)

Remark. All the above definitions give rise to appropriate categories: for instance, the fibrewise cocartesian functors between fibrewise cofibrations F∈ FIB/$\mathfrak{B}$(E',E) and G∈ FIB/$\mathfrak{B}$(E",E) should be the functors H∈ FIB/$\mathfrak{B}$(E',E"), such that F=GH *and* H_I∈ COFIB/$\mathcal{E}_I$(F_I, G_I). But we shall only need the following categories:

|BIFIB/E|	:=	the fibrewise bifibrations over E in BIFIB/$\mathfrak{B}$,
BIFIB/E(F,G)	:=	{H∈ (BIFIB/$\mathfrak{B}$)/E | ∀I.H_I∈ BIFIB/$\mathcal{E}_I$(F_I, G_I)}
|TRIFIB/E|	:=	the fibrewise trifibrations over E in TRIFIB/$\mathfrak{B}$,
TRIFIB/E(F,G)	:=	{H∈ (TRIFIB/$\mathfrak{B}$)/E | ∀I.H_I∈ TRIFIB/$\mathcal{E}_I$(F_I, G_I)}

and the full subcategories

$BIFIB_!$/E ⊆ BIFIB/E and

HYP/E ⊆ TRIFIB/E

spanned respectively by the fibrewise bifibrations and trifibrations F with the components F_I which have the Beck-Chevalley property. (In other words, all the F_I have small coproducts in the first case, and they are hyperfibrations in the second case.)

Moreover, we denote by

$\underline{\text{BIFIB}}_{v(c)}/\mathcal{E} \subseteq \underline{\text{BIFIB}}/\mathcal{E}$

$\underline{\text{TRIFIB}}_{v(c)}/\mathcal{E} \subseteq \underline{\text{TRFIB}}/\mathcal{E}$

the categories of bifibrations resp. trifibrations with the v(c)BC-property.

Propositions.

84. i) $\underline{\text{BIFIB}}/E = \underline{\text{BIFIB}}_{vc}/\mathcal{E}$.
ii) $\underline{\text{BIFIB}}_{!}/E = \underline{\text{BIFIB}}_{v}/\mathcal{E}$.

• We only show the correspondence on objects. The correspondence on arrows follows immediately from proposition 2.3.

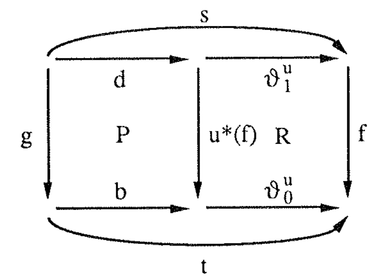
←: If F is a bifibration, then all the F_I are bifibrations, E'-inverse images preserve the F_I-cartesian arrows, and E'-direct images preserve the F_I-cocartesian arrows - by proposition 2.3.

(i) The F_I-cocartesian arrows are just the F-cocartesian liftings of arrows from $\mathcal{E}_I$; the E'-cartesian arrows are the F-cartesian liftings of E-cartesian arrows. The preservation of the F_I-cocartesian arrows under the E'-inverse images is just the vcBC-property of F.

(ii) If F has the vBC-property, then it also satisfies the BC-condition over all the pullback squares which lie in $\mathcal{E}_I$. By proposition 2.24, the inclusion $\mathcal{E}_I \hookrightarrow \mathcal{E}$ preserves the pullbacks. Since every F_I is obtained from F by pulling back along $\mathcal{E}_I \hookrightarrow \mathcal{E}$, proposition 71 applies: all F_I have small coproducts.

→: If F is a fibrewise bifibration, both F and F^o are fibrewise fibrations, and by 2.3 again, both F and F^o are global fibrations. We now derive (ii) the vBC-property of $F \in \underline{\text{BIFIB}}_{!}/E$. (The vcBC-property of $F \in \underline{\text{BIFIB}}/E$ is an obvious part of the same argument.)

Take in $\mathcal{E}'$ a commutative square Q such that FQ is a pullback, t and s are F-cartesian, and f is F-cocartesian. Let Ff and Fg be vertical; hence $E't = E's = u \in \mathcal{B}(I,J)$. Consider the E'-vertical-cartesian decompositions of s and t.



Since both t and ϑ_0^v are F-cartesian, b must be F_I-cocartesian. Idem for d. Since E'-inverse images preserve F_J-cocartesian arrows, u*(f) is F_I-cocartesian. The square P is a pullback because both R, and Q=P+R are (by 2.22). Thus F_IP is a pullback (2.22 again). The BC-property of F_I says that g is F_I-cocartesian, thus F-cocartesian too.•

85. i) $\underline{\text{TRIFIB}}/E = \underline{\text{TRIFIB}}_{vc}/\mathcal{E}$.
ii) $\underline{\text{HYP}}/E = \underline{\text{TRIFIB}}_{v}/\mathcal{E}$.

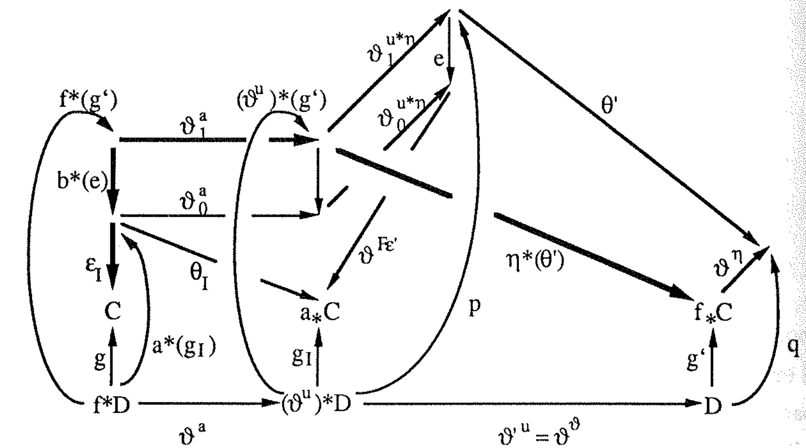
• ←: Given an $F \in |\underline{\text{TRIFIB}}_{vc}/\mathcal{E}|$, E':=EF is a trifibration by 73; $F \in \underline{\text{TRIFIB}}/\mathcal{B}(E',E)$ is immediate. F_I^{op} is obtained from F^{op} by pulling back - so that the conclusion that F_I is (i) a trifibration or (ii) a hyperfibration follows just as in the preceding proposition.

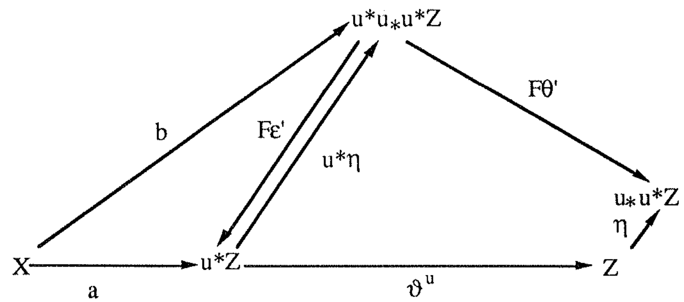
→: We know from the previous proposition that $F \in |\underline{\text{TRIFIB}}/E|$ is a bifibration over $\mathcal{E}$ with (i) the vcBC-property; and if the components F_I have the BC-property, F has (ii) the vBC-property . In view of fact 3, it is now sufficient to show that F^{op} is a cofibration.

Given $f \in \mathcal{E}(X,Z)$, $Ef = u \in \mathcal{B}(I,J)$, with a decomposition $f = \vartheta^u \circ a$, and $C \in |\mathcal{E}'_X|$ such that FC=X, we define

an F-opcartesian lifting $\psi = \theta/\varepsilon$ of f at C, using
an F_I-opcartesian lifting $\psi_I = \theta_I/\varepsilon_I$ of a at C, and
an E'-opcartesian lifting $\psi' = \theta'/\varepsilon'$ of u at a_*C.

Since F is a morphism of trifibrations, it preserves the E'-opcartesian arrows. $F\psi'$ is thus E-opcartesian. So it forms in $\mathcal{E}$ the lower diagram on the following picture.





For b := $u^*\eta \circ a$ holds $a = F\varepsilon' \circ b$, because $F\varepsilon' \circ u^*\eta = id$.

The upper diagram is in $\mathcal{E}'$. The E'-vertical arrow ε' decomposes

1) $\varepsilon' = \vartheta^{F\varepsilon'} \circ e$

with respect to F. $\eta^*(\theta')$ is the unique factorisation over ϑ^u of θ' through ϑ^η. Define:

$\varepsilon := \varepsilon_I \circ b^*(e)$

$\theta := \eta^*(\theta') \circ \vartheta^a_1$.

Given a $g \in \mathcal{E}'_X(f^*D, C)$, by the assumption on ψ_I there is a unique

$g_I \in \mathcal{E}'_{u^*Z}\big((\vartheta^u)^*D, a_*C\big)$, such that

2) $g = \varepsilon_I \circ a^*(g_I)$.

Of course, g_I is E'-vertical, so by the assumption on ψ', there is a unique E'-vertical arrow q such that

3) $g_I = \varepsilon' \circ p$

holds for the E'-inverse image p of q along u. Since F is a cartesian functor, Fp is an inverse image of Fq. Since F is opcartesian, $F\theta'/F\varepsilon'$ is an E-opcartesian arrow. Putting these two facts together, we conclude that $Fq \in \mathcal{E}_J(Z, u_*u^*Z)$ corresponds to $F\varepsilon' \circ Fp = id \in \mathcal{E}_I(u^*Z, u^*Z)$. By the uniqueness, $Fq = \eta$. Let g' be the unique arrow in $\mathcal{E}'_Z$ such that $q = \vartheta^\eta \circ g'$. Clearly

4) $p = \vartheta_1^{u^*\eta} \circ (\vartheta^u)^*(g')$.

From the calculation:

$$\begin{aligned}
\theta_I \circ a^*(g_I) &= g_I \circ \vartheta^a \overset{(3)}{=} \varepsilon' \circ p \circ \vartheta^a \overset{(4)}{=} \\
&\varepsilon' \circ \vartheta_1^{u^*\eta} \circ (\vartheta^u)^*(g') \circ \vartheta^a = \\
&\varepsilon' \circ \vartheta_1^{u^*\eta} \circ \vartheta_1^a \circ f^*(g') \overset{(1)}{=} \\
&\vartheta^{F\varepsilon'} \circ e \circ \vartheta_1^{u^*\eta} \circ \vartheta_1^a \circ f^*(g') = \\
&\vartheta^{F\varepsilon'} \circ \vartheta_0^{u^*\eta} \circ \vartheta_0^a \circ b^*(e) \circ f^*(g') = \\
&\theta_I \circ b^*(e) \circ f^*(g')
\end{aligned}$$

follows $a^*(g_I) = b^*(e) \circ f^*(g')$, and by (2) and the definition of ε

$g = \varepsilon \circ f^*(g')$.•

Asymmetries. Given fibrations $E:\mathcal{E} \longrightarrow \mathcal{B}$ and $F:\mathcal{E}' \longrightarrow \mathcal{E}$, the cartesian arrows of E'=EF are easily obtained: they are the F-cartesian liftings of E-cartesian liftings. Idem for E'-cocartesian arrows. But it wasn't so simple to get the E'-opcartesian arrows: this was the contents of proposition 73.

Given fibrations $E:\mathcal{E} \longrightarrow \mathcal{B}$ and $E':\mathcal{E}' \longrightarrow \mathcal{B}$, and a fibrewise fibration $F:E' \longrightarrow E$, it was easy to obtain the cartesian arrows of $F:\mathcal{E}' \longrightarrow \mathcal{E}$ by composing E'-cartesian arrows and the F_I-cartesian arrows (proposition 2.3). Idem for F-cocartesian arrows. And again, the F-opcartesian arrows demand more: the last proposition.

The preservation of the F_I-cocartesian arrows by the E'-inverse images takes globally the form of the vBC-property. The fact that the F_I-cartesian arrows are preserved by the E'-inverse images corresponds to the closure of the F-cartesian arrows under the composition (2.31); the preservation of the F_I-cocartesian arrows under the left E'-

direct images corresponds to the closure of the F-cocartesian arrows under the composition. *However*, we could not requre that the F_I-opcartesian arrows are preserved under the right E'-direct images, because the right E'-direct images don't preserve the F_I^{op}-*arrows* in the first place. Namely, they needn't preserve the cartesian arrows.

3a. The Beck-Chevalley condition without direct images

1. Interpolation condition.

In the preceding section (part 3) the Beck-Chevalley condition was introduced as a categorical expression of the logical concept of independence of variables - in the form: "Every variable is invariant under quantifying over other variables". There is, however, another way to express the independence of variables:

$\alpha(x,y) \vdash \gamma(y,z) \Leftrightarrow$

there is an *interpolant* $\beta(y)$, such that $\alpha(x,y) \vdash \beta(y) \vdash \gamma(y,z)$.

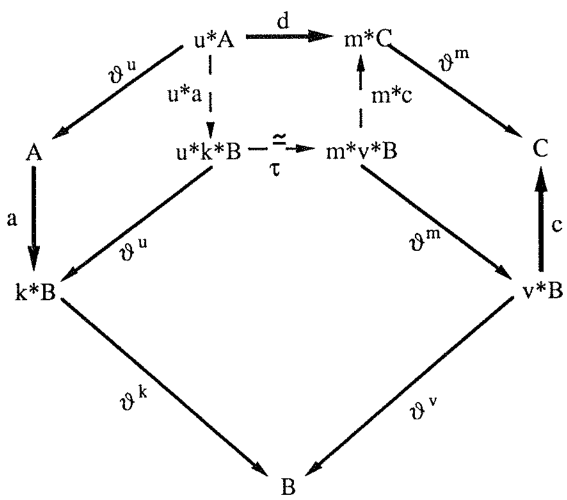
(I.e.: "The different variables do not interfere with each other in a proof". This means that x cannot play any role in a proof of $\gamma(y,z)$; and z cannot play a role in a proof from $\alpha(x,y)$.)

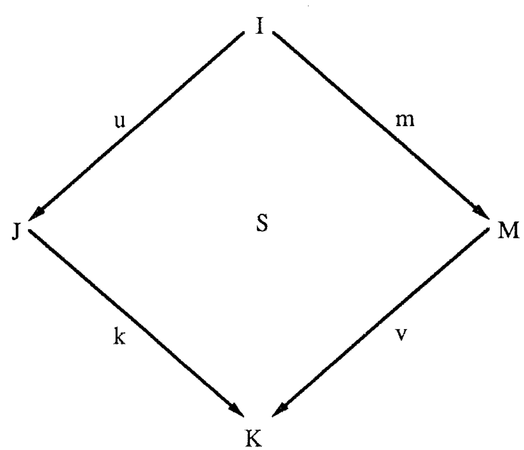
At the first sight, this seems to be a *different idea* of independence of variables. Surprisingly, it is not: in the logic with quantifiers, the two forms of the independence of variables are equivalent. Lifting logic in category theory, we get a *direct-image free* characterisation of the Beck-Chevalley condition. In other words, there is a property of fibrations which a bifibration will have if and only if it has the Beck-Chevalley property. (This characterisation will be applied in III.4.2.)

Definitions. Let $E : \mathcal{E} \longrightarrow \mathcal{B}$ be a fibration, and S a commutative square in $\mathcal{B}$. An (S-)*interpolant* of an arrow $d \in \mathcal{E}_I(u^*A, m^*C)$ is a triple $\langle a, B, c\rangle$, where $B \in |\mathcal{E}_K|$, $a \in \mathcal{E}_J(A, k^*B)$, $c \in \mathcal{E}_M(v^*B, C)$, such that

$$d = m^*(c) \circ \tau \circ u^*(a).$$

(The unique factorisation τ of $\vartheta^k_B \circ \vartheta^u_{k^*B}$ through $\vartheta^v_B \circ \vartheta^m_{v^*B}$ is a vertical iso because these two arrows are cartesian liftings of $ku = vm$.)





An interpolant $\langle a_0, B_0, c_0\rangle$ is *initial* if for any other interpolant $\langle a, B, c\rangle$ (of the same arrow, over the same square) there is a unique arrow $b\in \mathcal{E}_K(B_0, B)$, such that

$$a = k^*(b)\circ a_0.$$

An interpolant $\langle a_0, B_0, c_0\rangle$ is *strong* if for any other interpolant $\langle a, B, c\rangle$ (of the same arrow, over the same square)

$a = a_0$ <u>implies</u> $c = c_0$.

A fibration E satisfies the *interpolation condition* over the square S if there is an S-interpolant for every $d\in \mathcal{E}_I(u^*A, m^*C)$. It satisfies the *strong* interpolation condition if, besides, every initial interpolant is strong.

Propositions. Let a bifibration E and a square S (as above) be given.

11. E satisfies the interpolation condition over S <u>iff</u> the canonical vertical arrow $\rho=\rho_A : m_!u^*A \longrightarrow v^*k_!A$ (i.e. such that $\vartheta^v_{k_!A}\circ\rho\circ\sigma^v_{u^*A} = \sigma^k_A\circ\vartheta^u_A$) is a split mono for every $A\in|\mathcal{E}_J|$.

12. E satisfies the strong interpolation condition over S <u>iff</u> it satisfies the Beck-Chevalley condition over S.

Proofs.

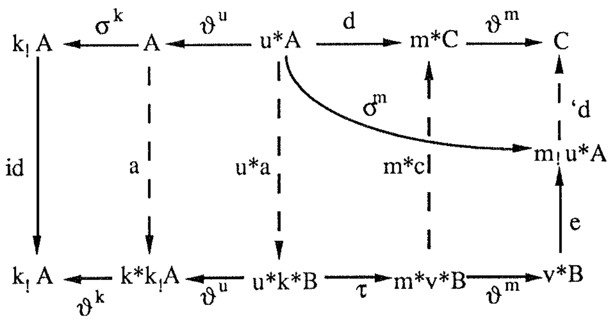
111. <u>If</u>: Suppose $e\circ\rho = id$. Define an interpolant:

$B := k_!A$;

$a := \eta : A \longrightarrow k^*k_!A$ is the unique vertical arrow by which $\sigma^k : A \longrightarrow k_!A$ factorizes through $\vartheta^k : k^*k_!A \longrightarrow k_!A$;

$c := {}'d\circ e : v^*B \longrightarrow C$, where ${}'d : m_!u^*A \longrightarrow C$ is the unique vertical arrow by which $\vartheta^m\circ d : u^*A \longrightarrow m^*C \longrightarrow C$ factorizes through $\sigma^m : u^*A \longrightarrow m_!u^*A$.

(When E is cloven, a is a component of the unit of $k_! \dashv k^*$, while ${}'d$ is the left transpose of $d : u^*A \longrightarrow m^*C$. But even without a cleavage, all the usual adjunction tricks go through: cf. lemmas 3 below.)



To prove that $d = m^*(c)\circ\tau\circ u^*(a)$, it is sufficient to show $\vartheta^m_C\circ d = \vartheta^m_C\circ m^*(c)\circ\tau\circ u^*(a)$. But $\vartheta^m_C\circ d = \text{‘}d\circ\sigma^m_{u^*A}$ (by the definition of ‘d), and $\vartheta^m_C\circ m^*(c) = \text{‘}d\circ e\circ\vartheta^m_{v^*B}$ (by the definitions of c and m*), so that it is enough to prove

$$\sigma^m_{u^*A} = e\circ\vartheta^m_{v^*B}\circ\tau\circ u^*(a).$$

But this is follows from $e\circ\rho = id$ and lemma 31 below.

112. Then: Let $\langle a_\eta, B_\eta, c_\eta\rangle$ be an interpolant of the "unit"

$$\eta : u^*A \longrightarrow m^*m_!u^*A,$$

defined as above, and let $\text{‘}a_\eta : k_!A \longrightarrow B$ be the "left transpose" of a_η. The arrow

$$e := c_\eta\circ v^*(\text{‘}a_\eta) : v^*k_!A \longrightarrow v^*B \longrightarrow m_!u^*A$$

is then a left inverse of ρ by lemma 33.

121. If: Suppose that every ρ is an iso. As we saw in 111,

$$\langle \eta{:}A \longrightarrow k^*k_!A,\ k_!A,\ \text{‘}d\circ\rho^{-1}{:}v^*k_!A \longrightarrow C\rangle$$

is an interpolant of d. It is initial: if ⟨a, B, c⟩ is another interpolant of d, then a factorizes through η by $k^*(\text{‘}a)$. If $\langle a_0, B_0, c_0\rangle$ is another initial interpolant of d, then $a_0 \simeq \eta$ (by the uniqueness of their factorisations through each other). So it is sufficient to prove that $\langle \eta, k_!A, \text{‘}d\circ\rho^{-1}\rangle$ is a strong interpolant. But this follows from the ⇐-part of lemma 34, and the assumption that ρ is an epi (indeed, an iso).

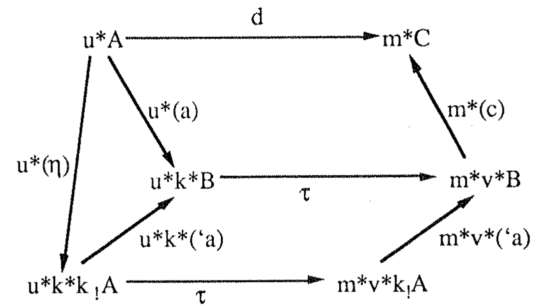
122. Then: We proved in 112 that every ρ is a split mono if there is an interpolant for each $d\in \mathcal{E}_I(u^*A, m^*C)$. On the other hand, every triple

$$\langle \eta{:}A \longrightarrow k^*k_!A,\ k_!A,\ c{:}v^*k_!A \longrightarrow C\rangle$$

is an initial interpolant. By assumption, it must be strong. From the ⇒-part of lemma 34 it follows now that ρ is an epi.

Since ρ is an epi and a split mono, it must be iso.•

Remark. In a bifibration, from any interpolant ⟨a, B, c⟩ of $d\in\mathcal{E}_I(u^*A, m^*C)$ an initial interpolant of the same arrow can be obtained, namely $\langle\eta, k_!A, c\circ v^*(\text{‘}a)\rangle$.



2. Uniform interpolation.

In logic with quantifiers, the interpolation condition can be expressed in the following way:

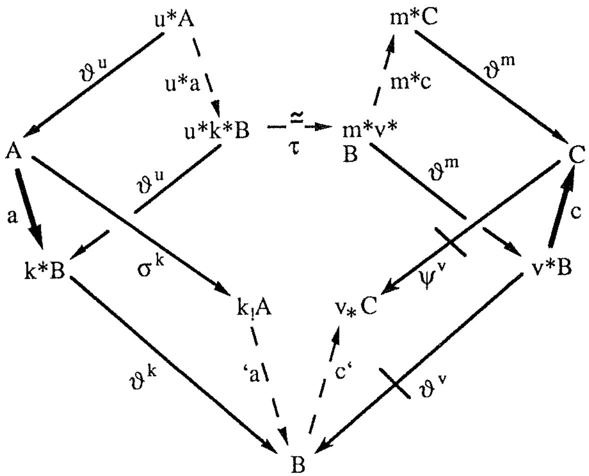
$$\alpha(x,y) \vdash \gamma(y,z) \Leftrightarrow$$

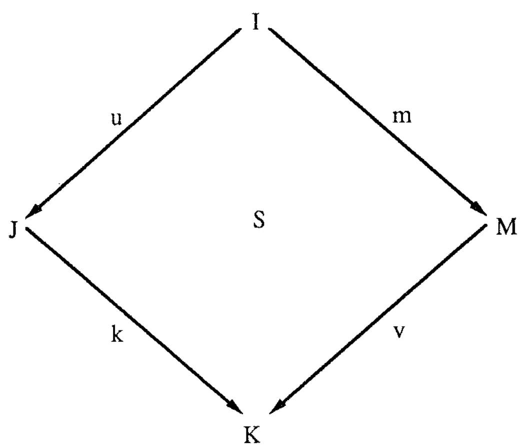
there is an interpolant β(y), such that $\exists x.\alpha(x,y) \vdash \beta(y) \vdash \forall z.\gamma(y,z)$.

An initial interpolant $\exists x.\alpha(x,y)$ is now given, as well as a terminal one, $\forall z.\gamma(y,z)$.

Definitions. Let E be a trifibration, the other data as above. We say that the interpolants of d are *uniform* if for each two of them, say ⟨a, B, c⟩ and $\langle\tilde{a}, \tilde{B}, \tilde{c}\rangle$, holds

$$c\text{‘}\circ\text{‘}a = \tilde{c}\text{‘}\circ\text{‘}\tilde{a}.$$

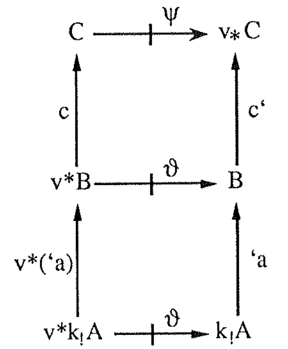




E satisfies the *uniform interpolation condition* over S if it satisfies the interpolation condition and the interpolants are uniform for every d.

Proposition. A trifibration E satisfies the uniform interpolation condition over S iff it satisfies the Beck condition over this square.

• If: The following diagram in $\mathcal{E}^{op}$



shows that in every trifibration

$$c\circ v^*('a) = \tilde{c}\circ v^*('\tilde{a}) \Leftrightarrow c'\circ 'a = \tilde{c}'\circ '\tilde{a}.$$

But lemma 33 tells that

$$c\circ v^*('a) = \tilde{c}\circ v^*('\tilde{a}) \Leftrightarrow m^*(c)\circ\tau\circ u^*(a) = m^*(\tilde{c})\circ\tau\circ u^*(\tilde{a})$$

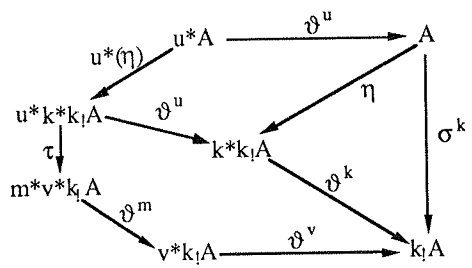
holds whenever ρ is epi. Hence, when the Beck-Chevalley condition is satisfied, i.e. when ρ is an iso, the interpolants are uniform.

Then: Suppose $c\circ\rho = \tilde{c}\circ\rho$ for some c and $\tilde{c}$. By lemma 34, $\langle\eta, k_!A, c\rangle$ and $\langle\eta, k_!A, \tilde{c}\rangle$ are interpolants of the arrow $m^*(c)\circ\rho' = m^*(\tilde{c})\circ\rho'$. Since $'\eta = id_{k_!A}$, from the uniformity follows $c' = \tilde{c}'$. Since c and $\tilde{c}$ were arbitrary, this means that ρ is an epi. But ρ is certainly a split mono by proposition 1. Hence ρ must be an iso, and the Beck-Chevalley condition is satisfied over S.•

3. Lemmas. The following statements are true for any bifibration E.

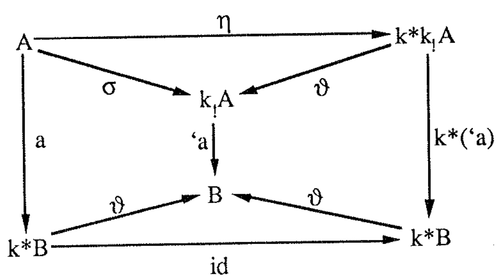
31. $\rho\circ\sigma^m = \vartheta^m\circ\tau\circ u^*(\eta)$.

• By the definition of ρ, the left side is the unique factorisation over m of $\sigma^k_A \circ \vartheta^u_A$ through $\vartheta^v_{k!A}$. But the right side is such a factorisation too, as the following diagram shows.

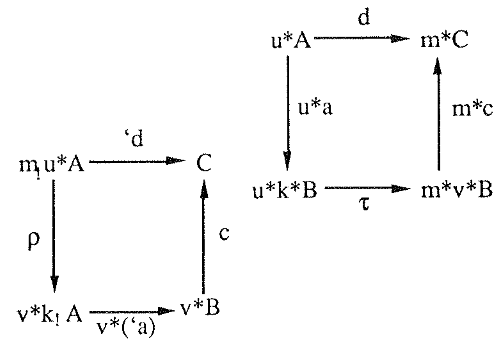


32. a = k*(‘a)∘η.

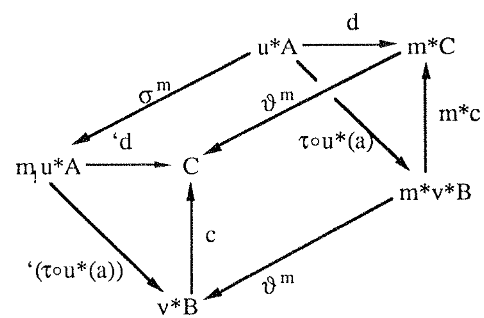
•



33. Each of two squares below commutes <u>iff</u> the other one does.

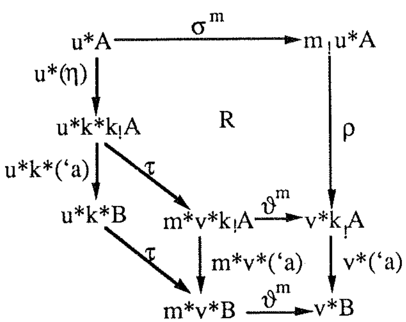


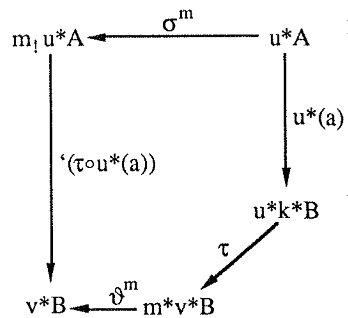
• Clearly, each of the triangles in the following diagram commutes <u>iff</u> the other one does.



Thus we are done if we prove v*(‘a)∘ρ = ‘(τ∘u*(a))

As for this equality, compare the following diagrams:

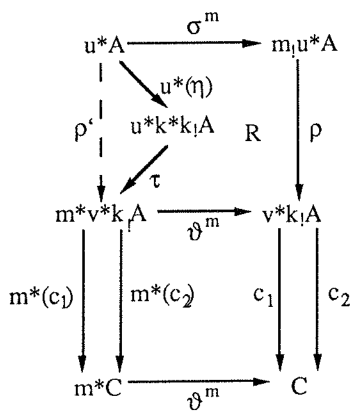




The pentangle R commutes by lemma 31, the rest by definitions. From lemma 32, it follows that $u^*(a) = u^*k^*(‘a)\circ u^*(\eta)$. Hence $‘(\tau\circ u^*(a))$ and $v^*(‘a)\circ\rho$ are vertical factorizations of the same arrow $\vartheta^m\circ\tau\circ u^*(a) = \vartheta^m\circ\tau\circ u^*k^*(‘a)\circ u^*(\eta)$ through σ^m. By the uniqueness, they must be equal.•

34. $c_1\circ\rho = c_2\circ\rho \Leftrightarrow m^*(c_1)\circ\tau\circ u^*(\eta) = m^*(c_2)\circ\tau\circ u^*(\eta)$.

•



Again, R commutes by lemma 31. Since the upper square commutes by the definition of ρ‘, we have $\tau\circ u^*(\eta) = \rho‘$. But

$$c_1\circ\rho = c_2\circ\rho \Leftrightarrow m^*(c_1)\circ\rho‘ = m^*(c_2)\circ\rho‘$$

holds by "adjunction": $c_1\circ\rho$ is the unique vertical factorisation of $\vartheta^m\circ m^*(c_1)\circ\rho‘$ through σ^m; $m^*(c_1)\circ\rho‘$ is the factorisation of $c_1\circ\rho\circ\sigma^m$ through ϑ^m.•

4. Families of arrows

1. Examples.

The simplest fibred category of arrows is $\nabla\underline{Set}$ = Cod: $\underline{Set}/\underline{Set} \to \underline{Set}$. It is a hyperfibration. If we use the correspondence $\int_I$ (from 1.1) to represent the fibre $(\underline{Set}/\underline{Set})_I = \underline{Set}/I$ as the category $\underline{Set}^I$ of I-indexed sets (i.e. if we regard $\nabla\underline{Set}$ as a family fibration), the horizontal structure over a function $u \in \underline{Set}(I,J)$ will be

$$u^*: \underline{Set}^J \to \underline{Set}^I : \{\varphi_y \mid y \in J\} \mapsto \{\varphi_{u(x)} \mid x \in I\}$$

$$u_! : \underline{Set}^I \to \underline{Set}^J : \{\gamma_x \mid x \in I\} \mapsto \left\{ \sum_{x \in u^{-1}(y)} \gamma_x \mid y \in J \right\}$$

$$u_* : \underline{Set}^I \to \underline{Set}^J : \{\gamma_x \mid x \in I\} \mapsto \left\{ \prod_{x \in u^{-1}(y)} \gamma_x \mid y \in J \right\}.$$

The subfibration $\underline{Mon} \subseteq \nabla\underline{Set}$ spanned by the indexed sets $\{\gamma_x\}$ in which every γ_x has at most one element - is a hyperfibration too. The subfibration $\underline{Epi} \subseteq \nabla\underline{Set}$, consisting of $\{\gamma_x\}$ where every γ_x has at least one element - has small coproducts; it has small products iff the axiom of choice is true. Without the axiom of choice, Epi has only the products along the elements of

$$\text{Fset} := \{u \in \underline{Set} \mid \forall y \in \text{Cod}(u).\ u^{-1}(y) \text{ is finite}\}.$$

This family of arrows also spans a full subcategory $\text{Fset}/\underline{Set} \subseteq \underline{Set}/\underline{Set}$, which is stable under all pullbacks, which means that $\nabla\text{Fset} := \nabla\underline{Set}\restriction_{\text{Fset}/\underline{Set}}$ is a subfibration. This fibration has the products and coproducts only along the elements of the family Fset. (Its elements can be viewed as sets $\{\gamma_x\}$ where every γ_x is finite.)

We shall say that categories of arrows Epi and Fset/Set, fibred over Set, are *relative* hyperfibrations with respect to the family of arrows $\text{Fset} \subseteq \underline{Set}$, or that they are Fset-hyperfibrations.

$\underline{Dfib} \to \underline{Cat}$ and $\underline{Fib} \to \underline{Cat}$ are also fibred categories of arrows, and subfibrations of a basic fibration, namely $\nabla\underline{Cat}$. They are both trifibrations, with the direct images given by the Kan extensions; neither of them has the Beck-Chevalley property, since they contain the counterexample 3.3. Dfib is the full subcategory Dfib/Cat spanned in Cat/Cat by the class of arrows Dfib, while Fib is not full in Cat/Cat.

∇Cat itself is, of course, a bifibration. It is, moreover, a Fib-trifibration: $\nabla\mathrm{Cat}^{op}$ has the direct images along fibrations. (But the construction seems quite complicated.)

2. Relativisation.

Notation. Let $\mathcal{A} \subseteq \mathcal{B}$ be a class of arrows, and $C:\mathcal{C} \to \mathcal{B}$ a functor. The category $\mathcal{A}/C$ is then defined to be the full subcategory of the comma category $\mathcal{B}/C$, spanned by the objects $\langle a:I \to CX, X\rangle$, where $a \in \mathcal{A}$. The definition of the category $C/\mathcal{A}$ is analogous.

We shall mostly consider $\mathcal{A}/\mathcal{B} := \mathcal{A}/id_{\mathcal{B}}$, with the projection

$$\nabla\mathcal{A} : \mathcal{A}/\mathcal{B} \to \mathcal{B} : \langle a:I \to J, J\rangle \mapsto J.$$

The fibres will be $\mathcal{A}{\downarrow}J := (\mathcal{A}/\mathcal{B})_J$. (When $\mathcal{A}$ is a category, there is also $\nabla\mathcal{A}:\mathcal{A}/\mathcal{A} \to \mathcal{A}$, with fibres $\mathcal{A}/J := (\mathcal{A}/\mathcal{A})_J$.)

Definitions. A functor $E : \mathcal{E} \to \mathcal{B}$ is a fibration *relative* to a family $\mathcal{A}$ of arrows, or an $\mathcal{A}$*-fibration*, if every $q \in \mathcal{B}(I,J) \cap \mathcal{A}$ has at every $Z \in |\mathcal{E}_J|$ a *cartesian* lifting ϑ^u such that for every $v \in \mathcal{B}(K,I)$ and every f over qv there is a unique g over v, $f = \vartheta^u \circ g$. (Cf. 1.31(b).)

E is an $\mathcal{A}$*-bifibration* if it is a fibration and a relative $\mathcal{A}$-cofibration (i.e. E^o is an $\mathcal{A}$-fibration). An $\mathcal{A}$-bifibration is said to have the $\mathcal{A}$*-coproducts* if it satisfies the Beck-Chevalley condition over the pullback squares with two opposite sides belonging to $\mathcal{A}$.

E is an $\mathcal{A}$-trifibration if E and E^{op} are $\mathcal{A}$-bifibrations. It is an $\mathcal{A}$-hyperfibration if it has $\mathcal{A}$-products and $\mathcal{A}$-coproducts.

Remark. The relativized versions of many propositions from the preceding sections are easily obtained - provided that $\mathcal{A}$ satisfies appropriate closure conditions. For instance:

1) 1.31(f). (AC) E is an $\mathcal{A}$-fibration iff the functor $\mathcal{A}^*(E_{\mathcal{E}})$, obtained by pulling back $E_{\mathcal{E}}$ along $\mathcal{A}/E \hookrightarrow \mathcal{B}/E$, has a right adjoint right inverse.

2) 3.4. Pull back along $\mathcal{A}/\mathcal{B} \hookrightarrow \mathcal{B}/\mathcal{B}$ the whole diagram used in this proposition; keep the same names. $\mathcal{B}^*$ becomes the category of $\mathcal{A}$-arrows with pullbacks as arrows between them. The proposition now reads: An $\mathcal{A}$-bifibration E has $\mathcal{A}$-coproducts iff $E_!$ is a fibration and U is a cartesian functor.

3) 3.6. (Bénabou 1975b) Let $\mathcal{A}$ be a calibration. (The definition follows below.) Given a fibration $E:\mathcal{E} \to \mathcal{B}$, the fibration $Fam_{\mathcal{A}}(E) : E/\mathcal{A} \to \mathcal{B}$ and the cartesian functor $\Delta_{\mathcal{A}} : E \to Fam_{\mathcal{A}}(E)$ are defined as in proposition 3.6, *but* with $\mathcal{A}/\mathcal{B}$ in place of $\mathcal{B}/\mathcal{B}$. E has $\mathcal{A}$-coproducts iff $\Delta_{\mathcal{A}}$ has a cartesian left adjoint.

3. Intrinsic structures and closure conditions for families.

Definitions. A *calibration* (Bénabou 1975a) on a category $\mathcal{B}$ is a family of arrows $\mathcal{A} \subseteq \mathcal{B}$, which satisfies the following conditions:

- C0) $\mathcal{A}$ is stable under the composition with all isomorphisms (i.e. $f \in \mathcal{A}$ implies $i \circ f \circ j \in \mathcal{A}$ for all appropriately composable isos $i,j \in \mathcal{B}$)
- C1) $\mathcal{A}$ contains all the isomorphisms of $\mathcal{B}$;
- C2) the pullbacks of $\mathcal{A}$-arrows along arbitrary arrows exist, and $\mathcal{A}$ is stable;
- C3) $\mathcal{A}$ is closed under composition, and
- C4) under left division.

A *saturated family* of arrows $\mathcal{A}$ satisfies satisfies (C0) and (C1). A *stable family* satisfies (C1) and (C2) (and (C0) a fortiori); a *saturated subcategory* satisfies (C0), (C1) and (C3).

If $\mathcal{B}$ has a terminal object $\top$, we say that $\mathfrak{a}$ is a *display family* if it satisfies the *display condition* (Taylor 1986):

D) $\mathfrak{a}$ contains $K \to \top$ for every $K \in |\mathcal{B}|$.

Terminology. The arrow fibrations, cofibrations, etc. are generally considered in the form of codomain functors. A *codomain functor* $A: \mathcal{A} \to \mathcal{B}$ is a restriction of the basic functor

$$\nabla\mathcal{B} = \mathrm{Cod}: \mathcal{B}/\mathcal{B} \to \mathcal{B}.$$

to a subcategory $\mathcal{A} \hookrightarrow \mathcal{B}/\mathcal{B}$. Thus, "arrow fibration A" always means that $A = \nabla\mathcal{B}\restriction_{\mathcal{A}}$.

A (fibrewise or horizontal) structure in $\mathcal{A}$ is called *intrinsic* if it is preserved by the inclusion in $\mathcal{B}/\mathcal{B}$. E.g., the intrinsic inverse images are given by pullbacks, the intrinsic direct images by composition (i.e. $u_!(a) = u \circ a$), the intrinsic (fibrewise) terminal objects are identities. $\nabla\mathcal{B}$ itself is always a split cofibration with terminal objects. If it is a fibration, it always has the small coproducts (i.e. the Beck-Chevalley property). But even if $\nabla\mathcal{B}$ is not a fibration, we consider its subfibrations, just as we consider subgroups of a monoid. An intrinsic (co-, bi-,...)fibration is a sub(co-,...)fibration of $\nabla\mathcal{B}$. Intrinsic structure is a partial structure in $\nabla\mathcal{B}$.

Facts. For a class of arrows $\mathfrak{a} \subseteq \mathcal{B}$, consider the functor

$$A := \nabla\mathfrak{a} : \mathfrak{a}/\mathcal{B} \to \mathcal{B},$$

the restriction of $\nabla\mathcal{B}$ on the full subcategory $\mathfrak{a}/\mathcal{B} \subseteq \mathcal{B}/\mathcal{B}$.

31. $\mathfrak{a}$ satisfies (C1) iff A has intrinsic terminal objects.

32. $\mathfrak{a}$ satisfies (C2) (and (C0)) iff A is an intrinsic fibration.

33. $\mathfrak{a}$ satisfies (C3) iff A is a split intrinsic $\mathfrak{a}$-cofibration. A is a split intrinsic cofibration iff $\mathfrak{a}$ is a left ideal in $\mathcal{B}$ (i.e. for every appropriately composable $u \in \mathcal{B}$, $f \in \mathfrak{a}$ holds $u \circ f \in \mathfrak{a}$).

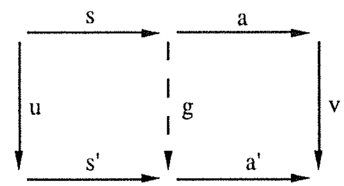
34. $\mathfrak{a}$ is a stable subcategory iff A has terminal objects, inverse images and $\mathfrak{a}$-coproducts, everything intrinsic. Iff, besides, $\mathfrak{a}/\mathcal{B} = \mathfrak{a}/\mathfrak{a}$, then $\mathfrak{a}$ is a calibration. The only calibration satisfying the display condition is $\mathfrak{a} = \mathcal{B}$.

4. Arrow cofibrations and factorisations.

41. Definition. (Relativisation of Freyd-Kelly 1972, 2.2.) Let $\mathfrak{d} \subseteq \mathcal{B}$ be a saturated family. A *factorisation system relative* to $\mathfrak{d}$, or a $\mathfrak{d}$*-factorisation system* is a pair $\langle \mathfrak{e}, \mathfrak{m} \rangle$ of saturated families $\mathfrak{e}, \mathfrak{m} \subseteq \mathfrak{d}$ such that the following conditions are satisfied:

F1) for every $u \in \mathfrak{d}$ there are $a \in \mathfrak{m}$, $s \in \mathfrak{e}$, such that $u = a \circ s$;

F2) for all $a, a' \in \mathfrak{m}$, $s, s' \in \mathfrak{e}$ (and u,v arbitrary), $v \circ a \circ s = a' \circ s' \circ u$ implies that there is a unique arrow g which makes the following diagram commutative:



The elements of the class $\mathfrak{e}$ are called *epis of the factorisation*, the elements of $\mathfrak{m}$ are its *monos*.

A relative factorisation system is *stable* if class $\mathfrak{e}$ is stable under pullbacks.

Lemmas. Let $\langle \mathfrak{e}, \mathfrak{m} \rangle$ be a $\mathfrak{d}$-factorisation system.

42. The assertions $(x_{\mathfrak{j}})$ are all equivalent, for $x \in \{a,b,c\}$, $\mathfrak{j} \in \{\mathfrak{e}, \mathfrak{m}\}$.

$a_{\mathfrak{j}}$) $\mathfrak{j}$ is closed in $\mathfrak{d}$ under composition: $p, q \in \mathfrak{j}$ and $pq \in \mathfrak{d}$ implies $pq \in \mathfrak{j}$.

$b_{\mathfrak{j}}$) $\mathfrak{j} := \left\{ \begin{matrix} \mathfrak{e} \\ \mathfrak{m} \end{matrix} \right\}$ is closed in $\mathfrak{d}$ under the $\left\{ \begin{matrix} \text{right} \\ \text{left} \end{matrix} \right\}$ division: p', p'', $\left\{ \begin{matrix} qp' \\ p''q \end{matrix} \right\} \in \mathfrak{j}$ and $q \in \mathfrak{d}$ implies $q \in \mathfrak{j}$.

$c_{\mathfrak{j}}$) If the condition (F2) is satisfied for some $f \in \mathfrak{d}$ in place of $\left\{ \begin{matrix} s \in \mathfrak{e} \\ a' \in \mathfrak{m} \end{matrix} \right\}$, then $f \in \mathfrak{j} := \left\{ \begin{matrix} \mathfrak{e} \\ \mathfrak{m} \end{matrix} \right\}$.

If any of the above conditions is satisfied, the closure properties of d (in $\mathcal{B}$) are reflected on e and m:

i) If d is closed under composition, <u>then</u> both e and m are.

ii) If d is closed under $\left\{\begin{matrix}\text{right}\\ \text{left}\end{matrix}\right\}$ division, <u>then</u> $\left\{\begin{matrix}\mathsf{e}\\ \mathsf{m}\end{matrix}\right\}$ is closed too.

• $(a_{\mathsf{m}})\Rightarrow(b_{\mathsf{m}})$: Suppose that $\mathring{a}\circ f\in\mathsf{m}$ for some $\mathring{a}\in\mathsf{m}$ and $f\in\mathsf{d}$. Let $f=a_f\circ s_f$ be a factorisation of f given by (F1). Apply (F2) on $a:=a_f$, $s:=s_f$, $a':=\mathring{a}\circ f$, $s':=id$, $u:=id$, $v:=\mathring{a}$. So there is g such that $g\circ s_f=id$, $\mathring{a}\circ f\circ g=\mathring{a}\circ a_f$. Since $a\circ a_f\in\mathsf{m}$, from $\mathring{a}\circ a_f\circ s_f\circ g=\mathring{a}\circ a_f$ follows $s_f\circ g=id$, by the uniqueness of the factorisation. Since s_f is an iso, and m is saturated, $f\in\mathsf{m}$.

$(b_{\mathsf{m}})\Rightarrow(c_{\mathsf{m}})$: Putting in (F2) $a:=a_f$, $s:=s_f$, $a':=f$, and the remaining arrows identities, we get g such that $g\circ s_f=id$ and $a_f\circ s_f\circ g=a_f$. By the left division, the second equation gives $s_f\circ g=id$. So s_f is an isomorphism again.

$(c_{\mathsf{m}})\Rightarrow(a_{\mathsf{m}})$: Suppose that for $a_0,a_1,a_2\in\mathsf{m}$, $s_0,s_1\in\mathsf{e}$ and arbitrary, u',v' holds $v'\circ a_0\circ s_0=a_2\circ a_1\circ s_1\circ u'$. First apply (F2) to $a:=a_0$, $s:=s_0$, $a':=a_2$, $s':=id$, $u:=a_1\circ s_1\circ u'$, $v:=v'$, to get g' such that $a_2\circ g'=v'\circ a_0$ and $g'\circ s_0=a_1\circ s_1\circ u'$; then set $a:=id$, $s:=s_0$, $a':=a_1$, $s':=s_1$, $u:=u'$, $v:=g'$ to get g such that $g'=a_1\circ g$ and $g\circ s_0=s_1\circ u'$. This g satisfies the condition (F2) for $a:=a_0$, $s:=s_0$, $a':=a_2\circ a_1$, $s':=s_1$, $u:=u'$, $v:=v'$. By (c_{m}), $a_2\circ a_1\in\mathsf{m}$.

$(a_{\mathsf{e}})\Rightarrow(c_{\mathsf{m}})$: (Freyd-Kelly 1972, proposition 2.2.1.) Consider again this arrow g for which $g\circ s_f=id$ and $f\circ g=a_f$. Since $(a_{\mathsf{e}})\Rightarrow(b_{\mathsf{e}})$, the first equation implies $g\in\mathsf{e}$. But then $s_f\circ g\in\mathsf{e}$, and from $f\circ g=a_f\circ s_f\circ g=a_f$ (by the uniqueness of the factorisation) follows $s_f\circ g=id$. So g is an iso.•

43. If d is stable under $\left\{\begin{matrix}\text{pushouts}\\ \text{pullbacks}\end{matrix}\right\}$ in $\mathcal{B}$, <u>then</u> $\left\{\begin{matrix}\mathsf{e}\\ \mathsf{m}\end{matrix}\right\}$ is stable too.

• Let S be a pullback of $a'\in\mathsf{m}\subseteq\mathsf{d}$ along v. Denote by u the arrow opposite to v, and let f be opposite to a'. Since $f\in\mathsf{d}$, there is a factorisation $a\circ s=f$. Condition (F2) (with $s':=id$) now gives an arrow g so that $a'\circ g=v\circ a$ and $u=g\circ s$. Since S is a pullback, the

first equation induces a unique arrow r such that $a=f\circ r=a\circ s\circ r$, $g=u\circ r$. $r\circ s=id$ is the unique factorisation of S through itself, while $s\circ r=id$ follows by another application of condition (F2).•

44. Proposition. Let d be a stable family in $\mathcal{B}$, $\mathsf{a}\subseteq\mathsf{d}$, and let d be closed under the left division by a (i.e. $aq\in\mathsf{d}$ <u>and</u> $a\in\mathsf{a}$ <u>implies</u> $q\in\mathsf{d}$). Then the statements below are related as follows

(c) $\Rightarrow$ (a) $\Leftrightarrow$ (b) $\overset{AC}{\Rightarrow}$ (c).

All these implications remain valid when all the outlined parts are omitted.

a) $\nabla\mathsf{a}:\mathsf{a}/\mathcal{B}\longrightarrow\mathcal{B}$ is a d-bifibration with intrinsic inverse images, intrinsic terminal objects, and the BC-property.

b) a is a family of monos of a stable d-factorisation system.

c) a is a stable family <u>and</u> the inclusion $\mathsf{a}/\mathcal{B}\hookrightarrow\mathsf{d}/\mathcal{B}$ has a cartesian left adjoint.

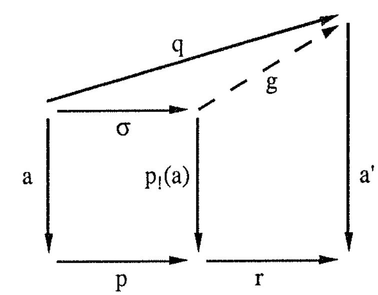
If any of these conditions is fulfilled, the following is true:

i) a is closed under the composition in d and under left division in $\mathcal{B}$. (I.e., (a_{a}), (b_{a}) and (c_{a}) from 42 are all true).

ii) If d is a stable subcategory in $\mathcal{B}$, a is a calibration.

• (a)$\Rightarrow$(b): The statement "$\langle\sigma,p\rangle\in\mathsf{a}/\mathcal{B}(a,p_!(a))$ is $\nabla\mathsf{a}$-cocartesian" <u>means</u>:

"$\forall a'\in\mathsf{a}\ \forall q,r\in\mathcal{B}$. $a'q=rpa$ <u>then</u> $\exists!g$ such that the diagram



commutes".

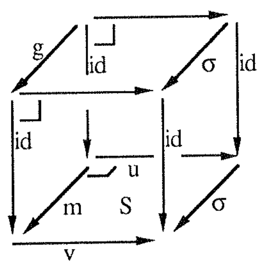
(Note that $p_!(a) \simeq (pa)_!(id)$, and $\sigma^p_a \simeq \sigma^{pa}_{id}$.)

The pair $\langle \boldsymbol{\sigma}, \mathcal{A} \rangle$ is a $\mathcal{D}$-factorisation system, where

$$\boldsymbol{\sigma} := \{\sigma \in \mathcal{B} \mid \langle \sigma, p\rangle \text{ is } \nabla\mathcal{A}\text{-cocartesian}, p \in \mathcal{D}\}.$$

(F1) follows from $p = p_!(id) \circ \sigma^p_{id}$. Since $\mathcal{D}$ is closed under the left division by $\mathcal{A}$, this implies $\boldsymbol{\sigma} \subseteq \mathcal{D}$. (F2) is obtained if we take $q := s'u$, $r := v$.

It remains to derive the stability of $\boldsymbol{\sigma}$ under pullbacks from the BC-property of $\mathcal{A}/\mathcal{B} \longrightarrow \mathcal{B}$. So consider a pullback square S with $\sigma \in \boldsymbol{\sigma}$.



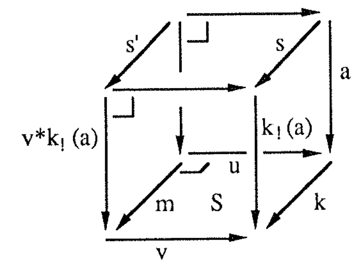
Clearly, the right hand square is a cocartesian lifting of σ, while the back and the front squares are cartesian liftings of u and v. By the BC-property, the left hand square must be cocartesian, i.e. $g \in \boldsymbol{\sigma}$. But m=g.

(b)⇒(a): Suppose that $\langle \mathcal{E}, \mathcal{A} \rangle$ is a $\mathcal{D}$-factorisation system. By (F1) there are (some arrows, which we suggestively denote) $p_!(a) \in \mathcal{A}$ and $\sigma \in \mathcal{E}$ such that

$$p \circ a = p_!(a) \circ \sigma$$

With $v := r$, $u := q$, $s' := id$, the condition (F2) tells that $\langle \sigma, p\rangle$ is a cocartesian lifting of p at a, and that $p_!(a)$ is a direct image.

We further prove that the stability of the epis of the factorisation implies the BC-property of $\mathcal{A}/\mathcal{B} \longrightarrow \mathcal{B}$. So consider a pullback square S, with a cocartesian lifting (=factorisation) of k, and with cartesian liftings of u and v.



Since the back, front and bottom squares on this diagram are pullbacks, the top square must be one. By the stability, s' must then be an epi of the factorisation. Since $v^*k_!(a) \in \mathcal{A}$, the left hand square is cocartesian.

(c)⇒(b): The pair $\langle \eta, \mathcal{A} \rangle$ is a $\mathcal{D}$-factorisation system, where η is the (family of the components of the) unit of the adjunction.

(a)⇒(c): $H \dashv I : \mathcal{A}/\mathcal{B} \hookrightarrow \mathcal{B}/\mathcal{B}$ is defined by

$$H(p) := p_!(id).$$

((a)⇔(c) can also be obtained from proposition 3.6.)

i) Note that for the $\mathcal{D}$-factorisation system $\langle \boldsymbol{\sigma}, \mathcal{A} \rangle$ the condition (c_σ) from lemma 42 is satisfied. The lemma now tells that $(a_\mathcal{A})$, $(b_\mathcal{A})$ and $(c_\mathcal{A})$ are true. For the left division, $aq, a \in \mathcal{A}$ implies $q \in \mathcal{D}$ by the assumption about $\mathcal{D}$; but now $(b_\mathcal{A})$ says that from $aq, a \in \mathcal{A}$ and $q \in \mathcal{D}$ follows $q \in \mathcal{A}$.

ii) If $\mathcal{D}$ is closed under the composition in $\mathcal{B}$, then $\mathcal{A}$ is closed too, by (i) and 42(i).•

45. Remark. When $\mathcal{B}$ has a terminal object 1, and $\mathcal{D} = \mathcal{B}$, the reflection in (c) restricts to a (full) reflection of $\mathcal{A}$ in $\mathcal{B}$. It is the result of Cassidy-Hébert-Kelly (1985) that there is a Galois connection between (full replete) reflective subcategories and $\mathcal{B}$-factorisation systems: reflective subcategories of $\mathcal{B}$ correspond exactly to those factorisation systems in which epis are closed under the left division. Translating through the last proposition, we conclude:

- Reflective subcategories of $\mathcal{B}$ exactly correspond to the arrow bifibrations over $\mathcal{B}$ in which the class of cocartesian arrows is closed under the left division.

- A reflective subcategory of $\mathcal{B}$ is a localisation - i.e. its reflection preserves the finite limits - exactly when this corresponding arrow bifibration has the Beck-Chevalley property.

46. Corollary. A finitely complete category $\mathcal{B}$ is regular <u>iff</u> the fibration

$$\nabla\mathrm{Mon} : \mathrm{Mon}/\mathcal{B} \longrightarrow \mathcal{B},$$

(where $\mathrm{Mon} \subseteq \mathcal{B}$ is the family of the monomorphisms) has the small coproducts. (*Regular* categories are those in which every arrow factors as a mono followed by a regular epi - i.e. a coequalizer of some pair - and regular epis are stable under pullbacks.)

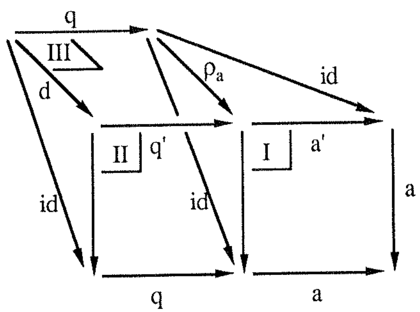
• Street 1984, theorem 3 (attributed to Joyal): The epis of a factorisation system $\langle \mathcal{E}, \mathrm{Mon}\rangle$ in a finitely complete category $\mathcal{B}$ are coequalisers of their kernel pairs.•

47. Lemma. (Probably Bénabou.) Let $\mathfrak{d}$ be a stable subcategory, and $\mathfrak{a} \subseteq \mathfrak{d}$. Then

$$(aq \in \mathfrak{d} \text{ and } a \in \mathfrak{a} \text{ implies } q \in \mathfrak{d}) \Leftrightarrow \forall a \in \mathfrak{a}.\ \rho_a \in \mathfrak{d},$$

where $\rho_a := \langle id, id\rangle \in \mathfrak{d}\!\downarrow\! J(a, a \times_J a)$ is the diagonal arrow in the fibrewise cartesian structure.

• $\Leftarrow$: Suppose $aq \in \mathfrak{d}$ <u>and</u> $a \in \mathfrak{a}$.

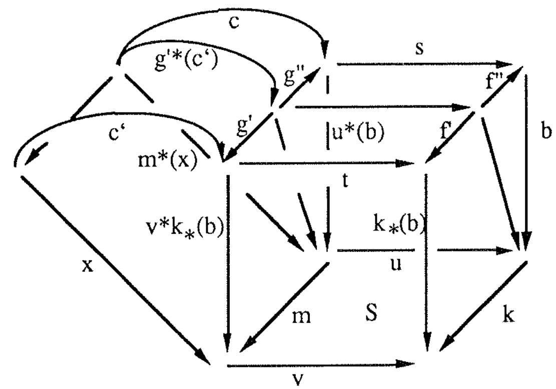


I+II pullback, $aq \in \mathfrak{d} \Rightarrow a'q' \in \mathfrak{d}$.
II, II+III pullback <u>then</u> III pullback.
III pullback, $\rho_a \in \mathfrak{d}$ <u>then</u> $d \in \mathfrak{d}$.
$a'q'$, $d \in \mathfrak{d}$ <u>then</u> $q = a'q'd \in \mathfrak{d}$.•

5. Right arrow bifibrations.

Lemma. Every (relative) right bifibration with intrinsic inverse images has the Beck-Chevalley property.

• Consider a pullback square S in $\mathcal{B}$, cartesian liftings $\langle s,u\rangle$, $\langle t,v\rangle$, and opcartesian $\langle f'/f'', k\rangle$. Op-arrow $\langle g'/g'', m\rangle$ is induced as a unique factorisation of $\langle f'/f'', k\rangle \circ \langle s,u\rangle$ by $\langle t,v\rangle$: g' by pulling m back, and g'' as a factorisation through the pullback on the back side of the cube. On the following diagram everything commutes, and all the squares are pullbacks.



We must show that $\langle g'/g'', m\rangle$ is opcartesian. (We use description II.3.2.) An arbitrary vertical arrow $c : m^*(x) \longrightarrow u^*(b)$ induces $sc : k^*(vx) \longrightarrow b$, since $k^*(vx) \simeq u \circ m^*(x)$. Since $\langle f'/f'', k\rangle$ is opcartesian, there is a unique arrow $d : vx \longrightarrow k_*(b)$ such that $f'' \circ f'^*(d) = sc$.

But the front square is cartesian, and there is unique arrow c' such that $d = tc'$ and $x = v^*k_*(b) \circ c'$. The equality $g'' \circ g'^*(c') = c$ is now obtained by chasing the diagram.•

Proposition. (Streicher 1988, chapter 1) In every right bifibration $\nabla\mathcal{A}:\mathcal{A}/\mathcal{B}\longrightarrow\mathcal{B}$ (relative to a stable family) the right direct images must be intrinsic <u>if</u> the inverse images and terminal objects are intrinsic.

• Let $k_*(b)$ be a right direct image with respect to $\mathcal{A}/\mathcal{B}$, and $v\in\mathcal{B}(M,K)$ an arbitrary arrow. Let S be a pullback as before, i.e. $u\in\mathcal{B}(I,J)$ is a pullback of v along k.

$$\mathcal{B}/K\big(v,\ k_*(b)\big)\simeq\mathcal{B}/M\big(id,\ v^*k_*(b)\big)=$$
$$\mathcal{A}{\downarrow}M\big(id,\ v^*k_*(b)\big)\simeq\mathcal{A}{\downarrow}M\big(id,\ m_*u^*(b)\big)\simeq\mathcal{A}{\downarrow}I(id,\ u^*(b)=$$
$$\mathcal{B}/I(id,\ u^*(b))\simeq\mathcal{B}/J(u,\ b)=\mathcal{B}/J(k^*(v),\ b)\bullet$$

Corollary. Let $\mathcal{D}$ be a stable subcategory, and $\mathcal{A}\subseteq\mathcal{D}$ a saturated family. If $\nabla\mathcal{A}:\mathcal{A}/\mathcal{B}\longrightarrow\mathcal{B}$ is a $\mathcal{D}$-trifibration, then it is a $\mathcal{D}$-hyperfibration; the right direct images over $\mathcal{D}$ are intrinsic. If $\mathcal{D}$ is closed under left division by $\mathcal{A}$ (e.g. if $\mathcal{D}=\mathcal{B}$), then the left direct images over $\mathcal{A}$ are intrinsic too; $\nabla\mathcal{A}:\mathcal{A}/\mathcal{A}\longrightarrow\mathcal{A}$ is then a subhyperfibration (i.e. a locally cartesian closed category: see below).

• The first sentence follows from the preceding proposition and fact 3.3. The second one from proposition 44.•

6. rccc

Definition. A category $\mathcal{B}$ is category:*locally cartesian closed* (or *lccc*) if $\nabla\mathcal{B}:\mathcal{B}/\mathcal{B}\longrightarrow\mathcal{B}$ is a fibrewise cartesian closed category (cf. 2.1). $\mathcal{B}$ is *relatively cartesian closed* with respect to a stable subcategory $\mathcal{A}\subseteq\mathcal{B}$ (or: $\mathcal{B}$ is an $\mathcal{A}$*-rccc*) if $\nabla\mathcal{A}:\mathcal{A}/\mathcal{B}\longrightarrow\mathcal{B}$ is a fibrewise cartesian closed category and each exponentiation functor $a\rightarrow\_:\mathcal{A}{\downarrow}J\longrightarrow\mathcal{A}{\downarrow}J$ preserves the $\mathcal{A}$-arrows (i.e. restricts to $a\rightarrow\_:\mathcal{A}/J\longrightarrow\mathcal{A}/J$).

References. The last notion has been introduced in Taylor 1986, IV.3, and in Hyland-Pitts 1987, definition 2.7. The original definitions are in terms of the characterisation which we formulate below. In fact, for type-theoretical reasons, the original definitions are slightly stronger: Hyland and Pitts require that $\mathcal{B}$ has finite

products, Taylor moreover that $\mathcal{A}$ satisfies the display condition, i.e. that it contains all the projections.

The following characterisation itself is proposition 2.6 in Hyland-Pitts 1987. In fact, it can be traced back to Freyd (1972, 1.34) and Day (1974, 4.1) (who both considered only locally cartesian closed categories - but the proof is essentially the same). We quote it for completeness.

Proposition. $\mathcal{B}$ is a relative cartesian closed category with respect to a stable subcategory $\mathcal{A}$ <u>iff</u> $\nabla\mathcal{A}:\mathcal{A}/\mathcal{B}\longrightarrow\mathcal{B}$ is an intrinsic $\mathcal{A}$-trifibration with intrinsic terminal objects. The cartesian closed structure is then intrinsic, and $\nabla\mathcal{A}$ is an $\mathcal{A}$-hyperfibration.

• <u>If</u>: It is immediate to check that for every $a\in|\mathcal{A}{\downarrow}J|$

$$a\times\_ := a_!\circ a^*(\_):\mathcal{A}{\downarrow}J\longrightarrow\mathcal{A}{\downarrow}J,\text{ and}$$
$$a\rightarrow\_ := a_*\circ a^*(\_):\mathcal{A}{\downarrow}J\longrightarrow\mathcal{A}{\downarrow}J$$

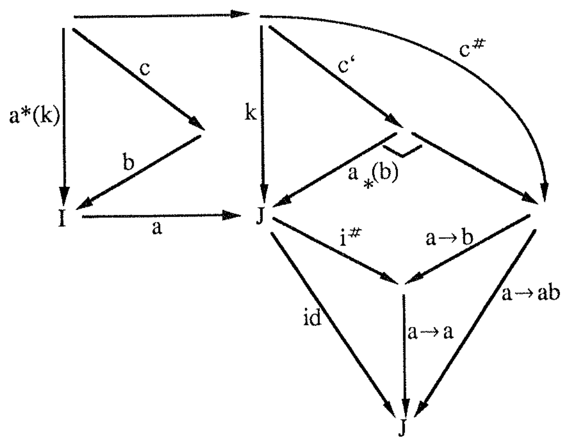
define the required vertical structure. It is preserved under the inverse images by the BC-property. The proof that a_* (and therefore $a\rightarrow\_$) preserves the $\mathcal{A}$-arrows can be found in Hyland-Pitts (1987, 2.5.)

<u>Then</u>: We show that a right direct image of $b\in|\mathcal{A}/I|$ along an $\mathcal{A}$-arrow $a:I\longrightarrow J$ is the following $\nabla\mathcal{A}$-inverse image (i.e. pullback):

$$a_*(b) := \big(i^\#\big)^*(a\rightarrow b),$$

where

$i^\#\in\mathcal{A}{\downarrow}J(id,a\rightarrow a)$ is the right transpose of $i:=id_I\in\mathcal{A}{\downarrow}J(a,a)$, and
$a\rightarrow b\in\mathcal{A}/J(a\rightarrow ab,a\rightarrow a)$ is the image of $b\in\mathcal{A}/J(ab,a)$ by $a\rightarrow\_:\mathcal{A}/J\longrightarrow\mathcal{A}/J$.



For arbitrary $k \in |\mathsf{a}{\downarrow}J|$, we calculate

$$\begin{aligned}\mathsf{a}{\downarrow}I(a^*(k),b) &\simeq \{c \in \mathsf{a}{\downarrow}J(a\times k, ab) \mid b\circ c = \pi : a\times k \longrightarrow a\} \simeq \\ &\simeq \{c^{\#} \in \mathsf{a}{\downarrow}J(k, a\to ab) \mid (a\to b)\circ c^{\#} = \pi^{\#} : k \longrightarrow (a\to a)\} \overset{\ddagger}{=} \\ &\simeq \{c^{\#} \in \mathsf{a}{\downarrow}J(k, a\to ab) \mid (a\to b)\circ c^{\#} = i^{\#}\circ k\} = \\ &= \mathsf{a}{\downarrow}i^{\#}(k, a\to b) \simeq \\ &\simeq \mathsf{a}{\downarrow}J\Big(k, \big(i^{\#}\big)^*(a\to b)\Big).\end{aligned}$$

The step (‡) follows from the fact that $k \in \mathsf{a}{\downarrow}J(k, id_J)$ is the terminal arrow, so that $\pi^{\#} = i^{\#}\circ k : k \longrightarrow id \longrightarrow (a\to a)$. $\mathsf{a}{\downarrow}i^{\#} := (\mathsf{a}/\mathcal{B})_{i^{\#}}$ denotes the set of arrows over $i^{\#}$.•

Remark. The functors $a\times\_ : \mathsf{a}{\downarrow}J \longrightarrow \mathsf{a}{\downarrow}J$ always preserve the a-arrows, and restrict to $a\times\_ : \mathsf{a}/J \longrightarrow \mathsf{a}/J$. But lemma 47 implies that the $\mathsf{a}{\downarrow}J$-products $a\times a'$ are products in a/J <u>iff</u> a is a calibration. (• The projections from the $\mathsf{a}{\downarrow}J$-products are certainly a/J-arrows. So it is necessary and sufficient that the diagonals ρ_a to the $\mathsf{a}{\downarrow}J$-products are a/J-arrows.•) We conclude that a is a sub-lccc of an a-rccc $\mathcal{B}$ <u>iff</u> a is a calibration. (A statement to this effect appears in chapter 2 of Streicher 1988.)

Note, moreover, that for a calibration a, $\mathcal{B}$ is an a-rccc <u>iff</u> $\nabla\mathsf{a} : \mathsf{a}/\mathcal{B} \longrightarrow \mathcal{B}$ is an fccc (•since the functors $a\to\_$ then certainly preserve the a-arrows•).

7. Fibrewise rccc.

In a category of predicates the predicates will be fibred over sets. The quantification will be interpreted by the horizontal structure: categories of predicates will be hyperfibrations. Within this framework - inside each fibre - constructive logic in the form of Martin-Löf type theory will be implemented. So each fibre will be a relatively cartesian closed category. (Cf. Seely 1984, Cartmell 1986, Hyland-Pitts 1987.)

Given a fibred category $E : \mathcal{E} \longrightarrow \mathcal{B}$, we consider families $\mathsf{r} \subseteq \mathcal{E}$ of *vertical* arrows. Definition 3 is adapted for this situation simply by putting "vertical isomorphisms" in place of "isomorphisms" in (C0) and (C1); and "vertical arrows $X \longrightarrow \top EX$" in place of "arrows $K \longrightarrow \top$" in condition (D).. Instead of a calibration, saturated family etc. on $\mathcal{B}$, we obtain a *vertical* calibration, *vertical* saturated family etc. on a fibred category $\mathcal{E}$. Equivalently, we could have defined that a family $\mathsf{r} \subseteq \mathcal{E}$ of vertical arrows is a vertical calibration (vertical saturated...) if it is stable under inverse images <u>and</u> every $\mathsf{r}_I := \mathsf{r} \cap \mathcal{E}_I$ is a calibration (...).

A fibred category $\mathcal{E}$ will be *fibrewise relatively cartesian closed* with respect to a vertical stable subcategory r (should we call it r-frccc?) if every fibre $\mathcal{E}_I$ is relatively cartesian closed with respect to r_I, and the inverse images of

$$E' := E\circ\nabla\mathsf{r} : \mathsf{r}/\mathcal{E} \longrightarrow \mathcal{E} \longrightarrow \mathcal{B}$$

preserve the fibrewise cartesian closed structure of all $\nabla\mathsf{r}_I : \mathsf{r}_I/\mathcal{E}_I \longrightarrow \mathcal{E}_I$.

Applying proposition 6, we conclude that $\mathcal{E}$ is an r-frccc <u>iff</u> every $\nabla\mathsf{r}_I$ is an r_I-hyperfibration with intrinsic terminal objects, <u>and</u> the E'-inverse images preserve the cartesian, cocartesian and opcartesian arrows of all $\nabla\mathsf{r}_I$.

Since $\nabla\mathsf{r}_I$ are split intrinsic cofibrations, their cocartesian arrows over u are in the form $\langle id, u\rangle$. Obviously these $\nabla\mathsf{r}_I$-cocartesian arrows will be preserved by the E'-direct images, if they exist. Thus, when $E : \mathcal{E} \longrightarrow \mathcal{B}$ is a hyperfibration, $\nabla\mathsf{r}$ is a fibrewise hyperfibration (definitions 3.8) relative to r. The obvious relativized version of proposition 3.85(ii) now implies that $\mathcal{E}$ is an r-frccc <u>iff</u> $\nabla\mathsf{r} : \mathsf{r}/\mathcal{E} \longrightarrow \mathcal{E}$ is an r-

hyperfibration with intrinsic terminal objects. But by proposition 6 again, this last fact is equivalent with $\mathcal{E}$ being an honest $\mathtt{r}$-rccc. To resume,

$$\begin{aligned} \mathcal{E} \text{ is } \mathtt{r}\text{-frccc} \quad & \overset{6}{\Leftrightarrow} \nabla \mathtt{r}_I \text{ are } \mathtt{r}_I\text{-hyperfibrations} + \ldots \\ & \Leftrightarrow \nabla \mathtt{r} \in |\mathtt{r}\text{-}\underline{HYP}/E| \overset{3.85(ii)}{\Leftrightarrow} \\ & \Leftrightarrow \nabla \mathtt{r} \in |\mathtt{r}\text{-}\underline{HYP}/\mathcal{E}| \overset{6}{\Leftrightarrow} \\ & \Leftrightarrow \mathcal{E} \text{ is } \mathtt{r}\text{-rccc}. \end{aligned}$$

III. Notions of size

In chapter II we regarded fibrations as variable categories. Discrete fibrations were variable sets; and the base category could be thought of as a category of truth values and constructive proofs (in the style of Brouwer-Heyting-Kolmogorov), or of possible worlds and causal connections (in the style of Kripke). In the first case, for a variable set $E:\mathcal{E} \longrightarrow \mathcal{B}$, an object EX gives the truth value of the statement "$X \in \mathcal{E}$"; in the second case, a fibre $\mathcal{E}_K$ represents the set $\mathcal{E}$ at the moment K of history.

A different conception is that the base stands for the category of sets and functions. Every fibration is just a category, given with all the set-indexed families of its objects and arrows. The family fibrations (II.3.6) are paradigmatic. Discrete fibrations now become *classes*. Such a class is *small* if it is *representable* in the base category.

This second point of view, propagated principally by Bénabou (1985), leads us into a new circle of notions.

chapter II				chapter III
variable categories	←	fibrations	→	categories
variable sets	←	discrete fibrations	→	classes
propositions	←	objects in base	→	sets

In section 1 we consider *small fibrations* – those coming from internal categories in the base $\mathcal{B}$ – and the ways in which their functorial behaviour is represented in $\mathcal{B}$. Internal presheaves and descent data appear as objects of Yoneda-type representations.

Section 2 discusses the meaning of size for some important discrete fibrations as classes which can be derived from fibrations as categories. In particular, the question of *comprehension* for fibrations is considered. An idea of a *constructive extent operation* is given a categorical formulation. *Comprehensive fibrations* are introduced: they will be applied in the next chapter as a setting for the constructive comprehension principle as presented in the theory of predicates. The underlying concept connects, in a sense, Bénabou's (1985) *definability* with Lawvere's (1970) *comprehension scheme*.

Several characterisations of comprehension are given in section 3. In section 4 we study the relation of a comprehensive fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ and the *extent fibration* $\iota E:\iota\mathcal{E} \longrightarrow \mathcal{B}$ associated to it. ιE is an arrow fibration which tells "how $\mathcal{B}$ sees $\mathcal{E}$".

1. Yoneda

1. Representable fibrations.

The first category theoretical approximation of the slogan

All mathematics can be done in set theory,

is:

Most of mathematics can be done *over* the category of sets.

Every category $\mathcal{C}$ appears as a family fibration $\nabla\mathcal{C}$ over Set (cf. II.3.6); in particular, every class C gives rise to a *discrete* fibration $\nabla\mathsf{C}$. Class C is small, i.e. a set, iff $\nabla\mathsf{C}$ is *representable in* Set.

A second approximation is:

A bit of mathematics can be done over *a* category of sets.

In other words, we take an abstract base category $\mathcal{B}$ and think of its objects as sets. Fibrations over $\mathcal{B}$ are to be thought of as categories, discrete fibrations are classes. The notion of size is determined by representability in $\mathcal{B}$.

Definition. For $I \in |\mathcal{B}|$, the discrete fibration

$$\nabla I : \mathcal{B}/I \longrightarrow \mathcal{B} : u \longmapsto \mathrm{Dom}(u)$$

is called *representable*. The functor

$$\nabla : \mathcal{B} \longrightarrow \underline{\mathrm{FIB}}/\mathcal{B} : I \longmapsto \nabla I,\ u \longmapsto u\circ(\_)$$

is the *Yoneda embedding*.

Yoneda lemma. (cf. Bénabou 1983, 2.8.3) For every fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ and $I \in |\mathcal{B}|$, the functor

$$\mathbf{Y}_{I,E} : \underline{\mathrm{FIB}}/\mathcal{B}(\nabla I, E) \longrightarrow \mathcal{E}_I : F \longmapsto F(id_I)$$

is full and faithful. The family of functors $\mathbf{Y}$ is natural in I and E.

When E is cloven, $\mathbf{Y}_{I,E}$ restricts to an isomorphism $\underline{\mathrm{CLEAV}}_{\mathcal{B}}(\nabla I, E) \cong \mathcal{E}_I$.

• For every cartesian functor $F:\nabla I \longrightarrow E$ and every $u \in \mathcal{B}/I(u,id_I)$ the arrow $F(u):u^*F(id_I) \longrightarrow F(id_I)$ is cartesian. Every cartesian natural transformation $\varphi:F \longrightarrow G$ is uniquely determined by $\varphi_{id}:F(id_I) \longrightarrow G(id_I)$, since $\varphi_u = u^*(\varphi_{id}):F(u) \longrightarrow G(u)$.

Using a cleavage or the axiom of choice, we can define

$$\mathbf{Y}^{-1}(X) : \nabla I \longrightarrow E : u \mapsto u^*(X).•$$

The **Yoneda embedding** is full and faithful (i.e. an embedding). • Since ∇I is cloven, the Yoneda lemma gives $\underline{FIB}/\mathcal{B}(\nabla I, \nabla J) \cong (\mathcal{B}/I)_J = \mathcal{B}(I,J)$.•

Representables generate. For every cloven fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ holds

$$E = \underline{\text{lax}}\nabla E,$$

(where $\nabla E : \mathcal{E} \longrightarrow \mathcal{B} \longrightarrow \underline{FIB}/\mathcal{B}$ is regarded as a strong diagram). E is discrete iff

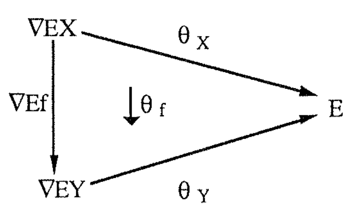
$$E = \underline{\text{lim}}\nabla E.$$

• (Some references on lax limits are listed in Gray 1980.) The X-component of the colimit cocone $\theta:\nabla E \longrightarrow E$ will be the cartesian functor $\theta_X := \mathbf{Y}^{-1}(X)$ which corresponds by the Yoneda lemma to $X \in |\mathcal{E}_{EX}|$. We need a cleavage to choose such a functor:

$$\theta_X : \nabla EX \longrightarrow E \ : \ u \mapsto u^*(X), (s:v \longrightarrow u) \mapsto \vartheta^{s}_{u^*(X)} \circ (c^{us})^{-1}.$$

(i.e. $\theta_X = Dom \circ \Theta_X$ - where Θ_X is from II.1.31(e) - is a "constant functor" $\ulcorner X \urcorner$, which we used in II.2.1, to define exponents).

The f-component of θ is obtained by the naturality part of the Yoneda lemma. For $f \in \mathcal{E}(X,Y)$, we have $\nabla Ef = Ef \circ (\_)$, and $\theta_Y \circ \nabla Ef(u) = (Ef \circ u)^*(Y)$.



Define

$$\theta_f : \theta_X \longrightarrow (\theta_Y)(\nabla Ef) : u \mapsto \left(c^{Ef,u} \circ u^*(\bar{f}) : \theta_X(u) \longrightarrow \theta_Y \circ \nabla Ef(u)\right),$$

where $\bar{f}$ is the unique vertical factorisation of f (i.e. $f = \vartheta^{Ef}_{Y} \circ \bar{f}$). (Checking that θ_f is a natural transformation is an exercise using the properties of c.)

Let a lax cocone $\varphi : \nabla E \longrightarrow E'$ be given. We must define a cleavage preserving functor $F:E \longrightarrow E'$, such that $\varphi_X = F\theta_X$, and $\varphi_f(u) = F(\theta_f(u))$. Since $\theta_X(id_{EX}) = X$, the first equality holds iff

$$FX := \varphi_X(id_{EX}).$$

In a similar way, the equality $\theta_f(id_{EX}) = \bar{f}$, together with the requirement that F preserves cleavages, determines

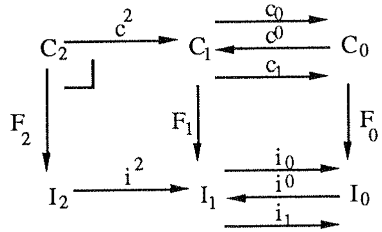
$$Ff \ := \vartheta^{Ef}_{FY} \circ \ \varphi_f(id_{EX}).•$$

2. Small fibrations.

21. Notation, references. An internal category $\mathbb{C}$ consists in principle of an *object of objects* C_0, an *object of arrows* C_1, arrows $\partial_0, \partial_1:C_1 \longrightarrow C_0$ representing *domain* and *codomain*, and $\eta:C_0 \longrightarrow C_1$ representing *identity* arrows of $\mathbb{C}$; and then there is an *object of composable pairs* C_2, obtained as a pullback of ∂_0 and ∂_1, and a *composition* $\mu:C_2 \longrightarrow C_1$. The intended meanings of these data are expressed by a set of equations imposed on them. Unless specified otherwise, p_0 will denote the arrow obtained by pulling back ∂_0 along ∂_1; p_1 is obtained by pulling back ∂_1 along ∂_0. Moreover, we shall denote the arrows ∂_0, ∂_1, μ, η belonging to a category $\mathbb{C}$ alternatively by c_0, c_1, c_2, c^0 respectively. $\underline{cat}_{\mathcal{B}}$ is the category of internal categories in $\mathcal{B}$. An introduction to internal categories can be found in chapter 2 of Johnstone 1977.

The notion of a small fibration is due to Bénabou (1975b). A recent reference is Hyland-Robinson-Rossolini 1988.

22. Example. Small categories are just the internal categories in Set. A functor $F:\mathbb{C} \longrightarrow \mathbb{I}$ consists of an *object part* $F_0:C_0 \longrightarrow I_0$, and an *arrow part* $F_1:C_1 \longrightarrow I_1$, with the obvious commutativity conditions.



By proposition II.1.61(d), the functor F will be a discrete fibration iff the arrow c_1 is a pullback of i_1. If we regard the arrows F_0 and F_1 as objects of $\underline{Set}/\underline{Set}$, the preceding diagram becomes an internal category $\mathbb{F}$ in $\underline{Set}/\underline{Set}$. This category is projected by the fibration $\nabla\underline{Set}$ = Cod on the internal category $\mathbb{I}$ in $\underline{Set}$. We can say that $\mathbb{F}$ is a *category over* $\mathbb{I}$ with respect to the fibration $\nabla\underline{Set}$. Discrete fibrations - or *presheaves* - over $\mathbb{I}$ are exactly those categories $\mathbb{F}$ over $\mathbb{I}$ (with respect to $\nabla\underline{Set}$) in which the codomain arrow $f_1:=\langle c_1,i_1\rangle:F_1\to F_0$ is cartesian (i.e. a pullback). The categories $\mathbb{F}$ over $\mathbb{I}$ in which both domain and codomain arrows are cartesian correspond to the discrete *bi*fibrations (or *descent data*) over $\mathbb{I}$. They can be viewed as the functors $\mathbb{I}\to\underline{Bij}$, where $\underline{Bij}$ is the category of sets and bijections (i.e. the largest groupoid contained in $\underline{Set}$). (• Namely, discrete bifibrations must satisfy the Beck-Chevalley condition over all the commutative squares. By II.3.5, the inverse and direct image functors must be equivalences; in fact - bijections, since fibres are sets.•)

These notions can be generalized from the fibration $\nabla\underline{Set}$ to an arbitrary fibration $E:\mathcal{E}\to\mathcal{B}$ as soon as the categories $\mathcal{E}$ and $\mathcal{B}$ allow internal category theory, and functor E projects categories on categories.

23. Definitions. Let $\underline{LEX}$ be the category of finitely complete categories and left exact functors. Let $\underline{LEXFIB}/\mathcal{B}$ be the category of fibrations $E:\mathcal{E}\to\mathcal{B}$, where $\mathcal{E},\mathcal{B}\in|\underline{LEX}|$, and $E\in\underline{LEX}(\mathcal{E},\mathcal{B})$; with the left exact cartesian functors as arrows between them. (Cf. II.2.2). The functor

$$\underline{cat}:\underline{LEX}\to\underline{LEX}:\mathcal{B}\mapsto\underline{cat}_{\mathcal{B}}$$

lifts to

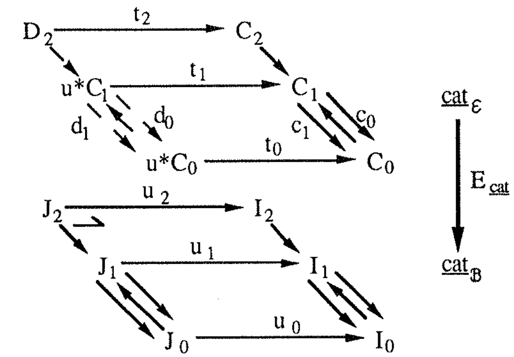
$$\underline{cat}:\underline{LEXFIB}/\mathcal{B}\to\underline{LEXFIB}/\underline{cat}_{\mathcal{B}}:$$
$$(E:\mathcal{E}\to\mathcal{B})\mapsto(E_{\underline{cat}}:\underline{cat}_{\mathcal{E}}\to\underline{cat}_{\mathcal{B}}),$$

where $E_{\underline{cat}}$ takes the categories in $\mathcal{E}$ into their E-images: since E is left exact, these images are categories in $\mathcal{B}$.

For $\mathbb{I}\in|\underline{cat}_{\mathcal{B}}|$ and a fibration $E:\mathcal{E}\to\mathcal{B}$, an $\mathbb{I}$-*presheaf* in $\mathcal{E}$ is an internal category $\mathbb{C}\in|(\underline{cat}_{\mathcal{E}})_{\mathbb{I}}|$ (i.e. $E\mathbb{C}=\mathbb{I}$) in which the domain arrow ∂_1 is cartesian. (This implies that η and μ are cartesian too, the latter because of the equation $\partial_1\mu=\partial_1 p_1$.) $\underline{psh}_{\mathcal{E}}(\mathbb{I})$ is the full subcategory of $(\underline{cat}_{\mathcal{E}})_{\mathbb{I}}$ spanned by the presheaves.

Descent data over $\mathbb{I}$ is an internal category $\mathbb{C}\in|(\underline{cat}_{\mathcal{E}})_{\mathbb{I}}|$ which consists of cartesian arrows. $\underline{des}_{\mathcal{E}}(\mathbb{I})$ is the full subcategory of $(\underline{cat}_{\mathcal{E}})_{\mathbb{I}}$ spanned by the descent data.

24. Comments. $E_{\underline{cat}}$-cartesian liftings of $\mathbb{u}\in\underline{cat}_{\mathcal{B}}(\mathbb{I},\mathbb{J})$ at $\mathbb{C}\in|(\underline{cat}_{\mathcal{E}})_{\mathbb{I}}|$ consist of E-cartesian liftings of the components of $\mathbb{u}$ at the appropriate components of $\mathbb{C}$.



We define t_0 and t_1 to be cartesian liftings of u_0 and u_1 respectively ; d_0 and d_1 are induced as factorisations. D_2 can now be defined either as an inverse image of C_2 along u_2, or as a pullback of d_0 and d_1: lemma 25 says that both ways give the same result. So we get a category $\mathbb{u}^*(\mathbb{C})$.

Clearly, $\underline{psh}_{\mathcal{E}}$ and $\underline{des}_{\mathcal{E}}$ are subfibrations of $\underline{cat}_{\mathcal{E}}$.

A presheaf $\mathbb{C}$ over $\mathbb{I}$ is determined by an object C over I_0 and an arrow $\gamma:i_1^*C\to i_0^*C$ over I_1. C is the object of objects of $\mathbb{C}$, i_1^*C is the object of arrows, γ is the vertical part of c_0. There is an equivalence of categories

$$\underline{psh}_{\mathcal{E}}(\mathbb{I})\to\underline{psh}^{\Diamond}_{\mathcal{E}}(\mathbb{I}),$$

where

$$|\underline{psh}^{\Diamond}_{\mathcal{E}}(\mathbb{I})|:=\left\{\langle C,\gamma\rangle\in\sum_{X\in|\mathcal{E}_{I_0}|}\mathcal{E}_{I_1}(i_1^*X,i_0^*X)\;\middle|\; i^{0*}\gamma\simeq id,\ i_2^*\gamma\simeq p_0^*\gamma\circ c\circ p_1^*\gamma\right\}$$

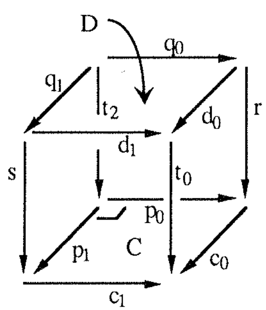
$$\underline{psh}^{\langle\rangle}{}_{\mathcal{E}}(\mathbb{I})(\langle C,\gamma\rangle, \langle D,\delta\rangle) := \left\{ f\in \mathcal{E}_{I_0}(C,D) \mid i_0{}^*f\circ\gamma = \delta\circ i_1{}^*f \right\}.$$

(c is the canonical isomorphism $p_1{}^*i_0{}^*\gamma \longrightarrow p_0{}^*i_1{}^*\gamma$. In general, for arrows f and g, "$f\simeq g$" means "$\exists$ isos c_0,c_1. $f=c_0\circ g\circ c_1$".)

For the sake of simplicity, we shall freely move between $\underline{psh}$ and $\underline{psh}^{\langle\rangle}$; but this could be avoided, and no assertion will really depend on representation.

A presheaf $\langle C,\gamma\rangle$ represents a descent data $\underline{\text{iff}}$ γ is an isomorphism. If $\mathbb{I}$ is a groupoid, $\underline{\text{then}}$ all the $\mathbb{I}$-presheaves are descent data. (• $\mathbb{I}$ is a groupoid $\underline{\text{iff}}$ there is an involution $\tau:I_1\longrightarrow I_1$, $i_0=i_1\tau$, which takes each arrow of $\mathbb{I}$ to its inverse: $i_2«\tau,id»=i^0i_0$ $\underline{\text{and}}$ $i_2«id,\tau»=i^0i_1$ (where $«id,\tau»:I_1\longrightarrow I_2$ is the factorisation induced by $i_0=i_1\tau$, while $«\tau,id»:I_1\longrightarrow I_2$ is induced by $i_0\tau=i_1$). For arbitrary $\mathbb{C}\in \underline{psh}_{\mathcal{E}}(\mathbb{I})$, consider the unique arrow $\upsilon:C_1\longrightarrow C_1$ over τ such that $c_0=c_1\upsilon$. This equality induces $«id,\upsilon»:C_1\longrightarrow C_2$, which is cartesian since $p_1«id,\upsilon»=id$, and the pullback p_1 of c_1 along c_0 is cartesian. Therefore, the equality $c_2«id,\upsilon»=c^0c_1a$ must hold for some vertical automorphism a, since both sides are cartesian liftings of the same arrow at the same object. But now $c_1\upsilon\upsilon=c_0\upsilon=c_0c_2«id,\upsilon»=c_0c^0c_1a=c_1a$ $\underline{\text{implies}}$ that $\upsilon\upsilon=a$. This means that υ is an isomorphism, and c_0 must be cartesian.•)

25. Lemma. Consider the following commutative diagram in $\mathcal{E}$



where arrows r and s are cartesian, while squares C in $\mathcal{E}$ and ED in $\mathcal{B}$ are pullbacks. The following implications are true:

i) If EC is a pullback $\underline{\text{and}}$ t_2 is cartesian $\underline{\text{then}}$ D is a pullback.

ii) If D is a pullback $\underline{\text{and}}$ t_0 is cartesian $\underline{\text{then}}$ t_2 is cartesian.

• We use the name of each data to abbreviate the assumption about it: (r) stands for "r is cartesian", (C) for "C is a pullback" and so on.

i) $d_0a=d_1b \Rightarrow c_0ra=c_1sb \overset{(C)}{\Rightarrow} \exists!x.\ ra=p_0x$ $\underline{\text{and}}$ $sb=p_1x$:(α)

$d_0a=d_1b \Rightarrow Ed_0\circ Ea=Ed_1\circ Eb \overset{(ED)}{\Rightarrow} \exists!h.\ Ea=Eq_0\circ h$ $\underline{\text{and}}$ $Eb=Eq_1\circ h$:(β)

(α) $\underline{\text{and}}$ (β) $\Rightarrow Ep_0\circ Ex=Ep_0\circ Et_2\circ h$ $\underline{\text{and}}$ $Ep_1\circ Ex=Ep_1\circ Et_2\circ h$

$\overset{(EC)}{\Rightarrow} Ex=Et_2\circ h \overset{(t_2)}{\Rightarrow} \exists!y.\ Ey=h$ $\underline{\text{and}}$ $x=t_2y$:(γ)

(α) $\underline{\text{and}}$ (γ) $\Rightarrow ra=rq_0y$ $\underline{\text{and}}$ $sb=sq_1y \overset{(r,s,\beta)}{\Rightarrow} a=q_0y$ $\underline{\text{and}}$ $b=q_1y$.

ii) $Ex=Et_2\circ h \Rightarrow E(p_0x)=Er\circ Eq_0\circ h$ $\underline{\text{and}}$ $E(p_1x)=Es\circ Eq_1\circ h$

$\overset{(r,s)}{\Rightarrow} \exists!a,b.\ ra=p_0x$ $\underline{\text{and}}$ $sb=p_1x$ $\underline{\text{and}}$:(α)

$Ea=Eq_0\circ h$ $\underline{\text{and}}$ $Eb=Eq_1\circ h$:(β)

Since (α) $\Rightarrow t_0d_0a=t_0d_1b$ and (β) $\Rightarrow E(d_0a)=E(d_1b)$,

(α) $\underline{\text{and}}$ (β) $\overset{(t_0)}{\Rightarrow} d_0a=d_1b \overset{(D)}{\Rightarrow} \exists!y.\ a=q_0y$ $\underline{\text{and}}$ $b=q_1y$:(γ)

(α) $\underline{\text{and}}$ (γ) $\Rightarrow p_0x=p_0t_2y$ $\underline{\text{and}}$ $p_1x=p_1t_2y \overset{(C)}{\Rightarrow} x=t_2y$

(β) $\underline{\text{and}}$ (γ) $\Rightarrow Eq_0\circ Ey=Eq_0\circ h$ $\underline{\text{and}}$ $Eq_1\circ Ey=Eq_1\circ h \overset{(ED)}{\Rightarrow} Ey=h$.•

26. Definition. The *externalisation* of $\mathbb{I}\in \underline{cat}_{\mathcal{B}}$ is the split fibration

$$\nabla\mathbb{I} : \mathcal{B}/\mathbb{I}\longrightarrow \mathcal{B}$$

where

$$|\mathcal{B}/\mathbb{I}| := |\mathcal{B}/I_0|$$

$$\mathcal{B}/\mathbb{I}(k,m) := \left\{\langle f,\varphi\rangle\in \mathcal{B}(K,M)\times\mathcal{B}(K,I_1) \mid \partial_0\varphi=k,\ \partial_1\varphi=m\circ f\right\},$$

for $k:K\longrightarrow I_0$, $m:M\longrightarrow I_0$. The *externalisation functor* is

$$\nabla : \underline{cat}_{\mathcal{B}}\longrightarrow \underline{FIB}/\mathcal{B} : \mathbb{I}\longmapsto \nabla\mathbb{I},$$

with the arrow part induced by the composition.

A fibration is *small* if it is fibrewise equivalent with one in the form $\nabla\mathbb{I}$.

27. Yoneda lemma. Given $\mathbb{I}\in |\underline{cat}_{\mathcal{B}}|$ denote by $\mathbb{I}^o$ the category obtained by interchanging ∂_0 and ∂_1; consider $\nabla\mathbb{I}^o:=\nabla(\mathbb{I}^o)\cong(\nabla\mathbb{I})^{op}$. For every fibration $E:\mathcal{E}\longrightarrow\mathcal{B}$ and $\mathbb{I}\in|\underline{cat}_{\mathcal{B}}|$, the following functor is full and faithful:

$$\mathbf{Y}_{\mathbb{I},E} : \underline{FIB}/\mathcal{B}\left(\nabla\mathbb{I}^o, E\right)\longrightarrow \underline{psh}_{\mathcal{E}}(\mathbb{I}) : F\longmapsto \left\langle F(id_{I_0}), F(\upsilon)\right\rangle,$$

where $\upsilon:=\langle id_{I_1},id_{I_1}\rangle\in \mathcal{B}/\mathbb{I}^o(\partial_0,\partial_1)=\mathcal{B}/\mathbb{I}(\partial_1,\partial_0)$. The family of functors $\mathbf{Y}$ is natural in $\mathbb{I}$ and E.

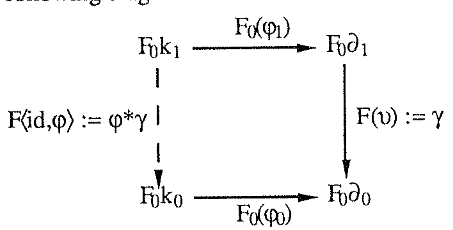
When E is cloven, $\mathbf{Y}_{\mathbb{I},E}$ restricts to an isomorphism $\underline{\mathrm{CLEAV}}_{\mathcal{B}}\left(\nabla\mathbb{I}^{o}, E\right) \cong \underline{\mathrm{psh}}_{\mathcal{E}}(\mathbb{I})$. (Therefore, when the axiom of choice is assumed, $\mathbf{Y}_{\mathbb{I},E}$ is surjective on objects; hence an equivalence of categories.)

• We only prove the surjectivity of $\mathbf{Y}_{\mathbb{I},E}$ for a cloven E. Given $\langle C,\gamma\rangle \in \underline{\mathrm{psh}}_{\mathcal{E}}(\mathbb{I})$, the Yoneda lemma for representable fibrations gives $F_0 \in \underline{\mathrm{FIB}}/\mathcal{B}(\nabla I_0, E)$ such that $F_0(\mathrm{id}_{I_0})=C$. We now extend F_0 to $F \in \underline{\mathrm{FIB}}/\mathcal{B}\left(\nabla\mathbb{I}^{o}, E\right)$ such that $F(\mathrm{id}_{I_0})=C$ and $F(\upsilon)=\gamma$.

The subcategory $\mathcal{B}/I_0$ of $\mathcal{B}/\mathbb{I}^{o}$ contains all the objects and canonical cartesian arrows of the latter. We define $F \in \underline{\mathrm{FIB}}/\mathcal{B}\left(\nabla\mathbb{I}^{o}, E\right)$ to restrict to F_0 along $\mathcal{B}/I_0 \hookrightarrow \mathcal{B}/\mathbb{I}^{o}$ (i.e. $F\restriction\nabla I_0 := F_0$). It is now sufficient to supply the definition of F on vertical arrows $\langle \mathrm{id},\varphi\rangle$, and then set

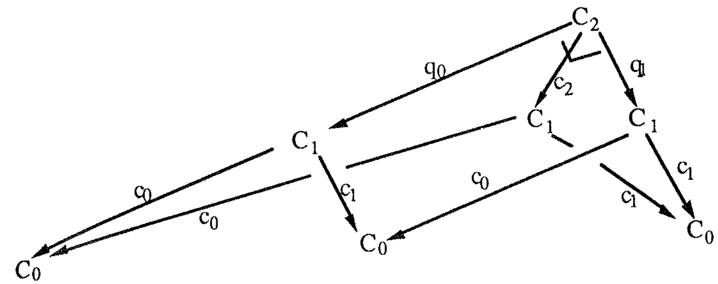
$$F\langle f,\varphi\rangle := F_0(f)\circ F\langle \mathrm{id},\varphi\rangle.$$

Every vertical arrow $\langle \mathrm{id},\varphi\rangle \in \mathcal{B}/\mathbb{I}(k_0,k_1)$ appears as $\varphi=\varphi_0 \in \mathcal{B}/I_0(k_0,\partial_0)$ and as $\varphi=\varphi_1 \in \mathcal{B}/I_0(k_1,\partial_1)$. Since $F_0(\varphi_0)$ and $F_0(\varphi_1)$ are cartesian arrows, $F\langle \mathrm{id},\varphi\rangle$ can be defined by the following diagram.

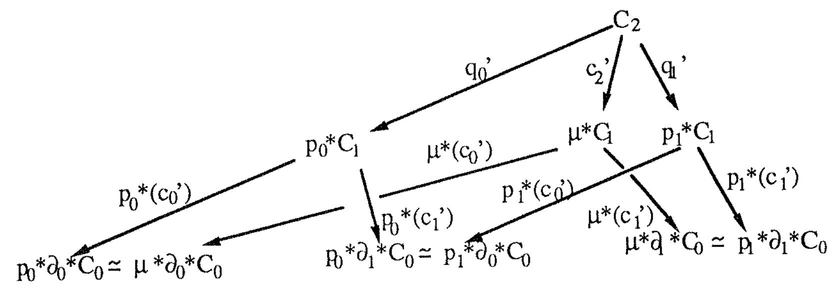


It remains to prove that F is a functor. We first show that its vertical components preserve the composition.

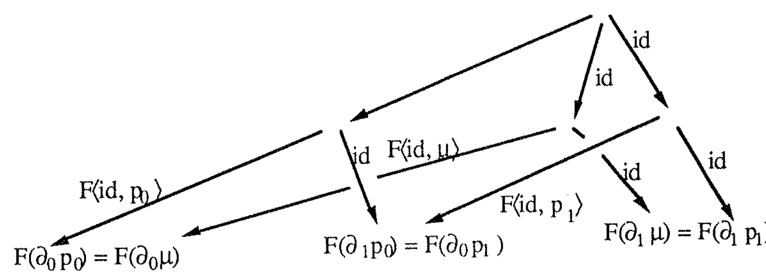
For every internal category $\mathbb{C}$ the following diagram commutes.



For $\mathbb{C} \in \underline{\mathrm{cat}}_{\mathcal{E}}(\mathbb{I})$, this diagram induces in $\mathcal{E}_{I_2}$ the following one:



where f' denotes the vertical component of f (while $p_i{:}I_2 \longrightarrow I_1$ is, as before, obtained by pulling ∂_i back along ∂_j, $j\neq i\in 2$). If $\mathbb{C} \in \underline{\mathrm{psh}}_{\mathcal{E}}(\mathbb{I})$, the short arrows in the first diagram are cartesian. As defined above, F determines a choice of cartesian liftings, so that the second diagram reduces to



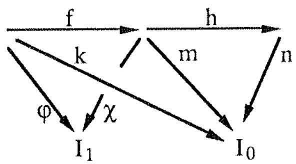
(because c_1' and c_2' are identities, while $\gamma=c_0$'). Thus $F\langle \mathrm{id},p_0\rangle\circ F\langle \mathrm{id},p_1\rangle = F\langle \mathrm{id},\mu\rangle$.

Every composable pair $\psi_0,\psi_1{:}K \longrightarrow I_1$ of K-indexed families of arrows in $\mathbb{I}$ induces in $\mathcal{B}$ an arrow $«\psi_0,\psi_1»{:}K \longrightarrow I_2$ so that $\psi_i=p_i«\psi_0,\psi_1»$, $i\in 2$. The proof that

$F\langle id,\psi_0\rangle \circ F\langle id,\psi_1\rangle = F\langle id,\mu«\psi_0,\psi_1»\rangle$

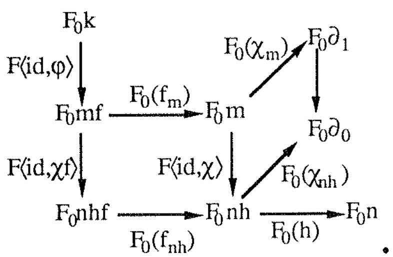
follows directly from the preceding arguments. But $\mu«\psi_0,\psi_1»$ is the internal composition of ψ_0 and ψ_1. So F preserves the composition of vertical arrows.

F preserves the composition in general because for



holds

$$
\begin{aligned}
F\langle h,\chi\rangle \circ F\langle f,\varphi\rangle &= F_0(h)\circ F\langle id,\chi\rangle \circ F_0(f)\circ F\langle id,\varphi\rangle = \\
&= F_0(h)\circ F_0(f)\circ F\langle id,\chi\circ f\rangle \circ F\langle id,\varphi\rangle = \\
&= F_0(h\circ f)\circ F\langle id,\mu«f^*\chi,\varphi»\rangle = \\
&= F\langle h\circ f,\ \mu«f^*\chi,\varphi»\rangle.
\end{aligned}
$$



28. Yoneda embedding. Externalisation is full and faithful:

$\underline{Cleav}_{\mathcal{B}}\left(\nabla\mathbb{I}, \nabla\mathbb{J}\right) \cong \underline{cat}_{\mathcal{B}}(\mathbb{I},\mathbb{J})$.

• If we forget the $(\_)^o$-busyness, then the preceding proposition gives an isomorphism between $\underline{Cleav}_{\mathcal{B}}(\nabla\mathbb{I},\nabla\mathbb{J})$ and an appropriate category of pairs

$\left\langle c\in|(\mathcal{B}/\mathbb{J})_{I_0}|, \langle id,\gamma\rangle\in(\mathcal{B}/\mathbb{J})_{I_1}(i_0{}^*c,i_1{}^*c)\right\rangle$.

Now $c\in\mathcal{B}(I_0,J_0)$ will be the object part, and $\gamma\in\mathcal{B}(I_1,J_1)$ the arrow part of a functor $\mathbb{I}\longrightarrow\mathbb{J}$, since $(\mathcal{B}/\mathbb{J})_{I_1}(i_0{}^*c,i_1{}^*c)\cong\mathcal{B}/J_0\times J_0\big(\langle ci_0,ci_1\rangle,\langle j_0,j_1\rangle\big)$.•

2. Fibred classes and predicates

1. Examples.

11. Every category $\mathcal{C}$ carries some classes with it. For instance

1) a class of objects $OB_{\mathcal{C}}$ (or $|\mathcal{C}|$);
2) classes of arrows $HOM_{\mathcal{C}}(X,Y)$ (or $\mathcal{C}(X,Y)$);
3) classes of isomorphisms $ISO_{\mathcal{C}}(X,Y)$ $(\subseteq HOM_{\mathcal{C}}(X,Y))$;
4) classes of cones $CONE_{\mathcal{C}}(X,\Delta)$, where $\Delta:\mathcal{D}\longrightarrow\mathcal{C}$ is a diagram.

In a similar fashion, from an aribtrary fibration $E:\mathcal{E}\longrightarrow\mathcal{B}$ we can derive some discrete fibrations as *fibred classes*. (For simplicity, we present these discrete, hence split fibrations as functors to the category SET of classes.)

Ad 1) "Morally", the fibred class OB_E should have the class of $|\mathcal{E}_I|$ as its fibre over $I\in|\mathcal{B}|$. However, while the arrows of a category $\mathcal{C}$ can always be removed to uncover the class $OB_{\mathcal{C}}$ of its objects, removing the vertical arrows from a fibred category $\mathcal{E}$ will result in a fibration *only if* $E:\mathcal{E}\longrightarrow\mathcal{B}$ is a split fibration. There are two extremal ways to "force" E to split: one is to replace each fibre $\mathcal{E}_I$ by its skeleton $[\mathcal{E}_I]$ (i.e. a quotient of $\mathcal{E}_I$, in which isomorphic objects are identified); the other way is to put $\underline{FIB}/\mathcal{B}(\nabla I,E)$ in place of $\mathcal{E}_I$ (remembering the Yoneda lemma). Hence two fibred classes of objects assigned to each fibration:

$OB_E : \mathcal{B}^o\longrightarrow\underline{SET} : I\longmapsto|[\mathcal{E}_I]|$, and

$OB'_E : \mathcal{B}^o\longrightarrow\underline{SET} : I\longmapsto|\underline{FIB}/\mathcal{B}(\nabla I,E)|$.

The arrow part of OB_E is determined by the unique object parts of inverse image functors; the arrow part of OB'_E is induced by composition. (OB'_E appeared in Bénabou 1983, exercise 11.)

Ad 2) The definition of fibred hom-classes goes easier. (It is due to Giraud 1971, 2.6.) For $I\in|\mathcal{B}|$ and $X,Y\in|\mathcal{E}_I|$

$HOM_E(X,Y) : (\mathcal{B}/I)^o\longrightarrow\underline{SET} : \big(v:K\longrightarrow I\big)\longmapsto\mathcal{E}_K(v^*X,v^*Y)$.

(It only remains to check that nothing depends on the choice of inverse images here.) If $\mathcal{B}$ has binary products, a fibred class HOM'_E can be defined also for objects which are not in the same fibre. If X is over J_0 and Y over J_1, then

$$HOM'_E(X,Y):=HOM_E(\pi_0{}^*X,\ \pi_1{}^*Y) : (\mathcal{B}/J_0\times J_1)^o \longrightarrow \underline{SET},$$

where $\pi_i : J_0 \times J_1 \longrightarrow J_i$, $i \in 2$, are projections.

Ad 3) The obvious

$$ISO_E(X,Y) : (\mathcal{B}/I)^o \longrightarrow \underline{SET} : (v{:}K \longrightarrow I) \longmapsto \{a \in \mathcal{E}_K(v^*X, v^*Y) \mid a \text{ is iso}\}$$

can equivalently be defined to map

$$v \longmapsto \{f \in \mathcal{E}_v(v^*X,Y) \mid f \text{ is cartesian}\}.$$

Ad 4) Let $\Delta{:}\mathcal{D} \longrightarrow \mathcal{E}_I$ be a diagram.

$$CONE_E(X,\Delta) : (\mathcal{B}/I)^o \longrightarrow \underline{SET} : v \longmapsto CONE_v(v^*X,\Delta),$$

(where $CONE_v(v^*X,\Delta)$ denotes, of course, the class of cones $\delta{:}v^*X \longrightarrow \Delta$ over v, i.e. such that $E\delta=v$).

12. Some predicates on a category $\mathcal{C}$:

5) $iso_{\mathcal{C}}(X,Y)$:= "there is an iso $X \longrightarrow Y$";

6) $iso_{\mathcal{C}}(f)$:= "f is an iso"

7) $cone_{\mathcal{C}}(X,\Delta)$:= "there is a cone $X \longrightarrow \Delta$".

In an abstract category of sets - a topos - the truth values appear as subobjects of the terminal object. In $\underline{FIB}/\mathcal{B}$ they are the subfibrations of the terminal fibration $id{:}\mathcal{B} \longrightarrow \mathcal{B}$ - i.e. discrete fibrations which contain at most one object per fibre. They can be presented as *cribles*, "downward" closed families of objects of $\mathcal{B}$: $X \in$ crible <u>and</u> $\mathcal{B}(Y,X) \neq \emptyset$ <u>imply</u> $Y \in$ crible. More generally, just as predicates are viewed as subobjects of sets, *fibred predicates* are taken to be the subfibrations of representable fibrations $\nabla I{:}\mathcal{B}/I \longrightarrow \mathcal{B}$. And they are just cribles in $\mathcal{B}/I$.

Ad 5) The fibred predicate $iso_E(X,Y)$ can be derived from $ISO_E(X,Y)$ by saying that its fibre over v contains one element <u>iff</u> the fibre of $ISO_E(X,Y)$ is inhabited. As a crible, it is

$$iso_E(X,Y) := \left\{ v \in |\mathcal{B}/I| \mid \text{there is an iso } v^*X \longrightarrow v^*Y \right\}$$

Ad 6) $iso_E(f) := \left\{ v \in |\mathcal{B}/I| \mid v^*(f) \text{ is iso (i.e. } f\circ\vartheta^v \text{ is cartesian)} \right\}.$

Ad 7) $cone_E(X,\Delta) := \left\{ v \in |\mathcal{B}/I| \mid \text{there is a cone of cartesian arrows } v^*X \longrightarrow \Delta \right\}.$

13. And now, the idea is to lift some set theoretical and logical concepts among fibrations by postulating that

a fibred $\left\{\begin{matrix}\text{class}\\ \text{predicate}\end{matrix}\right\}$ is $\left\{\begin{matrix}\text{small}\\ \text{definable}\end{matrix}\right\}$ if it is representable.

Besides technical problems, like that of formulating OB_E, some deeper conceptual problems arise in the realisation of this idea. In the sequel, we first briefly survey a development of category theory in $\underline{FIB}/\mathcal{B}$, using fibred classes; and then we turn to the concept of comprehension for fibrations, viewed, in particular, as categories of predicates varying over a category of sets.

2. Locally and globally small.

Definitions. A fibration $E{:}\mathcal{E} \longrightarrow \mathcal{B}$ is *globally small* if the class OB_E is representable.

E is *locally small* if all classes $HOM_E(X,Y)$ are representable.

Examples. A family fibration $\nabla\mathcal{C}{:}\underline{Set}/\mathcal{C} \longrightarrow \underline{Set}$ is locally small <u>iff</u> the category $\mathcal{C}$ is, i.e. if its hom-classes are small. $\nabla\mathcal{C}$ is globally small <u>iff</u> $\mathcal{C}$ is equivalent to a category with small class of objects.

Every small fibration $\nabla\mathbb{C}{:}\mathcal{B}/\mathbb{C} \longrightarrow \mathcal{B}$ is locally small: a representant of the discrete fibration $HOM_{\nabla\mathbb{C}}(k_0,k_1)$ is an arrow $\mathbb{C}^K(k_0,k_1) := K \times_{C_0} C_1 \longrightarrow K$, obtained by pulling back $\langle\partial_0,\partial_1\rangle$ along $\langle k_0,k_1\rangle{:}K \longrightarrow C_0 \times C_0$. If $\mathcal{B}$ is a topos, $\nabla\mathbb{C}$ is globally small too: the skeleton $[\mathbb{C}]$ can be constructed, and its object of objects $[\mathbb{C}]_0$ is a representant of $OB_{\nabla\mathbb{C}}$.

An arrow fibration $\nabla\mathfrak{a}{:}\mathfrak{a}/\mathcal{B} \longrightarrow \mathcal{B}$ is locally small <u>iff</u> $\mathcal{B}$ is relatively cartesian closed with respect to $\mathfrak{a}$: the exponent $k_0 \rightarrow k_1 \in |\mathfrak{a}\downarrow K|$ represents $HOM_{\nabla\mathfrak{a}}(k_0,k_1)$. The statement that $\nabla\mathfrak{a}$ is globally small means that the class $\mathfrak{a}$ consists of the pullbacks of a generic arrow $\xi \in |\Omega \downarrow \Omega|$. This arrow can be regarded as an Ω-indexed family of objects of $\mathcal{B}$; or as a *small full subcategory* of $\mathcal{B}$ spanned by these objects. (Cf. Johnstone 1977, 2.38; the idea attributed to Bénabou). A globally small arrow fibration over a locally cartesian closed category must be small (assuming the axiom of choice: cf. Hyland-Robinson-Rossolini 1990, lemma 3.2.). A basic fibration $\nabla\mathcal{B}$ is both globally and locally small iff $\mathcal{B}$ is equivalent with 1 (Pitts-Taylor 1989).

A simple non-example. Consider the category

$$|\underline{Set}/\!\!/2| := |\underline{Set}/2|$$
$$\underline{Set}/\!\!/2(\alpha{:}I \to 2, \beta{:}J \to 2) := \{\langle u,f\rangle \in \underline{Set}(I,J) \times \underline{Set}(2,2) \mid f\alpha = \beta u\}$$

fibred by

$$\nabla\!\!\!/ 2{:}\underline{Set}/\!\!/2 \to \underline{Set} : \alpha \mapsto Dom(\alpha),\ \langle u,f\rangle \mapsto u.$$

Take $\alpha{:}2 \to 2$ to be the constant function $\ulcorner 0 \urcorner$, and $\beta{:}2 \to 2$ the identity. For $u_0{:}\{0\} \hookrightarrow 2$, $u_1{:}\{1\} \hookrightarrow 2$, the sets $(\underline{Set}/\!\!/2)_{\{i\}}(u_i{}^*\alpha, u_i{}^*\beta)$ are both inhabited (for $i \in 2$). If $\nabla\!\!\!/2$ is locally small, then both sets $\underline{Set}/\{i\}(u_i, \iota(\alpha,\beta))$ must be inhabited, i.e. $\iota(\alpha,\beta)$ must be surjective. This contradicts the fact that $\underline{Set}/2(id, \iota(\alpha,\beta))$ must be empty, since $(\underline{Set}/\!\!/2)_2(\alpha,\beta)$ is empty.

Representants. For a globally small fibration E there is a representant $\Omega \in |\mathcal{B}|$ and a cartesian isomorphism

$$H : OB_E \to \nabla\Omega.$$

Since every $q \in (\mathcal{B}/\Omega)_I = \mathcal{B}(I,\Omega)$ is $q = q^*id_\Omega$, every object $Q \in |[\mathcal{E}_I]|$ must be the inverse image of $H^{-1}(id_\Omega) \in [\mathcal{E}_\Omega]$ along $q := H(Q)$:

$$Q = H^{-1}(q) = H^{-1}(q^*id_\Omega) = q^*\big(H^{-1}(id_\Omega)\big).$$

Each element $\xi \in |\mathcal{E}_\Omega|$ of the equivalence class $H^{-1}(id_\Omega)$ (of isomorphic objects from $\mathcal{E}_\Omega$) is a *generic object* for the fibred category $\mathcal{E}$, in the sense that for every object $A \in |\mathcal{E}_I|$ there is a unique arrow $\ulcorner A \urcorner{:}I \to \Omega$ such that

$$A \simeq \ulcorner A \urcorner^* \xi.$$

The mapping $\ulcorner \_ \urcorner : |\mathcal{E}| \to |\mathcal{B}/\Omega|$ is obtained by composing the isomorphism H with the obvious surjection $|\mathcal{E}| \to |[\mathcal{E}]|$.

The representants of $HOM_E(X,Y)$ will be generically denoted $\iota(X,Y){:}D(X,Y) \to I$. For any $v \in \mathcal{B}(K,I)$, each $a \in \mathcal{E}_K(v^*X, v^*Y)$ is an inverse image of a *generic arrow* $\gamma(X,Y) \in \mathcal{E}_{D(X,Y)}(\iota^*X, \iota^*Y)$ along a unique $\ulcorner a \urcorner : K \to D(X,Y)$.

The arrows $\iota(X,Y)$ display each fibre $\mathcal{E}_I$ of a locally small fibration $E{:}\mathcal{E} \to \mathcal{B}$ as a $\mathcal{B}/I$-enriched category. (Standard reference for enriched category theory: Kelly 1982.) The isomorphism $\iota(v^*X, v^*Y) \simeq v^*\big(\iota(X,Y)\big)$ means on one hand that the inverse image functors appear as enriched fully faithful. On the other hand, this implies that if E is locally small, the pullback of every representant $\iota(X,Y)$ along any arrow in $\mathcal{B}$ must exist. Bearing this in mind, we obtain the enriched structure as follows. For every X, Y, $Z \in |\mathcal{E}_I|$ the transformation

$$\mathcal{B}/I\big(v,\ \iota(X,Y) \underset{I}{\times} \iota(Y,Z)\big) \simeq \mathcal{E}_K(v^*X, v^*Y) \times \mathcal{E}_K(v^*Y, v^*Z) \overset{\circ}{\to}$$
$$\to \mathcal{E}_K(v^*X, v^*Z) \simeq \mathcal{B}/I\big(v,\ \iota(X,Z)\big),$$

natural in $v \in \mathcal{B}(K,I)$, induces by the Yoneda lemma a *composition arrow* in $\mathcal{B}/I$

$$\mu(X,Y,Z) : \iota(X,Y) \underset{I}{\times} \iota(Y,Z) \to \iota(X,Z)$$

such that for $p_0{:}\ \iota(X,Y) \underset{I}{\times} \iota(Y,Z) \to \iota(X,Y)$ and $p_1{:}\ \iota(X,Y) \underset{I}{\times} \iota(Y,Z) \to \iota(Y,Z)$

$$p_1{}^*\big(\gamma(Y,Z)\big) \circ \tau \circ p_0{}^*\big(\gamma(X,Y)\big) = \tau' \circ \mu^*\big(\gamma(X,Z)\big) \circ \tau'',$$

where τ, τ', τ'' are canonical vertical isos. On the other hand, the *identities arrow*

$$\eta(X) : id_I \to \iota(X,X)$$

corresponds by the representation to $id \in \mathcal{E}_I(X,X)$, so that

$$\eta^*\big(\gamma(X,X)\big) = id.$$

The bulk of category theory over $\mathcal{B}$ can be expressed in terms of this structure.

Comments. The notion of a locally small fibration is due to Bénabou (1975b) again. The corresponding pseudofunctors have been introduced by Penon (1974), under the name "locally internal categories"; see also Appendix of Johnstone 1977. Recently, locally small fibrations have been studied as $\underline{span}_{\mathcal{B}}$-enriched categories (namely, the representants of HOM'_E are spans in $\mathcal{B}$) in a series of papers by Betti and Walters (1987, 1989) and by Betti alone (1989).

And while locally small fibrations reappear so often, the notion of a globally small fibration may seem a bit dubious. Why did we choose to require the representability of OB_E, and not OB'_E? To produce an essentially surjective cartesian functor $OB_E \to E$ (for the role of $OB_{\mathcal{E}} \hookrightarrow \mathcal{E}$), one needs the axiom of choice, while the cartesian functor $OB'_E \to E{:}\ F \mapsto F(id)$ is canonical and surjective on objects.

Well, the point is that OB'_E tends to be very large. For instance, if $\mathbb{G}$ is a groupoid with two objects and one arrow in each hom-set, the class $OB'_{\nabla\mathbb{G}}(I)$ of cartesian functors $\nabla I \longrightarrow \nabla\mathbb{G}$ is proper for every inhabited set I. (The fibre $(\underline{Set}/\mathbb{G})_I$, on the other hand, has just 2^I objects!)

Moreover, if the goal of constructing a fibred class of objects is to realize a representation of $A \in |\mathcal{E}_I|$ by $\ulcorner A \urcorner \in \mathcal{B}(I,\Omega)$, so that $A \simeq \ulcorner A \urcorner^*\xi$, then it is reasonable to identify in this fibred class the isomorphic objects of fibres - as we did in OB_E - since they cannot be distinguished as inverse images of ξ anyway (unless a cleavage is given).

The unpleasant fact remains that some cocompleteness of $\mathcal{B}$ is needed to make a small fibration $\nabla\mathbb{C}$ globally small. The definition of globally small should perhaps be relaxed to the requirement that OB_E is just *weakly* representable, i.e. that there is a *weak representant* $\Omega_w \in |\mathcal{B}|$, equipped with a natural surjection $\nabla\Omega_w \twoheadrightarrow OB_E$. (A weak representant for $OB_{\nabla\mathbb{C}}$ is the object of objects C_0.) With this weaker notion of globally small, the following proposition would extend to "globally+locally small $\Leftrightarrow$ small". However, we shall need the stronger notion in chapter IV.

Proposition. Let $\mathcal{B}$ be a finitely complete category. If a cloven fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ is globally and locally small then it must be small: there is an internal category $\mathbf{\Omega}$ in $\mathcal{B}$ and a fibrewise equivalence $\nabla\mathbf{\Omega} \longrightarrow E$.

• With generic data of E denoted as above, we shall also write

$$\xi_i := \pi_i^*(\xi) \text{ for } \pi_i:\Omega^2 \longrightarrow \Omega,\ i \in 2;$$

$$\tilde{\xi}_i := \pi_i^*(\xi) \text{ for } \pi_i:\Omega^3 \longrightarrow \Omega,\ i \in 3;$$

$$\pi_{ij} := \langle \pi_i, \pi_j \rangle : \Omega^3 \longrightarrow \Omega^2,\ i,j \in 3.$$

Of course, $\Omega^2 := \Omega\times\Omega$, $\Omega^3 := \Omega\times\Omega\times\Omega$.

The hypothesis that E is globally and locally small means for every $X, Y \in |\mathcal{E}_I|$

$$\mathcal{E}_I(X, Y) \simeq \mathcal{E}_I(\ulcorner X \urcorner^*\xi, \ulcorner Y \urcorner^*\xi) \simeq \mathcal{B}/\Omega^2\big(\langle \ulcorner X \urcorner, \ulcorner Y \urcorner \rangle, \iota(\xi_0, \xi_1)\big).$$

The internal category $\mathbf{\Omega}$ is defined as follows:

$\Omega_0 := \Omega$;

$\Omega_1 := D(\xi_0, \xi_1)$;

$\langle \partial_0, \partial_1 \rangle := \iota(\xi_0, \xi_1) : \Omega_1 \longrightarrow \Omega_0 \times \Omega_0$;

$\eta:\Omega_0 \longrightarrow \Omega_1$ is the element of $\mathcal{B}/\Omega_0\times\Omega_0\big(\langle id, id\rangle, \iota(\xi_0, \xi_1)\big) \simeq \mathcal{E}_{\Omega_0}(\xi,\xi)$ which corresponds to id_ξ;

Ω_2, a pullback of ∂_0 and ∂_1, is also (isomorphic with and can be chosen to be equal to) the domain of

$$\iota(\tilde{\xi}_0, \tilde{\xi}_1) \underset{\Omega\times\Omega\times\Omega}{\times} \iota(\tilde{\xi}_1, \tilde{\xi}_2) = \big(\langle \partial_0, \partial_1 \rangle \times \Omega\big) \underset{\Omega\times\Omega\times\Omega}{\times} \big(\Omega \times \langle \partial_0, \partial_1 \rangle\big);$$

and then

$\mu:\Omega_2 \longrightarrow \Omega_1$ will be $p_{02} \circ \mu(\tilde{\xi}_0, \tilde{\xi}_1, \tilde{\xi}_2)$, where

$$\mu(\tilde{\xi}_0, \tilde{\xi}_1, \tilde{\xi}_2) : \iota(\tilde{\xi}_0, \tilde{\xi}_1) \times \iota(\tilde{\xi}_1, \tilde{\xi}_2) \longrightarrow \iota(\tilde{\xi}_0, \tilde{\xi}_2)$$

is as defined above, while

$$p_{02}: \pi_{02} \circ \iota(\tilde{\xi}_0, \tilde{\xi}_2) \longrightarrow \iota(\xi_0, \xi_1)$$

is obtained by pulling back π_{02} along $\iota(\xi_0, \xi_1)$.

It is obvious from the definitions that $\langle \partial_0, \partial_1 \rangle \eta = \langle id, id \rangle$ and $\langle \partial_0, \partial_1 \rangle \mu = \langle \partial_0 p_0, \partial_1 p_1 \rangle$, where p_0 is still the pullback of ∂_0 along ∂_1, p_1 of ∂_1 along ∂_0.

Let us check one of the nontrivial commutativity conditons required from the internal category $\mathbf{\Omega}$. Consider the arrow $«id, \eta\partial_1»:\Omega_1 \longrightarrow \Omega_2$, induced by $\partial_0 \circ \eta\partial_1 = \partial_1 \circ id$. It should satisfy

$$\mu \circ «id, \eta\partial_1» = id_{\Omega_1}$$

(which means "f∘id=f" in $\mathbf{\Omega}$). First note that we actually have

$$«id, \eta\partial_1» \in \mathcal{B}/\Omega^3\big(\langle \partial_0, \partial_1, \partial_1 \rangle, \iota(\tilde{\xi}_0, \tilde{\xi}_1) \times \iota(\tilde{\xi}_1, \tilde{\xi}_2)\big).$$

This arrow clearly comes from

$$\big\langle \langle id, \partial_1 \rangle, \langle \partial_0, \eta\partial_1 \rangle \big\rangle \in \mathcal{B}/\Omega^3\Big(\langle \partial_0, \partial_1, \partial_1 \rangle, \pi_{01}^*\big(\iota(\xi_0, \xi_1)\big)\Big)$$

$$\times\, \mathcal{B}/\Omega^3\Big(\langle \partial_0, \partial_1, \partial_1 \rangle, \pi_{12}^*\big(\iota(\xi_0, \xi_1)\big)\Big) \simeq$$

$$\mathcal{E}_{\Omega_1}(\partial_0^*\xi, \partial_1^*\xi) \times \mathcal{E}_{\Omega_1}(\partial_1^*\xi, \partial_1^*\xi).$$

In $\mathcal{E}_{\Omega_1}(\partial_0{}^*\xi,\partial_1{}^*\xi)$ the same arrow corresponds to $\langle id,\partial_1\rangle\in\mathcal{B}/\Omega^3$ and to $id\in\mathcal{B}/\Omega^2\big(\langle\partial_0,\partial_1\rangle,\ \iota(\xi_0,\xi_1)\big)$; to the latter corresponds $\gamma(\xi_0,\xi_1)$ by definition.

In $\mathcal{E}_{\Omega_1}(\partial_1{}^*\xi,\partial_1{}^*\xi)$, the same arrow corresponds to $\langle\partial_0,\eta\partial_1\rangle\in\mathcal{B}/\Omega^3$ and to $\eta\partial_1\in\mathcal{B}/\Omega^2\big(\langle\partial_1,\partial_1\rangle,\iota(\xi_0,\xi_1)\big)$; by the definition of η, this arrow must be $id\in\mathcal{E}_{\Omega_1}(\partial_1{}^*\xi,\partial_1{}^*\xi)$.

So when we compose these arrows in $\mathcal{E}_{\Omega_1}$, the result will be $\gamma(\xi_0,\xi_1)$.

The mapping

$$\mathcal{B}/\Omega^3\big(\langle\partial_0,\partial_1,\partial_1\rangle,\ \iota(\tilde\xi_0,\tilde\xi_1)\times\iota(\tilde\xi_1,\tilde\xi_2)\big)\longrightarrow\mathcal{B}/\Omega^3\big(\langle\partial_0,\partial_1,\partial_1\rangle,\ \iota(\tilde\xi_0,\tilde\xi_2)\big)$$

induced by the composition in $\mathcal{E}_{\Omega_1}$ is represented in $\mathcal{B}$ by composing with $\mu(\tilde\xi_0,\tilde\xi_1,\tilde\xi_2)$. (Such is the definition of $\mu(X,Y,Z)$.) This means that we have above actually concluded that the arrow

$$\mu(\tilde\xi_0,\tilde\xi_1,\tilde\xi_2)\circ\langle\!\langle id,\eta\partial_1\rangle\!\rangle\in\mathcal{B}/\Omega^3\big(\langle\partial_0,\partial_1,\partial_1\rangle,\ \iota(\tilde\xi_0,\tilde\xi_2)\big)$$

corresponds to $\gamma(\xi_0,\xi_1)\in\mathcal{E}_{\Omega_1}$. If we transpose $\mu(\tilde\xi_0,\tilde\xi_1,\tilde\xi_2)\circ\langle\!\langle id,\eta\partial_1\rangle\!\rangle$ along the adjunction

$$\mathcal{B}/\Omega^3\big(\langle\partial_0,\partial_1,\partial_1\rangle,\ \pi_{02}{}^*\big(\iota(\xi_0,\xi_1)\big)\big)\simeq\mathcal{B}/\Omega^2\big(\pi_{02}\langle\partial_0,\partial_1,\partial_1\rangle,\ \iota(\xi_0,\xi_1)\big)$$

by postcomposing $p_{02}:\pi_{02}\circ\iota(\tilde\xi_0,\tilde\xi_2)\longrightarrow\iota(\xi_0,\xi_1)$, the resulting arrow in $\mathcal{B}/\Omega^2$ will still correspond to $\gamma(\xi_0,\xi_1)\in\mathcal{E}_{\Omega_1}$. Hence

$$\mu\circ\langle\!\langle id,\eta\partial_1\rangle\!\rangle=id_{\Omega_1}.$$

So we defined $\Omega\in\underline{cat}_{\mathcal{B}}$. The cleavage preserving functor

$$F:\nabla\Omega\longrightarrow E$$

will be the one represented by $\big\langle\xi,\ \gamma(\partial_0{}^*\xi,\partial_1{}^*\xi)\big\rangle\in\underline{psh}_{\mathcal{E}}(\Omega)$. Using the cleavage, we first define

$$F_0:\nabla\Omega\longrightarrow E:\ u\longmapsto u^*\xi;$$

and then proceed as in 1.27. It remains to prove that F is a fibrewise equivalence.

F is fibrewise essentially surjective because ξ is a generic object: every $A\in|\mathcal{E}_K|$ is $A\simeq F(\ulcorner A\urcorner)$.

The vertical parts $F_K:(\mathcal{B}/\Omega)_K\longrightarrow\mathcal{E}_K$ are full and faithful because the natural isomorphism

$$\mathcal{B}/\Omega(k_0,\ k_1)=\mathcal{B}/\Omega\times\Omega\big(\langle k_0,k_1\rangle,\ \langle\partial_0,\partial_1\rangle\big)\simeq\mathcal{E}_K(k_0{}^*\xi,\ k_1{}^*\xi)=$$
$$=\mathcal{E}_K(F_K(k_0),\ F_K(k_1))$$

is realized by

$$\varphi\longmapsto\big(\varphi^*\big(\gamma(\partial_0{}^*\xi,\partial_1{}^*\xi)\big):k_0{}^*\xi\longrightarrow k_1{}^*\xi\big),$$

which is just

$$\langle id,\varphi\rangle\longmapsto\big(F_K\langle id,\varphi\rangle:F_K(k_0)\longrightarrow F_K(k_1)\big).$$

But a cartesian functor is full and faithfull <u>iff</u> its vertical components are.•

3. Definability.

Motivation. As everybody knows, *comprehension* is the assignment

$$\Lambda(x)\longmapsto\{x|\ \Lambda(x)\},$$

where Λ is any given *description*. In set theory, descriptions are just the definable classes, i.e. those given by a formula. The *comprehension principle* says that for every set X the class

$$\Lambda\cap X:=\{x\in X|\ \Lambda(x)\}$$

must be a set. *Every definable subclass of a set must be a set.*

The question is: *Which fibred classes should be considered as definable, so that the comprehension principle is satisfied in* <u>FIB</u>/$\mathcal{B}$*?* Certainly not all of them: e.g., the crible $R:=\big\{f\in|\underline{Set}/2|;\ im(f)\subseteq\{0\}\ \underline{or}\ im(f)\subseteq\{1\}\big\}$ is a nonrepresentable subfibration of a representable fibration. Bénabou (1985) offered an answer again. (But he avoided technicalities very consequently, and gave just a six lines long definition, in the glossary to his article. The explanations which follow here are completely apocryphal.)

For every object X of a fibred category $\mathcal{E}$ define $\mathcal{E}\cap X$ to be the full subcategory of $\mathcal{E}/X$ spanned by the cartesian arrows to X. The functor

$$\mathcal{E}\cap X\longrightarrow\mathcal{B}/EX:\ f\longmapsto Ef$$

is an equivalence of categories. Being thus essentially representable, the fibred category

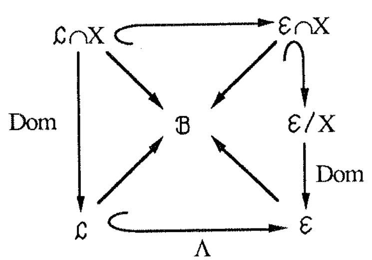
$$\mathcal{E}\cap X \longrightarrow \mathcal{B}$$

can be thought of as a "subset", a "small part" of $\mathcal{E}$. As we saw in 1.1, $\mathcal{E}$ is a lax colimit - "union" - of such small parts. The faithful cartesian functor

$$\theta_X : \mathcal{E}\cap X \hookrightarrow \mathcal{E}/X \longrightarrow \mathcal{E}$$

can be taken as a "canonical inclusion".

The discrete subfibrations of $\mathcal{E}$ are, of course, its *subclasses*. For every subclass $\Lambda:\mathcal{C} \hookrightarrow \mathcal{E}$, the pullback along θ_X gives the subcategory $\mathcal{C}\cap X \subseteq \mathcal{E}\cap X$, consisting of the cartesian arrows to X with the domain in $\mathcal{C}$.



Since $\mathcal{C} \longrightarrow \mathcal{B}$ is a discrete fibration, the fibration

$$\mathcal{C}\cap X \longrightarrow \mathcal{C} \longrightarrow \mathcal{B}$$

must be discrete too. The mapping $\mathcal{C}\cap X \hookrightarrow \mathcal{E}\cap X \longrightarrow \mathcal{B}/EX$ is therefore an injection, and we may assume (for simplicity) that $\mathcal{C}\cap X$ is a subfibration of $\mathcal{B}/EX$, and not of $\mathcal{E}\cap X$, as above.

Definition. A discrete subfibration $\mathcal{C}$ of a fibred category $\mathcal{E}$ is *definable* if for every object $X \in |\mathcal{E}|$, the crible

$$\mathcal{C}\cap X = \{ v \in |\mathcal{B}/EX| \mid v^*X \in |\mathcal{C}| \}.$$

is representable. (Or in terms of 13: A fibred subclass $\mathcal{C}$ of $\mathcal{E}$ is definable if all the fibred predicates $\mathcal{C}\cap X$ are.)

Examples. Let $\mathbb{H}$ be a complete Heyting algebra and $E:A \longrightarrow \mathbb{H}$ an $\mathbb{H}$-set (example II.1.54). An $\mathbb{H}$-*subset* is a subset $L \subseteq A$ such that $x \in L$ <u>implies</u> $x\restriction p \in L$ for all $p \in \mathbb{H}$. Such a subset L is definable <u>iff</u> every $\mathbb{H}$-subset $L\cap c = \{x \in L \mid x = c\restriction Ex\}$ contains a join, i.e. an element x_c such that $L\cap c = \{z \in A \mid z = x_c\restriction Ez\}$.

A discrete subfibration $\mathcal{C}$ of a family fibration $\nabla\mathcal{C}$ is a class of set-indexed families of objects of $\mathcal{C}$, closed under reindexing (i.e. with every $(C_j \mid j \in J)$, the reindexing $(C_{u(i)} \mid i \in I)$ must be in $\mathcal{C}$ too, for all functions $u:I \longrightarrow J$). $\mathcal{C}$ is definable <u>iff</u> for every object $X = (C_j \mid j \in J)$ there is a set $\iota_{\mathcal{C}}(X) \subseteq J$ such that exactly those reindexings of X are in $\mathcal{C}$ which consist of C_i, $i \in \iota_{\mathcal{C}}(X)$.

Fact. If $\mathcal{E}$ is a cloven globally small fibration (with a generic object ξ), its discrete subfibration $\mathcal{C}$ is definable <u>iff</u> $\mathcal{C}\cap\xi$ is representable. If Ω represents OB_E and if $\iota:D \rightarrowtail \Omega$ represents $\mathcal{C}\cap\xi$, then a definable subclass $\mathcal{C}$ must consist exactly of the inverse images of $\iota^*(\xi)$, with the cartesian arrows between them.

4. Constructive comprehension.

Motivation. The extent $\{x \in K \mid \varphi(x)\}$ of a predicate φ over a set K collects the elements of K on which φ is satisfied. The idea for a *constructive extent* is that it should collect the pairs $\langle x, p(x)\rangle$, where p(x) is a proof of $\varphi(x)$, i.e.

$$\{x \in K \mid \varphi(x)\} := \sum_{x \in K} \iota\varphi(x), \text{ where}$$

$$\iota\varphi(x) := \text{the set of proofs of } \varphi(x).$$

This extent is equipped with a canonical projection

$$\iota\varphi : \{x \in K \mid \varphi(x)\} \longrightarrow K : \langle x, p(x)\rangle \longmapsto x.$$

If there is at most one proof for every $\varphi(x)$, the projection $\iota\varphi$ is reduced to the inclusion $\{x \in K \mid \varphi(x)\} \subseteq K$. Otherwise, it can be regarded as representing a K-indexed set $\{\iota\varphi(x) \mid x \in K\}$.

In order to find a categorical presentation for the notion of a constructive extent, let us take up the paradigm of a variable category of predicates again: imagine that the fibre $\mathcal{E}_K$ consists of predicates over the "set" $K \in |\mathcal{B}|$, with proofs as arrows between them.

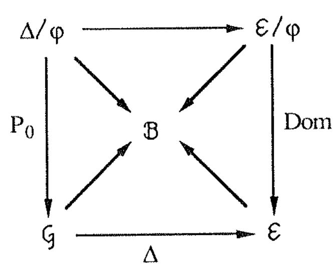
Observe first that the notion of definability has still a natural interpretation. A class-as-discrete-subfibration can be understood as *a class of predicates*. Discrete fibrations $\mathcal{L}\cap X$ (for $X \in |\mathcal{E}_K|$) could now better be written as

$$\Lambda\wedge\varphi : (\mathcal{B}/K)^o \longrightarrow 2 : (v{:}I \longrightarrow K) \mapsto \begin{cases} 1 & \underline{\text{if}}\ \forall y{:}I.\Lambda(v(y))\wedge\varphi(v(y)) \\ 0 & \underline{\text{otherwise}} \end{cases}$$

This logical picture seems open to various generalisations. But to make a link between the definability and the constructive extent, we shall make a detour from both.

If logic is to be fully constructive, the notion of a description cannot be reduced to definable classes any more: *a description must take the constructive proofs into account.* We shall look for descriptions not just among discrete subfibrations $\Lambda{:}\mathcal{L} \hookrightarrow \mathcal{E}$, but among more general *diagrams* $\Delta{:}\mathcal{G} \longrightarrow \mathcal{E}$.

For a diagram/description $\Delta{:}\mathcal{G} \longrightarrow \mathcal{E}$ and a predicate $\varphi \in |\mathcal{E}_K|$, the derivations of φ from Δ are contained in the comma category Δ/φ.



A logical picture of a fibration "measuring" the Δ-part of $\{x \in K|\ \varphi(x)\}$ will be

$$\Delta\to\varphi : (\mathcal{B}/K)^o \longrightarrow \underline{\mathrm{SET}} : (v{:}I \longrightarrow K) \mapsto \{\text{proofs of } \forall y \in I.\Delta(v(y)) \to \varphi(v(y))\}$$

The idea is that a "proof of $\forall y \in I.\Delta(v(y)) \to \varphi(v(y))$" should be a *cocone*: a derivation of φ from Δ should respect the proofs contained in the description Δ (i.e. commute with the arrows of the diagram Δ).

This idea suggests which diagrams could be considered as descriptions on a category of predicates.

Definitions. Let $E{:}\mathcal{E} \longrightarrow \mathcal{B}$ and $G{:}\mathcal{G} \longrightarrow \mathcal{B}$ be fibrations. A *description* (on E) is a cartesian functor $\Delta{:}\mathcal{G} \longrightarrow \mathcal{E}$, satisfying the following conditions:

i) for every $A \in |\mathcal{G}|$ and every E-cartesian arrow $\vartheta \in |\mathcal{E}/\Delta A|$ there must be a G-cartesian arrow $\vartheta' \in |\mathcal{G}/A|$ such that $\vartheta = \Delta\vartheta'$;
ii) for every $v \in \mathcal{B}(I,K)$ and every cocone $\beta{:}\Delta_K \longrightarrow X$ in $\mathcal{E}_K$ there is a unique cocone $\tilde{\beta} : \Delta_I \longrightarrow X$ over v, such that for every G-cartesian arrow $\vartheta' \in \mathcal{G}_v(v^*A, A)$ holds

$$\tilde{\beta}_{v^*A} = \beta_A \circ \Delta\vartheta'.$$

(As before, $\Delta_I{:}\mathcal{G}_I \longrightarrow \mathcal{E}_I$ denote restrictions of Δ on fibres.)

A description Δ is *comprehensible* if for every object $X \in |\mathcal{E}_K|$ the discrete fibration

$$\Delta\to X{:}\ (\mathcal{B}/K)^o \longrightarrow \underline{\mathrm{SET}}{:}\ (v{:}I \longrightarrow K) \mapsto \mathrm{COCONE}_v(\Delta_I, X)$$
$$(\simeq \mathrm{COCONE}_I(\Delta_I, v^*X))$$

is representable. (To define the arrow part of $\Delta\to X$, use condition (ii).)

Fibration E is *comprehensive* if the functor $id{:}\mathcal{E} \longrightarrow \mathcal{E}$ is a comprehensible description.

Comments. It follows from (i) that the diagram $v^*\Delta_K : \mathcal{G}_K \longrightarrow \mathcal{E}_I$ is contained in the diagram $\Delta_I{:}\mathcal{G}_I \longrightarrow \mathcal{E}_I$. Condition (ii) tells that the cocone $\beta'{:}\ v^*\Delta_K \longrightarrow X$, obtained from β, has a unique extension $\tilde{\beta} : \Delta_I \longrightarrow X$.

What do these conditions mean in the logical perspective of a fibration-as-category-of-predicates? Condition (i) just demands that a *description should be stable under substitution.* This condition generalizes to arbitrary cartesian functors $\Delta{:}\mathcal{G} \longrightarrow \mathcal{E}$ the requirement that a subfibration $\mathcal{L} \hookrightarrow \mathcal{E}$ should be stable under inverse images. Another way to express (i) is to say that Δ must be a ϑ-fibration over $\mathcal{E}$. Condition (ii), on the other hand, tries to capture the idea that a *description should be uniformly applicable to the elements of all sets*: e.g., $\Delta(x)$:="x is a red apple" should allow x to be anything, and pick the red apples absolutely everywhere. This is at least in part conveyed by demanding that the derivations/cocones from fibrewise parts of Δ are invariant under inverse images - that nothing from Δ is lost when substitutions are performed.

(Some people will undoubtedly prefer to forget this "explanation", and regard condition (ii) just as necessary to define the arrow part of $\Delta\to X$.)

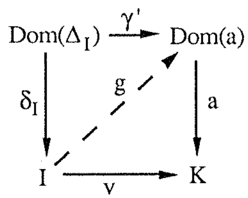
Notation, terminology. Generic letters for representants are again ι and D: a representant of $\Delta \to X$ will be $\iota_\Delta X : D_\Delta X \longrightarrow EX$. For Δ=id we omit the subscript and write $\iota X{:}DX \longrightarrow EX$. The arrows ιX are *extents* of X.

$\underline{\mathrm{CFIB}}/\mathfrak{B}$ will be the category of comprehensive fibrations with the extent preserving cartesian functors (i.e. $F{:}E' \longrightarrow E$ must satisfy $\iota^E(FX) \simeq \iota^{E'}X$).

Facts. If a locally small fibration $E{:}\mathcal{E} \longrightarrow \mathfrak{B}$ has fibrewise terminal objects $\top{:}\mathfrak{B} \longrightarrow \mathcal{E}$, then it is comprehensive, with the extents $\iota Z{:}=\iota(\top EZ, Z)$.

A fibrewise cartesian closed fibration (II.2.1) is locally small iff it is comprehensive: a representant of the fibred class HOM(X,Y) is $\iota(X,Y){:}=\iota(X \to Y)$.

Examples. A diagram $\Delta_I{:}\ \mathcal{G}_I \longrightarrow \mathfrak{a}{\downarrow}I$ over a fibre of an arrow fibration $\mathfrak{a}/\mathfrak{B} \longrightarrow \mathfrak{B}$ (cf. II.4) can be regarded as a cocone δ_I: $\mathrm{Dom}(\Delta_I) \longrightarrow I$. If a description $\Delta{:}\ \mathcal{G} \longrightarrow \mathfrak{a}/\mathfrak{B}$ is such that each δ_I is a colimit cocone in $\mathfrak{B}$, then Δ is comprehensible, and each $a \in \mathfrak{a}$ represents $\Delta \to a$. • Namely, the cocones γ: $\Delta_I \longrightarrow a$ over $v \in \mathfrak{B}(I,K)$ in $\mathfrak{a}/\mathfrak{B}$ are in one-to-one correspondence with the cocones γ': $\mathrm{Dom}(\Delta_I) \longrightarrow \mathrm{Dom}(a)$ in $\mathfrak{B}$ such that $a \circ \gamma'_L = u \circ \delta_{I,L}$ for every $L \in \mathrm{Dom}(\Delta_I)$.



But if δ_I is a colimit of $\mathrm{Dom}(\Delta_I)$ in $\mathfrak{B}$, the cocones γ' are in one-to-one correspondence with the arrows $g{:}I \longrightarrow \mathrm{Dom}(a)$. Since a colimit cocone is jointly epi, if $a \circ g \circ \delta_{I,L} = a \circ \gamma'_L = v \circ \delta_{I,L}$ holds for all $L \in \mathrm{Dom}(\Delta_I)$ then $a \circ g = v$. So we have a correspondence $\mathrm{COCONE}_v(\Delta_I,a) \simeq \mathfrak{B}/K(v,a)$.•

In particular, every arrow fibration with intrinsic terminal objects is comprehensive.

The empty class $\emptyset \hookrightarrow \mathcal{E}$ is comprehensible for every $\mathcal{E}$: the representants are $\iota_\emptyset X{:}=\mathrm{id}_X$, for $X \in |\mathcal{E}|$, • because $\mathrm{COCONE}_v(\emptyset,X)=1$•. On the other hand, $\emptyset$ can never

be definable, • since a crible $\emptyset \cap X=\emptyset$ cannot be representable: it should contain at least a representant•.

If E is a discrete fibration, $\mathrm{COCONE}_v(\mathcal{E}_I,X) \neq \emptyset$ iff $\mathcal{E}_I$ has at most one object. The representable fibrations $\nabla J{:}\underline{\mathrm{Set}}/J \longrightarrow \underline{\mathrm{Set}}$ are comprehensive: the extents are $\iota u{:}\emptyset \longrightarrow \mathrm{Dom}(u)$.

Let $\mathcal{G}$ be a fibred category over Set, and consider a diagram $\Delta{:}\ \mathcal{G} \longrightarrow \underline{\mathrm{Set}}/\mathcal{C}$. Its vertical part $\Delta_I{:}\ \mathcal{G}_I \longrightarrow \mathcal{C}^I$ is a family $(\Delta_{I,i}{:}\ \mathcal{G}_I \longrightarrow \mathcal{C} \mid i \in I)$ of diagrams in $\mathcal{C}$. The isomorphisms $\Delta_{I,i}(v^*A) \simeq \Delta_{K,v(i)}(A)$ (for all $v{:}I \longrightarrow K$, $i \in I$, $A \in |\mathcal{G}_K|$) tell that Δ is cartesian. Condition (i) now requires that (the image of) each $\Delta_{I,i}$ is closed under isos. Condition (ii) tells that every $\Delta_{K,v(i)}{:}\mathcal{G}_K \longrightarrow \mathcal{C}$ must have a unique extension to $\Delta_{I,i}{:}\mathcal{G}_I \longrightarrow \mathcal{C}$.

The most "uniform" instance of a description on $\underline{\mathrm{Set}}/\mathcal{C}$ is the cartesian functor $\Phi{:}\ \mathcal{F} \times \underline{\mathrm{Set}} \longrightarrow \underline{\mathrm{Set}}/\mathcal{C}$ consisting of

$$\Phi_I := (F{:}\mathcal{F} \longrightarrow \mathcal{C} \mid i \in I),$$

for some category $\mathcal{F}$ and a functor F, (the image of) which is closed under isos. If the classes

$$Q_C := \mathrm{COCONE}(F,C)$$

are small for all $C \in |\mathcal{C}|$, Φ is comprehensible, with representants

$$\iota_\Phi X : \sum_{k \in K} Q_{C_k} \longrightarrow K$$

for $X = (C_k \mid k \in K)$. If $A{:}=\underline{\lim}F$ exists in $\mathcal{C}$, then $Q_C \simeq \mathcal{C}(A,C)$, and $\iota_\Phi X$ is just a representant of the fibred class $\mathrm{HOM}'_{\nabla\mathcal{C}}(A,X)$.

Just apparently different is the description $\Psi : \underline{\mathrm{Set}}/\mathcal{F} \longrightarrow \underline{\mathrm{Set}}/\mathcal{C}$, where

$$\Psi_I := (F \circ \pi_i{:}\mathcal{F}^I \longrightarrow \mathcal{C} \mid i \in I).$$

It is comprehensible under the same condition and with the same representants as Φ. If we take $\mathcal{F} := \mathcal{C}$, and $\Psi :=$ id, we see that $\nabla\mathcal{C}$ is a comprehensive fibration iff the classes

$$Q_C := \mathrm{COCONE}(\mathcal{C},C)$$

are small for all $C \in |\mathcal{C}|$. If $\mathcal{C}$ has a terminal object $\top$, then of course $Q_C \simeq \mathcal{C}(\top,C)$. A locally small category $\mathcal{C}$ with a terminal object induces a comprehensive family fibration. The extent of a family of objects $X = (C_k \mid k \in K)$ will be

$$\iota X = \sum_{k\in K}\mathcal{C}(\top, C_k) \longrightarrow K.$$

5. Associated arrow fibrations.

Fact. For every definable subclass $\mathcal{L} \hookrightarrow \mathcal{E}$ and every comprehensible description $\Delta:\mathcal{G} \longrightarrow \mathcal{E}$, the classes of representants $\iota_\square \subseteq \mathcal{B}$, $\square \in \{\Delta, \mathcal{L}\}$, are stable under pullbacks.

• $\mathcal{B}/K\big(w, v^*(\iota_\square X)\big) \simeq \mathcal{B}/I\big(vw, \iota_\square X\big) \simeq$

$$\simeq \left\{\begin{array}{c} COCONE_{vw}(\Delta_M, X) \simeq COCONE_w(\Delta_M, v^*X) \\ \mathcal{L}\cap X(vw) \simeq \mathcal{L}\cap v^*X(w) \end{array}\right\} \simeq$$

$$\simeq \mathcal{B}/K\big(w, \iota_\square(v^*X)\big)$$

<u>implies</u> that $v^*(\iota_\square X) \simeq \iota_\square(v^*X)$ holds for $\square \in \{\Delta, \mathcal{L}\}$.•

Consequence. Given a fibration E, to each comprehensible description Δ and each definable class $\mathcal{L}$ arrow fibrations $\nabla\iota_\square : \iota_\square/\mathcal{B} \longrightarrow \mathcal{B}$, $\square \in \{\Delta, \mathcal{L}\}$, are naturally associated. An associated arrow fibration is like a shadow which a subclass or diagram casts on the base category: $\nabla\iota_\square$ tells how $\mathcal{B}$ sees $\square$.

The fibration $\nabla\iota : \iota/\mathcal{B} \longrightarrow \mathcal{B}$ spanned in $\mathcal{B}/\mathcal{B}$ by the extents of a comprehensive fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ is called *the extent of* E, and denoted $\iota E : \iota\mathcal{E} \longrightarrow \mathcal{B}$. It is, in a sense, *the best approximation of* E *by an arrow fibration*: each predicate $\varphi \in |\mathcal{E}|$ is approximated by the set $\iota\varphi$ of its proofs. How accurate is this approximation? Which parts of the structure of E are preserved in it? When is the logic carried by a category of predicates E *extensional*, in the sense that every predicate in it is completely determined by the set of its proofs? - We shall answer these questions in section 4.

3. Aspects of comprehension

1. Homogenous descriptions.

We begin by considering descriptions which satisfy a bit more than required by condition (ii) (from definition 2.4).

Definition. A description $\Delta:\mathcal{G} \longrightarrow \mathcal{E}$ is *homogenous* if for every $X \in |\mathcal{E}_K|$ and $v \in \mathcal{B}(I,K)$, every cocone $\beta : v^*\Delta_K \longrightarrow X$ over v can be extended in a unique way to a cocone $\tilde{\beta} : \Delta_I \longrightarrow X$ over v.

Proposition. Let $\Delta:\mathcal{G} \longrightarrow \mathcal{E}$ be a description on a locally small fibration $E:\mathcal{E} \longrightarrow \mathcal{B}$. If $\mathcal{B}$ is complete (to the maximal size of fibres $\mathcal{G}_I$, $I \in |\mathcal{B}|$), Δ is comprehensible.

• For every $X \in |\mathcal{E}_K|$, the diagram $\Delta_K:\mathcal{G}_K \longrightarrow \mathcal{E}_K$ induces diagram

$$\Psi:(\mathcal{G}_K)^o \longrightarrow \mathcal{B}/K : \quad A \longmapsto \iota(\hat{A}, X),$$
$$\big(f:A \to B\big) \longmapsto \iota(\hat{f}, X) \in \mathcal{B}/I\big(\iota(\hat{B},X), \iota(\hat{A},X)\big)$$

where $\hat{x} := \Delta_I(x)$. The arrow $\iota(\hat{f}, X)$ is defined to correspond by the representation of the fibred class $HOM_E(\hat{A}, X)$ to

$$\gamma(\hat{B},X)\circ\big(\iota(\hat{B},X)\big)^*(\hat{f}) \in \mathcal{E}_{D(\hat{B},X)}\Big(\big(\iota(\hat{B},X)\big)^*\hat{A},\ \big(\iota(\hat{B},X)\big)^*X\Big).$$

Using the correspondence $\mathcal{E}_v(v^*\hat{A}, X) \simeq \mathcal{B}/K\big(v, \iota(\hat{A},X)\big)$ for $v \in \mathcal{B}(I,K)$, it is routine to check that the cocones $v^*\Delta_K \longrightarrow X$ over v in $\mathcal{E}$ bijectively correspond to the cones $v \longrightarrow \Psi$ in $\mathcal{B}/K$. It follows that

$$COCONE_v(v^*\Delta_I, X) \simeq \mathcal{B}/K(v, \varprojlim \Psi)$$

holds naturally in v. But the hypothesis that Δ is homogenous means that the transformation

$$COCONE_v(\Delta_I, X) \longrightarrow COCONE_v(v^*\Delta_K, X) : \beta \longmapsto \beta\restriction v^*\Delta_K,$$

natural in v, is an isomorphism. Hence a natural iso

$$COCONE_v(\Delta_I, X) \simeq \mathcal{B}/I(v, \varprojlim \Psi),$$

which means that we can take

$$\iota_\Delta X := \varprojlim \Psi.\bullet$$

2. Descriptions with colimits.

Consider a fixed diagram $\Delta:\mathcal{G}\to\mathcal{E}$ on a fibred category $E:\mathcal{E}\to\mathcal{B}$. Suppose that $\mathcal{E}$ has the fibrewise colimits to the maximal size of all $\mathcal{G}_I$, $I\in|\mathcal{B}|$ (regarded as diagram schemes). (In other words, for every $J,K\in|\mathcal{B}|$, every diagram $F:\mathcal{G}_J\to\mathcal{E}_K$ has a colimit $\gamma: F\to\varinjlim F$ in $\mathcal{E}_K$, and for every $v\in\mathcal{B}(I,K)$, the cocone $v^*\gamma: v^*F\to v^*(\varinjlim F)$ is a colimit in $\mathcal{E}_I$. Cf. II.2.) If Δ now satisfies condition (i) (from definition 2.4), then condition (ii) is equivalent with the existence of a cartesian section

$$\varinjlim\Delta:\mathcal{B}\to\mathcal{E}: I\mapsto\varinjlim\Delta_I.$$

Of course, the axiom of choice is necessary for the then-direction. If we want to avoid the axiom of choice, we can consider the subcategory $\mathcal{L}_\Delta\hookrightarrow\mathcal{E}$, consisting of *all* the colimits $\varinjlim\Delta_I$ and all cartesian arrows between them. The diagram Δ (with colimits as above) is now a description iff it satisfies condition (i) and $\mathcal{L}_\Delta\hookrightarrow\mathcal{E}$ is a subfibration. ($\mathcal{L}_\Delta$ is a description on $\mathcal{E}$ as soon as it is its subfibration. $\mathcal{L}_\Delta$ is then equivalent with $\mathcal{B}$.)

A diagram Δ, characterized in any of these ways will be called *description with colimits*.

A description with colimits $\Delta:\mathcal{G}\to\mathcal{E}$ is comprehensible iff the corresponding description $\mathcal{L}_\Delta\hookrightarrow\mathcal{E}$ is (• since there is one-to-one correspondence between the cocones over some v from Δ and those from $\mathcal{L}_\Delta$•). In particular, if the category $\mathcal{E}$ has fibrewise terminal objects, then it is comprehensive iff the subfibration $\mathcal{T}\hookrightarrow\mathcal{E}$, spanned by the terminal objects, is comprehensible (• because the terminal objects are just the colimits of whole fibres•).

Applying the axiom of choice, for a description with colimits Δ, the discrete fibrations $\Delta\to X$ can equivalently be defined

$$\Delta\to X:(\mathcal{B}/EX)^o\to\underline{SET}:(v:I\to EX)\mapsto\mathcal{E}_v(\varinjlim\Delta_I,\ X)$$

(independently of the choice of representants $\varinjlim\Delta_I$), or

$$\Delta\to X:\varinjlim\Delta/X\to\mathcal{B}/EX: f\mapsto Ef.$$

Δ will be comprehensible iff every $\Delta\to X$ is isomorphic with a discrete fibration

$$«\iota_\Delta X»:\mathcal{B}/D_\Delta X\to\mathcal{B}/EX: v\mapsto\iota_\Delta X\circ v.$$

Putting all the $\Delta\to X$ together, discrete fibration

$$\Delta/\mathcal{E}:\varinjlim\Delta/\mathcal{E}\to\mathcal{B}/E:\langle I,\ f:\varinjlim\Delta_I\to X\rangle\mapsto\langle Ef: I\to EX,\ X\rangle$$

is obtained, which must be isomorphic to

$$«\iota_\Delta»:\mathcal{B}/D_\Delta\to\mathcal{B}/E:\langle v: I\to D_\Delta X,\ X\rangle\mapsto\langle\iota_\Delta X\circ v: I\to EX,\ X\rangle.$$

Proposition. Consider a description with colimits $\Delta:\mathcal{G}\to\mathcal{E}$, given together with a cartesian functor $\varinjlim\Delta:\mathcal{B}\to\mathcal{E}$. The following statements are equivalent.

a) Δ is comprehensible.

b) There is a right adjoint $D_\Delta:\mathcal{E}\to\mathcal{B}$ of $\varinjlim\Delta$ and a natural transformation $\iota_\Delta: D_\Delta\to E$ such that

$$\iota_\Delta X\circ f'=Ef$$

holds for every $f\in\mathcal{E}(\varinjlim\Delta_I, X)$ and its transpose $f'\in\mathcal{B}(I, D_\Delta X)$.

c) There is a right adjoint R to the functor

$$L:\mathcal{B}/\mathcal{B}\to E/\mathcal{B}:(v:I\to K)\mapsto\langle\varinjlim\Delta_I,\ v\rangle.$$

This adjointness is cartesian with respect to the functors $\nabla\mathcal{B}$ and Fam(E).

• (a)⇒(b): $$\mathcal{E}(\varinjlim\Delta_I, X)=\bigcup_{v\in\mathcal{B}(I,EX)}\mathcal{E}_v(\varinjlim\Delta_I, X)\simeq\bigcup_{v\in\mathcal{B}(I,EX)}\mathcal{B}/EX(v,\iota_\Delta X)=\mathcal{B}(I, D_\Delta X).$$

The statement "$\iota_\Delta X\circ f'=Ef$" just means that the adjointness isomorphism $\varinjlim\Delta/\mathcal{E}\simeq\mathcal{B}/D_\Delta$ commutes with the functors $\Delta/\mathcal{E}$ and $«\iota_\Delta»$ defined above.

(b)⇒(c): $R(X, w:EX\to K):=w\circ\iota_\Delta X$.

The transpose of

$$\langle f,u\rangle\in E/\mathcal{B}(Lv,\ \langle X,w\rangle)\ \text{(i.e. } f\in\mathcal{E}(\varinjlim\Delta_I, X)\text{, such that } w\circ Ef=u\circ v)$$

is

$$\langle f', u\rangle \in \mathfrak{B}/\mathfrak{B}(v,\ R(X,w)).$$

(c)$\Rightarrow$(a): $\mathcal{E}(\varinjlim \Delta_I, X) \simeq E/\mathfrak{B}\big(\langle \varinjlim \Delta_I, id_I\rangle, \langle X, id_{EX}\rangle\big) \simeq \mathfrak{B}/\mathfrak{B}\big(id_I,\ \iota_\Delta X\big)$.

Since the adjointness is cartesian, every two transposes must be projected to the same arrow. The correspondence above restricts to

$$\mathcal{E}_v(\varinjlim \Delta_I, X) \simeq (\mathfrak{B}/\mathfrak{B})_v\big(id_I,\ \iota_\Delta X\big) = \mathfrak{B}/K(v,\ \iota_\Delta X).\bullet$$

3. Comprehension structures.

Let $\Delta:\mathcal{G} \longrightarrow \mathcal{E}$ be a description. Assuming the axiom of choice, the property that Δ is comprehensible can be expressed as a structure. As one might have noticed in the preceding proposition, whenever we choose for each $X \in |\mathcal{E}|$ one representant $\iota_\Delta X : D_\Delta X \longrightarrow EX$ of $\Delta \to X$, the construct $D_\Delta : \mathcal{E} \longrightarrow \mathfrak{B}$ becomes functorial, and ι_Δ appears as a natural transformation $D_\Delta \longrightarrow E$.[1] • The D_Δ-image of $f \in \mathcal{E}(X,Y)$, is the arrow $D_\Delta f \in \mathfrak{B}/EY(Ef \circ \iota X, \iota Y)$ which corresponds to $f \circ \varepsilon_X \in \mathrm{COCONE}_{Ef\circ\iota X}(\Delta_{D_\Delta X}, Y)$; the generic arrow $\varepsilon_X \in \mathrm{COCONE}_{\iota X}(\Delta_{D_\Delta X},\ X)$ is the one that corresponds to $id \in \mathfrak{B}/EX(\iota X, \iota X)$.•

> A *comprehension transformation* of a comprehensive fibration E is a natural transformation
>
> $$\iota : D \longrightarrow E : \mathcal{E} \longrightarrow \mathfrak{B}$$
>
> which consists of extents $\iota X : DX \longrightarrow EX$.

On the other hand, by the couniversal property of comma categories (II.1.76), each natural transformation $\varphi: G \longrightarrow H : \mathcal{A} \longrightarrow \mathcal{C}$ corresponds to a unique functor $«\varphi»: \mathcal{A} \longrightarrow \mathcal{C}/\mathcal{C}$, such that $arr * «\varphi» = \varphi$, where

$$arr : Dom \longrightarrow Cod : \mathcal{C}/\mathcal{C} \longrightarrow \mathcal{C}$$

[1]This is not true for the representants induced by a definable class. Analogous "definability structures" cannot be obtained in this way.

has the components $arr_f := f$. Hence another presentation of the comprehension structure:

> A *comprehension functor* of a comprehensive fibration E is
>
> $$«\iota» : \mathcal{E} \longrightarrow \mathfrak{B}/\mathfrak{B} : X \longmapsto \iota X,\ \big(f: X \longrightarrow Y\big) \longmapsto \langle Df, Ef\rangle.$$

It follows from fact 2.5 that «ι» is a cartesian functor. The extent fibration $\iota E : \iota\mathcal{E} \longrightarrow \mathfrak{B}$ is (the subfibration of $\mathfrak{B}/\mathfrak{B}$ equivalent to) the image of «ι».

The morphisms in the category of fibrations over $\mathfrak{B}$ equipped with comprehension transformations would be, of course, the cartesian functors preserving ι (i.e. $F: E' \longrightarrow E$ such that $\iota^E * F = \iota^{E'}$). If $\nabla\mathfrak{B}$ is a fibration, it is the terminal object in this category. (Its comprehension transformation is $arr : Dom \longrightarrow Cod : \mathfrak{B}/\mathfrak{B} \longrightarrow \mathfrak{B}$.) The comprehension functors are the terminal arrows.

The question now arises: *Which transformations* $F \longrightarrow E$, *which functors* $\mathcal{E} \longrightarrow \mathfrak{B}/\mathfrak{B}$ *represent comprehension on a fibration* E*?* When E has some additional structure/properties, there are intrinsic characterisations of its comprehension structure(s) - by simple adjunctions. In terms of these characterisations - in special situations which they cover - both kinds of comprehension structure have actually been considered before. In his seminal paper about hyperdoctrines (1970), Lawvere introduced comprehension functors using coproducts and terminal objects. By means of terminal objects only, Ehrhard (1988, 1989) described comprehension transformations, and used them in his interpretation of the theory of constructions - although without any connection with the concept of comprehension, and under a different name. The following characterisations show that our notion of comprehension restricts to these two.

4. Ehrhard's comprehension.

Proposition. For a fibration $E:\mathcal{E} \longrightarrow \mathfrak{B}$ the statements below are related as follows:

$$(c) \Rightarrow (a) \overset{(AC)}{\Rightarrow} (b) \overset{(AC)}{\Rightarrow} (c) \Rightarrow (b)$$

a) E has fibrewise terminal objects and it is comprehensive.

b) There are functors S and C such that

$$E \dashv S \dashv C : \mathcal{E} \to \mathcal{B},$$

and S is full and faithful.

c) There are functors $\top$ and D such that

$$E \dashv \top \dashv D : \mathcal{E} \to \mathcal{B},$$

$\varepsilon^E : E\top \to id$ and $\eta^D : id \to D\top$ are identities ($E\top = D\top = id$), and

$$\iota = D*\eta^E = E*\varepsilon^D.$$

(Remember that * denotes the horizontal composition of 2-cells: II.1.74.)

• (c)⇒(a): (1) The right transpose of $f \in \mathcal{E}(\top I, X)$ is always $f' := D(f)\circ\eta^D_I \in \mathcal{B}(I, DX)$; since $\eta^D_I = id_I$, $f' = D(f)$.

(2) Since $\varepsilon^E = id$, certainly $\eta^E * \top = id$, and therefore $\eta^E_X \circ f = \top E(f) \circ \eta^E_{\top I} = \top E(f)$.

It follows that

$$\iota X \circ f' = D(\eta^E_X)\circ f' \overset{(1)}{=} D(\eta^E_X)\circ D(f) = D(\eta^E_X \circ f) \overset{(2)}{=} D\top E(f) = E(f).$$

Applying proposition 2, we conclude that E is comprehensive.

(b)⇒(c): $\top$ is obtained from S using corollary II.1.78, by lifting the natural iso $id \to ES$. D is obtained from C using the following diagram:

$$\begin{array}{ccccccc} C\top I & \xrightarrow{Cf} & CX & \xrightarrow{Cg} & CY & \xrightarrow{Ch} & C\top J \\ \varphi_{\top I}\big\downarrow\simeq & & \varphi_X\big\downarrow = & & \varphi_Y\big\downarrow = & & \varphi_{\top J}\big\downarrow\simeq \\ D\top I = I & \xrightarrow[Df]{} & DX & \xrightarrow[Dg]{} & DY & \xrightarrow[Dh]{} & D\top J = J \end{array}$$

where isomorphism $\varphi_{\top I} : C\top I \to CSI \to I$ is obtained by composing the inverse of the unit $\eta^C_I : I \to CSI$ and the C-image of the iso $\chi_I : \top I \to SI$, obtained by lifting $id \to ES$. Hence, for $X, Y \notin Im(\top)$, we have

$$D(q) := \begin{cases} \top^{-1}(q) & \underline{\text{for}}\ q : \top I \to \top J \\ \varphi_{\top J}\circ C(q) & \underline{\text{for}}\ q : Y \to \top J \\ C(q)\circ(\varphi_{\top I})^{-1} & \underline{\text{for}}\ q : \top I \to X \\ C(q) & \underline{\text{for}}\ q : X \to Y. \end{cases}$$

It is easy to see from the constructions that $E \dashv \top \dashv D$, and that ε^E and η^D are identities. The equation $D*\eta^E = E*\varepsilon^D$ then follows from lemma 43.•

Remark. Perhaps the simplest view of a comprehensive fibration E with terminal objects is that it *is* a triple $E \dashv S \dashv C$ as in (b) above. This is the structure used by Ehrhard. In this form it is obvious that comprehensive fibrations are closed under the composition; that $F \in \underline{FIB}/E$ is comprehensive <u>iff</u> it is comprehensive as $F \in \underline{FIB}/\mathcal{E}$ <u>and</u> its extents are E-vertical; and so on. (• Use lemma II.2.4 for the <u>if</u>-direction.•) For comprehensive fibrations as defined in 2.4, these and similar facts require much longer arguments.

Examples. Let $\mathbb{C}$ be a small category with a terminal object 1 and an initial object 0. Then there is a comprehensive fibration $\Lambda \dashv \Delta \dashv \Gamma : \underline{Set}^{\mathbb{C}^o} \to \underline{Set}$ defined

$$\Lambda : F \mapsto F(0),$$
$$\Delta : A \mapsto \big(\bar{A} : \mathbb{C}^o \to \underline{Set} : X \mapsto A,\ f \mapsto id_A\big),$$
$$\Gamma : F \mapsto F(1),$$
$$\iota F := F(0 \to 1).$$

On the other hand, for a set C there are $\Sigma \dashv \Delta \dashv \Pi : \underline{Set}^C \to \underline{Set}$:

$$\Sigma : F \mapsto \sum_{x\in C} F(x),$$
$$\Delta : A \mapsto \big(\bar{A} : C \to \underline{Set} : x \mapsto A\big),$$
$$\Pi : F \mapsto \prod_{x\in C} F(x).$$

Σ is a discrete fibration, and a comprehensive one, as we mentioned in examples 2.4: the extents are $\iota F : \emptyset \hookrightarrow \Sigma F$. But Δ is not full, not a terminal objects functor, and the triple $\Sigma \dashv \Delta \dashv \Pi$ has nothing to do with the comprehension.

It might be helpful to mention the trivial but paradigmatic example of comprehensive arrow fibrations once again. The comprehension structre on them now becomes

$$Cod \dashv Ids \dashv Dom : \mathcal{a}/\mathcal{B} \to \mathcal{B}$$

(where Ids: $X \mapsto id_X$).

Lemmas. Consider functors $E \dashv T \dashv D$; denote the data of the first adjointness by η^E, ε^E, the data of the second one by η^D, ε^D.

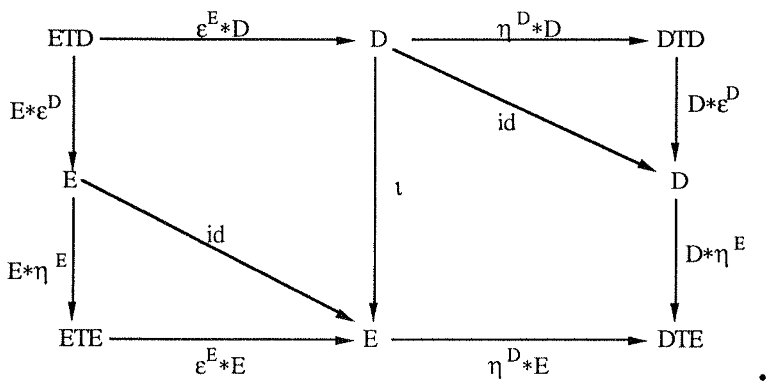
41. $id \simeq DT$ iff T is full and faithful iff $ET \simeq id$[2].

42. D is full and faithful iff $TD \simeq id$ iff $id \simeq TE$ iff E is full and faithful.

43. If T is full and faithful, then $D*\eta^E \simeq E*\varepsilon^D$. More precisely, there is a natural transformation $\iota : D \longrightarrow E$, such that

$$D*\eta^E = (\eta^D*E)\circ\iota,$$
$$E*\varepsilon^D = \iota\circ(\varepsilon^E*D).$$

• Define $\iota_X : DX \longrightarrow EX$ by the requirement that $T(\iota_X) = \eta^E_X \circ \varepsilon^D_X$. The equalities then follow by chasing the diagram:



•

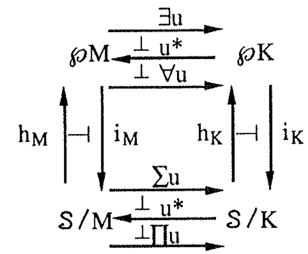
[2] $X \simeq Y$ means "$\exists f : X \simeq Y$". Lemma 1.3. in Johnstone-Moerdijk 1989 tells that any natural isomorphism $id \simeq DT$ forces the unit of the adjunction $T \dashv D$ to be an isomorphism.

5. Lawvere's comprehension.

Hyperdoctrines bis. Each topos $\mathcal{S}$ gives rise to two hyperdoctrines:

$$\nabla\mathcal{S} : \mathcal{S}^o \longrightarrow \underline{CAT} : K \mapsto \mathcal{S}/K, \text{ and}$$
$$\wp : \mathcal{S}^o \longrightarrow \underline{CAT} : K \mapsto \wp K := \mathcal{S}(K,\Omega).$$

The logic is implemented in $\mathcal{S}$ along the lines of the so called *doctrinal diagram*:



(Cf. II.3.1, I.1.6.) Comprehension is here represented by the functors

$$i_K : \wp K \longrightarrow \mathcal{S}/K : \alpha(y^K) \mapsto \big(\{y^K | \alpha(y^K)\} \rightarrowtail K\big),$$

which are right adjoint to

$$h_K : \mathcal{S}/K \longrightarrow \wp K : \big(w : M \longrightarrow K\big) \mapsto \exists x^M.v(x^M) \equiv y^K$$

Lawvere (1970) used this adjunction to define comprehension abstractly, in hyperdoctrines. We translate his definition into fibred categories.

Definition. Let $E:\mathcal{E} \longrightarrow \mathcal{B}$ be a cloven bifibration with chosen terminal objects. We say that E is *Lawvere comprehensive* if the functor

$$h : \mathcal{B}/\mathcal{B} \longrightarrow \mathcal{E} : k \mapsto k_!\big(\top Dom(k)\big)$$

has a right adjoint i, so that Cod(iX)=EX and the unit and counit of this adjunction are cartesian natural transformations.

Remarks. If $\langle u,v\rangle \in \mathcal{B}/\mathcal{B}(m,k)$ (i.e. ku=vm),

$$h\langle u,v\rangle : m_!\big(\top I\big) \longrightarrow k_!\big(\top J\big)$$

is defined to be the unique arrow over v such that $h\langle u,v\rangle \circ \sigma^m_{\top I} = \sigma^k_{\top J} \circ \top u$, where I:=Dom(m), J:=Dom(k). The functor h is cartesian if E has the Beck-Chevalley property.

By lemma I.2.4, the functor i must be cartesian if $\nabla\mathcal{B}$ is a fibration. In that case, the above definition can (• by the same lemma •) equivalently be expressed by demanding that i is cartesian and fibrewise right adjoint to h. (This is perhaps closer to the spirit of Lawvere's original definition in terms of indexed functors.)

Thus, when $\mathcal{B}$ has pullbacks, and E the Beck-Chevalley property, the definition just asks for a cartesian adjointness $h \dashv i$.

Proposition. A cloven bifibration with chosen terminal objects is comprehensive iff it is Lawvere comprehensive. The functor i is a comprehension functor.

• If E is Lawvere comprehensive, then it is comprehensive by proposition 4, and

$$D := Dom{*}i \dashv h{*}Ids \simeq \top .$$

(If $i{*}\top = Ids$, then $\eta^D: id \to D\top$ is identity.)

It remains to prove that $iX : DX \to EX$ is indeed an extent. First note that every $f \in \mathcal{E}(h(m),X)$ is projected to the same arrow as its transpose $f' = i(f)\circ\eta_m$, since

$$\nabla\mathcal{B}(f') = Cod\big(i(f)\circ\eta_m\big) = Cod(i(f)) = Ef.$$

Therefore, for $m \in \mathcal{B}(I,M)$ the correspondence

$$\bigcup_{v\in\mathcal{B}(M,EX)} \mathcal{E}_v(m_!(\top I), X) = \mathcal{E}(h(m), X) \simeq \mathcal{B}/\mathcal{B}(m, iX) = \; = \bigcup_{v\in\mathcal{B}(M,EX)} \mathcal{B}/v\ (m, iX)$$

restricts over each v separately:

$$\mathcal{E}_K\big(\top K, (vk)^*X\big) \simeq \mathcal{E}_v\big(m_!(\top I), X\big) \simeq \mathcal{B}/v\ (m, iX)) = \; = \mathcal{B}/EX\ (vm, iX).$$

Then: By reversing the last step - making the unions, instead of partitioning them - we see that the comprehension functor «ι» is a right adjoint of h.•

4. About extent fibrations

1. What do they inherit?

In this section we shall study comprehension functors as essentially surjective functors $«\iota»:\mathcal{E} \to \iota\mathcal{E}$. $\mathcal{E}$ is a comprehensive fibred category with terminal objects.

We shall first formulate some propositions in the form:

"If $\mathcal{E}$ has a property/structure P then $\iota\mathcal{E}$ has and «ι» preserves P."

In 2.5 we saw that this holds for P = "inverse images". It obviously holds for P = "terminal objects". And more?

Propositions. Let E be a comprehensive fibration with terminal objects. Let $\mathcal{d} \subseteq \mathcal{B}$ be a stable family (II.4.3).

11. $\iota\mathcal{E}$ has and $«\iota»:\mathcal{E} \to \iota\mathcal{E}$ preserves all kinds of fibrewise limits which exist in $\mathcal{E}$.

• Since

$$\eta : \mathcal{E}_K \to \mathcal{E}_K/\top K : X \mapsto \big(\eta_X : X \to \top K\big)$$

is an isomorphism of categories, $\lambda : L \to \Delta$ is a limit cone in $\mathcal{E}_K$ iff $\eta(\lambda):\eta_L \to \eta\Delta$ is a limit cone in $\mathcal{E}_K/\top K$. But the image of λ by «ι» is the image of $\eta(\lambda)$ by D, and D preserves limits, because it has a left adjoint $\top$.•

12. If $\mathcal{E}$ has $\mathcal{d}$-products, then $\iota\mathcal{E}$ has and «ι» preserves them.

$$\bullet\ \iota\mathcal{E}_J\big(\iota Y, \iota(u_*X)\big) = \mathcal{B}/J\big(\iota Y, \iota(u_*X)\big) \simeq \mathcal{E}_{DY}(\top, \iota Y^*u_*X) \overset{\#}{\simeq}$$
$$\simeq \mathcal{E}_{DY}\big(\top, (D\vartheta^u)_*\iota(u^*Y)^*X\big) \simeq$$
$$\simeq \mathcal{E}_{D(u^*Y)}\big((D\vartheta^u)^*\top, \iota(u^*Y)^*X\big) \simeq$$
$$\simeq \mathcal{E}_{D(u^*Y)}\big(\top, \iota(u^*Y)^*X\big) \simeq \mathcal{B}/I\big(\iota(u^*Y), \iota X\big) \simeq$$
$$\simeq \mathcal{B}/I(u^*\iota Y, \iota X) = \iota\mathcal{E}_I(u^*\iota Y, \iota X)$$

The step (#) follows from the Beck-Chevalley property of $\mathcal{E}$, and the fact that the square

$$\begin{array}{ccc} D(u^*Y) & \xrightarrow{D\vartheta^u} & DY \\ {\scriptstyle \iota(u^*Y)}\downarrow & \lrcorner & \downarrow{\scriptstyle \iota Y} \\ I & \xrightarrow[u]{} & J \end{array}$$

is a pullback (since «ι» is a cartesian functor).•

13. If $\mathcal{E}$ is globally small, $\iota\mathcal{E}$ is. • If $\xi \in |\mathcal{E}_\Omega|$ is the generic object of $\mathcal{E}$, $\iota\xi \in |\iota\mathcal{E}_\Omega|$ is the generic object of $\iota\mathcal{E}$.• (If, moreover, $\mathcal{B}$ is lccc, then $\iota\mathcal{E}$ is small. • See the last example in 2.2.•)

14. Corollary. If E is a globally small cloven fibration with $\mathfrak{d}$-products and finite fibrewise products, ιE is equivalent to a small arrow fibration with fibrewise cartesian closed structure, and with $\mathfrak{d}$-products.

• The cartesian closed structure is

$$\iota Y \times \iota Z := \iota(Y \times Z);$$
$$\iota Y \to \iota Z := \iota Y_* \iota Y^*(Z).$$

Using the diagram from 12, we derive

$$\begin{aligned} \mathcal{B}/J\big(u, \iota Y_* \iota Y^*(Z)\big) &\simeq \mathcal{E}_I\big(\top, u^* \iota Y_* \iota Y^*(Z)\big) \overset{\#}{\simeq} \\ &\simeq \mathcal{E}_I\big(\top, \iota(u^*Y)_*(D\vartheta^u)^* \iota Y^*(Z)\big) \simeq \\ &\simeq \mathcal{E}_{D(u^*Y)}\big(\top, (\iota Y \circ D\vartheta^u)^* Z\big) \simeq \\ &\simeq \mathcal{B}/J\big(u \circ u^*(\iota Y), \iota Z\big). \end{aligned}$$

Now put $u := \iota X$, and note that $\iota X \times \iota Y \simeq \iota X \circ \iota X^*(\iota Y)$.

By facts 2.4, $\iota\mathcal{E}$ is locally small. By propositions 13 and 2.2, it is then equivalent to a small fibration.•

2. Weak coproducts.

In general, the coproducts in $\mathcal{E}$ don't seem to induce coproducts in $\iota\mathcal{E}$. Let us see what they do induce.

Definitions. An arrow $s \in \mathcal{E}(X,Y)$ is *weakly cocartesian* with respect to a functor $E:\mathcal{E} \to \mathcal{B}$ if for every $f \in \mathcal{E}(X,Z)$ such that $Ef = r \circ Es$ (for some r), there is an arrow g (*not necessarily unique*) such that $Eg = r$ and $g \circ s = f$.

E is a *weak* $\mathfrak{d}$*-cofibration* if every $\mathfrak{d}$-arrow has a weakly cocartesian lifting. E is a *weak* $\mathfrak{d}$*-bifibration* if it is a fibration and a weak $\mathfrak{d}$-cofibration.

We say that E has *weak* $\mathfrak{d}$*-coproducts* if it is a weak $\mathfrak{d}$-bifibration with the strong interpolation property (relative to $\mathfrak{d}$).

Remark. Wouldn't it be simpler to say that a weak $\mathfrak{d}$-bifibration E has weak coproducts if it has the weak Chevalley property? Let us first spell out what would a weak Chevalley property be.

Consider a commutative square Q=(f,g,s,t) over S=(k,m,u,v) as in II.3.3; suppose $k,m \in \mathfrak{d}$. The *weak Chevalley condition* on Q is:

C_w) if s and t are cartesian <u>and</u> f is weakly cocartesian, <u>then</u> g is weakly cocartesian.

A weak $\mathfrak{d}$-bifibration E has the *weak Chevalley property* if it satisfies the weak Chevalley condition over all pullback squares S (with $k,m \in \mathfrak{d}$).

If we denote by $k_!A$ and $m_!A$ some weak direct images, and if $g : u^*A \to v^*k_!A$ is the unique arrow over m such that $tg = fs$[3], and $\sigma^m : u^*A \to m_!u^*A$ is a weakly cocartesian lifting of m, the weak Chevalley condition just says that there must be vertical arrows

[3]Look at the picture with remark II.3.3.

$\rho : m_! u^* A \longrightarrow v^* k_! A$ and $e : v^* k_! A \longrightarrow m_! u^* A$ such that $g = \rho \circ \sigma^m$ and $\sigma^m = e \circ g$. Nothing more. This would be the corresponding *weak Beck condition.*

Clearly, when restricted to strong $\mathfrak{d}$-bifibrations, the weak Beck-Chevalley property doesn't imply the strong one. In fact, on weak $\mathfrak{d}$-bifibrations, the weak Beck-Chevalley condition is <u>equivalent</u> with the interpolation condition.

• To prove this, first notice that lemmas 31 and 32 in II.3a go through with weakly cocartesian arrows instead of cocartesian arrows. The same weakening of the remaining two lemmas in II.3a gives:

33_w. $c \circ v^*({}^{\backprime}a) \circ \rho \circ \sigma^m = {}^{\backprime}d \circ \sigma^m \Leftrightarrow m^*(c) \circ \tau \circ u^*(a) = d$

34_w. $c_1 \circ \rho \circ \sigma^m = c_2 \circ \rho \circ \sigma^m \Leftrightarrow m^*(c_1) \circ \tau \circ u^*(\eta) = m^*(c_2) \circ \tau \circ u^*(\eta)$.

Given g and σ^m as above, the existence of a vertical arrow $\rho : m_! u^* A \longrightarrow v^* k_! A$, such that $g = \rho \circ \sigma^m$, is an immediate consequence of the fact that σ^m is weakly cocartesian.

<u>If</u> the weak BC-condition is satisfied - i.e. if there is $e : v^* k_! A \longrightarrow m_! u^* A$ such that $g = \rho \circ \sigma^m$ and $\sigma^m = e \circ g$ - an interpolant can be defined as in II.3a.111. The fact that $\sigma^m = e \circ \rho \circ \sigma^m$ is sufficient to let the argument given there go through.

<u>Conversely</u>, if the interpolation condition is satisfied, an arrow $e : v^* k_! A \longrightarrow m_! u^* A$ can be defined as in II.3a.112:

$$e := c_\eta \circ v^*({}^{\backprime}a_\eta),$$

where $\langle a_\eta, B_\eta, c_\eta \rangle$ is an interpolant of $\eta : u^* A \longrightarrow m^* m_! u^* A$. 33_w just says that $\sigma^m = e \circ g$, since ${}^{\backprime}\eta = \mathrm{id}$.•

The *strong* interpolation condition just means that for arrows ρ and e given by the weak Beck-Chevalley property holds $\rho \circ e = \mathrm{id}$.

• From $\sigma^m = e \circ \rho \circ \sigma^m$ <u>follows</u> $\vartheta^v \circ \rho \circ \sigma^m = \vartheta^v \circ \rho \circ e \circ \rho \circ \sigma^m$, and <u>then</u> $\rho \circ \sigma^m = \rho \circ e \circ \rho \circ \sigma^m$. 34_w now tells that the interpolation is strong <u>iff</u> $\rho \circ e = \mathrm{id}$.•

Propositions. Let $\mathfrak{d}$ be a stable subcategory of $\mathcal{B}$, such that $\iota \subseteq \mathfrak{d}$.

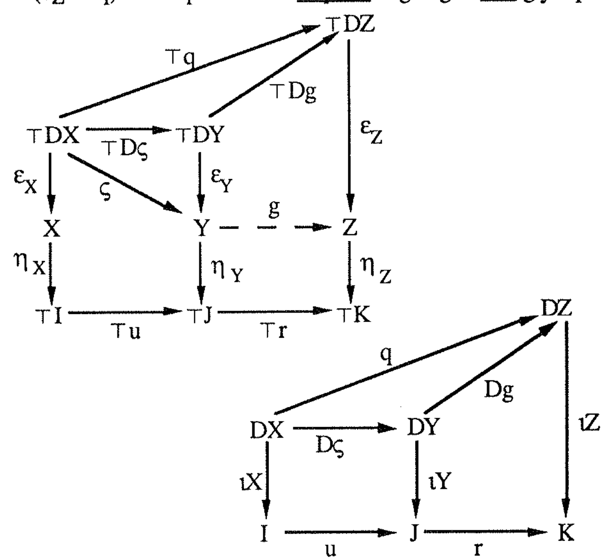
21. <u>If</u> $\mathcal{E}$ is a $\mathfrak{d}$-bifibration, $\iota\mathcal{E}$ is a weak $\mathfrak{d}$-bifibration.

• If $u \in \mathcal{B}(I,J) \cap \mathfrak{d}$ and $X \in |\mathcal{E}_I|$, and if $\varsigma \in \mathcal{E}(\top DX, Y)$ is an E-cocartesian lifting of $u \circ \iota X$, then $\langle D\varsigma, u \rangle \in \iota\mathcal{E}(\iota X, \iota Y)$ is a weakly ιE-cocartesian lifting of u at ιX. $\langle D\varsigma, u \rangle$ is indeed an arrow from ιX to ιY, since

$$\iota Y \circ D\varsigma = D(\eta_Y \circ \varsigma) = D(\top u \circ \eta_X \circ \varepsilon_X) = u \circ \iota X.$$

Consider arbitrary $r \in \mathcal{B}(J,K)$ and $\langle q, ru \rangle \in \iota\mathcal{E}(\iota X, \iota Z)$. If ${}^{\backprime}q \in \mathcal{E}(\top DX, Z)$ is the left transpose of q, then

$$E({}^{\backprime}q) = E(\varepsilon_Z \circ \top q) = \iota Z \circ q = r \circ u \circ \iota X \text{ \underline{implies} } \exists! g.\ Eg = r \text{ \underline{and} } g\varsigma = {}^{\backprime}q.$$



Now $\langle Dg, r \rangle \in \iota\mathcal{E}(\iota Y, \iota Z)$ is a factorisation over r of $\langle q, ru \rangle$ through $\langle D\varsigma, u \rangle$. $\langle Dg, r \rangle$ is an arrow in $\iota\mathcal{E}$ from ιY to ιZ because

$$\iota Z \circ Dg = D(\eta_Z \circ g) = D\top r \circ D\eta_Y = r \circ \iota Y.$$

And $\langle Dg, r \rangle \circ \langle D\varsigma, u \rangle = \langle q, ru \rangle$ because

$$Dg \circ D\varsigma = D({}^{\backprime}q) = q.\bullet$$

22. If $\mathcal{E}$ has the $\mathfrak{d}$-coproducts, <u>then</u> $\iota\mathcal{E}$ has the weak $\mathfrak{d}$-coproducts.

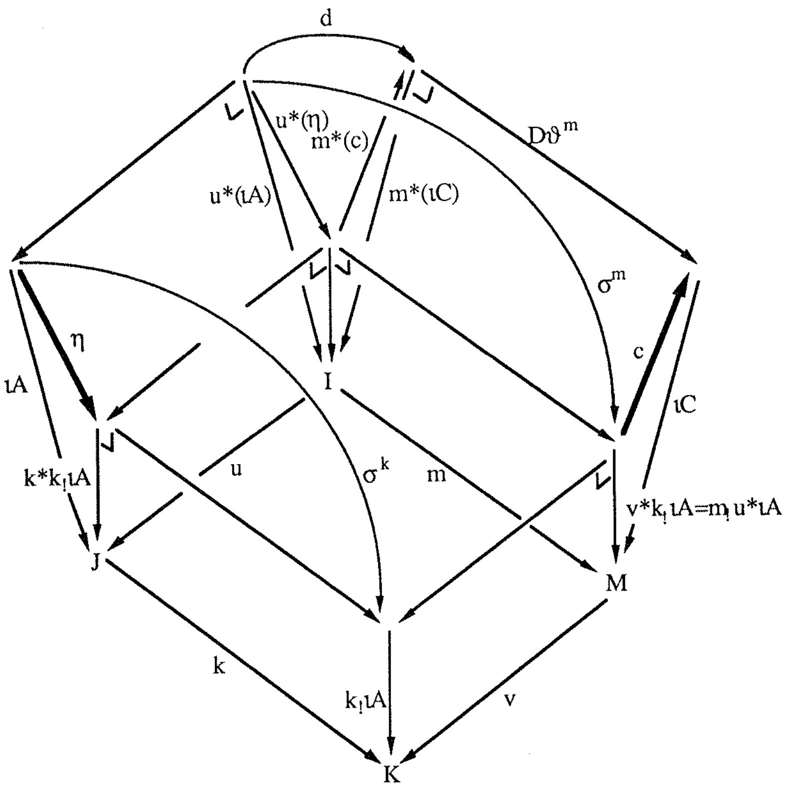
• Let us first introduce some notation (for the weak cocartesian liftings constructed above):

$$p_!(\iota X) := \iota\big((p \circ \iota X)_! \top DX\big),$$

$\sigma^p_{\iota X} := D\sigma^{p\circ\iota X}_{\top}$,

for arbitrary $p \in d$.

Consider a commutative square $S=(k,m,u,v)$ in $\mathfrak{B}$, with $k,m \in d$. The following diagram shows how we construct an interpolant $\langle \eta, k_!(\iota A), c\rangle$ for an arbitrary arrow $d \in \iota\mathcal{E}_I\big(u^*(\iota A), m^*(\iota C)\big)$. (In fact, we should write "$\langle d,id\rangle$" instead of "d".)



Since

$$v^*k_!(\iota A) = v^*\iota\big((k\circ\iota A)_!\top\big) \simeq \iota\big(v^*k_!\iota A_!\top\big) \overset{BC}{\simeq} \iota\big(m_!u^*\iota A_!\top\big) \overset{BC}{\simeq}$$
$$\simeq \iota\Big(m_!\big(\iota(u^*A)\big)_!\top\Big) = m_!\big(\iota(u^*A)\big) \simeq$$
$$\simeq m_!u^*(\iota A),$$

we can choose the pullback $v^*k_!(\iota A)$ so that the equality $v^*k_!(\iota A)=m_!u^*(\iota A)$ holds. The arrow

$$c \in \iota\mathcal{E}_M\big(v^*k_!(\iota A), \iota C\big)$$

is now induced as a factorisation of $\langle D\vartheta^m_C\circ d, m\rangle \in \iota\mathcal{E}\big(u^*(\iota A), \iota C\big)$ through the weakly cocartesian arrow $\langle \sigma^m_{u^*\iota A}, m\rangle \in \iota\mathcal{E}\ \big(u^*(\iota A), m_!u^*(\iota A)\big)$.

It remains to prove that every initial interpolant is strong.

It is easy to see that $\langle \eta, k_!(\iota A), c\rangle$ is an initial interpolant <u>iff</u> $\langle \sigma^k, k\rangle \in \iota\mathcal{E}\big(\iota A, k_!(\iota A)\big)$ is (strongly) initial, i.e. a cocartesian lifting. If this is the case, then there must be an iso $a_0 \simeq \eta$ for any other initial interpolant $\langle a_0, B_0, c_0\rangle$. To prove that the interpolation in $\iota\mathcal{E}$ is strong, it is sufficient to show that c, determined as above, is unique (up to isomorphism) <u>if</u> $\langle \sigma^k, k\rangle$ is (strongly) cocartesian.

By the lemma below, $\langle \sigma^k_{\iota A}, k\rangle$ is ιE-cocartesian <u>iff</u> $\varepsilon_Y \in \mathcal{E}_{k_!(\iota A)}(\top DY, Y)$ is E-cocartesian, for $Y:=(k\circ\iota A)_!\top DA$.

But if ε_Y is a cocartesian lifting of $k_!(\iota A)$, <u>then</u> $\varepsilon_{v^*Y} \simeq v^*\varepsilon_Y \in \mathcal{E}(\top Dv^*Y, v^*Y)$ must be a cocartesian lifting of $v^*k_!(\iota A)=m_!u^*(\iota A)$. Applying the lemma again, we conclude that the arrow $\langle \sigma^m_{u^*\iota A}, m\rangle$ is ιE-cocartesian, i.e. strongly initial. Therefore, the factorisation $c \in \iota\mathcal{E}_M\big(m_!u^*(\iota A), \iota C\big)$ through it must be unique.•

Lemma. The lifting $\langle D\varsigma, u\rangle$ constructed in proprosition 21 is strongly initial - i.e. a ιE-cocartesian arrow - <u>iff</u> the counit $\varepsilon_Y \in \mathcal{E}_{\iota Y}(\top DY, Y)$ is E-cocartesian.

• To every $\langle h,r\rangle \in \iota\mathcal{E}(\iota Y, \iota Z)$ corresponds a unique $`h \in \mathcal{E}_r(\top DY, Z)$. <u>If</u> $\varepsilon_Y$ is cocartesian, then $`h = g\circ\varepsilon_Y$; thus $h=Dg$.

<u>Conversely</u>, if for every $\langle q, ru\rangle \in \iota\mathcal{E}(\iota X, \iota Z)$, $\langle q, ru\rangle = \langle h,r\rangle\circ\langle D\varsigma, u\rangle$ <u>implies</u> $h=Dg$, <u>then</u> $`h=g\circ\varepsilon_Y$ holds for every arrow $`h \in \mathcal{E}_r(\top DY, Z)$. This g is unique as a factorisation of $`h$ through $\varepsilon_Y$, because it is also the factorisation of $`q = `h\circ\top D\varsigma$ through cocartesian $\varsigma = \varepsilon_Y\circ\top D\varsigma$.•

A final remark. The weak coproducts are to be used for an interpretation of the type theoretical weak Σ. (Cf. IV.1.2.) And nothing less than the full Beck-Chevalley property can allow a sound interpretation of variables (in view of the explanations in II.3.3). This is why we *had to* use the strong interpolation condition in the definition of coproducts. (And this is why we went out of our way in II.3a, to characterize the Beck-Chevalley condition without direct images.)

3. Which comprehensive fibrations are arrow fibrations?

In other words, when is the comprehension functor «ι»:$\mathcal{E} \longrightarrow \iota\mathcal{E}$ an equivalence of categories? - Exactly when $\mathcal{E}$ has and «ι» preserves the ι-coproducts of terminal objects!

Definition. Let $E:\mathcal{E} \longrightarrow \mathcal{B}$ be a fibration with terminal objects. For every $X,Y \in |\mathcal{E}_I|$, each arrow $f \in \mathcal{E}_I(X,Y)$ induces the functions

$$\varphi_v : \mathcal{E}_v(\top,X) \longrightarrow \mathcal{E}_v(\top,Y) : q \longmapsto f \circ q$$

naturally in $v \in |\mathcal{B}/I|$. Hence the mapping

$$\mathcal{G}_{XY} : \mathcal{E}_I(X,Y) \longrightarrow \mathrm{Nat}\left(\mathrm{HOM}_E(\top,X), \mathrm{HOM}_E(\top,Y)\right)$$

We say that $\mathcal{E}$ is *generated by its terminal objects* if all $\mathcal{G}_{XY}$ are injections. It is *fully generated by its terminal objects* if all $\mathcal{G}_{XY}$ are bijections.

Proposition. For every comprehensive fibration E with fibrewise terminal objects the conditions listed below are related:

(a)⇔(b)⇔(c).

If E has fibrewise cartesian closed structure, then

(a)⇔(d)

holds too. If E has ι-products, then

(a)⇔(e).

a) «ι»:$\mathcal{E} \longrightarrow \iota\mathcal{E}$ is an equivalence of categories.

b) The counits $\varepsilon_X \in \mathcal{E}(\top DX,X)$ of the adjunction $\top \dashv D$ are cocartesian.

c) $\mathcal{E}$ is fully generated by its terminal objects.

d) «ι» preserves the fibrewise exponents.

e) $\mathcal{E}$ has fibrewise exponents

$$X \to Y := \iota X_* \iota X^*(Y).$$

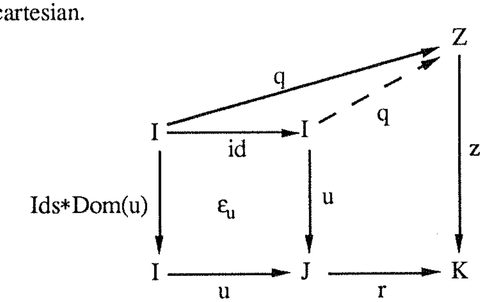
(If it has fibrewise binary products, they are

$$X \times Y \simeq \iota X_! \iota X^*(Y).)$$

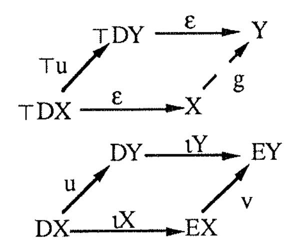
• (a)⇒(b): If $\mathcal{E}$ is (equivalent to) and arrow fibration with terminal objects, then

$$\varepsilon_u := \langle id,u \rangle \in \mathcal{E}(\mathrm{Ids}^*\mathrm{Dom}(u), u)$$

is clearly cocartesian.



(b)⇒(a): If $\varepsilon_X \in \mathcal{E}(\top DX,X)$ is cocartesian <u>then</u> every $\langle u,v \rangle \in \iota\mathcal{E}(\iota X,\iota Y)$ induces a unique arrow $g \in \mathcal{E}(X,Y)$, over v such that $\varepsilon_Y \circ \top u = g \circ \varepsilon_X$.



From the last equality follows that $Dg = u$ (since $D^*\varepsilon$=id); thus $\langle u,v \rangle$ = «ι»(g).

(a)⇔(c): ιX is a representant of $\mathrm{HOM}_E(\top,X)$. By the Yoneda lemma

$$\mathrm{Nat}\left(\mathrm{HOM}_E(\top,X), \mathrm{HOM}_E(\top,Y)\right) \simeq \mathcal{B}/I(\iota X,\iota Y).$$

This correspondence is realized by the mapping

$$\varphi \mapsto \varphi_{\iota X}(\varepsilon_X) \in \mathcal{E}_{\iota X}(\top DX, Y) \simeq \mathcal{B}/I(\iota X, \iota Y).$$

($\varepsilon_X \in \mathcal{E}_{\iota X}(\top DX, X)$ is here the generic arrow, i.e. the one which corresponds to $id_{\iota X}$.)
Postcomposing on $\mathcal{G}_{XY}$, we get

$$\begin{array}{ccccccc} \mathcal{E}_I(X,Y) & \longrightarrow & \mathrm{Nat}\big(HOM_E(\top,X), HOM_E(\top,Y)\big) & \longrightarrow & \mathcal{E}_{\iota X}(\top DX, Y) & \simeq & \mathcal{B}/I(\iota X, \iota Y) \\ f & \mapsto & \mathcal{G}_{XY}(f) & \mapsto & f \circ \varepsilon_X & \mapsto & Df \end{array}$$

($Df \in \mathcal{B}/I(\iota X, \iota Y)$ was defined to be the arrow corresponding to $f \circ \varepsilon_X$. Cf. 3.3.) Since $\langle\langle\iota\rangle\rangle_I(f) = Df$, we see that

$$\langle\langle\iota\rangle\rangle_I \text{ is full and faithful } \underline{\text{iff}}\ \mathcal{G}_{XY} \text{ are bijections for all } X, Y \in |\mathcal{E}_I|.$$

But $\langle\langle\iota\rangle\rangle$ is full and faithful <u>iff</u> all $\langle\langle\iota\rangle\rangle_I$ are.

(d)$\Rightarrow$(a): $\langle\langle\iota\rangle\rangle_I : \mathcal{E}_I \longrightarrow \iota\mathcal{E}_I$ is full and faithful because the correspondence

$$\mathcal{E}_I(X,Y) \simeq \mathcal{E}_I(\top, X \to Y) \simeq \mathcal{B}/I(id, \iota(X \to Y)) \simeq \iota\mathcal{E}_I(id, \iota X \to \iota Y) \simeq$$
$$\simeq \iota\mathcal{E}_I(\iota X, \iota Y)$$

is again realized by $f \mapsto Df$.

(e)$\Rightarrow$(a):
$$\mathcal{E}_I(X,Y) \simeq \mathcal{E}_I(\top, X \to Y) = \mathcal{E}_I\big(\top, \iota X_* \iota X^*(Y)\big) \simeq$$
$$\simeq \mathcal{E}_{DX}\big(\iota X^* \top, \iota X^* Y\big) \simeq \mathcal{E}_{\iota X}(\top, Y) \simeq \mathcal{B}/I(\iota X, \iota Y) =$$
$$= \iota\mathcal{E}_I(\iota X, \iota Y).$$

(a)$\Rightarrow$(e): A fibrewise equivalence preserves and reflects fibrewise cartesian closed structure and horizontal structure. Thus

$$\mathcal{E} \text{ has } \iota\text{-products} \Rightarrow \iota\mathcal{E} \text{ has } \iota\text{-products} \overset{\#}{\Rightarrow} X \to Y \simeq \iota X_* \iota X^*(Y) \text{ in } \iota\mathcal{E} \Rightarrow$$
$$\Rightarrow X \to Y \simeq \iota X_* \iota X^*(Y) \text{ in } \mathcal{E}.$$

The step (#) is sound because $a_* a^*(b)$ is the exponent $a \to b$ in every arrow fibration $\mathfrak{a}/\mathcal{B}$ with terminal objects and $\mathfrak{a}$-products (all intrinsic):

$$\mathfrak{a}{\downarrow}I(a,b) \simeq \mathfrak{a}{\downarrow}K(id, a^*(b)) \simeq \mathfrak{a}{\downarrow}K(a^*(id), a^*(b)) \simeq \mathfrak{a}{\downarrow}I(id, a_* a^*(b)).\bullet$$

IV. Semantics

In this chapter we should capitalize the investments in "abstract nonsense", and interprete the theory of predicates. In section 1, all the previously described concepts are put together, and a categorical meaning is formally assigned to each operation of the theory of predicates. The notion of a *category of predicates* is introduced: it is a small hyperfibration, with the rccc structure in the base as well as in the fibres. The theory of predicates is the natural logical syntax for categories of predicates.

In section 2 we show how a given theory of predicates generates a category of predicates. This semantical construction is then proved to be complete. The proof has been built according to a standard scheme (cf. Lambek-Scott 1986, I.11, II.13-16), and upon the standard completeness result for the Martin-Löf type theory (Seely 1984).

In the last section, a first effort has been made to put the syntax and semantic together at work - to *speak* a natural language of predicates. At the end of this section, we show how to produce "mathematical" examples of proper categories of predicates. A category of *constructive* internal presheaves is constructed in an arbitrary category of predicates. (In particular, this can be done in each of the well known "mathematical" models for the theory of constructions.)

1. Conceptual elements of the interpretation.

Sorts, types and terms. The idea for a categorical interpretation of type theories was conceived in late sixties, in Lambek's papers on deductive systems (cf. bibliography in Lambek-Scott 1986). It points out the *basic analogy*:

types	$\mapsto$	objects,
terms	$\mapsto$	arrows,
substitution	$\mapsto$	composition.

Of course, this works only for *simple* type theories, with no variable types. A *model assignment* of such a type theory Λ in a cartesian category $\mathcal{C}$ is a mapping

$$[\![\_]\!] : \Lambda \to \mathcal{C},$$

which respects the basic analogy, and satisfies the conditions:

$$[\![X{:}P \Rightarrow q(X){:}Q]\!] \in \mathcal{C}\big([\![P]\!], [\![Q]\!]\big),$$

$$[\![q[X{:=}p]]\!] = [\![q]\!] \circ [\![p]\!].$$

The central result at this level is the correspondence of simple typed λ-calculi and cartesian closed categories (Lambek-Scott 1986, chapter I).

The base, and most of the superstructure needed for the categorical interpretation of *variable* types, was contained in Lawvere's articles on hyperdoctrines (1969, 1970). In principle, variable types are interpreted as objects of variable categories. The basic analogy is now extended:

sorts	$\mapsto$	categories
variations	$\mapsto$	fibrations
types varying over a type P (and their terms)	$\mapsto$	objects (and arrows) of the fibre over $[\![P]\!]$
substitution in a variable type	$\mapsto$	inverse images

To produce a *model* for a type theory Λ, one first chooses a category $\llbracket\Delta\rrbracket$ for each of its sorts Δ. For every Δ, the class Λ_Δ of Δ-types and terms, will be interpreted according to the basic analogy by a model assignment

$$\llbracket\_\rrbracket_\Delta : \Lambda_\Delta \longrightarrow \llbracket\Delta\rrbracket.$$

A variation $\Delta'\Delta''$ of types, which is allowed in Λ, will be represented by a variation of categories: a fibration

$$\llbracket\Delta'\Delta''\rrbracket : \llbracket\Delta''\rrbracket \longrightarrow \llbracket\Delta'\rrbracket$$

must be chosen. The model assignments must now satisfy the conditions:

$$\llbracket X:P:\Delta' \Rightarrow Q(X):\Delta''\rrbracket \in |\llbracket\Delta''\rrbracket_{\llbracket P\rrbracket}|,$$

$$\llbracket X:P:\Delta', x:Q_0:\Delta'' \Rightarrow q(x):Q_1:\Delta''\rrbracket \in \llbracket\Delta''\rrbracket_{\llbracket P\rrbracket}\big(\llbracket Q_0\rrbracket, \llbracket Q_1\rrbracket\big),$$

Substitution is interpreted by inverse images: for

$$\llbracket p\rrbracket = \llbracket Y:R:\Delta' \Rightarrow p(Y):P:\Delta'\rrbracket \in \llbracket\Delta'\rrbracket\big(\llbracket R\rrbracket, \llbracket P\rrbracket\big),$$

$$\llbracket Y:R:\Delta' \Rightarrow Q[X:=p(Y)]:\Delta''\rrbracket := \llbracket p\rrbracket^*\llbracket Q\rrbracket \in |\llbracket\Delta''\rrbracket_{\llbracket R\rrbracket}|.$$

(N.B. To add a dummy variable Y in Q, means to substitute $Q[X:=\pi_0\langle X,Y\rangle]$ (also written $Q(\not{Y})$), i.e. to take an inverse image along a projection. In semantics, this is done all the time: types and terms must be brought under the same context - in the same fibre. It is therefore helpful to regard a type Q together with all its instances with dummies, i.e. to think of $\llbracket Q\rrbracket$ *together with all its inverse images along projections.*)

The inheritance of variation is interpreted by composing fibrations: the contexts with more than one layer are represented using towers of fibrations. For instance, if besides variation $\Delta''\Delta'$, theory Λ allows a variation $\Delta''\Delta'''$, a fibration $\llbracket\Delta''\Delta'''\rrbracket : \llbracket\Delta'''\rrbracket \longrightarrow \llbracket\Delta''\rrbracket$ will be used to assign

$$\llbracket X:P:\Delta',\ Y:Q(X):\Delta'' \Rightarrow S(X,Y):\Delta'''\rrbracket \in |\llbracket\Delta'''\rrbracket_{\llbracket Q\rrbracket}|.$$

A particular case are variations $\Delta\Delta$. Each of them is interpreted using an (intrinsic) *arrow* fibration $\llbracket\Delta\Delta\rrbracket$ over category $\llbracket\Delta\rrbracket$. This is the well known representation of dependent types by display arrows - the well ploughed ground of categorical semantics for Martin-Löf type theories. The main sources are: Seely 1984, Cartmell 1986, Hyland-Pitts 1987; the literature is quite extensive.

The complex contexts of dependent types of sort Δ can now be interpreted *within* the category $\llbracket\Delta\rrbracket$, as trees built of arrows (since a fibrewise arrow fibration over an arrow fibration is just another arrow fibration).

To resume - for a categorical interpretation of the theory of predicates, we shall need two categories:

$\mathcal{B} = \llbracket\Theta\rrbracket$ for sets, and

$\mathcal{E} = \llbracket\Omega\rrbracket$ for propositions.

The three variations which this theory allows will demand three fibrations:

$E \quad = \llbracket\Theta\Omega\rrbracket : \mathcal{E} \longrightarrow \mathcal{B}$ for predicates,

$\nabla a \quad = \llbracket\Theta\Theta\rrbracket : a/\mathcal{B} \longrightarrow \mathcal{B}$ for dependent sets, and

$\nabla r \quad = \llbracket\Omega\Omega\rrbracket : r/\mathcal{E} \longrightarrow \mathcal{E}$ for dependent propositions.

Sums and products. Type theoretical operations are, of course, interpreted by some adjunctions.

It is implicite in the basic analogy that the type theoretical machinery of variables must use finite products:

$$\llbracket X:P, Y:R \Rightarrow q(X,Y):Q\rrbracket \in \mathcal{C}\big(\llbracket P\rrbracket\times\llbracket R\rrbracket, \llbracket Q\rrbracket\big).$$

λ-abstraction is then interpreted by the exponents - right adjoints of the product functors:

$$\llbracket Y:R \Rightarrow \lambda X.q(X,Y):P\to Q\rrbracket := \llbracket q\rrbracket` \in \mathcal{C}\big(\llbracket R\rrbracket, \llbracket P\rrbracket\to\llbracket Q\rrbracket\big),$$

(As always, f‘ is the right transpose of f.)

Just a step further is Lawvere's observation (1969) that the universal quantifier is the right adjoint to the substitution, while the existential quantifier is its left adjoint. See II.3.1. The quantifiers, as presented in our type theory, will thus be interpreted by the horizontal structure of hyperfibration $E = \llbracket\Theta\Omega\rrbracket$.

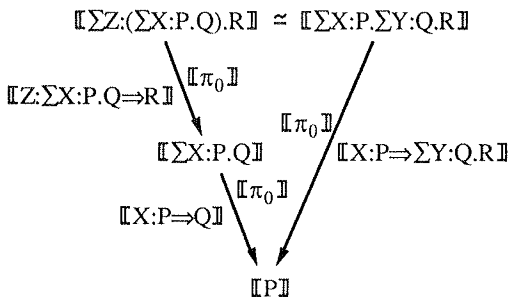
With the restrictions from definition I.1.5, the quantifier rules just express an adjointness (see below). The Σ-rules for $\Theta\Theta$ and $\Omega\Omega$, on the other hand, are essentially stronger, and their interpretations demand more than just adjunctions. Namely, they allow a first projection $\pi_0 := \lambda Z.\nu(Z,(X,Y).X)$ to be formed; for every type P there is a bijective correspondence between the types depending on P and the projections π_0 on P:

$$(X:P\Rightarrow Q) \mapsto (Y:\Sigma X:P.Q \Rightarrow \pi_0 Y:P).$$

On the side of semantics, a dependent type becomes *synonymous* with a particular term - the projection from its sum: this type and this term are interpreted by the same arrow[1]. (The basic analogy "types$\mapsto$objects" is preserved by changing the point of view: while $q = [\![Y:\Sigma X:P.Q\Rightarrow\pi_0 Y]\!]$ is regarded as an arrow, $q = [\![X:P\Rightarrow Q]\!]$ is an object of arrow category.) This determines the interpretation of $\Sigma\Delta\Delta$:

$$[\![\Sigma X:P.Q]\!] = \mathrm{Dom}[\![Y:\Sigma X:P.Q\Rightarrow\pi_0 Y:P]\!] = \mathrm{Dom}[\![X:P\Rightarrow Q]\!].$$

Moreover, since a composition of two first projections is (isomorphic to) a first projection,



the distinguished class of arrows which interprete the projections and the dependent types must be closed under the composition. Hence, $\Sigma\Delta\Delta$ will be interpreted by composition. - Recalling II.4, the sums and products of sets and of propositions will be interpreted by the *intrinsic* horizontal structures of arrow hyperfibrations $\nabla\mathrm{a}$ and $\nabla\mathrm{r}$ respectively.

Polymorphism is type-theoretically expressed by the axiom $\Omega:\Theta$, i.e. by the fact that *every proposition is a type* (of the sort Ω) *and a term* (of the type Ω) at the same time. A category theoretical expression of this "impredicativity" seems to be the requirement that

---

[1]In terms of II.1.1, the indexed set $\{\gamma_x \mid x\in B\}$ is identified with the projection $\sum_{x\in B}\gamma_x \longrightarrow B:\langle x,c\rangle\mapsto x$.

the fibred category E of propositions over sets is globally small –. If a proposition $\alpha(X^K)$ *as a type* is interpreted by an object $A=[\![\alpha]\!]\in|\mathcal{E}_{[\![K]\!]}|$ of a globally small fibred category $\mathcal{E}$, the corresponding arrow $\ulcorner A\urcorner\in\mathcal{B}([\![K]\!],\Omega)$ appears as an interpretant of the same proposition *as a term*. (The object $\Omega\in|\mathcal{B}|$, which represents the fibred class OB_E, is assigned to the type $\Omega:\Theta$.)

Extent operation of the theory of predicates will be interpreted by the extents of the fibred category $E=[\![\Theta\Omega]\!]$. So E will have to be comprehensive. Since it is fibrewise cartesian closed, it will be locally small (fact III.2.4).

Remark. This parallelism of type theory and category theory suggests that they are like two languages which refer to the same things. The point is, as we explained in the introduction, that they approach these things differently: type theory studies some operations as *structure*, while they arise in category theory from some *properties*. This difference is, of course, the reason why it is worth-while to speak both languages. But it is also the source of various problems.

For instance, the mapping $\ulcorner\_\urcorner: |\mathcal{E}_I|\longrightarrow\mathcal{B}(I,\Omega)$, which represents objects of a globally small fibration by arrows (as defined in III.2.2), is generally not injective. This means that two propositions could be interpreted differently "as types", but equally "as terms"! The solution of this problem comes from an unexpected direction. The requrements that a fibred category of predicates is globally and locally small, add up (by proposition III.2.2) to tell that it must be small, i.e. equivalent to one in the form $\nabla\Omega: \mathcal{B}/\Omega\longrightarrow\mathcal{B}$. Moving along this equivalence solves the problem: the objects of $(\mathcal{B}/\Omega)_I$ are the arrows from $\mathcal{B}(I,\Omega_0)$, and $\ulcorner\_\urcorner$ can now be taken to be identity. Moral: For the interpretation of type theory, *it is not inessential which of the equivalent "copies" of a category is taken.*

This is emphasised even stronger by the fact that *type theory demands split fibrations.* Namely, one of the basic principles of substitution is that that P[X:=u(Y)][Y:=v(Z)] is identical with P[X:=u[Y:=v(Z)]]. Interpreting substitution by inverse images, this means that $v^*u^*P=(uv)^*P$ must hold.

Taken in the form $\nabla\Omega$, the fibration $E=[\![\Theta\Omega]\!]$ is, of course, split. On the other hand, the splittings of the arrow fibrations $\nabla\mathrm{a}=[\![\Theta\Theta]\!]$ and $\nabla\mathrm{r}=[\![\Omega\Omega]\!]$ should have to be explicite: in fact, they should be *contextual categories* (Cartmell 1986, 14, Streicher 1988, 1.1). It has, however, become a tradition in interpreting Martin-Löf type theories to neglect this splitting requirement, and to consider stable families instead of contextual categories. The interpretants of P[X:=u(Y)][Y:=v(Z)] and P[X:=u[Y:=v(Z)]] may not be identical, but just isomorphic. Quietly, the semantics seems to have relaxed a syntactical principle in our minds.

2. Categories of predicates.

Definitions. Let a category $\mathcal{B}$ with a terminal object be relatively cartesian closed with respect to a display family $\mathfrak{a}$. A *category of predicates* over (*a category of sets*) $\mathcal{B}$ is a small $\mathfrak{a}$-hyperfibration $E:\mathcal{E} \longrightarrow \mathcal{B}$ with terminal objects, such that its class of extents ι is contained in $\mathfrak{a}$. Moreover, the category $\mathcal{E}$ must be relatively cartesian closed with respect to a vertical display family $\mathfrak{r}$.

A category of predicates which is equivalent to an arrow fibration is called a *category of constructions*.

Remarks. Since $\mathfrak{a}$ is a display family, $\mathcal{B}$ must be cartesian closed. Since $\mathfrak{r}$ is a vertical display family, $\mathcal{E}$ is fibrewise cartesian closed (• using II.4.7•).

Corollary III.4.14 tells that the extent $\iota\mathcal{E}$ of a category of predicates $\mathcal{E}$ must be equivalent to a small fibrewise cartesian closed category too. The same corollary further says that $\iota\mathcal{E}$ has $\mathfrak{a}$-products. Proposition III.4.22 tells that it has weak $\mathfrak{a}$-coproducts.

From the lemma below, it follows that the requirement

$$\iota \subseteq \mathfrak{a},$$

imposed upon the class ι of extents by the definition of a category of predicates $\mathcal{E}$, can equivalently be expressed by demanding

$$D(\mathfrak{r}) \subseteq \mathfrak{a},$$

where $D:\mathcal{E} \longrightarrow \mathcal{B}$ is any extent functor of $\mathcal{E}$. Moreover, when $\mathcal{E}$ is an arrow fibration, the lemma implies that $\iota = \mathfrak{r}$. The definitions can now be resumed:

category of predicates :=
- $\mathfrak{a}$-rccc $\mathcal{B}$ +
- $\mathfrak{r}$-rccc $\mathcal{E}$ +
- small $\mathfrak{a}$-hyperfibration $E:\mathcal{E} \longrightarrow \mathcal{B}$, such that
  - $E(\mathfrak{r}) \subseteq \mathrm{id}$,
  - $D(\mathfrak{r}) \subseteq \mathfrak{a}$;

category of constructions := $\mathfrak{r} \subseteq \mathfrak{a} \subseteq \mathcal{B}$, such that
- $\mathcal{B}$ is $\mathfrak{a}$-rccc +
  - $\mathfrak{r}$-rccc +
- $\nabla\mathfrak{r}$ is a small $\mathfrak{a}$-hyperfibration.

$\mathcal{B}$ is an elementary topos iff it constitutes a category of constructions with $\mathfrak{a}=\mathcal{B}$ and $\mathfrak{r}$=monics.

Some characterisations of categories of constructions among categories of predicates can be found in III.4.3.

From III.4.1 and III.4.2, it follows that the extent fibration ιE of an arbitrary category of predicates E misses being a category of constructions by very little: it has all the structure as it should, *except* that its coproducts may be weak. In terms of the interpretation which we are about to give, this means that ιE will support the theory of constructions, with the exception of the rule $\eta\Sigma\Theta\Omega$; or the strong theory of predicates without $\eta\exists$.

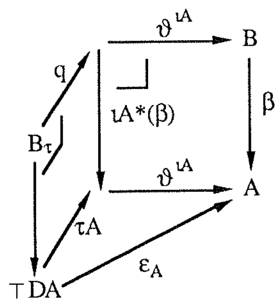
Sources. The structure of a category of constructions has been described in detail by Hyland and Pitts (1987); they only did not give it a name. Ehrhard's (1989) *dictos* is equivalent to a category of constructions with $\mathfrak{a}=\mathcal{B}$. Streicher's (1988, 1.16) *doctrine of constructions*, on the other hand, conceptually corresponds to a category of constructions without any left direct images (i.e., replace the words "rccc" and "hyperfibration" in the definition by "right bifibration"). However, this correspondence is not precise, since Streicher is working with contextual categories, which carry more structure. (A *contextual category* is a split arrow fibration $\nabla\mathfrak{a}:\mathfrak{a}/\mathcal{B}\to\mathcal{B}$, such that the mapping $\mathrm{Dom}:|\mathfrak{a}/\mathcal{B}|\to|\mathcal{B}|$ has an inverse p, and the endomorphism $f:=\mathrm{Cod}\circ p:|\mathcal{B}|\to|\mathcal{B}|$ induces a tree structure on $\mathcal{B}$, with the terminal object 1 as root - i.e. for every $I\in|\mathcal{B}|$, the orbit $\{I, f(I), f^2(I), f^3(I)...\}$ is finite and contains $1=f(1)$.)

Lemma. Let $E:\mathcal{E}\to\mathcal{B}$ be a comprehensive fibration with terminal objects. (ι is the class of its extents, and $D(\dashv\top\dashv E)$ is an extent functor.)

i) $D(\mathcal{r})\subseteq\iota$ holds for every stable family $\mathcal{r}\subseteq\mathcal{E}$.

ii) If $\mathcal{r}$ is a display family, ι is the smallest saturated family containing $D(\mathcal{r})$.

• i) Consider $\beta\in\mathcal{E}_I(B,A)\cap\mathcal{r}$. Let $\tau A:\top DA\to\iota A^*(A)$ be the vertical factorisation of $\varepsilon_A\in\mathcal{E}_{\iota A}(\top DA,A)$, and define B_τ by the pullback of $\iota A^*(\beta)$ along τA.



This pullback must exist in $\mathcal{E}_{DA}$ because $\iota A^*(\beta)\in\mathcal{r}_{DA}$, and $\mathcal{r}_{DA}=\mathcal{E}_{DA}\cap\mathcal{r}$ is a stable family. From $D\vartheta^{\iota A}_A = D\varepsilon_A = id_{DA}$, follows $D\vartheta^{\iota A}_B\simeq id_{DB}$, because a square with two vertical and two cartesian arrows (at the opposite sides) must be a pullback (by II.2.22), and D preserves pullbacks. Hence

$$\iota B_\tau = D\eta_{B\tau} = D(\eta_{B\tau}\circ\varepsilon_A) = D(\beta\circ\vartheta^{\iota A}_B\circ q) \simeq D\beta.$$

ii) If $\mathcal{r}$ contains all the fibrewise terminal arrows $\eta_X\in\mathcal{E}_{EX}(X,\top EX)$, then $D(\mathcal{r})$ spans ι, because $\iota X\simeq D\eta_X$.•

3. Simplifying contexts.

One of the crucial philosophical problems in the categorical semantics of type theories is "explaining away the variables" (cf. Lambek 1980, section 1). Variables are an eminently syntactical part of logic ("universalia"); an honest category doesn't seem to be a natural environment for them.

However, if a type theory is strong enough, the variables which occur in a type or term can be bound, and then unbound without any loss. In the meantime, an interpretation can be defined, not having to cope with them. For instance, every extensional typed λ-calculus Λ can be recovered from its class Λ_0 of closed terms, and a notion of application on them. A model assignment $[\![\_]\!]:\Lambda\to\mathcal{C}$ is thus uniquely determined by its restriction to Λ_0, assuming that the interpretation of the application is known.

By another sort of binding, using the surjective pairing, *any context can be reduced to one variable* in a simple single-sorted type theory (with no dependent types):

$$X_0:P_0,\dots X_{n-1}:P_{n-1} \Rightarrow f(X_0,\dots,X_{n-1}) : Q$$
$$\downarrow$$
$$Z:\prod_{i\in n}P_i \Rightarrow f(\pi_0 Z,\dots,\pi_{n-1}Z) : Q.$$

Assuming that the pairing and projections are interpreted in a cartesian category $\mathcal{C}$ by the appropriate cartesian operations, a model assignment will be determined by its restriction to the class of terms with a single variable. Since $[\![\times_{i\in n} P_i]\!]:=\times_{i\in n}[\![P_i]\!]$, from

$$[\![f]\!]\in\mathcal{C}\left([\![\times_{i\in n}P_i]\!],\ [\![Q]\!]\right)$$

we get e.g.

$$[\![\lambda X_0.f]\!]:=[\![f]\!]`\in\mathcal{C}\left(\times_{i\neq 0}[\![P_i]\!],\ [\![P_0]\!]\to[\![Q]\!]\right);$$

and for $[\![s]\!]\in\mathcal{C}([\![S]\!],[\![P_0]\!])$, there is

$$[\![f[X_0:=s]]\!]:=[\![f]\!]\circ\left([\![s]\!]\times_{i\neq 0}[\![P_i]\!]\right)\in\mathcal{C}\left([\![S]\!]\times_{i\neq 0}[\![P_i]\!],\ [\![Q]\!]\right).$$

By the lemmas I.1.31-2, all this can immediately be extended to dependent types:

$$X_0{:}P_0, X_1{:}P_1(X_0), \ldots, X_n{:}P_n(X_{n-1}) \Rightarrow f(X_0,\ldots,X_n) : Q(X_0,\ldots,X_n)$$

$$\downarrow$$

$$Z{:}\Sigma X_0{:}P_0.\big(\Sigma X_1 \ldots (\Sigma X_{n-1}{:}P_{n-1}.P_n)\big) \Rightarrow f(\pi_0 Z,\ldots,\pi_n Z) : Q(\pi_0 Z,\ldots,\pi_n Z).$$

Note that the sum in this last context is isomorphic with

$$\Sigma Z_{n-1}{:}\Big(\Sigma Z_{n-2}\Big(\ldots\big(\Sigma Z_1{:}(\Sigma X_0{:}P_0.P_1).P_2\big)\ldots P_{n-2}\Big).P_{n-1}\Big).P_n.$$

Top-down, bottom-up, and all the mixed applications of Σ on a sequential single-sorted context lead to isomorphic results.

Terminology. To *bind a context* Γ means to apply the operation $\Sigma\Delta\Delta$ somewhere in it. We say that a context is *bound* when this cannot be done (any more). A bound context obtained from Γ will be denoted by $\Sigma(\Gamma)$.

We say that a type or term is *packed* when it is presented with a bound context. To deal with the terms more naturally, we shall sometimes pack them in two parts. Given

$\Gamma \Rightarrow R{:}\Delta$, and

$\Gamma,\Psi \Rightarrow r{:}R{:}\Delta$,

where all the elements of Ψ are in the sort Δ, the term r is *packed in two parts* when it is presented in the form

$\Sigma(\Gamma), \Sigma(\Psi) \Rightarrow r{:}R{:}\Delta$.

($\Sigma(\Psi)$ is of course the result of applying $\Sigma\Delta\Delta$ in Ψ as long as possible.)

Packed types and terms. Clearly, every context can be bound. Different approaches to the sequentialisation and binding produce isomorphic results.

A sequentialized context in a theory of predicates must be in the form

$\Gamma = \Gamma_\Theta, \Gamma_\Omega$

where

$\Gamma_\Theta := (X_0{:}K_0,\ X_1{:}K_1, \ldots,\ X_m{:}K_m)$,

$\Gamma_\Omega := (x_0{:}\alpha_0,\ x_1{:}\alpha_1,\ \ldots,\ x_n{:}\alpha_n)$,

$m,n \in \omega$. Clearly $\Sigma(\Gamma_\Theta)$ is in the form $(X{:}K)$; $\Sigma(\Gamma)$ is $(X{:}K,\ x{:}\alpha)$.

A packing of a set $\Gamma_\Theta \Rightarrow M_1$ will be in the form

$X{:}K \Rightarrow M_1$;

a packing of a function $\Gamma_\Theta, \Psi_\Theta \Rightarrow m{:}M_1$ in two parts will be

$X{:}K,\ Z{:}M_0 \Rightarrow m{:}M_1$;

a packing of a proposition $\Gamma_\Theta, \Gamma_\Omega \Rightarrow \gamma_1$ will be

$X{:}K,\ x{:}\alpha \ \Rightarrow \gamma_1$;

and a packing in two parts of a proof $\Gamma_\Theta, \Gamma_\Omega, \Psi_\Omega \Rightarrow c{:}\gamma_1$ will be

$X{:}K,\ x{:}\alpha,\ z{:}\gamma_0 \ \Rightarrow c{:}\gamma_1$.

Unpacking. A set

$X{:}K, Y{:}L(X) \Rightarrow M(X,Y)$

will be packed to

$Z{:}\Sigma X{:}K.L(X) \Rightarrow M(\pi_0 Z, \pi_1 Z)$

and then interpreted by an a-arrow

$\llbracket M \rrbracket \in |\mathsf{a}\downarrow\llbracket \Sigma X{:}K.L(X)\rrbracket|$.

Before this, an interpretation of $X{:}K \Rightarrow L(X)$

$\llbracket L \rrbracket \in |\mathsf{a}\downarrow\llbracket K \rrbracket|$

must have been known, since $\llbracket \Sigma X{:}K.L(X)\rrbracket$ has been derived from it.

To *unpack* M means to look at it as depending on L and K, and not on $\Sigma X{:}K.L$. To unpack $\llbracket M \rrbracket$ means to view it as an arrow to $\llbracket L \rrbracket$, and not to $\llbracket \Sigma X{:}K.L \rrbracket$.

$$\begin{array}{ccc} \llbracket \Sigma Z{:}(\Sigma X{:}K.L).M \rrbracket & & \\ {\scriptstyle \llbracket M \rrbracket}\downarrow & & \\ \llbracket \Sigma X{:}K.L \rrbracket & \xrightarrow{\llbracket L \rrbracket} & \llbracket K \rrbracket \end{array}$$

I.e.,

$\llbracket M \rrbracket \in |(\mathsf{a}\downarrow\llbracket K \rrbracket)\downarrow\llbracket L \rrbracket|$,

where $(\mathsf{a}\downarrow\llbracket K \rrbracket)\downarrow\llbracket L \rrbracket$ is the $\llbracket L \rrbracket$-fibre of the full subfibration $\mathsf{a}/(\mathsf{a}\downarrow\llbracket K \rrbracket) \longrightarrow (\mathsf{a}\downarrow\llbracket K \rrbracket)$ of $\nabla(\mathsf{a}\downarrow\llbracket K \rrbracket)$ spanned by a. (Remember that $\mathsf{a}\downarrow\llbracket K \rrbracket$ is the $\llbracket K \rrbracket$-fibre of the full subfibration $\mathsf{a}/\mathcal{B} \longrightarrow \mathcal{B}$ of $\nabla\mathcal{B}$, spanned by a; in other words, of $\mathsf{a}/(\mathsf{a}\downarrow 1) \longrightarrow (\mathsf{a}\downarrow 1)$, since $\mathcal{B} \cong \mathsf{a}\downarrow 1$.)

Remark. Packing makes no sense in the theory of constructions, since a packed context can still be arbitrarily long there. A formal interpretation of the theory of constructions and a description of its term models tend to be quite a bit more complicated.

4. Definition of the interpretation.

The essence of the interpretation is the fixed correspondence between the operations of the theory of predicates, and those which constitute the structure of a category of predicates. Every model assignment

$$\llbracket \_ \rrbracket : \Lambda \longrightarrow E$$

of a category of predicates E to a theory of predicates Λ is defined by a structural recursion which follows this correspondence of operations. The ground case of the recursion - the interpretation of the generators of Λ - must be chosen, respecting some conditions (413 below).

Note, however, that the generators are added dynamically in type theory: a basic type or term may vary over derived types. Of course, a generator can be assigned a meaning in a model only when all the elements of its context have been interpreted. But clearly, this is a well founded process.

In order to simplify the assignment of an interpretation to a generator, we shall always pack it. If the interpretation of a context Γ is known, the interpretation of $\Sigma(\Gamma)$ can be readily obtained. And when the interpretation of a packed type or term is known, it can easily be unpacked in the model.

Let us now fix a theory of predicates Λ, and a category of predicates E, and list the items of the interpretation.

41. Types and terms. To simplify notation, we shall often use the same name for a type/term and its interpretant:

$$\llbracket K \rrbracket = K,$$
$$\llbracket X{:}K \Rightarrow \alpha \rrbracket = \alpha,$$
$$\llbracket X{:}K, x{:}\alpha \Rightarrow f(X,x){:}\beta(X) \rrbracket = f, \text{ etc.}$$

411. Constants.

$$\llbracket 1{:}\Theta \rrbracket := 1 \in |\mathcal{B}|$$
$$\llbracket X{:}K \Rightarrow \emptyset{:}1 \rrbracket := \emptyset_{\llbracket K \rrbracket} \in \mathcal{B}\big(\llbracket K \rrbracket, 1\big)$$
$$\llbracket \top{:}\Omega \rrbracket := \top \in |\mathcal{E}|$$
$$\llbracket X{:}K, x{:}\alpha \Rightarrow \emptyset{:}\top \rrbracket := \emptyset_{\llbracket \alpha \rrbracket} \in \mathcal{E}_{\llbracket K \rrbracket}\big(\llbracket \alpha \rrbracket, \top\llbracket K \rrbracket\big)$$
$$\llbracket \Omega{:}\Theta \rrbracket := \Omega \in |\mathcal{B}|$$

412. Variables.

$$\llbracket X{:}P \rrbracket := id_{\llbracket P \rrbracket}$$
$$\llbracket \xi{:}\Omega \rrbracket := \xi \in |\mathcal{E}_{\Omega}|$$

413. Generators (atoms).

$$\llbracket X{:}K \Rightarrow M_1 \rrbracket \in |\mathsf{a}\downarrow\llbracket K \rrbracket|$$
$$\llbracket X{:}K, Z{:}M_0 \Rightarrow m{:}M_1 \rrbracket \in \mathsf{a}\downarrow\llbracket K \rrbracket\,\big(\llbracket M_0 \rrbracket, \llbracket M_1 \rrbracket\big)$$
$$\llbracket X{:}K, x{:}\alpha \Rightarrow \gamma_1 \rrbracket \in |\mathsf{r}_{\llbracket K \rrbracket}\downarrow\llbracket \alpha \rrbracket|$$
$$\llbracket X{:}K, x{:}\alpha, z{:}\gamma_0 \Rightarrow b{:}\gamma_1 \rrbracket \in \mathsf{r}_{\llbracket K \rrbracket}\downarrow\llbracket \alpha \rrbracket\,\big(\llbracket \gamma_0 \rrbracket, \llbracket \gamma_1 \rrbracket\big)$$

42. Substitution. If $\Gamma \Rightarrow \mathbb{r}{:}R$, and $X^P \in \Gamma$, the substitution of $p(Z^Q){:}P$ for X^P produces $\tilde{\Gamma} \Rightarrow \mathbb{r}{:}R[X^P{:=}p]$, with Z^Q replacing X^P in $\tilde{\Gamma}$. It is routine to show that p induces a unique term from $\Sigma(\tilde{\Gamma})$ to $\Sigma(\Gamma)$. Just as in 2 above, it is sufficient to take into account the substitution in packed types and terms. We now suppose that they are packed in one part.

421. A function

$$u := \llbracket Z{:}H \Rightarrow u{:}K \rrbracket \in \mathcal{B}(H,K),$$

is substituted for the element variable X^K in $\llbracket X{:}K \Rightarrow \mathbb{m}{:}M \rrbracket$ by the $\nabla\mathsf{a}$-inverse images, i.e. pullbacks:

$$\llbracket M[X^K{:=}u(Z^H)] \rrbracket \quad := u^*\llbracket M \rrbracket \in |\mathsf{a}\downarrow H|,$$
$$\llbracket m[X^K{:=}u(Z^H)] \rrbracket \quad := u^*\llbracket m \rrbracket \in \mathsf{a}\downarrow H\big(id, u^*\llbracket M \rrbracket\big).$$

422. Before u is substituted for X^K in $\llbracket X{:}K, x{:}\alpha \Rightarrow \mathbb{c}{:}\gamma \rrbracket \in |\mathsf{r}_K\downarrow\llbracket \alpha \rrbracket|$, it must be substituted in $\llbracket \alpha \rrbracket \in |\mathcal{E}_K|$. The E-inverse images do this.

$$\llbracket \alpha[X^K{:=}u(Z^H)] \rrbracket \quad := u^*\llbracket \alpha \rrbracket \in |\mathcal{E}_H|,$$
$$\llbracket \gamma[X^K{:=}u(Z^H)] \rrbracket \quad := u^*\llbracket \gamma \rrbracket \in |\mathsf{r}_H\downarrow u^*\alpha|,$$

$[\![c[X^K{:=}u(Z^H)]]\!] \quad := u^*[\![c]\!] \in r_H{\downarrow}u^*\alpha\left(\mathrm{id},\, u^*[\![\gamma]\!]\right).$

423. Let a proof

$f := [\![X{:}K,Y{:}L,z{:}\psi \Rightarrow f{:}\alpha]\!] \in \mathcal{E}_{\Sigma L}(\psi, L^*\alpha),$

be given, where

$L := [\![X{:}K \Rightarrow L]\!] \in |\mathfrak{a}{\downarrow}K|.$

$\Sigma L := [\![\Sigma X{:}K.L]\!] = \mathrm{Dom}(L),$

$\psi := [\![\psi]\!] \in |\mathcal{E}_{\Sigma L}|.$

To substitute f for x^α in $[\![X{:}K,x{:}\alpha \Rightarrow c{:}\gamma]\!]$, we first add dummy variable Y^L in $c{:}\gamma$, i.e. take the E-inverse image of $[\![c{:}\gamma]\!]$ along L. When all the types and terms are in the same fibre $\mathcal{E}_{\Sigma L}$, vertical pullbacks, i.e. $\nabla r_{\Sigma L}$-inverse images are used to interprete substitution.

$[\![\gamma[x^\alpha{:=}f(z^\psi)]]\!] \quad := f^*\left(L^*[\![\gamma]\!]\right) \in |r_{\Sigma L}{\downarrow}\psi|$

$[\![c[x^\alpha{:=}f(z^\psi)]]\!] \quad := f^*\left(L^*[\![c]\!]\right) \in r_{\Sigma L}{\downarrow}\psi\left(\mathrm{id},\, f^*[\![\gamma]\!]\right).$

(By II.2.25, inverse image along L + $\mathcal{E}_{\Sigma L}$-pullback along f = $\mathcal{E}$-pullback along $\vartheta^L \circ f \in \mathcal{E}_L(\psi,\alpha)$.)

43. Quantifiers. A variable which is to be bound must be unpacked. We shall now consider a partially unpacked proposition

$X{:}K, Y{:}L(X) \Rightarrow \gamma,$

with an interpretant

$\gamma := [\![\Sigma X{:}K.L \Rightarrow \gamma]\!] \in |\mathcal{E}_{\Sigma L}|,$

Furthermore, we shall need

$X{:}K,Y{:}L,y{:}\beta(\not{Y}) \Rightarrow c{:}\gamma,$

$X{:}K,y{:}\beta \Rightarrow d{:}\forall Y{:}L.\gamma$

$X{:}K,Y{:}L,z{:}\gamma \Rightarrow f{:}\varphi(\not{Y},\not{z}),$

$X{:}K,w{:}\exists Y{:}L.\gamma \Rightarrow g{:}\varphi(\not{w})$

($Y^L \notin DV(\beta)$ means that the condition $Y^L \in MIN(c{:}\gamma)$ is still satisfied, although there is $y^\beta \in DV(c)$. $Y^L, z^\gamma \notin DV(\varphi)$ is the familiar condition on E∃.) Given β, $\varphi \in |\mathcal{E}_K|$, these restrictions just mean

$c := [\![c]\!] \in \mathcal{E}_{\Sigma L}(L^*\beta, \gamma),$

$d := [\![d]\!] \in \mathcal{E}_K(\beta, L_*\gamma)$

$f := [\![f]\!] \in \mathcal{E}_{\Sigma L}(\gamma, L^*\varphi),$

$g := [\![g]\!] \in \mathcal{E}_K(L_!\gamma, \varphi).$

The quantification will be interpreted by:

431. the right bifibration structure of E:

$[\![\forall Y{:}L.\gamma]\!] \quad := L_*\gamma \in |\mathcal{E}_K|,$

$[\![\lambda Y.c]\!] \quad := c` \in \mathcal{E}_K(\beta, L_*\gamma),$

$[\![dY]\!] \quad := `d \in \mathcal{E}_{\Sigma L}(L^*\beta, \gamma);$

432. the left bifibration structure of E:

$[\![\exists Y{:}L.\gamma]\!] \quad := L_!\gamma \in |\mathcal{E}_K|,$

$[\![\nu(w,(Y,z).f]\!] := `f \in \mathcal{E}_K(L_!\gamma, \varphi)$

$[\![\langle Y,z\rangle]\!] \quad := \eta \in \mathcal{E}_{\Sigma L}(\gamma, L^*L_!\gamma).$

To check the soundness, note that

β∀: $(\lambda Y.c)Y = c$ means $`(c`) = c$, and

η∀: $\lambda Y.(dY) = d$ is $(`d)` = d$;

β∃: $\nu(\langle Y,z\rangle,(Y,z).f) = f$ is translated in $L^*(`f)\circ\eta = f$, and

η∃: $\nu(w,(Y,z).g(\langle Y,z\rangle)) = g(w)$ in $`(L^*(g)\circ\eta) = g$.

44. Sums and products. The Martin-Löf theories of sets and of propositions, contained in the theory of predicates - its ΘΘ- and ΩΩ-fragments - are interpreted by the relatively cartesian closed structures in the base $\mathfrak{B}$, and in the fibres $\mathcal{E}_K$ respectively - in the standard way, exhaustively treated in the literature (referred to in part 1 above). Propositions must be brought in the same fibre (under the same context of sets) using the E-inverse images (i.e. adding dummy variables).

45. Extents. For $\alpha = [\![X{:}K \Rightarrow \alpha]\!] \in |\mathcal{E}_K|$ and $a = [\![X{:}K,Y{:}L \Rightarrow a{:}\alpha]\!] \in \mathcal{E}_{\Sigma L}(\top, L^*\alpha)$ define

$[\![X{:}K \Rightarrow \iota\alpha]\!] \quad := \quad \iota\alpha \in |\mathfrak{a}{\downarrow}K|,$

$[\![X{:}K,Y{:}L \Rightarrow \delta a{:}\iota\alpha]\!] \quad := \quad D(\vartheta^L \circ a) \in \mathfrak{a}{\downarrow}K(L,\iota\alpha).$

For arbitrary $u = [\![X{:}K,Y{:}L \Rightarrow u{:}\iota\alpha]\!] \in \mathfrak{a}{\downarrow}K(L,\iota\alpha),$

$[\![X{:}K,Y{:}L \Rightarrow \tau u{:}\alpha]\!] \quad := \quad u^*(\tau\alpha) \in \mathcal{E}_{\Sigma L}(\top, L^*\alpha),$

where $\tau\alpha \in \mathcal{E}_{D\alpha}\big(\top D\alpha, \iota\alpha^*(\alpha)\big)$ is the vertical component of the counit $\varepsilon_\alpha \in \mathcal{E}_{\iota\alpha}(\top D\alpha, \alpha)$.

The arrow $[\![\delta a]\!]$ is indeed an element of $\mathcal{A}\!\downarrow\!K(L,\iota\alpha)$ because

$$\iota\alpha \circ D(\vartheta^L_\alpha \circ a) = D(\eta_\alpha \circ \vartheta^L_\alpha \circ a) = D\big(\top E(\vartheta^L_\alpha) \circ \eta_{L*\alpha} \circ a\big) = D\top L = L.$$

To check the interpretation of the conversion rules, note that the adjunction $\top \dashv D$ induces

$$\mathcal{E}(\top\Sigma L, \alpha) \ni \vartheta^L \circ a \quad \mapsto \quad D(\vartheta^L \circ a) \quad \in \mathcal{B}(\Sigma L, D\alpha)$$
$$\mathcal{B}(\Sigma L, D\alpha) \ni u \quad \mapsto \quad \vartheta^L \circ u^*(\tau\alpha) \quad \in \mathcal{E}(\top\Sigma L, \alpha).$$

The rule

$$\beta\iota: \qquad \delta(\tau u) = u$$

is translated to the equation $D(\vartheta^L \circ u^*(\tau\alpha)) = u$, which just says that $(`u)` = u$; while

$$\eta\iota: \qquad \tau(\delta a) = a$$

boils down to the requirement that $\big(D(\vartheta^L \circ a)\big)^*\tau\alpha = a$, which is true, because the left side is the vertical component of $`\big((\vartheta^L \circ a)`\big) = \vartheta^L \circ a$.

Remark. Translated to the notation from 432, proposition I.1.52 tells that a theory of predicates does not become stronger if the restriction $w \notin DV(\varphi)$ is dropped. This means that our categorical interpretation must remain sound if we consider

$$X{:}K, w{:}\exists Y{:}L.\gamma \Rightarrow \varphi'(w), \text{ and}$$
$$X{:}K, Y{:}L, z{:}\gamma \Rightarrow f'(X,Y,z){:}\varphi'(\langle Y,z\rangle),$$

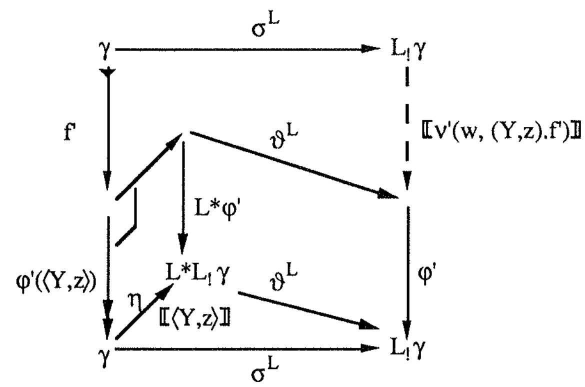
i.e. the interpretants

$$\varphi' \in |\mathcal{r}_K\!\downarrow\! L_!\gamma|.$$
$$\varphi'(\langle Y,z\rangle) := \eta^*\big(L^*\varphi'\big) \in |\mathcal{r}_{\Sigma L}\!\downarrow\!\gamma|, \text{ and}$$
$$f' \in \mathcal{r}_{\Sigma L}\!\downarrow\!\gamma\big(\mathrm{id}, \varphi'(\langle Y,z\rangle)\big).$$

The term

$$X{:}K, w{:}\exists Y{:}L.\gamma \Rightarrow \nu'(w,(Y,z).f') : \varphi',$$

constructed in I.1.52 will be interpreted by the vertical factorisation shown at the following diagram.



Is this interpretation well-defined? It has been given by an induction along the derivability relation ($\vdash$); while only a type or a term (with its context) is actually being interpreted. Different derivations might result in different interpretants. To prove that this will not be the case, one should show that any two operations which commute in the theory (so that they could be applied in various orders and produce different derivations of a type or term) are interpreted by operations which commute in all models. For instance, the Beck-Chevalley condition interpretes the commutativity of the quantifiers with substitution (as we explained in II.3.3). The fact that the extents are stable under the inverse images (III.2.5) reflects the commutation of the extent operation and substitution in the theory. And it follows from the propositions III.4.1-2 that the relation of the extents and quantifiers, sums and products is the same as in the theory, as described in I.1.8.

But a detailed proof of this waits to be written down.

5. Internal language of a category of predicates.

By this interpretation, to every category of predicates E corresponds a theory of predicates $\underline{\Lambda}(E)$ in a natural way:

- the packed sets and functions of $\underline{\Lambda}(E)$ are the objects and arrows of $\mathcal{A}/\mathcal{B}$;

- the packed propositions and proofs of $\underline{\Lambda}(E)$ are the objects and arrows of $\mathfrak{r}/\mathcal{E}$;
- the operations of $\underline{\Lambda}(E)$ are defined by the structure of E, as indicated in the interpretation.

Using projections, the variable types and terms are unpacked. They could have been defined directly too, as appropriate well-founded diagrams in $\mathfrak{a}$ and in $\mathfrak{r}$.

The canonical model assignment $[\![\_]\!]:\underline{\Lambda}(E) \longrightarrow E$ is the identical mapping: to each packed type/term, it assigns that same type/term, regarded as object/arrow. For every theory of predicates Λ' and each model assignment $[\![\_]\!]':\Lambda' \longrightarrow E$, there is a unique translation $\Phi:\Lambda' \longrightarrow \Lambda$, which preserves all the operations of the theory of predicates, and such that

$$[\![\_]\!]' \simeq [\![\_]\!]\circ\Phi.$$

To interprete a theory of predicates in E means to translate it into $\underline{\Lambda}(E)$.

In this sense, $\underline{\Lambda}(E)$ is the *internal language of* category of predicates E.

2. Term models

1. From a theory to a category of predicates.

In the preceding section we saw how to produce a theory of predicates, given a category of predicates. Now we go the other way round.

Let Λ be a theory of predicates. The "free" category of predicates $\underline{E}(\Lambda):\mathcal{E} \longrightarrow \mathfrak{B}$ generated by Λ consists of the following data.

11. Base category $\mathfrak{B}$:

$|\mathfrak{B}|$:= closed sets K,

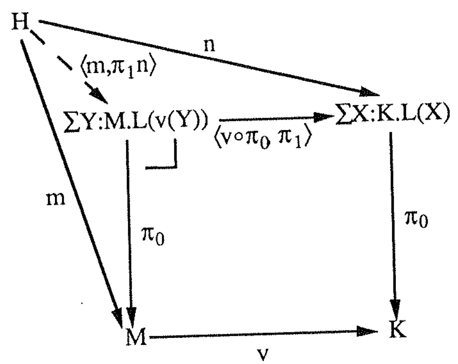
$\mathfrak{B}(I,J)$:= closed functions $u:I\to J$.

Of course, the functions are taken modulo conversion (=). Identities, composition, and the canonical cartesian closed structure of $\mathfrak{B}$ are recognized by the notation in the theory of predicates. (Cf. I.1.2)

12. The class $\mathfrak{a} \subseteq \mathfrak{B}$ of display maps consists of all the terms (modulo conversion) isomorphic to some first projection $\pi_0 \in \mathfrak{B}(\Sigma X{:}K.L,\ K)$:

$$\mathfrak{a} := \{u:I\to J \mid \exists \text{ sets } K, L(X^K)\ \exists \text{ isos } i:I\to\Sigma X^K.L,\ j:K\to J.\ u=j\circ\pi_0\circ i\}$$

$\mathfrak{a}$ is a stable display subcategory (cf. II.4.3). It is saturated, and satisfies the display condition because the identities and the terminal arrows are special projections. It is stable because projections are.



The diagram showing that the projections are closed under composition has been drawn in 1.1. Putting all this together, the functor

$$\nabla\mathfrak{a} : \mathfrak{a}/\mathfrak{B} \longrightarrow \mathfrak{B}$$

is an intrinsic left $\mathfrak{a}$-bifibration. It is easy to check that it is a right $\mathfrak{a}$-bifibration too. For

$$L \quad := \pi_0 \in \mathfrak{B}(\Sigma X{:}K.L,\ K)$$
$$M \quad := \pi_0 \in \mathfrak{B}\big(\Sigma Z{:}(\Sigma X{:}K.L).M,\ \Sigma X{:}K.L\big),$$

the right direct image is

$$L_*M \quad := \pi_0 \in \mathfrak{B}\big(\Sigma X{:}K.\Pi Y{:}L.M,\ K\big).$$

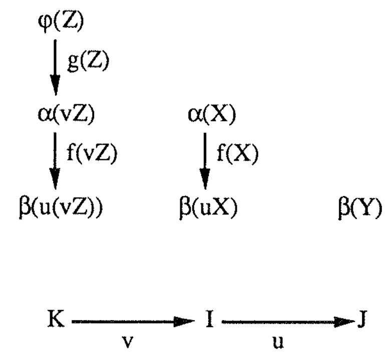
So $\mathfrak{B}$ is an $\mathfrak{a}$-rccc (by proposition II.4.6).

13. The fibred category of predicates $\mathcal{E}$ consists of

$$|\mathcal{E}| \quad := \text{the predicates } \varphi(X^K),\ K \text{ closed}$$
$$\mathcal{E}\big(\alpha(X^I),\ \beta(Y^J)\big) \quad := \Big\{\big\langle u{:}I\to J,\ f{:}\alpha(X^I)\to\beta(uX^I)\big\rangle \mid \begin{matrix}\text{u closed, } X^I \text{ the}\\ \text{only variable of f}\end{matrix}\Big\}$$

(The pairing operation in this definition comes, of course, from the metalanguage.) The composition in $\mathcal{E}$ is (just like in the Grothendieck construction):

$$\langle u,f\rangle\circ\langle v,g\rangle := \langle u\circ v,\ v^*(f)\circ g\rangle.$$



$\mathcal{E}$ is fibred over $\mathfrak{B}$ by the obvious projection

$$E{:}\mathcal{E} \longrightarrow \mathfrak{B}{:}\ \varphi(X^K) \longmapsto K,\ \langle u,f\rangle \longmapsto u.$$

Since the substitution induces its inverse images and cartesian arrows:

$$u^*\big(\beta(Y^J)\big) \quad := \beta(uX^I)$$
$$\vartheta^u_\beta \quad := \langle u,\ id_\beta\rangle,$$

this fibration is split (and normal). Each fibre $\mathcal{E}_K$ has a canonical cartesian closed structure, preserved by the inverse images.

14. The class $\mathfrak{r}_K \subseteq \mathcal{E}_K$ of display maps will be

$$\mathfrak{r}_K := \Big\{\big\langle id_K,\ f{:}\gamma(X^K)\to\varphi(X^K)\big\rangle \mid \exists \text{ isos } i{:}\gamma\to\Sigma x^{\alpha}.\beta,\ j{:}\alpha\to\varphi.\ f=j\circ\pi_0\circ i\Big\}$$

$\mathcal{E}_K$ is an $\mathfrak{r}_K$-rccc by the same argument as in 12. Since substitution commutes with all the operations which constitute the rccc structure of $\mathcal{E}_K$, the category $\mathcal{E}$ is fibrewise $\mathfrak{r}$-rccc, for

$$\mathfrak{r} := \bigcup_{K\in|\mathfrak{B}|} \mathfrak{r}_K$$

By II.4.7, $\mathcal{E}$ is an $\mathfrak{r}$-rccc.

15. E is an $\mathfrak{a}$-hyperfibration. Consider a projection $L=\pi_0 \in \mathfrak{B}(\Sigma X{:}K.L,\ K)$, and an object $\gamma\in|\mathcal{E}_{\Sigma X:K.L}|$. The predicate $Z{:}\Sigma X{:}K.L \Rightarrow \gamma(Z)$ can be unpacked as

$$X{:}K,\ Y{:}L(X) \Rightarrow \gamma(\langle X,Y\rangle).$$

The direct images are now

$$L_!\gamma \quad := \exists Y{:}L.\gamma(\langle X,Y\rangle), \text{ and}$$

$$L_* \gamma \quad := \forall Y{:}L.\gamma(\langle X,Y\rangle).$$

The canonical cocartesian and opcartesian liftings are:

$$\sigma^L_\gamma \quad := \big\langle L,\ \lambda x.\langle \pi_1 Z,\ x\rangle{:}\ \gamma(Z) \to \exists Y{:}L.\gamma(\langle \pi_0 Z, Y\rangle)\big\rangle,$$

$$\psi^L_\gamma \quad := \vartheta^{\ L}_{L_*\gamma} \,/\, \varepsilon^L_\gamma, \text{ \underline{where}}$$

$$\varepsilon^L_\gamma := \big\langle \mathrm{id},\ \lambda y.y(\pi_1 Z){:}\ \forall Y{:}L.\gamma(\langle \pi_0 Z, Y\rangle) \to \gamma(Z)\big\rangle.$$

An inverse image functor along an isomorphism i must always be a strong equivalence of categories: the functor $(i^{-1})^*$ is left and right adjoint to i^*; in the case of a split normal fibration, i^* is an isomorphism of categories, with $(i^*)^{-1} = (i^{-1})^*$. Thus, for an arbitrary display map $a = j \circ L \circ i \in \mathcal{a}$, where i,j are isos and L is a projection, the direct images will be

$$a_\square \gamma \quad := (j^{-1})^* \circ L_\square \circ (i^{-1})^*,\ \square \in \{!, *\}.$$

16. E is globally small. Each propositional variable $\xi{:}\Omega$ is a generic object $\xi \in |\mathcal{E}_\Omega|$. A predicate $\alpha(X^K) \in |\mathcal{E}_K|$ is classified by the function

$$\ulcorner \alpha \urcorner := \lambda X^K.\alpha(X^K){:}K \to \Omega$$

(i.e. α is an inverse image of ξ along $\ulcorner \alpha \urcorner$; cf. III.2.2.)

17. E is comprehensive. This follows by proposition III.3.4 from the fact that E has a full and faithful right adjoint

$$\top : \mathcal{B} \longrightarrow \mathcal{E} : K \longmapsto (X{:}K \Rightarrow \top),\ (u{:}I \to J) \longmapsto \langle u,\ \mathrm{id}_{\top I}\rangle,$$

which has a right adjoint

$$C : \mathcal{E} \longrightarrow \mathcal{B} : \quad \varphi(X^K) \quad \longmapsto \Sigma X{:}K.\iota\varphi,$$

$$\big(\langle u,f\rangle{:}\alpha(X^I) \longrightarrow \beta(Y^J)\big) \longmapsto \lambda Z.\big\langle (u\circ\pi_0)Z,\ (\delta\circ f(\pi_0 Z)\circ\tau\circ\pi_1)Z\big\rangle,$$

(where $f(\pi_0 Z){:}{=}f[X^I{:}{=}\pi_0 Z]$ is an instance of $f(X){:}\alpha(X) \to \beta(uX)$).

• Is C a functor? C(id,id) = id, follows from the rule $\beta\iota$. For arrows as above in 13, it holds

$$C\langle u,f\rangle \circ C\langle v,g\rangle =$$

$$= \lambda Z.\big\langle (u\circ\pi_0)Z,\ (\delta\circ f(\pi_0 Z)\circ\tau\circ\pi_1)Z\big\rangle \circ \lambda W.\big\langle (v\circ\pi_0)W,\ (\delta\circ g(\pi_0 W)\circ\tau\circ\pi_1)W\big\rangle =$$

$$= \lambda WZ.\big\langle (u\circ\pi_0)Z,\ (\delta\circ f(\pi_0 Z)\circ\tau\circ\pi_1)Z\big\rangle\Big(\big\langle (v\circ\pi_0)W,\ (\delta\circ g(\pi_0 W)\circ\tau\circ\pi_1)W\big\rangle\Big) =$$

$$= \lambda W.\big\langle (u\circ v\circ\pi_0)W,\ (\delta\circ f((v\circ\pi_0)W)\circ\tau\circ\delta\circ g(\pi_0 W)\circ\tau\circ\pi_1)W\big\rangle =$$

$$= \lambda W.\big\langle (u\circ v\circ\pi_0)W,\ (\delta\circ f((v\circ\pi_0)W)\circ g(\pi_0 W)\circ\tau\circ\pi_1)W\big\rangle =$$

$$= C\langle u\circ v,\ v^* f\circ g\rangle = C(\langle u,f\rangle \circ \langle v,g\rangle).$$

Now we check $\top \dashv C$.

$$\eta_K := \lambda X^K.\langle X^K, \varnothing\rangle \qquad \in \mathcal{B}(K, C\top K),$$

$$\varepsilon_{\varphi(X^K)} := \langle \pi_0{:}C\varphi \to K,\ \tau(\pi_1 Z^{C\varphi}){:}\varphi\rangle \in \mathcal{E}(\top C\varphi, \varphi).$$

$$\varepsilon_{\top K} \circ \top\eta_K = \langle \pi_0{:}\Sigma X{:}K.\top \to K,\ \tau\varnothing{:}\top\rangle \circ \langle \lambda X^K.\langle X^K, \varnothing\rangle, \mathrm{id}\rangle =$$
$$= \langle \mathrm{id}_K, \mathrm{id}_\top\rangle = \mathrm{id}_{\top K}.$$

$$C\varepsilon_\varphi \circ \eta_{C\varphi} = \lambda Z^{C\top C\varphi}.\big\langle (\pi_0\circ\pi_0)Z,\ (\delta\circ\tau((\pi_1\circ\pi_0)Z)\circ\tau\circ\pi_1)Z\big\rangle \circ \lambda X^{C\varphi}.\langle X, \varnothing\rangle =$$

$$= \lambda X^{C\varphi} Z^{C\top C\varphi}.\big\langle (\pi_0\circ\pi_0)Z,\ (\delta\circ\tau((\pi_1\circ\pi_0)Z)\circ\tau\circ\pi_1)Z\big\rangle\langle X, \varnothing\rangle =$$

$$= \lambda X^{C\varphi}.\big\langle \pi_0 X,\ \delta\circ\tau(\pi_1 X)\circ\tau(\varnothing)\big\rangle =$$

$$= \lambda X^{C\varphi}.\ \langle \pi_0 X, \pi_1 X\rangle = \mathrm{id}_{C\varphi}. \bullet$$

Note that the extent of $\varphi(X^K)$ is

$$\iota\varphi := E(\varepsilon_\varphi) = \pi_0{:}C\varphi \to K,$$

so that the requirement

$$\iota \subseteq \mathcal{a}$$

is satisfied.

18. Since E is fibrewise cartesian closed and comprehensive, it is locally small, by fact III.2.4. Since it is globally small too, E is a small fibration. This follows from proposition III.2.2, with the following two adjustments. For a *split* locally and globally small fibration E, this proposition gives not just a fibrewise equivalence, but an isomorphism $\nabla\Omega \longrightarrow E$ (as we already noticed at the end of 1.1). On the other hand, it goes through with a weaker assumption than that of finite completeness of $\mathcal{B}$: it is sufficient that the representants $\iota(X,Y)$ used in the construction belong to a stable display subcategory $\mathcal{a} \subseteq \mathcal{B}$ - and this is the case here.

$$\Omega_0 := \Omega;$$

$$\Omega_1 := \Sigma\xi_0\xi_1.\iota(\xi_0 \to \xi_1);$$

$\Omega_2 := \Sigma\xi_0\xi_1\xi_2.\iota(\xi_0\to\xi_1)\times\iota(\xi_1\to\xi_2)$

$\partial_i := \lambda Z.\pi_i Z : \Omega_1\to\Omega_0,\ i\in 2;$

$\eta := \lambda\xi.\langle\xi,\ \xi,\ \delta(id_\xi)\rangle : \Omega_0\to\Omega_1;$

$\mu := \lambda Z.\Big\langle\pi_0 Z,\ \pi_2 Z,\ \delta\Big(\big(\tau(\pi_4 Z)\big)\big(\tau(\pi_3 Z)\big)\Big)\Big\rangle : \Omega_2\to\Omega_1.$

The isomorphism is

$\ulcorner\_\urcorner : E \longrightarrow \nabla\Omega : \varphi(X^K)\longmapsto \ulcorner\varphi\urcorner : K\to\Omega_0,$

$\big\langle u:I\to J,\ f(X):\alpha(X)\to\beta(uX)\big\rangle \longmapsto \big\langle u:I\to J,\ \ulcorner f\urcorner : I\to\Omega_1\big\rangle,$

where

$\ulcorner\varphi\urcorner := \lambda X^K.\varphi(X^K)$

$\ulcorner f\urcorner := \lambda X^I.\Big\langle\alpha(X^I),\ \beta(uX^I),\ \delta\big(f(X^I)\big)\Big\rangle.$

The inverse functor is obtained using

$\varphi = \ulcorner\varphi\urcorner * \xi,$

$f = \ulcorner f\urcorner * \gamma,$

where generic arrow $\gamma\in\mathcal{E}_{\Omega_1}(\partial_0 * \xi,\ \partial_1 * \xi)$ is the term

$Z:\Omega_1 \Rightarrow \tau(\pi_2 Z):\pi_0 Z\to\pi_1 Z.$

19. If Λ is a strong theory of predicates, i.e. if it obeys the rule δab, <u>then</u> $E:\mathcal{E}\longrightarrow\mathcal{B}$ is a category of constructions. • Using the rule δab, lemma I.1.87 gives an isomorphism

$a:\ \alpha \simeq \iota\alpha\times\top\ :b.$

It is easy to see that for the E-cocartesian arrow $\sigma_\alpha^{\iota\alpha}$ and counit ε_α of $\top\dashv$ C - both defined above - holds

$\sigma_\alpha^{\iota\alpha} = a\circ\varepsilon_\alpha \in \mathcal{E}_{\iota\alpha}(\top,\ \iota\alpha\times\top)$

$\varepsilon_\alpha = b\circ\sigma_\alpha^{\iota\alpha} \in \mathcal{E}_{\iota\alpha}(\top,\ \alpha);$

ε_α is thus cocartesian. By proposition III.4.3, the fibred category $\mathcal{E}$ must be equivalent with $\iota\mathcal{E}$.•

(N.B. This is the semantical version of proposition I.2.33.)

The term model is a model. There is an obvious model assignment

$[\![\_]\!] : \Lambda\longrightarrow \underline{E}(\Lambda),$

which (just as $[\![\_]\!] : \underline{\Delta}(E)\longrightarrow E$ in the preceding section) does not move the "material", but changes the point of view: every type and term is interpreted by itself.

If $E':\mathcal{E}'\longrightarrow\mathcal{B}'$ is another category of predicates, every model assignment

$[\![\_]\!]' : \Lambda\longrightarrow E'$

induces functors

$F_\Theta : \mathcal{B}\longrightarrow\mathcal{B}' : K\longmapsto [\![K]\!],\ (u:I\to J)\longmapsto [\![uX^I]\!]$, and

$F_\Omega : \mathcal{E}\longrightarrow\mathcal{E}' : \varphi(X^K)\longmapsto [\![\varphi]\!],$

$\big\langle u:I\to J,\ f(X):\alpha(X)\to\beta(uX)\big\rangle \longmapsto \vartheta^{[\![uX]\!]}[\![fx^\alpha{:}\beta(uX)]\!],$

which preserve the rccc-structure of their respective domains, and the horizontal structure of $E:\mathcal{E}\longrightarrow\mathcal{B}$:

$F_\Omega(L_\square\gamma) = (F_\Theta L)_\square(F_\Omega\gamma)$, for $\square\in\{*,!\}$.

So $F = \langle F_\Theta, F_\Omega\rangle$ preserves all the structure of categories of predicates; it can be regarded as a morphism $F:E\longrightarrow E'$ of categories of predicates. As such, F is the unique morphism $E\longrightarrow E'$ by which the model assignment $[\![\_]\!]'$ factorizes:

$[\![\_]\!]'_\Theta = F_\Theta\circ[\![\_]\!]_\Theta,$

$[\![\_]\!]'_\Omega = F_\Omega\circ[\![\_]\!]_\Omega.$

This factorisation will be written:

$[\![\_]\!]' = F\circ[\![\_]\!].$

2. Semantical completeness.

A semantical construction $\underline{E}(\_)$ is said to be *complete* if every theory Λ can be recovered from its model $\underline{E}(\Lambda)$. For logical theories, a completeness theorem has traditionally been in the form

$\forall P.\ \Lambda\vdash P \Leftrightarrow \underline{E}(\Lambda)\models P,$

where "$\Lambda\vdash P$" means " the formula P is provable in Λ" and "$\underline{E}(\Lambda)\models P$" asserts that "P is true in $\underline{E}(\Lambda)$", i.e. an element $[\![P]\!]\in\underline{E}(\Lambda)$, assigned to P, has some property which we call truth in $\underline{E}(\Lambda)$. When Λ is a type theory - regarded as a logical theory with constructive proofs - while $\underline{E}(\Lambda)$ is a category E - with terms-as-arrows - this correspondence should be refined to a bijection between proofs and their semantical realizations

$$\forall P \forall Q.\ \{X{:}Q \Rightarrow p{:}P\} \simeq E([\![Q]\!],[\![P]\!])$$

(as we already remarked in I.2.4). The *constructive* completeness of $\underline{E}(\_)$ thus means that the model assignment

$$[\![\_]\!] : \Lambda \longrightarrow \underline{E}(\Lambda)$$

is "full and faithful" for every Λ.

To express this precisely, we shall now describe the connection of theories and models, realized by model assignments, as an adjointness between a semantical functor $\underline{E}$ and a syntactical functor $\underline{\Lambda}$.

Terminology. Let $E{:}\mathcal{E} \longrightarrow \mathcal{B}$ and $E'{:}\mathcal{E}' \longrightarrow \mathcal{B}'$ be fibrations. We say that a pair of functors

$$F = \langle F_0{:}\mathcal{B} \longrightarrow \mathcal{B}',\ F_1{:}\mathcal{E} \longrightarrow \mathcal{E}'\rangle, \text{ such that } E'F_1 = F_0E,$$

preserves a property P if F_1 preserves this property. (Thus, F is a hyperfibration functor if the arrow $F_1(f)$ is (co-, op-)cartesian whenever f is.)

Categories. By definition, the category $\underline{TOP}$ consists of theories of predicates, with the translations which preserve all the structure (defined in I.1).

Objects of the category $\underline{CAP}$ are the categories of predicates. A morphism $F \in \underline{CAP}(E,E')$ is a pair of functors

$$F = \langle F_\Theta{:}\mathcal{B} \longrightarrow \mathcal{B}',\ F_\Omega{:}\mathcal{E} \longrightarrow \mathcal{E}'\rangle, \text{ such that}$$

$F_\Theta(\mathfrak{a}) \subseteq \mathfrak{a}'$, so that $F_\mathfrak{a} : \mathfrak{a}/\mathcal{B} \longrightarrow \mathfrak{a}'/\mathcal{B}' : u \longmapsto F_\Theta(u)$ is induced

$F_\Omega(\mathfrak{r}) \subseteq \mathfrak{r}'$, hence $F_\mathfrak{r} : \mathfrak{r}/\mathcal{E} \longrightarrow \mathfrak{r}'/\mathcal{E}' : f \longmapsto F_\Omega(f)$;

$$\left.\begin{array}{l} F = \langle F_\Theta, F_\Omega\rangle \\ F_\mathcal{B} = \langle F_\mathfrak{a}, F_\Theta\rangle \\ F_\mathcal{E} = \langle F_\mathfrak{r}, F_\Omega\rangle \end{array}\right\} \begin{array}{l}\text{must be hyperfibration functors}\\ \text{preserving terminal objects}\\ \text{and extents.}\end{array}$$

Since $E = \nabla\Omega$ and $E' = \nabla\Omega'$, it is not hard to see that F_Ω is completely determined by an internal functor $f_\Omega{:}\ F_\Theta\Omega \longrightarrow \Omega'$ in $\mathcal{B}'$ (cf. III.1.28).

Functors. A definition of the arrow part of the functor

$$\underline{E} : \underline{TOP} \longrightarrow \underline{CAP} : \Lambda \longmapsto \underline{E}(\Lambda)$$

follows from the observation that $\underline{E}(\Lambda)$ is universal among the models of Λ. An interpretation $\Phi \in \underline{TOP}(\Lambda,\Lambda')$ induces a model assignment

$$[\![\_]\!]^\circ := [\![\_]\!]' \circ \Phi : \Lambda \longrightarrow \Lambda' \longrightarrow \underline{E}(\Lambda'),$$

and this assignment induces, as we have seen above, a unique morphism $\underline{E}(\Phi) \in \underline{CAP}\big(\underline{E}(\Lambda),\ \underline{E}(\Lambda')\big)$, such that

$$\underline{E}(\Phi) \circ [\![\_]\!] = [\![\_]\!]' \circ \Phi.$$

The arrow part of

$$\underline{\Lambda} : \underline{CAP} \longrightarrow \underline{TOP} : E \longmapsto \underline{\Lambda}(E)$$

has been implicitely defined in the remark about the internal language, at the end of the preceding section. Every morphism $F \in \underline{CAT}(E,E')$ induces a model assignment

$$[\![\_]\!]^\circ := F \circ [\![\_]\!] : \Lambda \longrightarrow \underline{E}(\Lambda) \longrightarrow \underline{E}(\Lambda'),$$

and this assignment induces a unique translation $\underline{\Lambda}(F) \in \underline{TOP}(\Lambda,\Lambda')$, such that

$$F \circ [\![\_]\!] = [\![\_]\!]' \circ \underline{\Lambda}(F).$$

Comment. The equations we used here are equations of model assignments. A model assignment is just a mapping from a theory to its model (or from two sorts to two categories); it does not live in any of our categories, but "in between".

Adjointness. The unit and counit of $\underline{E} \dashv \underline{\Lambda}$ are induced as follows:

$\eta : \Lambda \longrightarrow \underline{\Lambda E}(\Lambda)$ is the unique translation by which $[\![\_]\!]' : \Lambda \longrightarrow \underline{E}(\Lambda)$ factorizes through $[\![\_]\!] : \underline{\Lambda E}(\Lambda) \longrightarrow \underline{E}(\Lambda)$;

$\varepsilon : \underline{E\Lambda}(E) \longrightarrow E$ is the unique $\underline{CAP}$-morphism by which $[\![\_]\!]' : \underline{\Lambda}(E) \longrightarrow E$ factorizes through $[\![\_]\!] : \underline{\Lambda}(E) \longrightarrow \underline{E\Lambda}(E)$.

η is injective; hence $\underline{E}$ is a faithful functor. ε is a split mono; $\underline{\Lambda}$ is thus a full functor.

In fact, ε is an equivalence of categories. It is not an isomorphism only because the internal language has made $\nabla\mathfrak{a}$ and $\nabla\mathfrak{r}$ cloven. On the other hand, each term isomorphic in Λ to a first projection has became in $\underline{\Lambda E}(\Lambda)$ a type. η is not an isomorphism because of this surplus[2]. - But already the injectivity of η means the completeness of the construct E(_).

At this point, our story about constructive logic may seem to be close to a happy ending: the marriage of TOP and CAP looks stable and it can be expected that it will lead to love some day. It may be so, but it is dubious if this will help predicates to surmount the difficulties of a constructive life. The last section will uncover some of these difficulties.

[2]It does not occur if the class of first projections is closed under isos already in Λ: e.g. in the presence of equality types.

3. First steps

1. Generating small complete categories.

In order to really do mathematics in type theory, one needs to represent equality in it. Externally, of course, equality has been there all the time - generated by the conversion relation. In the Martin-Löf type theory, this relation can be immediately internalized, by introducing for every pair of terms $p(X^P)$, $q(Y^Q)$:R an *equality type* I(p,q), so that

$\exists t$: I(p,q) :$\Leftrightarrow$ p=q.

But which conversion rules should be imposed on the terms of an equality type? Translating this question into: "Which proofs of the statement p=q are equivalent?" does not seem to help much.

Martin-Löf (1984, "Propositional equality") has stipulated that there is at most one term of a type I(p,q). Since it fails to reflect even the constructive contents of various derivations of p=q (cf. Troelstra-van Dalen 1988, 11.1.7.), such an equality type I(p,q) could better be thought of as the set $\{\langle X,Y\rangle \in P\times Q \mid p(X)=q(Y)\}$, than as a constructive predicate. Interpreted categorically, I(p,q) becomes the pullback of p and q - and the unique term of this type corresponds to the unique factorisation through this pullback. - Martin-Löf type theories with equality types correspond exactly to locally cartesian closed categories. (• Without equality types, they correspond to relatively cartesian closed categories. But if the display family $\mathfrak{a} \subseteq \mathfrak{B}$ of an rccc $\mathfrak{B}$ must contain the arrows $[\![I(X^R,Y^R)]\!] = pb(id_R,id_R) = \rho: R \longrightarrow R\times R$ for all $R\in|\mathfrak{B}|$, then lemma II.4.6 and fact II.4.34 imply $\mathfrak{a}=\mathfrak{B}$. •) Seely (1984) has given a detailed account of this correspondence.

The idea of the *theory of predicates with equality types* (in both sorts) seems rather appealing. A term model of such a theory is a small lccc with small products and coproducts - over an lccc. In particular, this small category is small complete, since it is fibrewise finitely complete, and has small products. An argument of Peter Freyd (MacLane 1971, proposition V.2.3) shows that such a category in the setting of classical sets must be a preorder (with small meets). With constructive sets from categories of predicates, this clearly need not be the case.

One concrete example of such a small complete category of predicates (indeed, of constructions) is the category of modest sets fibred over the category of separable objects of the effective topos (as described in Hyland 1988, Hyland-Robinson-Rossolini 1988, Longo-Moggi 1988 etc.). This example shows that adding equality types to the theory of predicates does not lead to paradoxes. Small complete categories obtained as term models of theories of predicates are thus not degenerate.

2. Equality predicates.

Notation. For an arbitrary predicate $\alpha \in |\mathcal{E}_K|$ (i.e. $\alpha(X^K)$), we write

$\models \alpha$ <u>when</u> $\exists f \in \mathcal{E}_K(\top,\alpha)$ (i.e. $\exists f{:}\alpha{:}\Omega$);

and say that α *is provable*. We shall also use

$\alpha \models \beta$ for $\models \alpha \to \beta$.

By $Æ(u,v)$ will be denoted the equality type which can be formed *only it* u and v have the same context (and not just the same type). The operations I and Æ are obviously derivable from each other (• by adding dummies, and by substituting along the diagonal •). Interpreted in a category, $I(p,q)$ is a pullback of $[\![p]\!]$ and $[\![q]\!]$, while $Æ(u,v)$ is an equaliser of $[\![u]\!]$ and $[\![v]\!]$.

Note that we sometimes combine categorical and type theoretical notation, forgetting $[\![\_]\!]$, and confuse

$\langle u,v\rangle^*\alpha$ and $\alpha(u,v)$, or

$\alpha(X,Y)$ and $\langle \pi_0,\pi_1\rangle^*\alpha$.

Definition. An *equality predicate* $\alpha(X^K,Y^K)$ must satisfy:

$\models \alpha(X^K,X^K)$ and

$\models \alpha(X^K,Y^K) \wedge \varphi(X^K,Z_0,\ldots,Z_n) \to \varphi(Y^K,Z_0,\ldots,Z_n)$, for every φ.

Generically, such a predicate $\alpha(X^K,Y^K)$ will be written $X^K \equiv Y^K$.

Facts. For every equality predicate $\equiv$ holds

$\models X \equiv Y \leftrightarrow Y \equiv X$,

$\models X \equiv Y \wedge Y \equiv Z \to X \equiv Z$,

$\models X \equiv Y \to \forall \varphi{:}K \to \Omega.\varphi(X) \leftrightarrow \varphi(Y)$.

Examples. It follows from the last fact that the *Leibniz equality*, defined for arbitrary terms $u(X^I)$, $v(X^I) : K$ (i.e. for arrows $u,v\colon I \longrightarrow K$)

$u \,\Xi\, v := \forall \varphi{:}K \to \Omega.\varphi(u) \leftrightarrow \varphi(v)$,

is weakly terminal among all the equality predicates over a set K.

If the equality type $Æ(X^K,Y^K)$ is given on K, then there is a weakly initial equality predicate too, namely the *Lawvere equality*:

$u \,\theta\, v := \exists Z{:}Æ(u,v).\top$.

In categorical notation, this is $u\theta v := æ_!\top$, where $æ{:}Æ \rightarrowtail I$ is an equaliser of u and v. Lawvere (1970, p.6) has shown that θ is an equality predicate. To show that it is weakly initial, consider an arbitrary equality predicate $\equiv$ over the same set. Since $uæ=væ$, the reflexivity of $\equiv$ implies that there is a proof of $uæ \equiv væ$, i.e. a vertical arrow $\top \longrightarrow æ^*\langle u,v\rangle^*(\equiv)$. Hence

$u\theta v = æ_!\top \longrightarrow \langle u,v\rangle^*(\equiv) = u \equiv v$.

In other words, writing

$æ(u,v) := \pi_0\colon \Sigma X{:}I.Æ(u(X),v(X)) \to I$,

from

$\models u(æ(Y)) \equiv v(æ(Y))$

we obtain

$\exists Z{:}Æ(u,v).\top \models u(X) \equiv v(X)$.

(The idea behind the Lawvere equality becomes perhaps clearer if we consider $X\theta Y$ written categorically $\pi_0\theta\pi_1 = \rho_!(\top)$ (where $\rho{:=}\langle id,id\rangle\colon K \longrightarrow K{\times}K$ is an equaliser of $\pi_0,\pi_1\colon K{\times}K \longrightarrow K$). If $w_!(\gamma)$ can be thought of as $\exists Z.w(Z){=}X \wedge \gamma$ (II.3.1), then $X\theta Y$ is just $\exists Z.\langle Z,Z\rangle{=}\langle X,Y\rangle \wedge \top$.)

In a topos - which is a category of constructions by I.1.6 - the Leibniz equality and the Lawvere equality coincide: cf. Lambek-Scott 1986, II.2. This means that every topos has a unique equality predicate, • since at most one proof from a predicate to another can exist there•.

Propositions. Let $E:\mathcal{E} \longrightarrow \mathcal{B}$ be a category of predicates, $K\in|\mathcal{B}|$ a set in it.

21. For every equality predicate $\equiv\in|\mathcal{E}_{K\times K}|$ with an extent $\iota(\equiv)=\langle e_0,e_1\rangle$, the following arrows exist:

- *reflexivity* $m\in\mathcal{B}(K, D(\equiv))$, such that $\langle e_0,e_1\rangle\circ m = \langle id,id\rangle$;
- *symmetry* $a\in\mathcal{B}(D(\equiv), D(\equiv))$, such that $\langle e_0,e_1\rangle\circ a = \langle e_1,e_0\rangle$;
- *transitivity* $c\in\mathcal{B}(P, D(\equiv))$, such that $\langle e_0,e_1\rangle\circ c = \langle e_0\circ p_0, e_1\circ p_1\rangle$, where $p_i\in\mathcal{B}(P, D(\equiv))$ is obtained by pulling back e_i along e_j for $i\neq j\in 2$.

• A reflexivity arrow m exists because the set

$$\mathcal{B}/K^2(\langle id,id\rangle, \iota(\equiv))\simeq\mathcal{E}_K(\top, \langle id,id\rangle^*(\equiv))$$

must be inhabited, since $\langle id,id\rangle^*(\equiv) = [\![X\equiv X]\!]$.

A symmetry arrow a corresponds by the isomorphism

$$\mathcal{B}/K^2(\langle e_1,e_0\rangle, \langle e_0,e_1\rangle) \simeq \mathcal{E}_{D(\equiv)}(\top, e_1\equiv e_0))$$

to a proof of $e_1\equiv e_0$ derived using $\models X\equiv Y\to Y\equiv X$ from the generic proof of $e_0\equiv e_1$ (i.e. the one which corresponds to $id\in\mathcal{B}/K^2(\iota(\equiv),\iota(\equiv))$).

A transitivity arrow is constructed in a similar way.•

22. For every equality predicate $\equiv\in|\mathcal{E}_{K\times K}|$, and arbitrary $u,v\in\mathcal{B}(I,K)$, the extent $\iota(u\equiv v)$ is a weak equaliser of u and v. Hence

$$u=v \Leftrightarrow \models u\equiv v.$$

• Given an arrow h such that $u\circ h=v\circ h$, consider $t := m\circ u\circ h=m\circ v\circ h$, where m is a reflexivity map. Then $\langle e_0,e_1\rangle\circ t = \langle u,v\rangle\circ h$, and h must factorize through $\iota(u\equiv v)= \langle u,v\rangle^*\langle e_0,e_1\rangle$.

$\Leftarrow$: A proof of $u\equiv v$ gives by the correspondence $\mathcal{E}_I(\top,u\equiv v)\simeq\mathcal{B}/I(id,\iota(u\equiv v))$ an arrow s, such that $\iota(u\equiv v)\circ s=id$. From $\models u\equiv v$ <u>and</u> $u\circ\iota(u\equiv v) = v\circ\iota(u\equiv v)$ thus <u>follows</u> u=v.•

23. Suppose that $\rho_K\in\mathcal{B}(K, K\times K)$ is a display arrow in $\mathcal{B}$, so that the Lawvere equality θ can be defined on K. For all functions $u,v\in\mathcal{B}(I,K)$, the extent $\iota(u\theta v)$ is an equaliser of u and v. Moreover,

$$u=v \Leftrightarrow u\theta v\simeq\top I.$$

• The $\Leftarrow$-direction of the last assertion follows from 22; and $\Rightarrow$ directly from the definitiion, • because the equaliser æ=id <u>if</u> u=v•.

Since $u\theta v = v\theta u$, the identity on $D\theta$ is a symmetry arrow for the Lawvere equality and $\iota\theta$ must be in the form $\langle e,e\rangle$. But now

$$e = \langle id,id\rangle^*\langle e,e\rangle = \iota(id\theta id) = \iota(\top K) = id_K.•$$

Comment. The Lawvere equality predicate inherits from the equality types their nonconstructive strength: unique proofs. Consequently, any other equality seems to be a better choice for constructive logic.

For every arrow $h\in\mathcal{B}(H,I)$ such that uh=vh, the factorisations of h through $\iota(u\equiv v)$ exactly correspond to proofs of $uh\equiv vh$. The constructive contents of predicate $u\equiv v$ are reflected precisely in the *weakness* of $\iota(u\equiv v)$ as equaliser. Remarkably, the nonunique factorisations through $\iota(u\equiv v)$ do not appear as a point of disorder (as it is usually the case with weak universal constructions): they are *positively structured by internal constructive logic*. Informally, an element of a weak equaliser of u and v can be thought of as a pair $\langle X, p(X)\rangle$, such that u(X)=v(X) is true, and p(X) proves this fact. Using the constructive extent operation, this set can be presented as $\sum X{:}I.\iota(uX\equiv vX)$. (Cf. III.2.4.)

3. Describing functions.

But will a constructive equality predicate - the Leibniz equality in particular - not be too weak to carry mathematics? For instance, is it strong enough to allow the usual description of functions, characterization of monics, epis and so on? How much of the power, with which "the internal language at work" governs a topos (cf. Lambek-Scott 1986, II.5-6.), subsists in our more general case of categories of predicates?

31. Notation. $\{X{:}K \mid \alpha(X)\} := D\alpha = \Sigma X{:}K.\iota\alpha$.

$\iota\alpha := \pi_0 : \{X{:}K \mid \alpha(X)\} \to K$.

$\alpha \wedge \beta := \alpha \times \beta$.

$\exists! Y{:}K.\alpha(Y) := \exists Y{:}K.\alpha(Y) \wedge \forall YY'{:}K.\alpha(Y) \wedge \alpha(Y') \to Y \equiv Y'$.

32. Definition. Let $E{:}\mathcal{E} \longrightarrow \mathcal{B}$ be a category of predicates, $\alpha \in |\mathcal{E}_{I\times K}|$, $\langle q_0,q_1\rangle := \iota\alpha$. We say that the predicate α is *functional* if the following two conditions are satisfied:

$\exists$) q_0 is a retraction <u>and</u>

!) $q_0h=q_0k$ <u>implies</u> $q_1h=q_1k$, for each pair of functions h,k.

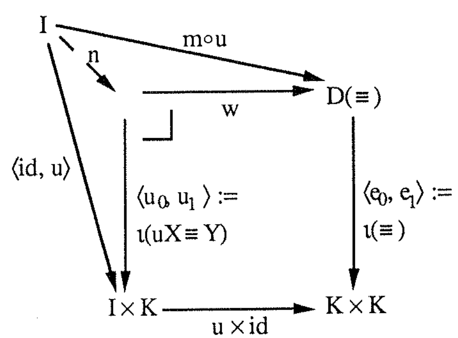
Propositions. Consider a category of predicates E, with equality predicates on sets K and I.

331. For every function $u \in \mathcal{B}(I,K)$, the predicate $uX \equiv Y$ is functional.

• Consider $\langle u_0,u_1\rangle := \iota(uX \equiv Y)$. The equation

1) $u \circ u_0 = u_1$

follows from proposition 22, since $u(u_0Z) \equiv u_1Z = \langle u_0,u_1\rangle^*(uX \equiv Y)$ is a provable proposition (as every $\iota\alpha^*(\alpha)$ is: its proof $\tau Z{:}\alpha(Z)$ is the generic one). Further define an arrow n by the following pullback



(where $m \in \mathcal{B}(K, D(\equiv))$ is a reflexivity map, from 21). We have

2) $u_0 \circ n = id$, and

3) $u_1 \circ n = u$.

The equation (2) just says that the predicate $uX \equiv Y$ satisfies condition ($\exists$). It satisfies condition (!) because

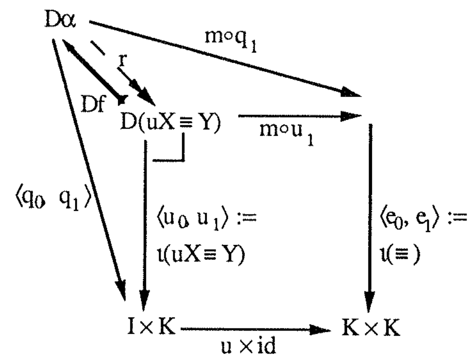
$u_1 \overset{(1)}{=} uu_0 \overset{(3)}{=} u_1nu_0$.•

332. If $\alpha \in |\mathcal{E}_{I\times K}|$ is a functional predicate, <u>then</u> there is a unique function $u \in \mathcal{B}(I,K)$ such that $\iota(uX \equiv Y)$ is a retract of $\iota\alpha$ (in $\mathcal{B}/I{\times}K$). In particular, every function u can be recovered from $\iota(uX \equiv Y)$.

• Given a functional predicate α, with $\langle q_0,q_1\rangle := \iota\alpha$ and a section p of q_0, take

$u := q_1p$.

Condition (!) implies that u does not depend on the choice of p; and that $uq_0=q_1$. The retraction $r : \iota\alpha \twoheadrightarrow \iota(uX \equiv Y)$ is obtained as a factorisation on the following diagram:



(because $(u \times id) \circ \langle q_0, q_1\rangle = \langle uq_0, q_1\rangle = \langle q_1, q_1\rangle = \langle e_0, e_1\rangle \circ m \circ q_1$). Df is the image of a proof $f : uX \equiv Y \to \alpha(X,Y)$. To construct this proof, use

$uX \equiv Y \models q_0pX \equiv X \wedge q_1pX \equiv Y$,

and the fact that

$\alpha(q_0pX, q_1pX) = \langle q_0p, q_1p\rangle^*(\alpha) \simeq p^*\iota\alpha^*(\alpha)$

is a provable predicate.

Suppose, finally, that there is another function $\tilde{u} \in \mathcal{B}(I,K)$ such that $\langle \tilde{u}_0,\tilde{u}_1\rangle := \iota(\tilde{u}X \equiv Y)$ is a retract of $\iota\alpha$: there is a retraction $\tilde{r} : \iota\alpha \twoheadrightarrow \iota(\tilde{u}X \equiv Y)$, with a splitting i (so that $\tilde{r}i=id$, and $\langle q_0,q_1\rangle \circ i = \langle \tilde{u}_0,\tilde{u}_1\rangle$). If $\tilde{n}$ is a section of $\tilde{u}_0$, then

$q_0 \circ i \circ \tilde{n} = \tilde{u}_0 \circ \tilde{n} = id_I = q_0 \circ p$ <u>implies</u> $q_1 \circ i \circ \tilde{n} = q_1 \circ p$.

But then

$$\tilde{u} = \tilde{u}_1 \circ \tilde{n} = q_1 \circ i \circ \tilde{n} = q_1 \circ p = u. \bullet$$

333. Consider a predicate $\alpha \in |\mathcal{E}_{I \times K}|$ in a category of predicates $\mathcal{E}$. There is a (unique) function $u \in \mathcal{B}(I,K)$, such that

$$\models uX \equiv Y \leftrightarrow \alpha(X,Y) \text{ (\underline{and} } \iota(uX \equiv Y) \text{ is a retract of } \iota\alpha)$$

<u>iff</u> α satisfies the following conditions:

i) $\models \forall X{:}I\ \exists! Y{:}K.\alpha(X,Y)$,

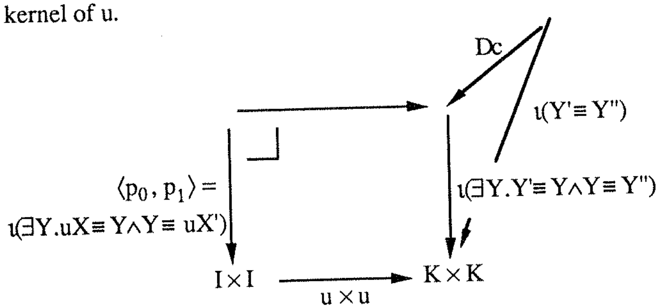
ii) there exists $s : \iota\big(\exists Y{:}K.\alpha(X,Y) \wedge \alpha(X',Y)\big) \longrightarrow \Sigma Y{:}K.\iota\big(\alpha(X,Y) \wedge \alpha(X',Y)\big)$.

• <u>Then</u>: We only show that condition (ii) is satisfied. We first prove it for $\alpha(X,Y) := uX \equiv Y$, $u \in \mathcal{B}(I,K)$.

Using a proof c: $Y' \equiv Y'' \to (\exists Y{:}K.Y' \equiv Y \wedge Y \equiv Y'')$, observe that

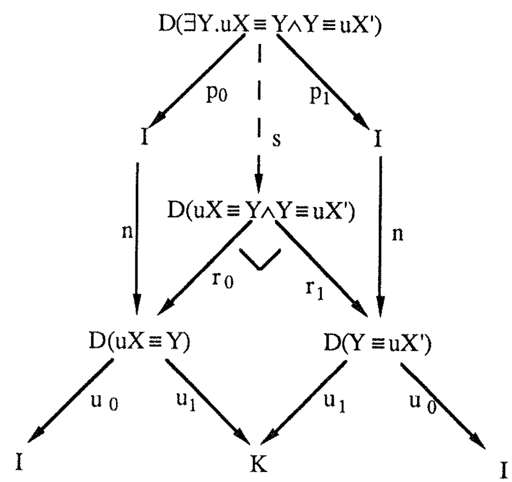
$$\langle p_0, p_1 \rangle := \iota\big(\exists Y{:}K.uX \equiv Y \wedge Y \equiv uX'\big)$$

is a weak kernel of u.



Namely, if $u \circ t_0 = u \circ t_1$, the arrow $\langle ut_0, ut_1 \rangle$ factorizes through $\iota(Y' \equiv Y'')$, by proposition 22. Hence, $\langle t_0, t_1 \rangle$ must factorize through $\langle p_0, p_1 \rangle$.

The required arrow s is now obtained as on the following diagram:



(because $u_1 \circ n \circ p_0 = u \circ p_0 = u \circ p_1 = u_1 \circ n \circ p_1$). Since

$$\Sigma Y{:}K.\iota\big(uX \equiv Y \wedge Y \equiv uX'\big) = \langle u_0 \circ r_0,\ u_0 \circ r_1 \rangle,$$

we have

$$\iota\big(\exists Y{:}K.uX \equiv Y \wedge Y \equiv uX'\big) = \langle p_0, p_1 \rangle = \langle u_0 n p_0, u_0 n p_1 \rangle = \langle u_0 r_0 s, u_0 r_1 s \rangle =$$
$$= \Sigma Y{:}K.\iota\big(uX \equiv Y \wedge Y \equiv uX'\big) \circ s.$$

For an arbitrary predicate α such that $\models uX \equiv Y \leftrightarrow \alpha(X,Y)$ holds for some function u, it is not hard to construct the arrows

$$\iota\big(\exists Y{:}K.\alpha(X,Y) \wedge \alpha(X',Y)\big) \longrightarrow \iota\big(\exists Y{:}K.uX \equiv Y \wedge Y \equiv uX'\big) \text{ and}$$
$$\Sigma Y{:}K.\iota\big(uX \equiv Y \wedge Y \equiv uX'\big) \longrightarrow \Sigma Y{:}K.\iota\big(\alpha(X,Y) \wedge \alpha(X',Y)\big).$$

Using them, the arrow s, required for α by condition (ii), is obtained from the one for $uX \equiv Y$.

<u>If</u>: From a proof of $\forall X{:}I\ \exists! Y{:}K.\alpha(X,Y)$, we can derive proofs

$$\eta : X \equiv X' \to \exists Y{:}K.\alpha(X,Y) \wedge \alpha(X',Y), \text{ and}$$
$$\varepsilon : \exists X{:}I.\alpha(X,Y) \wedge \alpha(X,Y') \to Y \equiv Y'.$$

(To derive η, define $\beta(X,X') := \exists Y{:}K.\alpha(X,Y) \wedge \alpha(X',Y)$, and use

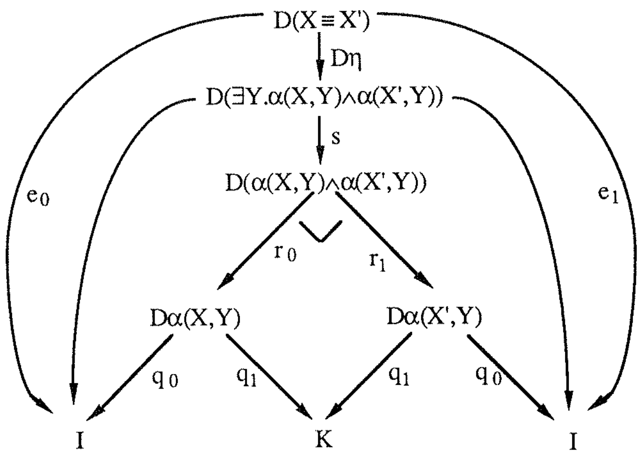
$$\exists Y{:}K.\alpha(X,Y) \models \beta(X,X), \text{ and}$$
$$X \equiv X' \wedge \beta(X,X) \models \beta(X,X').)$$

Consider the following arrows:

$D\eta : \iota(X\equiv X') \longrightarrow \iota(\exists Y{:}K.\alpha(X,Y)\wedge\alpha(X',Y))$,

$s : \iota(\exists Y{:}K.\alpha(X,Y)\wedge\alpha(X',Y)) \longrightarrow \Sigma Y{:}K.\iota(\alpha(X,Y)\wedge\alpha(X',Y))$,

$r_0 : \iota(\alpha(X,Y)\wedge\alpha(X',Y)) \longrightarrow \iota(\alpha(X,Y))$.



Note that the arrow $\Sigma Y{:}K.\iota(\alpha(X,Y)\wedge\alpha(X',Y)) : D \longrightarrow I\times I$ is obtained just by projecting away the middle component from $\iota(\alpha(X,Y)\wedge\alpha(X',Y)) : D \longrightarrow I\times K\times I$. Taking again $\langle q_0, q_1\rangle := \iota\alpha$, both $\Sigma Y{:}K.\iota(\alpha(X,Y)\wedge\alpha(X',Y))$ and $\iota(\alpha(X,Y)\wedge\alpha(X',Y))$ have the arrow $q_0\circ r_0$ as the first component. Hence

$q_0\circ r_0\circ s\circ D\eta = e_0$,

where $\iota(X\equiv X') = \langle e_0, e_1\rangle$. If m is a reflexivity arrow (i.e. a section of e_0), then

$p := r_0\circ s\circ D\eta\circ m$

is a section of q_0. Hence, α satisfies condition ($\exists$) for a functional predicate.

To prove that α satisfies condition (!), we use the fact that

$\vDash\alpha(u,v) \Leftrightarrow$ there is an arrow h, such that $\langle u, v\rangle = \langle q_0, q_1\rangle\circ h$.

For any pair of arrows $h,k : L \longrightarrow D(X\equiv X')$ holds

$\vDash \alpha(q_0\circ h, q_1\circ h)\wedge\alpha(q_0\circ k, q_1\circ k)$.

If $q_0\circ h = q_0\circ k$ <u>then</u>

$\vDash \exists X{:}I.\alpha(X, q_1\circ h)\wedge\alpha(X, q_1\circ k)$.

Using ε, we now derive

$\vDash q_1\circ h \equiv q_1\circ k$;

and therefore $q_1\circ h = q_1\circ k$ must be true, by proposition 22.

α is thus a functional predicate, and we define u as in proposition 332. A proof

$f : uX\equiv Y \to \alpha(X,Y)$

is constructed same as there. A converse proof

$g : \alpha(X,Y) \to uX\equiv Y$

is derived from ε, i.e. from a proof of $\alpha(X,uX)\wedge\alpha(X,Y)\to uX\equiv Y$. (Once again, note that $\alpha(X,uX) = \alpha(q_0pX, q_1pX)$.)•

34. u is a split monic $\overset{(a)}{\Rightarrow} \vDash\forall XY{:}I.uX\equiv uY \to X\equiv Y \overset{(b)}{\Rightarrow}$ u is monic.

• a) Since $uX\equiv uY \vDash e(uX)\equiv e(uY)$ always holds, $e\circ u=id$ <u>implies</u> $uX\equiv uY \vDash X\equiv Y$.

b) $uq=up \overset{(22)}{\Rightarrow} \vDash u(pX)\equiv u(qX) \Rightarrow \vDash pX\equiv qX \overset{(22)}{\Rightarrow} p=q$.•

351. $\vDash\forall X{:}I\ \exists Z{:}\{Y{:}K\mid \alpha(Y)\}.uX\equiv\iota\alpha Z \Leftrightarrow$ u factorizes through $\iota\alpha$.

• $\Rightarrow$: For clarity, we write out explicitly the dummies which must be added in u and $\iota\alpha$ when they are substituted in $\equiv$. Obviously, we have

$u(X,\not{Z})\equiv\iota\alpha(\not{X},Z) \vDash \ulcorner\alpha\urcorner\circ u(X,\not{Z})\equiv\ulcorner\alpha\urcorner\circ\iota\alpha(\not{X},Z)$.

But $\ulcorner\alpha\urcorner\circ\iota\alpha$ classifies (and is identical with) the provable proposition $\iota\alpha^*(\alpha)$. Using

$\ulcorner\beta\urcorner\equiv\ulcorner\gamma\urcorner \vDash \beta\leftrightarrow\gamma$,

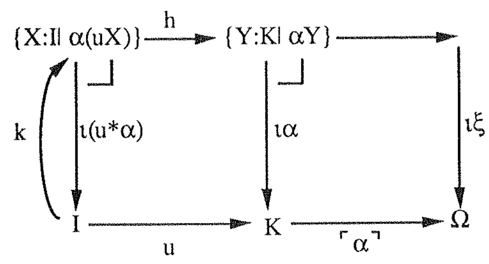
we derive

$u(X,\not{Z})\equiv\iota\alpha((\not{X},Z) \vDash \alpha(uX,\not{Z})$

(since $\ulcorner\alpha\urcorner\circ u(X,\not{Z})$ classifies $\alpha(uX,\not{Z})$). Hence

$\exists Z.u(X,\not{Z})\equiv\iota\alpha(\not{X},Z) \vDash \alpha(uX)$.

From the given hypothesis it thus follows that an arrow $k\in\mathcal{B}/I(id, \iota(\alpha(uX)))$ must exist.



Now

$g := hk \in \mathfrak{B}/K(u, \iota\alpha)$.•

352. Every function u factorizes through a display arrow

$im(u) := \iota(\exists X{:}I.uX \equiv Y)$.

• This follows from 351, using the fact that the proposition

$\forall X{:}I\ \exists Z{:}\{Y{:}K \mid \exists X{:}I.u(X) \equiv Y\}.uX \equiv im(u)Z$

is provable.•

36. u is a split epi $\overset{(a)}{\Rightarrow} \models \forall Y{:}K\exists X{:}I.u(X) \equiv Y \overset{(b)}{\Rightarrow}$ u is epi.

• a) Notice that $\models \forall Y{:}K\exists X{:}I.u(X) \equiv Y$ means exactly that im(u) is a split epi.
b) $ku=hu \overset{(22)}{\Rightarrow} \models k(uX) \equiv h(uX) \Rightarrow \models kY \equiv hY \overset{(22)}{\Rightarrow} k=h$.•

37. Remark. Why are propositions 34 and 36 so poor; why can I not prove that monics are monics and epics are epics? There is perhaps a deeper reason for this than my own incapability. Examples show that the base category of a category of predicates need not be balanced: an arrow can be epi and mono without being iso. (E.g. modest sets: Hyland 1988.) It can even be provably epi and mono (i.e. $\models \forall Y{:}K\exists ! X{:}I.uX \equiv Y$) and still lack a splitting. (Only im(u) must be iso then.) Perhaps something like condition (ii) from proposition 333 is needed to characterize epi, mono, iso functions. Or it might be that the theory shoud be improved at this point. Some additional requirements imposed on the class of extents (and expressed in the theory of predicates by some additional rules) could be useful. Perhaps there is some particular equality predicate which is better than others.

But let us not confuse the imperfections of the theory with the points at which only the *nonconstructive simplicity* is lost. For instance, the fact that each function has an image, but that this image is not a subset of its range may seem odd for a while, but it is certainly not a deficiency! According to the constructive conception of the extent operation, the image of a function $u{:}I \rightarrow K$ should consist of some points Y of the range K, *equipped with proofs* that $\exists X{:}I.uX \equiv Y$. If these proofs are constructive, there can be several of them for each Y. It is the constructiveness of proofs that spoils the inclusion of the image $\{Y{:}K \mid \exists X{:}I.uX \equiv Y\}$ into the range K.

4. Procreation of models.

Finally, we are in a position to show how the examples of categories of predicates come about in the "real world" - how some other models for the theory of predicates can be produced, besides term models. Starting from any category of predicates E and an internal category $\mathbb{I} \in \underline{cat}_{\mathfrak{B}}$ in it, we construct the category of presheaves over $\mathbb{I}$, which is a new category of predicates. In particular, each category of constructions gives in this way numerous categories of predicates. Since they are clearly not generated by terminal objects, these categories are (•by III.4.3•) not categories of constructions themselves. The well known mathematical models for the theory of constructions (modest sets: Hyland 1988, Hyland-Robinson-Rossolini 1988, Longo-Moggi 1988; algebraic toposes: Hyland-Pitts 1987; Girard-style domains: Coquand-Gunther-Winskel 1989) thus offer a source of relevant examples for the theory of predicates too.

It seems that not every category of predicates can arise in this way, i.e. over a category of constructions. • Namely, the extent fibration of each of those which do must be a category of constructions. As we remarked in 1.2, this is *almost* the case *except* that the induced coproducts may (it seems) remain weak.• But appropriate examples still wait to be found.

The theory of internal categories formulated in the internal language of a category of predicates - using some constructive equality - differs from the usual theory of internal categories (e.g. in Johnstone 1977) by the fact that the commutativity conditions are imposed by means of extents $\iota(u\equiv v)$, and the used *limits are weak*. However, in the spirit of comment 2, a *weak equalizer gives in this context not less than the strong one, but more*: $\iota(u\equiv v)$ not only equalizes u and v, but also issues some proofs that it does so. In this theory of internal categories, each performed construction carries a constructive proof of its own soundness. Quite involved already, the internal formulations of category theory become even more complicated. Our arguments in this part had to be severely truncated: completely written down, the constructed terms tend to be completely unreadable.[3]

Terminology. An *internal category* $\mathbb{I}$ in the category of sets $\mathfrak{B}$ underlying a category of predicates E will now be described in the internal language by the following types and terms:

- set of objects I_0,
- hom-sets $X,Y{:}I_0 \Rightarrow I_1(X,Y)$,
- "identity arrows" function $X{:}I_0 \Rightarrow \eta(X){:}I_1(X,X)$,
- "composition" function $X,Y,Z{:}I_0,\ g{:}I_1(X,Y),\ f{:}I_1(Y,Z) \Rightarrow \mu(f,g){:}I_1(X,Z)$.

The following equations must be satisfied:

$$\mu(\eta(Z), f) = f$$
$$\mu(f, \eta(Y)) = f$$
$$\mu(\mu(f,g), h) = \mu(f, \mu(g,h)).$$

We say that *all the arrows in* $\mathbb{I}$ *are retractions (split epis)* if there is a

- "splitting" function $f{:}I_1(X,Y) \Rightarrow \tau(f){:}I_1(Y,X)$,

such that

$$\mu(f,\tau(f)) = \eta(Y).$$

[3] Of course, I did perform the proofchecking here omitted. But the fact is that such things should be done by a computer.

Remarks. A category $\mathbb{I}\in \underline{cat}_{\mathfrak{B}}$, described in this way is distinguished only by the fact that its domain and codomain arrows constitute a display map: $\langle\partial_0,\partial_1\rangle\in\mathfrak{a}$. This assumption is, however, not necessary for the propositions below. All our constructions could be performed with arbitrary $\mathbb{I}\in \underline{cat}_{\mathfrak{B}}$ - but the descriptions would then resemble a bit less to what one does in ordinary category theory. Furthermore, if $\mathfrak{B}$ is finitely complete, Lawvere's equality can be used for $\equiv$, and then everything really boils down to the usual internal category theory.

As for the equations imposed on an internal category, it would perhaps look more constructive if we demanded an explicite proof for $\forall f.\mu(\eta(Z), f)\equiv id$, etc. to be given. But lemma 25 gives a canonical proof of this proposition whenever the equation $\mu(\eta(Z),f) = f$ is true.

Propositions. Let E be a category of predicates.

41. For arbitrary internal categories $\mathbb{I}$ and $\mathbb{D}$, there is an internal category $[\mathbb{I}, \mathbb{D}]$ (the "functor category") such that $[\mathbb{I}\times\mathbb{L}, \mathbb{D}] \simeq [\mathbb{L}, [\mathbb{I}, \mathbb{D}]]$ holds vor every internal category $\mathbb{L}$.

- $[I, D]_0 := \{F{:}\ \Sigma Z{:}I_0{\to}D_0\ \Sigma XY{:}I_0.\ I_1(X,Y){\to}D_1(ZX,ZY) \mid \underline{functor}(\pi_0F,\pi_1F)\}$
  $\underline{functor}(F_0,F_1) := \forall X{:}I_0.\ F_1(X,X)\circ\eta(X)\equiv\eta(F_0(X))$
  $\wedge\ \forall XYZ{:}I_0\forall g{:}I_1(X,Y)\forall f{:}I_1(Y,Z).\ F_1(X,Z)\mu(f,g)\equiv\mu(F_1(X,Y)f,\ F_1(Y,Z)g)$

  $[I, D]_1(F,G) := \{\psi{:}\Pi X{:}I_0.D_1(F_0X,G_0X) \mid \underline{natural}(\psi,F,G)\}$
  $\underline{natural}(\psi,F,G) := \underline{functor}(\pi_0F,\pi_1F)\wedge\underline{functor}(\pi_0G,\pi_1G)\ \wedge$
  $\wedge\ \forall XY{:}I_0\forall f{:}I_1(X,Y).\ \mu(\psi Y,F_1(X,Y)f)\equiv\mu(G_1(X,Y)f,\psi X)$

  $F{:}[I,D]_0 \Rightarrow \eta(F) : [I, D]_1(F,F)$ is defined:
  $\eta(F) := \langle\lambda X.\eta(F_0X),\ \delta a\rangle$

  $\varphi{:}[I,D]_1(G,H),\ \psi{:}[I,D]_1(F,G) \Rightarrow \mu(\varphi,\psi) : [I, D]_1(F,H)$ is:
  $\mu(\varphi,\psi) := \langle\lambda X.\mu(\varphi X,\psi X),\ \delta b\rangle$

(The task of deriving proofs
  $a : \underline{natural}(\lambda X.\eta(F_0X), F, F)$ and

$$b : \underline{natural}(\lambda X.\mu(\varphi X,\psi X), F, H)$$

from the proofs that $\mathbb{D}$ is a category is still an easy exercise.)

Let us now sketch the defionition of the component

$$j: [I \times L, D]_0 \longrightarrow [L, [I, D]]_0$$

of the isomorphism $[\mathbb{I} \times \mathbb{L}, \mathbb{D}] \simeq [\mathbb{L}, [\mathbb{I}, \mathbb{D}]]$. Given a functor $F:[I \times L, D]_0$, from its object part

$$F_0:=\pi_0 F : I_0 \times L_0 \to D_0$$

we derive

$$(jF)_0 := \lambda AX.F_0\langle X,A\rangle : L_0 \to (I_0 \to D_0).$$

The arrow part $F_1:=\pi_1 F$, i.e.

$$X,Y:I_0,\ A,B:L_0 \Rightarrow$$
$$F_1\big(\langle X,A\rangle,\langle Y,B\rangle\big): I_1(X,Y)\times L_1(A,B) \to D_1\big(F_0\langle X,A\rangle,F_0\langle Y,B\rangle\big),$$

gives

$$A,B:L_0 \Rightarrow (jF)_{10}: L_1(A,B) \to \prod X:I_0.D_1\big(F_0\langle X,A\rangle,F_0\langle X,B\rangle\big) \text{ as}$$
$$(jF)_{10} := \lambda hX.\Big(F_1\big(\langle X,A\rangle,\langle X,B\rangle\big)\langle \eta(X),h\rangle\Big).$$

For every given $h:L_1(A,B)$, the term $(jF)_{10}h$ is a natural transformation between the functors $\lambda X.F\langle X,A\rangle$ and $\lambda X.F\langle X,B\rangle$. Namely, for arbitrary $f:I_1(X,Y)$, a proof of

$$\mu\Big(F_1\big(\langle Y,A\rangle,\langle Y,B\rangle\big)\langle \eta(Y),h\rangle,\ F_1\big(\langle X,A\rangle,\langle Y,A\rangle\big)\langle f,\eta(A)\rangle\Big) \overset{\#}{=}$$
$$= F_1\big(\langle X,A\rangle,\langle Y,B\rangle\big)\langle f,h\rangle \overset{\#}{=}$$
$$= \mu\Big(F_1\big(\langle X,B\rangle,\langle Y,B\rangle\big)\langle f,\eta(B)\rangle,\ F_1\big(\langle X,A\rangle,\langle X,B\rangle\big)\langle \eta(X),h\rangle\Big)$$

is obtained from proof that F is a functor. (The step (#) is just composition with identities, using associativity.) Encoding this proof, we get

$$h:L_1(A,B) \Rightarrow (jF)_{11}h : \iota\underline{natural}((jF)_{10}h)$$

and define:

$$A,B:L_0 \Rightarrow (jF)_1: L_1(A,B) \to [I,D]_1((jF)_0A,\ (jF)_0B) \text{ as}$$
$$(jF)_1 := \langle (jF)_{10},\ (jF)_{11}\rangle.$$

Having encoded a proof that $(jF)_1$ is a functor, we use the corresponding element $(jF)_p:\iota\underline{functor}((jF)_1)$ to define

$$jF := \langle jF_0,\ jF_1,\ jF_p\rangle.\bullet$$

42. If $\nabla\mathbb{D} : \mathfrak{B}/\mathbb{D} \longrightarrow \mathfrak{B}$ is an $\mathfrak{a}$-hyperfibration, then $\nabla[\mathbb{I},\mathbb{D}]$ is an $\mathfrak{a}$-hyperfibration too.

• The idea for a proof comes from the fact that the fibration

$$\nabla\underline{psh}_{\mathcal{E}}(\mathbb{I}) : \mathcal{P} \longrightarrow \mathfrak{B}, \text{ with fibres}$$
$$\mathcal{P}_K := \underline{psh}_{\mathcal{E}}(\mathbb{I}\times K)$$

is an $\mathfrak{a}$-hyperfibration for every $\mathbb{I}$, whenever $E:\mathcal{E} \longrightarrow \mathfrak{B}$ is. Namely, for arbitrary $u \in \mathfrak{B}(K,J) \cap \mathfrak{a}$, and $\square \in \{!,*\}$, the direct images are:

$$u_\square : \underline{psh}(\mathbb{I}\times K) \longrightarrow \underline{psh}(\mathbb{I}\times J) : \langle F, \gamma\rangle \longmapsto \langle (I_0\times u)_\square F,\ (I_1\times u)_\square \gamma\rangle.$$

The fact that $u_\square$ preserves the presheaves and their morphisms, and that it gives the (co)products in $\mathcal{P}$ follows readily from the BC-property of E.

The propositions III.1.27-8 and proposition 1 above, imply that every fibre $(\mathfrak{B}/[\mathbb{I},\mathbb{D}])_K$ is isomorphic with a category consisting of

- objects $\langle C, \gamma, p\rangle$, where $\langle C,\gamma\rangle \in \underline{psh}_{\mathbb{D}}(\mathbb{I}^o\times K)$, and $p:\underline{functor}\langle C,\gamma\rangle$;
- arrows $\langle \psi, q\rangle : \langle C, \gamma, p\rangle \longrightarrow \langle C', \gamma', p'\rangle$, where ψ is a presheaf morphism $\langle C, \gamma\rangle \longrightarrow \langle C', \gamma'\rangle$ (cf.III.1.24), and $q:\underline{natural}(\psi)$.

Thus, to construct the direct images of $\nabla[\mathbb{I},\mathbb{D}]$, one needs to encode proofs

$$\tilde{u}_\square : \underline{functor}\langle C,\gamma\rangle \to \underline{functor}\big\langle (I_0\times u)_\square C,\ (I_1\times u)_\square \gamma\big\rangle \text{ and}$$
$$\hat{u}_\square : \underline{natural}(\psi) \to \underline{natural}\big((I_0\times u)_\square \psi\big),$$

for every $u \in \mathfrak{B}(K,J) \cap \mathfrak{a}$, $\square \in \{!,*\}$, and to append them in the above construction of the direct images for $\mathcal{P}$.•

43. If $\nabla\mathbb{D}$ is a fibrewise cartesian category, then $\nabla[\mathbb{I},\mathbb{D}]$ is. If each map of $\mathbb{I}$ is a retraction, and if $\nabla\mathbb{D}$ has exponents, then $\nabla[\mathbb{I},\mathbb{D}]$ has them. - If each map of $\mathbb{I}$ is a retraction, the fccc structure (II.2.1) on $\nabla\mathbb{D}$ induces the fccc structure on $\nabla[\mathbb{I},\mathbb{D}]$.

• A product of two functors is constructed pointwise:

$$\langle C, \gamma, p\rangle \times \langle C', \gamma', p'\rangle := \langle C\times C',\ \gamma\times\gamma',\ p''\rangle,$$

just like in ordinary category theory. An exponent of functors will be pointwise too, if every arrow of the source category is a retraction[4]. By an argument as in III.1.24, the splitting τ of the arrows in $\mathbb{I}$ is lifted in every presheaf $\mathbb{C}$ over $\mathbb{I}^o$: there is a splitting

$$g:C_1(X,Y) \Rightarrow \upsilon(g):C_1(Y,X),$$

[4]What is really needed here is that every arrow in the image of the functor splits. One way to ensure this is to demand that every arrow in the source category splits.

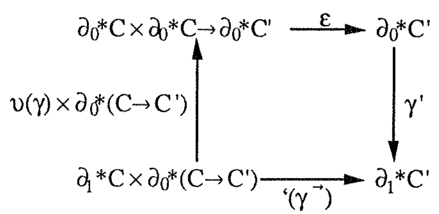
with

$\mu(g,\upsilon(g)) = \eta(Y)$.

Passing on the representation of presheaves as pairs $\langle C, \gamma\rangle$ (described in III.1.24), we conclude that for every presheaf over $\mathbb{I}$, its component γ must split: there is $\upsilon(\gamma)$ such that $\gamma\circ\upsilon(\gamma)=\mathrm{id}$. Hence the definition of an exponent:

$\langle C, \gamma, p\rangle \to \langle C', \gamma', p'\rangle := \langle C \to C', \gamma^{\to}, p^{\to}\rangle$,

with



$\gamma^{\to} := \left(\gamma' \circ \varepsilon \circ \left(\upsilon(\gamma) \times \partial_0{}^*(C \to C')\right)\right)^{\text{‘}} : \partial_0{}^*(C \to C') \longrightarrow \partial_1{}^*(C \to C')$,

$p^{\to}$: <u>functor</u>$\langle C \to C', \gamma^{\to}\rangle$.•

44. If each map of $\mathbb{I}$ is a retraction, <u>and</u> if $\nabla\mathbb{D}$ is a category of predicates, <u>then</u> $\nabla[\mathbb{I},\mathbb{D}]$ is a category of predicates.

• We already saw in 42 how the horizontal structure on $\nabla\mathbb{D}$ induces one on $\nabla[\mathbb{I},\mathbb{D}]$. Now we extend 43 and consider how the rccc-structure passes from $\nabla\mathbb{D}$ on $\nabla[\mathbb{I},\mathbb{D}]$.

A vertical display family $\mathfrak{r} \subseteq \mathfrak{B}/\mathbb{D}$ induces another such family

$\mathfrak{n} := \left\{\langle\psi,q\rangle \in \mathfrak{B}/[\mathbb{I},\mathbb{D}] \mid \psi \in \mathfrak{r}\right\}$

(with all the possible proofs q). $\mathfrak{B}/[\mathbb{I},\mathbb{D}]$ is an $\mathfrak{n}$-rccc <u>if</u> $\mathfrak{B}/\mathbb{D}$ is an $\mathfrak{r}$-rccc. The fibrewise exponents are constructed using the same idea as in 43. This time we must find an exponent of $\mathfrak{n}$-arrows

$\langle\psi, q\rangle : \langle A, \gamma_A, p_A\rangle \longrightarrow \langle X, \gamma_X, p_X\rangle$ and

$\langle\psi', q'\rangle : \langle A', \gamma_{A'}, p_{A'}\rangle \longrightarrow \langle X, \gamma_X, p_X\rangle$.

The definition will be

$\langle\psi,q\rangle \to \langle\psi',q'\rangle := \langle\psi \to \psi', q''\rangle : \langle A'', \gamma_{A''}, p_{A''}\rangle \longrightarrow \langle X, \gamma_X, p_X\rangle$.

$\psi \to \psi'$ is the exponent of $\mathfrak{r}$-arrows, and we denote its domain by A''. To obtain the arrow

$\gamma_{A''} : \partial_0{}^*(\psi \to \psi') \longrightarrow \partial_1{}^*(\psi \to \psi')$,

put ψ in place of C everywhere in diagram 43. $\langle\gamma_{A'}, \gamma_X\rangle : \partial_0{}^*\psi' \longrightarrow \partial_1{}^*\psi'$ will now replace γ', and $\langle\upsilon(\gamma_A), \upsilon(\gamma_X)\rangle : \partial_1{}^*\psi \longrightarrow \partial_0{}^*\psi$ will serve as $\upsilon(\gamma)$.

(Finding proofs

$p_{A''}$: <u>functor</u> $\langle A'', \gamma_{A''}\rangle$ and

q'' : <u>natural</u> $(\psi \to \psi')$

is a considerable exercise in encoding ordinary category theory in the theory of predicates.)•

References

Artin, M., Grothendieck, A., Verdier, J.L.

(1972) *Théorie des Topos et Cohomologie Etale des Schémas (SGA4)*, Lecture Notes in Mathematics 269 (Springer, Berlin)

Barendregt, H.P.

(1981) *The Lambda Calculus, Its Syntax and Semantic*, (North-Holland, Amsterdam)

(1989) Introduction to generalized type systems, *Proceedings of 3rd Italian Conference on Theoretical Computer Science* (World Scientific, Singapore) 1-37

Betti, R.

(1989) Change of base for locally internal categories, *J. Pure and Appl. Algebra* 61, 233-242

Betti, R., Walters, R.F.C

(1987) On completeness of locally-internal categories, *J. Pure Appl. Algebra* 47, 105-117

(1989) The calculus of ends over a base topos, *J. Pure Appl. Algebra* 56, 211-220

Bénabou, J.

(1975a) Théories relatives à un corpus, *C.R. Acad. Sc. Paris* t. 281, 831-834

(1975b) Fibrations petites et localement petites, *C.R. Acad. Sc. Paris* t. 281, 897-900

(1983) Fibered categories, Draft of lectures at Louvain-la-Neuve

(1985) Fibered categories and the foundations of naive category theory, *J. Symbolic Logic* 50(1), 10-37

Bénabou, J., Roubaud, J.
(1970) Monades et descente, *C.R. Acad. Sc. Paris* t. 270, 96-98

Cartmell, J.
(1986) Generalized algebraic theories and contextual categories, *Ann. Pure Appl. Logic* 32, 209-243

Cassidy, C., Hébert, M., Kelly, G.M.
(1985) Reflective subcategories, localisations and factorisation systems, *J. Austral. Math. Soc. (Series A)* 38, 287-329

Coquand, T., Gunther, C., Winskel, G.
(1989) Domain theoretic models of polymorphism, Manuscrtipt

Coquand, Th., Huet, G.
(1986) Constructions: A higher order proof system of mechanizing mathematics, *EUROCAL 85, Linz* , Lecture notes in Computer Science 203 (Springer, Berlin)
(1988) The Calculus of constructions, *Information and Computation* 76, 95-120

Dalen, D. van see *Troelstra, A.S., Dalen, D. van*

Day, B.
(1974) Limit spaces and closed span categories, in: *Category seminar, Sydney 1972/73*, Lecture Notes in Mathematics 420 (Springer, Berlin), 65-74

Ehrhard, T.
(1988) Une sémantique catégorique des types dépendants, Thesis, Université Paris VII
(1989) Dictoses, *Category theory in computer science,* Lecture Notes in Computer Science 389 (Springer, Berlin), 213-223

Freyd, P.
(1972) Aspects of topoi, *Bull. Austral. Math.* Soc. 7, 1-76

Freyd, P., Kelly, G.M.
(1972) Categories of continuous functors, I, *J. Pure Appl. Algebra* 2, 169-191

Giraud, J.
(1971) *Cohomologie non abelienne*, Die Grundlehren der mathematischen Wissenschaften, Bd. 179 (Springer, Berlin)

Gray, J.W.
(1966) Fibred and Cofibred Categories, *Proceedings of the Conference on Categorical Algebra, La Yolla 1965*, (Springer, Berlin), 21-84
(1974) *Formal Category Theory: Adjointness for 2-Categories*, Lecture Notes in Mathematics 391 (Springer, Berlin)
(1980) The existence and construction of lax limits, *Cah. de Topol. et de Géom. Differentielle*, XXI-3, 277-304

Grothendieck, A.
(1959) Technique de descente et théorémes d'existence en géométrie algébrique, I. Génétalités. Descente par morphismes fidélement plats, *Séminaire Bourbaki* 190
(1971) Catégories fibrées et descente, Exposé VI in: *Revêtements Etales et Groupe Fondamental (SGA 1)*, Lecture Notes in Mathematics 224 (Springer, Berlin) 145-194
see also *Artin, M., Grothendieck, A., Verdier, J.L.*

Gunther, C. see *Coquand, T., Gunther, C., Winskel, G.*

Hébert, M. see *Cassidy, C., Hébert, M., Kelly, G.M.*

Huet, G. see *Coquand, Th., Huet, G.*

Hyland, J.M.E.,
(1982) The effective topos, *The L.E.J. Brouwer Centenary Symposium.*, (North-Holland, Amsterdam) 165-216
(1988) A small complete category, *Ann. Pure Appl. Logic* 40, 135-165

Hyland, J.M.E., Johnstone, P.T., Pitts, A.M.
(1980) Tripos theory, *Math. Proc. Camb. Phil. Soc.* 88, 205-232

Hyland, J.M.E., Moerdijk, I.
(1990) An application of Beck's Theorem, Manuscript

Hyland, J.M.E., Pitts, A.M.
(1987) The theory of constructions: categorical semantics and topos-theoretic models, *Categories in Computer Science and Logic (Proc. Boulder 1987)*, Contemporary Math. (Amer. Math. Soc., Providence RI)

Hyland, J.M.E., Robinson, E.P., Rossolini, G.
(1990) The discrete objects in the effective topos, *Proc. London Math. Soc.*, 60 (1), 1-36

Johnstone, P.T.
(1977) *Topos theory*, L.M.S Monographs 10 (Academic Press, New York)
see also *Hyland, J.M.E., Johnstone, P.T., Pitts, A.M.*

Johnstone, P.T., Moerdijk, I.
(1989) Local maps of toposes, *Proc. London. Math. Soc.* (3)58, 281-305

Kelly, G.M.
(1982) *Basic Concepts of Enriched Category Theory*, London Mathematical Society Lecture Note Series 64 (Cambridge University Press, Cambridge)
see also *Cassidy, C., Hébert, M., Kelly, G.M.* and *Freyd, P., Kelly, G.M.*

Kelly, G.M., Street, R.
(1974) Review of the elements of 2-categories, *Category Seminar, Sydney 1972/3*, Lecture Notes in Mathematics 420 (Springer, Berlin) 75-103

Lambek, J.
(1980) From types to sets, *Advances in mathematics* 36, 113-164

Lambek, J., Scott, P.J.
(1986) *Introduction to higher order categorical logic* , Cambridge studies in advanced mathematics 7 (Cambridge University Press, Cambridge)

Lawvere, F.W.
(1969) Adjointness in Foundations, *Dialectica* 23, No 3/4, 281-296
(1970) Equality in hyperdoctrines and comprehension schema as an adjoint functor, *Applications of Category Theory*, Proceedings of A.M.S. Symposia on Pure Mathematics XVII (Amer. Math. Soc., Providence RI)
(1973) Continuously variable sets; algebraic geometry = geometric logic, in: *Logic Colloquium '73*, Studies in Logic and Foundations of Mathematics 80 (North Holland, Amsterdam)
(1976) Variable quantities and variable structures in topoi, in: *Algebra, Topology and Category theory* (Academic Press, New York)

Longo, G., Moggi, E.
(1988) Constructive natural deduction and its "modest" interpretation, Manuscript

MacLane, S.
(1971) *Categories for the Working Mathematician*, (Springer, Berlin)

Martin-Löf, P.
(1984) *Intuitionistic type theory*, Studies in Proof Theory (Bibliopolis, Napoli)

Moerdijk, I. see *Hyland, J.M.E., Moerdijk, I.* and *Johnstone, P.T., Moerdijk, I.*

Moggi, E. see *Longo, G., Moggi, E.*

Pavlović, D.
(1989) On structure of paradoxes, Preprint No. 570, Dept. of Mathematics, University Utrecht; submitted

Penon,J.
(1974) Catégories localement internes, *C. R. Acad. Sc. Paris*, t.278, 1577-1580

Pitts, A.M.
(1987) Polymorphism is Set Theoretic, Constructively, *Category Theory and Computer Science*, Lecture Notes in Computer Science 283 (Springer, Berlin), 12-39
see also *Hyland, J.M.E., Johnstone, P.T., Pitts, A.M.* and *Hyland, J.M.E., Pitts, A.M.*

Pitts, A.M., Taylor, P.
(1989) A note on Russell's paradox in locally cartesian closed categories, *Studia Logica* XLVIII(3), 377-387

Robinson, E.P. see *Hyland, J.M.E., Robinson, E.P., Rossolini, G.*

Rossolini, G. see *Hyland, J.M.E., Robinson, E.P., Rossolini, G.*

Roubaud, J. see *Bénabou, J., Roubaud, J.*

Scott, P.J. see *Lambek, J., Scott, P.J.*

Seely, R.A.G.
(1984) Locally cartesian categories and type theory, *Math. Proc. Camb. Phil. Soc.* 95, 33-48
(1987) Categorical semantics for higher order polymorphic lambda calculus, *J. Symbolic Logic* 52(4), 969-989

Spanier, E.H.
(1966) *Algebraic topology*, (McGraw-Hill, New York)

Street, R.
(1984) The family approach to total cocompleteness and toposes, *Trans. Amer. Math. Soc.* 284(1), 355-369

Street, R. see also *Kelly, G.M., Street, R.*

Streicher, Th.
(1988) Correctness and completeness of a categorical semantics of the calculus of constructions, Thesis, Univ. of Passau

(1989) Independence results for the calculi of dependent types, *Category theory in computer science*, Lecture Notes in Computer Science 389 (Springer, Berlin), 141-154

Taylor, P.
(1986) Recursive Domains, Indexed Category theory and Polymorphism, Thesis
see also *Pitts, A.M., Taylor, P.*

Troelstra, A.S., Dalen, D. van
(1988) *Constructivism in Mathematics. An Introduction*, Studies in Logic and Foundations of Mathematics 121, 123 (North-Holland, Amsterdam)

Verdier, J.L. see *Artin, M., Grothendieck, A., Verdier, J.L.*

Walters, R.F.C see *Betti, R., Walters, R.F.C*

Winskel, G. see *Coquand, T., Gunther, C., Winskel, G.*

Index

atoms 14
axioms 14

basic analogy 183
BC-property 88
Beck-Chevalley property 84
  vertical (cartesian) – 94

calculus
  – of constructions 29
  – of predicates 26
calibration 119
CAP 208
cartesian
  – adjointness 49, 77
  – arrow, functor, natural transformation 49
  co- 80
  op- 81, 103
  weakly co- 173
category
  contextual – 188, 190
  fibred – 52
  fibrewise cartesian closed – 71
  internal – 137, 224
  locally cartesian closed – 128
  locally internal – 149
  – of constructions 188, 189
  – of predicates 27, 188
  regular – 126
  relatively cartesian closed – 128, 197
    fibrewise – 132
  small complete – 211
cleavage 52
  co-, op- 82
COC 29
comprehension 153
  – functor 165, 178
  Lawvere – 169
  – principle 11, 153
  – transformation 164
constructions
  category of – 188
  doctrine of – 189
  theory, calculus of – 29
constructive proof 1
context 13
  bound – 192
convertibility 14
COP 26
crible 146

$\Delta \to X$ 157

definability 147, 154
dependent type 18
derivability 12
descent data 138, 139
description 153, 156
  comprehensible – 157
  homogenous – 161
  – with colimits 162
$\underline{\text{des}}_{\mathcal{E}}$ 139
diagonal 21
dictos 5, 189
display arrows, family 120, 184, 190, 201
doctrinal diagram 169
DV(T) 14

equality 80
  Lawvere's – 213
  Leibniz' – 213
  – predicate 212
  – type 211
expressions 12
extent 11, 26, 155, 158, 197, 223
  – fibration 160
externalisation 141, 144

factorisation system (relative) 121
Fam(E) 93
fccc 71
fibration 52
  arrow – 120, 184
  basic – 120
  cloven – 52
  co-, bi-, tri-, hyper- 80, 93
    fibrewise – 98
  comprehensive – 157
  discrete – 58
  extent – 160, 171, 189
  family – 93
  fibrewise – 75
  – generated by terminal objects 178
  globally small 187
  globally small – 147, 187
  Hurewicz – 60
  locally small – 147
  relative – 118
  representable – 135, 147
  small – 141
  split – 52, 187
  sub- 53
  – with (co)products 92
fibre 48
fibrewise
  – co-, bi-, tri-, hyperfibrations 98
  – equivalence 59
  – fibration, cartesian functor 75
  – structure/property 67
function 22

generator
  – in a fibration 178
  – in type theory 12, 14
generic
  – arrow 148

  – object 148

Grothendieck construction 54
groupoid 59, 150
  fundamental 60

HOM_{E} 145
homotopy 60
hyperdoctrine 79, 169, 183

image
  (left or right) direct – 81
  inverse – 49
  – of a function 222
interpolant 105
  initial, strong – 106
interpolation 105
interpolation condition 174
  strong – 107, 178
  uniform – 109
intrinsic structure 120, 186

lccc 128

MIN(T) 16
model assignment 13, 183, 184

OB_{E} 145, 149
opcartesian arrows 82

packing 192
polymorphism 23, 186
potential structure 53
predicate
  equality – 212
  functional – 216
predicates 23
  (strong) theory, calculus of – 26
  category of – 188
presheaf 57, 138, 139, 227
projection 21, 186
proof 22
proposition (truth value) 22
$\underline{\text{psh}}_{\mathcal{E}}$ 139, 227

quantifier 24, 80

rccc 128
right bifibration 82
rules
  conversion – 14, 20
  equality – 20
  formation – 12, 18
  structural – 16
  typing – 13, 19
  weakening and contraction – 21

saturated family 119
saturated subcategory 119
set 22

stable family 119
statements 12, 13
STOP 26
substitution 20, 79
system 33
  conservative over – 35
  equivalent – 35
  stronger – 41
  sub- 33

theory 13, 33
  – of constructions 29
  – of predicates 26
  retract of – 35
  sub- 33
TOC 29
TOP, TOP 26, 208
topos 5, 27, 189, 213
translation 35, 41
truth value (proposition) 22
type algebra 13

unit type 22

v(c)BC-property 94
variables 14, 18
  declared vs. free – 15
  dummy – 15, 184, 197
  – must not interfere 84, 105
variation 17, 184

vertical arrow 49

weak
  – Beck-Chevalley condition 173
  – cofibration, coproduct 173
  – equaliser 214, 215
  – equalizer 224

Yoneda embedding 135