Maps I: relative to a factorisation system

Duško Pavlović[1]

Department of Mathematics and Statistics, McGill University, Montreal, Que., Canada H3A 2K6

Abstract

Originally, the categorical calculus of relations was developed using the canonical factorisation in regular categories. More recently, relations restricted to a proper factorisation system have been studied by several authors. In the present paper, we consider the general situation, in which relations are induced by an arbitrary stable factorisation. This extension of the calculus of relations is necessary for a categorical development of strongly constructive (and computational) logic, where non-monic relations come about naturally.

In this setting, we analyse the correspondence of the *maps*, i.e. the total, single-valued relations, and the *functions*, as given by the arrows in the underlying category.

1. Introduction

In set theory, the functions are defined as the total and single-valued binary relations. In a category, the concept of function is primitive, given by the arrows, while relations are usually presented as subobjects. The two notions are thus defined separately, and it makes sense to ask whether functions correspond to the total, single-valued relations or not. A similar question arises, for instance, in formal arithmetic, where functions are defined by recursion, while relations are given as predicates. The statement that the total (and single-valued) relations exactly correspond to the recursive functions is a (restricted form) of Church's thesis [25, 4.3(3)].

However, while Church's thesis is not something one could prove, one of the standard theorems about *regular categories* (cf. [1] or [11, I.1.5.]) tells that every total and single-valued relation – a *map* – corresponds to a unique arrrow. Every regular category is thus isomorphic with its *category of maps*:

$$\mathscr{C} \cong \mathsf{Map}(\mathsf{Rel}\,\mathscr{C}). \tag{1}$$

[1] Current address: Department of Computing, Imperial College, London SW7 2BZ, United Kingdom. Email: pavlovic@doc.ic.ac.uk.

Recall that a finitely complete category is said to be regular if every arrow in it factorises as a regular epi (a coequaliser) followed by a monic, and if this factorisation is stable under pullbacks. Originally, the categorical calculus of relations was based on this canonical factorisation, usually denoted (Epi*, Mon).

In [22, 13], the calculus of relations has been extended from (Epi*, Mon), to stable *proper* factorisation systems $(\mathscr{E}, \mathscr{M})$. The word "proper" here means that every $\mathscr{M}$-arrow is monic, while every $\mathscr{E}$-arrow is epi. The relations are thus not given by arbitrary monics, but only by those from a smaller class $\mathscr{M}$. However, Kelly [13] has shown that the induced calculus reduces to the ordinary calculus of relations in a regular category. He has characterised the maps relative to $\mathscr{M}$, and proved that they still form a regular category. The calculus of Mon-relations in this new category coincides with the calculus of $\mathscr{M}$-relations in the old one.

Such a reduction will not be possible if we further generalise the concept of relation. Namely, *every* stable factorisation provides a sufficient basis for a calculus of relations. By this we mean that a *bicategory of relations* can be formed from it, in a similar fashion as it is formed from the canonical factorisation in a regular category. Although not in full generality, this fact has been checked long time ago, in J. Meisen's thesis [18]. But, in absence of applications, this general calculus of relations received little attention. It seemed difficult to make sense of a relation which would not be a monic in the first place.

In the eighties, computer science has revived the original idea of constructivism, by which a proposition-as-type could have many different proofs-as-constructions. Presented categorically, this *strongly constructive logic* naturally leads to the idea that a relation might be more than just a subobject [19, 20]. A generalised relation $r: R \to A \times B$ can be understood as an indexed family of constructive proofs: the fibre $r^{-1}(a, b)$ represents the set of proofs that $a \in A$ and $b \in B$ are r-related. According to this interpretation, the proofs are unique if and only if r is monic.[2]

And so, while the possible applications of the calculi of relations *restricted* to proper factorisations seem to lie in the realm of topology, quasi-toposes and similar structures, the *extended* calculi of relations might provide more insight in constructive logic and its models, such as the one described in Section 2.

Besides this model, in Section 2 we outline the construction of the bicategory of relations. Section 3 reviews the categorical concept of a map, due to Lawvere [16], and establishes its main properties in a generalised bicategory of relations. Section 4 sets up the background on which the relativised maps are to be compared with the arrows of the underlying category. In Section 5, finally, we determine a precise condition on the system $(\mathscr{E}, \mathscr{M})$, which guarantees an isomorphism like (1), and ensures that the maps relative to this factorisation exactly correspond to the

---

[2] D. Benson has pointed out to me that non-monic relations actually arise in computer science at a much larger scale: they come about, for instance, whenever one is searching tables with multiple entries for a key.

actual arrows. The condition is simply that all the elements of $\mathscr{E}$ must be regular epi.

Unfortunately, we must admit that the simplicity of this result conspicuously contrasts the relative complexity of its proof. In posetal bicategories, the notion of a map as a self-adjoint relation was a fairly transparent one – with the unit and the counit of the adjunction expressing the totality and the single-valuedness of the relation. In the non-posetal case, however, the adjunction equations, imposed on this unit and counit, do not hold automatically any more. Showing that they are satisfied involves chasing rather complicated diagrams. Moreover, we are lacking logical intuition for these commutativity conditions, supposed to tie together the proof that a map is total with the proof that it is single-valued.

Strongly constructive logic does not seem to have many other choices at this point, but to follow the advice of category theory. Chasing diagrams might be a good way to acquire some insight in the structure of proofs-as-constructions. This idea will be pursued in part II of this paper [21]. The logical meaning of the adjunction equations will unravel in a more general setting.

2. Bicategory of relations

Unless explicitly stated otherwise, throughout this paper we shall assume that $\mathscr{C}$ is a category with finite limits, and that $(\mathscr{E}, \mathscr{M})$ is a stable factorisation system in $\mathscr{C}$. The basic reference for factorisations is [10].

Every $(\mathscr{E}, \mathscr{M})$ induces in $\mathscr{C}$ a *bicategory of relations* $\mathscr{R}$. Some details can be found in [18]. Bicategories were introduced in [3]; see also [4, 6].

2.1. Relations, cells

The objects of the bicategory of relations $\mathscr{R} = \mathscr{R}_{(\mathscr{C}, \mathscr{E}, \mathscr{M})}$ are the same as the objects of $\mathscr{C}$. The hom-categories are

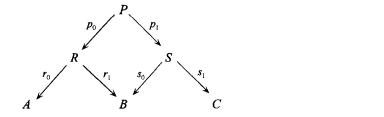$$\mathscr{R}(A, B) := \mathscr{M}/A \times B$$

i.e. the subcategories of $\mathscr{C}/A \times B$ spanned by the $\mathscr{M}$-arrows. The objects of $\mathscr{R}(A, B)$ are what we call the $(\mathscr{M}\text{-})$ *relations* from A to B. The arrows between them generalise the inclusion of one relation in another. Notice that not just the objects, but also the arrows of the category $\mathscr{M}/A \times B$ belong to $\mathscr{M}$, since this class is closed under the division on the left.

It is sometimes convenient to refer to the objects of a bicategory as to 0-cells, to its arrows, which are the objects of hom-categories, as to 1-cells, and to the arrows of its hom-categories as to 2-cells. The 1-cells, our *relations*, are usually denoted using the arrow symbols with a slash: instead of $(r: R \to A \times B) \in \mathscr{R}(A, B)$, we shall write $r: A \nrightarrow B$. The ordinary arrows will be kept for 2-cells. By $\alpha: r \to r'$ we shall denote an arrow in $\mathscr{M}/A \times B$, from $r: R \to A \times B$ to $r': R' \to A \times B$.

2.2. First examples

In any $\mathscr{C}$ with finite limits, the simplest calculi of relations are induced by the trivial factorisation systems (Iso, All) and (All, Iso), where All is the class of all arrows of $\mathscr{C}$, while Iso are the isos. The latter system induces uninteresting relations, but the former leads to the calculus of *spans* [3, 6, 30] in the bicategory $\mathbf{Span} = \mathscr{R}_{(\mathscr{C}, \mathsf{Iso}, \mathsf{All})}$. A span from A to B is an arbitrary arrow to $A \times B$. When $\mathscr{C}$ is the category of sets and functions, the spans can be represented by possibly infinite matrices of cardinal numbers; if we restrict to finite sets, they become ordinary matrices of natural numbers.

To compose spans $r = \langle r_0, r_1 \rangle : R \to A \times B$ and $s = \langle s_0, s_1 \rangle : S \to B \times C$, one forms a pullback of r_1 and s_0

$$(2)$$

and takes the span $\langle r_0 p_0, s_1 p_1 \rangle$ as the composite $s \otimes r : P \to A \times C$. In the category of sets, this operation exactly corresponds to the matrix composition.

Besides the two trivial factorisations, the category of sets allows only the epi-mono factorisation, which, of course, induces ordinary calculus of relations. Recall that binary relations on sets can be represented as matrices with entries 0 or 1. To compose them in this form, one first applies the ordinary matrix composition, and then reduces all the non-zero entries in the result to 1. Categorically, this last step means *taking the image* of the span $\langle r_0 p_0, s_1 p_1 \rangle$, constructed as in (2), but for monic spans r and s. The same procedure determines the composition of the Mon-relations in arbitrary regular category $\mathscr{C}$ as well as the composition of the relations relative to any stable factorisation.

2.3. Composition, identity

To compose relations, one thus manipulates with spans, using the adjunction

$$\mathscr{R}(A, B) = \mathscr{M}/A \times B \xrightarrow[\substack{\text{in}}]{\overset{\text{im}}{\longleftarrow}} \mathscr{C}/A \times B = \mathsf{Span}(A, B)$$

of the inclusion and the image functor. In principle, the bicategorical structure of $\mathscr{R}$ can be derived from the requirement that the global assignment $\mathrm{im} : \mathbf{Span} \to \mathscr{R}$ yields a homomorphism of bicategories. It determines that 1-cells $r : A \nrightarrow B$ and $s : B \nrightarrow C$ must compose to $s \otimes r = \mathrm{im} \langle r_0 p_0, s_1 p_1 \rangle : A \nrightarrow C$. Furthermore, any 2-cell

$\alpha : r \to r'$ in $\mathcal{R}(A, B)$ induces an arrow a from $\langle r_0 p_0, s_1 p_1 \rangle$ to $\langle r'_0 p'_0, s_1 p'_1 \rangle$ in $\mathcal{C}/A \times C$;
– with $\langle r'_0, r'_1 \rangle = r'$, and with p'_0 and p'_1 obtained from r' and s in a diagram like (2). The 2-cell $s \otimes \alpha : s \otimes r \to s \otimes r'$ is now the arrow im(a) from im$\langle r_0 p_0, s_1 p_1 \rangle$ to im$\langle r'_0 p'_0, s_1 p'_1 \rangle$ in $\mathcal{M}/A \times C$.

Given $\beta : s \to s'$ in $\mathcal{R}(B, C)$, the 2-cell $\beta \otimes r : s \otimes r \to s' \otimes r$ will be calculated in a similar fashion. One easily checks that the constructions $s \otimes (-)$ and $(-) \otimes r$ are functorial, and that α and β are "natural", in the sense that, for a suitable choice of images (or up to isomorphism)
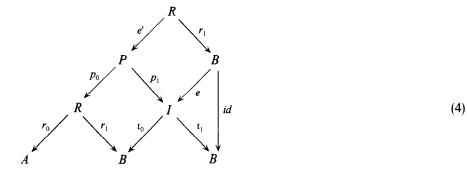
$$(\beta \otimes r')(s \otimes \alpha) = (s' \otimes \alpha)(\beta \otimes r) \tag{3}$$

holds. By definition, this 2-cell is taken to be the composite $\beta \otimes \alpha$.

The stability of the factorisation is necessary and sufficient for the associativity of $\otimes$. An early discussion about this can be found in [14].

The identities in $\mathcal{R}$ will be denoted $\iota_B : B \nrightarrow B$. By definition, ι_B is the $\mathcal{M}$-image of the diagonal $\delta_B := \langle id, id \rangle : B \to B \times B$, which is the identity in **Span**. Let us show, for instance, that $\iota_B \otimes r \cong r$, since this definition does not seem to be a part of common knowledge. (E.g. Meisen [18] assumes that all diagonals are in $\mathcal{M}$, apparently for no other reason than to use them as identities. This assumption then implies, for instance, that all $\mathcal{E}$-arrows are epi.)

Consider the diagram

$$
\begin{array}{c}
\text{diagram (4)}
\end{array}
\tag{4}
$$

Both squares are pullbacks. Let e be the coimage of δ_B, i.e. $\langle id, id \rangle = \langle \iota_0, \iota_1 \rangle e$. A suitable choice of pullbacks in (4) yields $r_0 p_0 e' = r_0$ and $\iota_1 p_1 e' = r_1$. By the stability, e' is an $\mathcal{E}$-arrow. With a suitable choice of images (or up to iso), we get

$$\iota_B \otimes r = \text{im} \langle r_0 p_0, \iota_1 p_1 \rangle = \text{im} \langle r_0 p_0, \iota_1 p_1 \rangle e' = \text{im} \langle r_0 p_0 e', \iota_1 p_1 e' \rangle = r.$$

2.4. Structure

In [8, 3.5.] and [11, II.2.154], bicategories of relations induced by the canonical factorisation in regular categories have been characterised and axiomatised. Our generalised bicategories of relations could be pinned down in a similar way. However,

the axioms appear to be significantly more complicated, due to the fact that we are not dealing with posetal bicategories.

Without going into details, let us just point out that a bicategory of relations relative to a factorisation still carries all the structure of a *'bicategory of relations'*[2] as defined by Carboni and Walters [8], *but* suitably generalised. It can also be regarded as a glorified *allegory* of Freyd and Scedrov [11].

For instance, every hom-category $\mathscr{R}(A, B) = \mathscr{M}/A \times B$ has finite products. Obviously, the product of relations $r: R \to A \times B$ and $r': R' \to A \times B$ will be their pullback in $\mathscr{C}$. We shall denote it by $r \wedge r'$. The canonical terminal object of $\mathscr{R}(A, B)$ is represented by the identity $id_{A \times B}$ in $\mathscr{C}$.

For the record, note that the composition (2) can be presented combining the local product $\wedge$ of $\mathscr{R}$ and the global product $\times$ of $\mathscr{C}$:

$$s \otimes r = \mathrm{im}(\pi((A \times s) \wedge (r \times C))), \tag{5}$$

where $\pi: A \times B \times C \to A \times C$ is the projection.

Finally, every bicategory of relations is self-dual: the bifunctor

$$(-)^{\circ}: \mathscr{R}^{\mathrm{op}} \longrightarrow \mathscr{R}$$

leaves the objects and the 2-cells unchanged, while it takes each $\mathscr{M}$-relation $r = \langle r_0, r_1 \rangle: R \to A \times B$ into the *opposite* relation $r^{\circ} := \langle r_1, r_0 \rangle: R \to B \times A$.

2.5. More examples

As observed in [9, 15], every category $\mathscr{C}$ generates a *free* factorisation system. It is couched in the arrow category $\mathscr{C}/\mathscr{C}$, and a telling notation is $(\triangle, \nabla)$. Namely, among the arrows

$$\binom{u}{v}: f \to g \quad \text{of } \mathscr{C}/\mathscr{C},$$

which are just the squares $vf = gu$, the abstract "epis" $\triangle$ can be recognised as the "upper triangles"

$$\triangle = \left\{ \binom{u}{v} \in \mathscr{C}/\mathscr{C} \mid u \text{ is iso} \right\},$$

while the abstract "monics" ∇ correspond to the "lower triangles"

$$\nabla = \left\{ \binom{u}{v} \in \mathscr{C}/\mathscr{C} \mid v \text{ is iso} \right\}.$$

---

[2] The quotation marks are due to the authors. [26] seems to be dealing with a related non-posetal structure.

A square thus decomposes into a ∇-part and a $\triangle$-part along the south-east diagonal. This factorisation is clearly stable, since the pullbacks in $\mathscr{C}/\mathscr{C}$ are created by the functor $\mathscr{C}/\mathscr{C} \to \mathscr{C} \times \mathscr{C}$, projecting an arrow to its domain and codomain.

Another factorisation in $\mathscr{C}/\mathscr{C}$ – stable for similar reasons – is obtained if we take ∇ as the class of abstract "epis". We get the system $(\nabla, \square)$, where $\square$ denotes the class of the pullback squares.

The calculus of relations with respect to $(\triangle, \nabla)$ can be *almost* reduced to the calculus of spans in $\mathscr{C}$. A ∇-relation is just a span, but equipped with an additional arrow coming out of its domain and another one coming out of its codomain. (Indeed, the category of ∇-relations is formed upon the objects of $\mathscr{C}/\mathscr{C}$.) Since a function out of a set induces a partition on it, a span with the arrows out of its domain and its codomain can be viewed as a matrix *divided in blocks*. The calculus of ∇-relations thus resembles algebra of blocked matrices.

The calculus of relations with respect to the factorisation $(\nabla, \square)$ can be interpreted in a similar fashion.

A particularly interesting factorisation, especially from the point of view of strongly constructive logic, can be found in the category $\mathscr{S}$ of *encoded sets* (also called *assemblies* [5], or *ω-sets* [17]). An encoded set can be presented as a function $K : |K| \to \Sigma^*$, where $|K|$ is an arbitrary set and Σ^* is the set of inhabited (nonempty) sets of natural numbers. An arrow $f : K \to L$ between two encoded sets in an ordinary function $f : |K| \to |L|$, but *effectively encodable*, in the sense that it can be traced on the codes by a partial recursive function. More precisely, for each $f \in \mathscr{S}(K, L)$ there must exist a natural number $n \in \mathbf{N}$ such that, for every $x \in |K|$, the partial recursive function $n^{\cdot} : \mathbf{N} {\rightharpoonup} \mathbf{N}$, encoded by n, is defined on all of $K(x) \in \Sigma^*$, and maps it into $L(f(x))$. This is summarised on the following picture

$$
\begin{array}{ccc}
|K| & \xrightarrow{\ f\ } & |L| \\
{\scriptstyle K}\downarrow & \subseteq & \downarrow{\scriptstyle L} \\
\Sigma^* & \xrightarrow[\ n^{\cdot\cdot}\ ]{} & \Sigma^*
\end{array}
\tag{6}
$$

where $n^{\,\cdot\cdot} : \Sigma^* {\rightharpoonup} \Sigma^*$ denotes the partial direct image of $n^{\cdot}$. This function is defined on $X \in \Sigma^*$ if and only if $n^{\cdot}$ is defined on all of X, and then

$$
n^{\,\cdot\cdot} X := \{ n^{\cdot}m \mid m \in X \}.
$$

Otherwise, $n^{\,\cdot\cdot} X$ is undefined. Therefore, to compose $n^{\cdot\cdot}$ and K on (6), we must require that the range of K is contained in the domain of $n^{\cdot\cdot}$.

Now we define the *modest* factorisation on $\mathscr{S}$. Each encoded set K comes with a symmetric, reflexive relation

$$
x \sim y \quad \Leftrightarrow \quad \exists n \in K(x) \cap K(y).
$$

The set 'K is said to be *modest* if $\sim$ is as small as possible:

$$x \sim y \;\Rightarrow\; x = y.$$

K is *immodest* if $|K|$ is inhabited and $\sim$ is large: every two elements of $|K|$ must be related in the transitive closure of $\sim$.

We say that an arrow of $\mathscr{S}$ is (im)modest if it is such fibrewise:

$$\mathsf{Mod} := \{ f \in \mathscr{S} \,|\, \forall y . f^{-1}(y) \text{ is modest} \},$$

$$\mathsf{Imm} := \{ f \in \mathscr{S} \,|\, \forall y . f^{-1}(y) \text{ is immodest} \}.$$

The pair $(\mathsf{Imm}, \mathsf{Mod})$ forms a stable factorisation in $\mathscr{S}$. It can be seen as the externalisation of the internal category of modest sets. The fact that we are obtaining a stable factorisation by externalising it means that this small category is cocomplete (in the fibrational sense). On the other hand, it is complete too. While contradicting the *classical* commonplace that the only small (co)complete categories are lattices, the category of modest sets can be thought of as the small category of strongly *constructive* truth values, with the small products and coproducts providing, respectively, the universal and the existential quantification. The encoded sets and the modest factorisation offer a basis for systematic semantical investigations, where the results of the present paper might be of use [20].

3. Maps

Definition 3.1 (following Lawvere [16]). In an arbitrary bicategory, a 1-cell is a *map* if it has a right adjoint.

Thus, for a map $r: A \nrightarrow B$, there must exist a 1-cell $r^*: B \nrightarrow A$ and two 2-cells $\eta: \iota_A \to r^* \otimes r$ and $\varepsilon: r \otimes r^* \to \iota_B$, satisfying the adjunction identities

$$(\varepsilon \otimes r)(r \otimes \eta) = r, \tag{7}$$

$$(r^* \otimes \varepsilon)(\eta \otimes r^*) = r^*, \tag{8}$$

In any bicategory, the maps form a subbicategory: they are closed under the composition and contain the identities. In the bicategory of relations as described in the preceeding section, they even form a subcategory. The reason is that they are quite rigid: two maps become isomorphic as soon as they are connected by a 2-cell which commutes with the adjunction structure. Some intuitive explanations will be provided at the end of this section.

Lemma 3.2. *Let $\mathcal{R}$ be any bicategory; let r, $s \in \mathcal{R}(A, B)$ be maps, with right adjoints r^* and s^*, units η_r, η_s and counits ε_r, ε_s; and let $\alpha : r \to s$ and $\alpha^* : r^* \to s^*$ be 2-cells satisfying*

$$(\alpha^* \otimes \alpha)\eta_r = \eta_s, \tag{9}$$

$$\varepsilon_s(\alpha \otimes \alpha^*) = \varepsilon_r. \tag{10}$$

Then r and s must be isomorphic: α must be an isomorphism. (Dually, the same holds for r^*, s^* and α^*.)

Proof. The following diagram shows that α must be a split monic.

$$\tag{11}$$

The right-hand square commutes by (10); the left-hand square, by the bicategorical naturality law (3). Since the adjunction $r \dashv r^*$ gives

$$(\varepsilon_r \otimes r)(r \otimes \eta_r) = id_r,$$

the commutativity of (11) tells that

$$\tilde{\alpha} := (\varepsilon_s \otimes r)(s \otimes \alpha^* \otimes r)(s \otimes \eta_r) : s \to r \tag{12}$$

is the left inverse of α.

A diagram dual to (11) yields a left inverse $\tilde{\alpha}^*$ of α^*. And then a third version of (11) yields a left inverse of $\tilde{\alpha}$. To get this third version, switch r and s everywhere in (11), put $\tilde{\alpha}$ in place of α and $\tilde{\alpha}^*$ in place of α^*. The left-hand square in this diagram will commute by naturality again. The commutativity of the right-hand square will boil down to condition (9) in a completely straightforward, though lengthy chase, using only the naturality and the definitions of $\tilde{\alpha}$ and $\tilde{\alpha}^*$.

Since it has both a left and a right inverse, $\tilde{\alpha}$ is an iso. Therefore, α must be an iso too. □

The preceding lemma allows a significant strengthening of the concept of map in bicategories of relations: every map must have a *canonical* right adjoint there. This is a clue for the "intended meaning" of Definition 3.1.

Proposition 3.3. *If $r : A \nrightarrow B$ is a map in the bicategory of relations $\mathcal{R}$, its right adjoint must be $r^o : B \nrightarrow A$.*

Proof. To simplify notation, let $s := r^* : B \rightarrowtail A$ be some given right adjoint of the map $r : A \rightarrowtail B$ now. Dualizing $r \dashv s$, we get $s^o \dashv r^o : B \rightarrowtail A$. So both r and s^o are maps. We need to construct 2-cells $\alpha : s^o \to r$ and $\alpha^* : r^o \to s$ satisfying (9) and (10). Applying Lemma 3.2, we shall then be able to conclude that $s^o \cong r$, and hence $r \dashv r^o$, as asserted.

The 2-cell $\alpha : s^o \to r$ will be obtained by dualising

$$\alpha^o := (r^o \otimes \varepsilon)(\kappa \otimes s) : s \to r^o \otimes r \otimes s \to r^o, \tag{13}$$

where $\varepsilon : r \otimes s \to \iota_B$ is the counit of the adjunction $r \dashv s$. The 2-cell $\kappa : \iota_A \to r^o \otimes r$ will be constructed below. In fact, we shall construct an arrow k from δ_A to $r^o \otimes r$ in $\mathscr{C}/A \times A$. By the universal property of the $\mathscr{M}$-image $\iota_A = \mathrm{im}(\delta_A)$, this arrow will induce κ from ι_A to $r^o \otimes r$ in $\mathscr{M}/A \times A$.

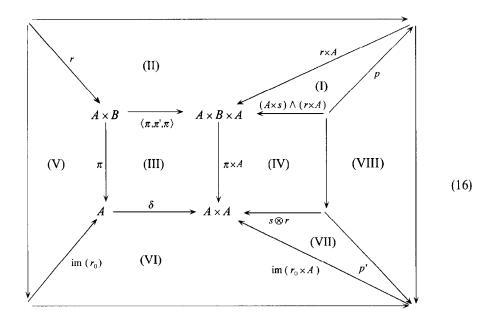First of all, consider the following diagram in $\mathscr{C}$.

The arrow d in (I) is induced by the fact that $r \wedge r$ is obtained by pulling back r along itself. Of course, d is a diagonal in $\mathscr{R}(A, B)$. Squares (II) and (III) are clearly pullbacks. Squares (IV) and (V) are obtained by forming the $\mathscr{M}$-images of the arrows $r_0 = \pi r$ and $(\pi \times A)((A \times r^o) \wedge (r \times A))$, respectively. Finally, the arrows d' and q' are induced by the universal property of the images.
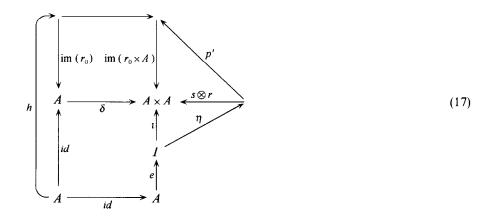
We shall now show that $\mathrm{im}(r_0)$ has a section h. Using it, we shall be able to define the arrow k in $\mathscr{C}/A \times A$ as

$$k := q'd'h : \delta_A \longrightarrow r^o \otimes r. \tag{15}$$

The following diagram will help us produce h.

$$(16)$$

The arrow p in (I) is this time the projection $(A \times s) \wedge (r \times A) \to r \times A$. Squares (II) and (III) are pullbacks again; and (IV) and (V) are again formed by factoring. The former is based on (5). Further decompose $(\pi \times A)(r \times A)$ to get $\mathrm{im}(r_0 \times A)$; then form (VI), (VII) and (VIII) using the universal property of the images. The stability of the factorisation implies that (VI) must be a pullback square, because (II) and (III) are, and (VI) is obtained by factorising two sides of their composite. Square (VI) is again displayed as the upper square on the next diagram. In the lower square, $\iota \circ e$ is the $\mathscr{E}\mathscr{M}$-decomposition of δ.

$$(17)$$

The equation

$$\operatorname{im}(r_0 \times A) \circ p' \circ \eta \circ e = \delta$$

induces the section h of $\operatorname{im}(r_0)$. Going back to (15), we get k. Hence κ for (13).

The other 2-cell, $\alpha^* : r^o \to s$, can be constructed in a similar fashion:

$$\alpha^* := (s \otimes \varepsilon^o)(\chi \otimes r^o) : r^o \longrightarrow s \otimes s^o \otimes r^o \longrightarrow s, \tag{18}$$

where $\chi : \iota_A \to s \otimes s^o$ is derived from the unit $\eta^o : \iota_A \to r^o \otimes s^o$ in exactly the same way as $\kappa : \iota_A \to r^o \otimes r$ was derived from $\eta : \iota_A \to s \otimes r$.

To complete all the conditions for Lemma 3.2, we must check that the constructed 2-cells satisfy (9) and (10), i.e. that we have

$$(\alpha^* \otimes \alpha)\eta^o = \eta,$$

$$\varepsilon(\alpha \otimes \alpha^*) = \varepsilon^o.$$

This is a completely "deterministic" task, which consists in "unfolding" the given and the omitted definitions. For the sake of brevity, we must leave it to the reader. $\square$

Corollary 3.4. *Let $\mathscr{R}$ be a bicategory of relations; let $r, s \in \mathscr{R}(A, B)$ be maps. If there is a 2-cell $r \to s$, then $r \cong s$.*

Proof. Since $r^* = r^o$ and $s^* = s^o$, every 2-cell $\alpha : r \to s$ automatically induces $\alpha^* := \alpha^o : r^o \to s^o$. It is easy to see that condition (10) is satisfied. But it is not easy at all to prove (9). Fortunately, in the bicategory of relations, we can show that α is an iso without referring to condition (9).

A left inverse $\tilde{\alpha}$ of α is constructed just as in Lemma 3.2: use diagram (11), substituting $(-)^o$ for $(-)^*$ everywhere. And now, in order to construct a left inverse of $\tilde{\alpha} : s \to r$, *just repeat the same procedure.*

This time, the right-hand square commutes without further ado, simply because (10) holds for *every* 2-cell β between maps and for its dual $\beta^* = \beta^o$. $\square$

We have proved that the maps in the bicategory of relations $\mathscr{R}$ are absolutely rigid: the only 2-cells between them are isomorphisms. The skeleton of the subcategory spanned in $\mathscr{R}(A, B)$ by maps is *discrete*: just a set. Taken modulo isomorphism, the maps of $\mathscr{R}$ thus form a *category* $\mathsf{Map}(\mathscr{R})$. By definition, the objects of $\mathsf{Map}(\mathscr{R})$ are the

same as the objects of $\mathscr{R}$. The arrows of **Map**$(\mathscr{R})$ are the equivalence classes of isomorphic maps from $\mathscr{R}$.

In fact, it will follow from [21, 5.3(b)] that a pair of maps in $\mathscr{R}(A, B)$ can be connected by *at most one* isomorphism. This means that the subcategory of $\mathscr{R}(A, B)$, spanned by the maps, is *equivalent* with its skeleton. The category **Map**$(\mathscr{R})$ is thus biequivalent with the subbicategory of $\mathscr{R}$ spanned by the maps.

Let us close this section by summarising some consequences of Proposition 3.3. On the technical side, this proposition greatly simplifies checking whether a relation r is a map or not. First of all, it is now sufficient to consider just one candidate for the right adjoint: the opposite relation r^o. Secondly, only one of the adjunction equations needs to be checked, since the other one will follow. Indeed, with $r^* := r^o$, identities (7) and (8) are dual to each other:

$$(r^o \otimes \varepsilon)(\eta \otimes r^o) = ((\varepsilon \otimes r)(r \otimes \eta))^o. \tag{20}$$

On the conceptual side, Proposition 3.3 explains how the categorical notion of map, as a relation with a right adjoint, captures the idea of a map as a total and single-valued relation. If each relation $r: A \leftrightarrow B$ is thought of as a predicate $r(x, y)$, the identity relation $\iota_A: A \leftrightarrow A$ becomes the equality $x \equiv x'$. On the other hand, the composition of relations $r: A \leftrightarrow B$ and $s: B \leftrightarrow C$ can be written in the form

$$s \otimes r = \exists y . r(x, y) \wedge s(y, z).$$

In this logical notation, the adjunction data for a map r appear as "proofs"

$$\eta : x \equiv x' \longrightarrow \exists y . r(x, y) \wedge r(x', y) \tag{21}$$

$$\varepsilon : \exists x . r(x, y') \wedge r(x, y) \longrightarrow y' \equiv y \tag{22}$$

Here we are neglecting the order of variables, so that both r and r^o correspond to $r(x, y)$. Evidently, the 2-cell ε "proves" that $r(x, y)$ is single-valued in y: if y and y' are both r-related to some x, then they must be equal. On the other hand, the 2-cell η tells that $r(x, y)$ is total in x: if we put $x' := x$, the domain of η becomes true and we can derive a "proof" of $\exists y . r(x, y)$.

Similar ideas motivate all the bicategorical constructions in this section. For instance, diagram (19) shows how a "proof" $\tilde{\alpha}$ from a total relation $s(x, y)$ to a single-valued relation $r(x, y)$ induces a "proof" in the opposite direction

$$r(x, y) \rightarrow \exists x'y' . r(x, y') \wedge s(x', y') \wedge s(x', y)$$

$$\rightarrow \exists x'y' . r(x, y') \wedge r(x', y') \wedge s(x', y)$$

$$\rightarrow s(x, y).$$

Construction (11) is a variation on the same theme. On the other hand, diagram (16) can be understood as deriving from $\eta : x \equiv x' \rightarrow \exists y . s(x, y) \wedge r(x, y)$ a "proof" of $\exists y . r(x, y)$, i.e. a splitting h of $\text{im}(r_0)$. From this, $\kappa : x \equiv x' \rightarrow \exists y . r(x, y) \wedge r(x', y)$ is then derived on diagram (14).

4. Convergence

Let us now address the central question of this paper: *What is the connection between the maps in the bicategory of relations $\mathscr{R} = \mathsf{Rel}(\mathscr{C})$ and the arrows of the underlying category $\mathscr{C}$?* The first observation in this direction is that every arrow of $\mathscr{C}$ induces a map in $\mathscr{R}$. We say that this map *converges* to the arrow which has induced it.

Lemma 4.1. (a) *For any arrow $f \in \mathscr{C}(A, B)$, the relation $\Gamma f := \mathrm{im}\langle id, f \rangle \in \mathscr{R}(A, B)$ is a map. It is* the *graph of the function f.*
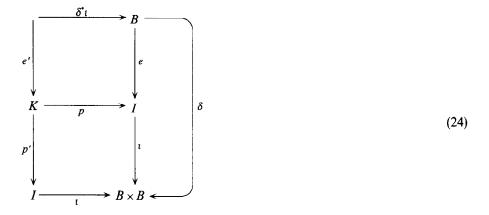
(b) *As an arrow in $\mathscr{C}$, a graph of a function is always a monic.*

(c) *For any relation $s = \langle c, d \rangle \in \mathscr{M}/C \times D$ holds $s = \Gamma d \otimes \Gamma c^{\circ}$.*

Proof. Parts (a) and (c) can be demonstrated by extending standard arguments about spans [6, Section 2].

Towards a proof of (b), observe that each graph of a function is a pullback of the identity relation ι

$$\Gamma f = (f \times B)^* \iota. \tag{23}$$

To prove that Γf is a monic, it suffices to prove that $\iota = \mathrm{im}(\delta)$ is a monic. We shall now show that the arrows p and p' on the next diagram are equal isomorphisms. Hence the result.

$$\tag{24}$$

Let both squares be pullbacks by definition. Since the factorisation is stable, we have

$$\delta^* \iota := \delta^* \mathrm{im}(\delta) = \mathrm{im}(\delta^* \delta).$$

So $\delta^* \iota$ must be iso, since $\delta^* \delta$ is one. But now it follows that

$$\iota \circ p = \mathrm{im}(\delta \circ \delta^* \iota) \cong \mathrm{im}(\delta) = \iota,$$

$$p \circ e' = e \circ \delta^* \iota \cong e.$$

Since ι is in $\mathcal{M}$, the first equation implies that p is in $\mathcal{M}$. Since e and e' are in $\mathcal{E}$, the second equation implies that p is in $\mathcal{E}$. Being both in $\mathcal{M}$ and in $\mathcal{E}$, p must be iso.

But p and p' constitute a kernel of ι. So they have a common right inverse. Since p is iso, they must be equal. □

All the notions that will now be defined are, of course, relative to the factorisation system $(\mathcal{E}, \mathcal{M})$.

Definition 4.2. A map $r: A \nrightarrow B$ in the bicategory of relations $\mathcal{R}$ is said to *converge* to a function $f: A \to B$ from $\mathcal{C}$ if $r \cong \Gamma f$.

Definition 4.3. An object $B \in \mathcal{R}$ is *separated* if a map to it can converge to at most one function; in other words, $\Gamma f \cong \Gamma f'$ implies $f = f'$ for all $f, f' \in \mathcal{C}(A, B)$.

Definition 4.4. B is said to be *Cauchy complete* if every map to it converges to a unique function.

Definition 4.5. The category $\mathcal{C}$ satisfies the principle of *function comprehension* if all its objects are Cauchy complete.

With the exception of the function comprehension, which stems back to Frege, Russell and Whitehead [29, Subsection *30], these concepts are developed along the lines proposed in Lawvere's seminal paper [16]. Hence somewhat metonymical terminology. Cauchy actually entered scene only through Lawvere's striking example of Cauchy complete spaces, which are exactly the Cauchy complete objects in the $\mathbf{R}^+$-enriched category of metric spaces. For more about the Cauchy completeness, see [12, 5.5.].

On the other hand, Walters [27, 28] has observed that the Cauchy completeness is the bicategorical version of the fundamental concept of *sheaf*. Thus, in some sense, we are studying *sheaf theory* relative to a factorisation here. (See also [23, 24].) Some propositions indeed sound familiar.

Proposition 4.6. *An object B is separated with respect to $(\mathcal{E}, \mathcal{M})$ if and only if the diagonal δ_B is in $\mathcal{M}$.*

Proof. ($\Leftarrow$) If δ_B is in $\mathcal{M}$, then $\langle id_A, f \rangle$ must be in $\mathcal{M}$ for every arrow $f: A \to B$, since $\langle id_A, f \rangle$ is a pullback of δ_B along $f \times B$. Therefore, $\Gamma f = \langle id_A, f \rangle$.

But $\langle id_A, f \rangle \cong \langle id_A, f' \rangle$ obviously implies $f = f'$.

($\Rightarrow$) Assume that B is separated. We first want to show that for the identity relation $\iota_B = \langle \iota_0, \iota_1 \rangle$ holds $\Gamma \iota_0 \cong \Gamma \iota_1$.

Consider the following diagram, consisting of pullbacks again.

$$
\begin{array}{ccc}
\xrightarrow{\ \ p\ \ } & \xrightarrow{\quad\quad} & I \\
\Big\downarrow{\scriptstyle q} & \Big\downarrow{\scriptstyle \Gamma_{\iota_0}} & \Big\downarrow{\scriptstyle \langle \iota_0, \iota_1 \rangle} \\
I \xrightarrow[\langle id, \iota_1 \rangle]{} & I \times B \xrightarrow[\iota_0 \times B]{} & B \times B
\end{array}
\tag{25}
$$

The external square is obviously a pullback of ι along itself. As a component of a kernel pair, the arrow q must be a split epi. By going along its section and then along p, we get an arrow $\langle id, \iota_1 \rangle \to \Gamma\iota_0$ in $\mathscr{C}/I \times B$. The factorisation induces $\Gamma\iota_1 \to \Gamma\iota_0$ in $\mathscr{M}/I \times B$.

We can now conclude that $\Gamma\iota_0 \cong \Gamma\iota_1$ either using Lemma 4.1(a) and Corollary 3.4, or by constructing an arrow $\Gamma\iota_0 \to \Gamma\iota_1$ and then applying Lemma 4.1(b). Since B is separated, we get $\iota_0 = \iota_1$ and hence $\iota_B = \langle \iota_0, \iota_0 \rangle$. Recalling the $\mathscr{E}\mathscr{M}$-decomposition $\delta_B = \iota_B \circ e$, which defines the identity relation ι_B, we get $\iota_0 \circ e = id_B$. On the other hand, by the uniqueness of decomposition, the diagonal in the square

$$
\begin{array}{ccc}
K \xrightarrow{\ \ e\ \ } & & I \\
\Big\downarrow{\scriptstyle e} \quad {\scriptstyle e}\nearrow B \searrow{\scriptstyle \delta} & & \Big\downarrow{\scriptstyle \langle i, i \rangle} \\
I \xrightarrow[\langle i, i \rangle]{} & & B \times B
\end{array}
\tag{26}
$$

must be identity. Hence $e \circ \iota_0 = id_I$. So ι_0 is an iso and $\iota_B = \langle \iota_0, \iota_0 \rangle$ is isomorphic to $\delta_B = \langle id_B, id_B \rangle$. So δ_B is in $\mathscr{M}$. $\quad\square$

The "global" appearance of the concepts defined in Definitions 4.2–4.5 is reflected in the properties of the *graph functor*

$$
G : \mathscr{C} \longrightarrow \mathsf{Map}(\mathsf{Rel}\,\mathscr{C}).
$$

By definition, it is the identity on the objects and takes an arrow f to the equivalence class Gf of all $\mathscr{M}$-maps isomorphic with Γf.

It follows immediately from the definitions that the graph functor G is faithful if and only if all objects of $\mathscr{C}$ are separated. But Proposition 4.6 allows some more informative characterisations. (Recall that Mon^* denotes the class of all *regular* monics in a category.)

Corollary 4.7. *For a stable factorisation* $(\mathscr{E}, \mathscr{M})$ *in finitely complete* $\mathscr{C}$, *the following conditions are equivalent*:
  (a) *The graph functor* $G : \mathscr{C} \to \mathsf{Map}(\mathsf{Rel}\,\mathscr{C})$ *is faithful*,
  (b) $\mathscr{E} \subseteq \mathsf{Epi}$,
  (c) $\mathscr{M} \supseteq \mathsf{Mon}^*$.

Proof (sketch). By definition, $Gf = Gf' \Leftrightarrow \Gamma f \cong \Gamma f'$. Condition (a) thus means that all objects $\mathscr{C}$ are separated. By Proposition 4.6, it further means that all diagonals are contained in $\mathscr{M}$. But it is well-known that this condition is equivalent with (b) and (c). (A proof can be found in [18, 2.4].) $\square$

Remark 4.8. When the graph functor is faithful, the "logic" of factorisation is, in a sense, faithful too. For instance, it is easy to see that the condition

> *an arrow f is epi whenever* $\text{im}(f)$ *is a split epi*

is equivalent to Corollary 4.7(b). What does it mean? If $\text{im}(f)$ is the predicate $\exists x . f(x) \equiv y$ and if a section of $\text{im}(f)$ is a "proof" of $\exists x . f(x) \equiv y$, this condition means that *f must really be epi if the factorisation says so.*

Analogous statement about the monics can be derived from the next lemma, which will also be useful later.

Lemma 4.9. *An object B is separated if and only if, for any pair $f, f' \in \mathscr{C}(A, B)$, the existence of an arrow $\text{im}\langle f, f' \rangle \to \iota_B$ implies $f = f'$.*

Proof. ($\Leftarrow$) Take $f = \iota_0$ and $f' = \iota_1$, to conclude $\iota_0 = \iota_1$, for $\iota_B = \langle \iota_0, \iota_1 \rangle$. Then proceed as in the proof of Proposition 4.6 to derive $\delta_B \in \mathscr{M}$.

($\Rightarrow$) If $\delta_B \in \mathscr{M}$, then $\iota_B = \delta_B$. The existence of an arrow $\varphi : \text{im}\langle f, f' \rangle \to \iota_B$ now implies that $\text{im}\langle f, f' \rangle = \langle \varphi, \varphi \rangle$. But then $f = \varphi e = f'$, where e is the coimage of $\langle f, f' \rangle$. $\square$

5. Function comprehension

To make the story shorter, we shall not derive a global characterisation of the function comprehension by extending the corresponding local property, the Cauchy completeness, like we derived Corollary 4.7 from Proposition 4.6. A characterisation of the Cauchy complete objects will be left implicit in our direct description of the function comprehension, the situation when *all* objects are Cauchy complete, and all maps converge to unique arrows. This is precisely when the graph functor G is full and faithful. Being by definition bijective on objects, G will actually be an isomorphism of categories in this case.

Theorem 5.1. *A finitely complete category $\mathscr{C}$ is function comprehensive with respect to a stable factorisation $(\mathscr{E}, \mathscr{M})$ if and only if $\mathscr{E} \subseteq \text{Epi}^*$.*

The proof of this theorem requires some prerequisites.

Recall, first of all, that the *kernel* of an arrow $f : A \to B$ is a pair of arrows $\ker(f) := \langle k_0, k_1 \rangle : K \to A \times A$, obtained by pulling back f along itself. More precisely, a kernel is an isomorphism class of such pairs, a *subobject* of $A \times A$. Viewed in

this way, kernels of the arrows with the same domain are partially ordered. Using the couniversal property of the pullbacks, one easily shows that the inclusion

$$\ker(a) \subseteq \ker(b)$$

is equivalent to the statement that

$$ax = ay \;\Rightarrow\; bx = by \quad \text{for all } x, y.$$

For reasons explained in the end of Section 3, let us call a relation $r : A \nrightarrow B$ *total* if there is a 2-cell $\iota_A \to r^o \otimes r$; and let us say that r is *single-valued* if there is a 2-cell $r \otimes r^o \to \iota_B$.

Lemma 5.2. *An $\mathcal{M}$-relation $r = \langle r_0, r_1 \rangle : R \to A \times B$ is*
(a) *total if and only if* $\mathrm{im}(r_0)$ *is a split epi;*
(b) *single-valued if* $\ker(r_0) \subseteq \ker(r_1)$; *the converse holds when B is separated.*

Proof. (a) $\eta : \iota \to r^o \otimes r$ induces a section j of $\mathrm{im}(r_0)$, because the upper square on the next diagram is a pullback. Conversely, j induces η because e is an $\mathcal{E}$-arrow and $r^o \otimes r$ is an $\mathcal{M}$-arrow.

$$(27)$$

(b) $\ker(r_0) \subseteq \ker(r_1)$ actually means $r_1 p_0 = r_1 p_1$, for $\langle p_0, p_1 \rangle = \ker(r_0)$. In other words, there is an arrow $\langle r_1 p_0, r_1 p_1 \rangle \to \delta_B$. Hence

$$r \otimes r^o = \mathrm{im}\langle r_1 p_0, r_1 p_1 \rangle \longrightarrow \mathrm{im}(\delta_B) = \iota_B.$$

The other way around, if there is a "proof" $\mathrm{im}\langle r_1 p_0, r_1 p_1 \rangle \to \iota_B$, then $r_1 p_0 = r_1 p_1$ must hold by Lemma 4.9, whenever B is separated. But the last equation means $\ker(r_0) \subseteq \ker(r_1)$. $\square$

Remark 5.3. It is well-known [6, Proposition 2] that in the calculus of spans all maps converge. In other words, every map is isomorphic to one in the form $\langle id, f \rangle$. On the other hand, by Lemma 5.2, a span $r = \langle r_0, r_1 \rangle$ is total and single-valued as soon as r_0 is a split epi and $\ker(r_0) \subseteq \ker(r_1)$. Clearly, such a span need not be isomorphic to one in the form $\langle id, f \rangle$; not all total, single-valued spans are maps. This shows that, in

general, maps cannot be reduced to the total, single-valued relations: the proofs of the totality and of the single-valuedness must be suitably connected. The adjunction equations do matter. Their meaning will be analysed in [21].

Proof of Theorem 5.1. ($\Rightarrow$) First suppose that every map from $\mathscr{R}$ converges to a unique arrow in $\mathscr{C}$. Let $e: X \to C$ be a member of $\mathscr{E}$ and let $g: X \to B$ be any arrow that equates the kernel pair $\langle h_0, h_1 \rangle$ of e:

$$gh_0 = gh_1. \tag{28}$$

We shall show that there is a unique arrow f such that $g = fe$. This means that e is the coequalizer of its kernel pair, a regular epi.

Let the relation $r = \langle c, b \rangle$ be defined as the $\mathscr{M}$-image of the pair $\langle e, g \rangle$. The arrow c must be in $\mathscr{E}$, because both $e = ce'$ and the coimage e' of $\langle e, g \rangle$ are in $\mathscr{E}$.

Further decompose $b = \mathrm{im}(b) \circ d$, and take $s := \langle c, d \rangle$. The arrow s is still in $\mathscr{M}$, because both $r = (id \times \mathrm{im}(b)) \circ s$ and $id \times \mathrm{im}(b)$ are in $\mathscr{M}$.
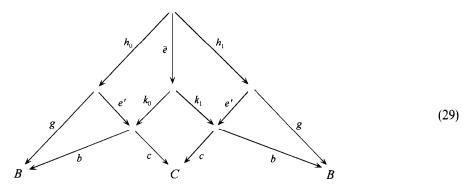
We claim that the relation s must be a map. In Lemma 5.4 below, this will be derived from the hypotheses that

(i) both c and d are in $\mathscr{E}$, while

(ii) $\ker(c) \subseteq \ker(d)$.

We shall prove (ii), since (i) is clear.

Let us first relate $\langle h_0, h_1 \rangle := \ker(e)$ and $\langle k_0, k_1 \rangle := \ker(c)$.

$$\tag{29}$$

The stability of the factorisation implies that the arrow $\tilde{e}$ from (29) must belong to $\mathscr{E}$. Hence

$$\mathrm{im}\langle bk_0, bk_1 \rangle = \mathrm{im}\langle gh_0, gh_1 \rangle. \tag{30}$$

But assumption (28) tells that there is an arrow from $\mathrm{im}\langle gh_0, gh_1 \rangle$ to ι_B. Hence there is one from $\mathrm{im}\langle bk_0, bk_1 \rangle$. Lemma 4.9, now yields

$$bk_0 = bk_1. \tag{31}$$

Indeed, the assumption that $\mathscr{C}$ is function comprehensive implies that B is Cauchy complete and, a fortiori, separated. Hence the hypothesis for Lemma 4.9.

Since c is in $\mathscr{E}$, the arrows k_0 and k_1 must be in $\mathscr{E}$ too: as its kernel pair, they are obtained from c in a pullback. On the other hand, d is by definition the coimage of b. The arrows dk_0 and dk_1 are thus in $\mathscr{E}$ and they must be the coimages of bk_0 and bk_1 respectively. (31) thus implies

$$dk_0 = dk_1. \tag{32}$$

And this is equivalent to (ii).

Now we can use Lemma 5.4 to conclude that s is a map. Since $\mathscr{C}$ is function comprehensive, s must be isomorphic to a unique graph $\langle id, d' \rangle$. The relation $r = (id \times \text{im}(b)) \circ s$ is thus isomorphic which the graph $\langle id, f \rangle$, where $f = \text{im}(b) \circ d'$. But r was defined as an $\mathscr{M}$-image of $\langle e, g \rangle$. If we choose $r = \langle id, f \rangle$, the decomposition becomes $\langle e, g \rangle = \langle id, f \rangle \circ e$, and we have $g = fe$, for unique f.

($\Leftarrow$) Now suppose $\mathscr{E} \subseteq \mathsf{Epi}^*$. To show that $\mathscr{C}$ is function comprehensive, we must, for any given map $r: A \mapsto B$, find a unique arrow $f: A \to B$ in $\mathscr{C}$, such that $r \cong \Gamma f$.

In fact, the uniqueness is easy, since we already know from Corollary 4.7 (and the remark preceding it) that $\mathscr{E} \subseteq \mathsf{Epi}^*$ implies that all objects are separated. Moreover it implies that all the diagonals, and hence all the arrows in the form $\langle id, f \rangle$ are in $\mathscr{M}$ (indeed, even all monics). The graphs of functions are thus in the form $\Gamma f = \langle id, f \rangle$. Given $r \in \mathscr{R}(A, B)$, we are thus looking for some $f \in \mathscr{C}(A, B)$ such that $r \cong \langle id, f \rangle$.

The desired arrow f will be obtained as the composition $f_1 \circ f_0$, according to the following scheme

$$\tag{33}$$

where $r = \langle a, b \rangle$, while c is the coimage of a. Note that $\langle c, b \rangle$ is still an $\mathscr{M}$-arrow, because both $r = (\text{im}(a) \times B) \circ \langle c, b \rangle$ and $\text{im}(a) \times B$ are in $\mathscr{M}$.

The section f_0 of $\text{im}(a)$ is obtained from Lemma 5.2(a), since a map r is, of course, total. The arrow f_1 is obtained using the fact that r is single-valued, and the assumption $\mathscr{E} \subseteq \mathsf{Epi}^*$.

Actually, we first use the fact that B is separated, a consequence of $\mathscr{E} \subseteq \mathsf{Epi}^*$. It allows us to apply Lemma 5.2(b) and express the single-valuedness of r in the form

$$\ker(a) \subseteq \ker(b). \tag{34}$$

Since $\ker(c) \subseteq \ker(a)$ obviously holds, we get $\ker(c) \subseteq \ker(b)$. In terms of $\langle k_0, k_1 \rangle = \ker(c)$, this is

$$bk_0 = bk_1. \tag{35}$$

Now we use the assumption that $c \in \mathscr{E}$ is a regular epi again: it coequalizes its kernel pair, i.e. the arrows k_0 and k_1. Eq. (35) thus induces a unique arrow f_1, such that $b = f_1 c$.

In fact, c must be an iso: it is in $\mathscr{M}$, because it satisfies $\langle id, f_1 \rangle \circ c = \langle c, b \rangle$, with $\langle id, f_1 \rangle$ and $\langle c, d \rangle$ in $\mathscr{M}$; and it is in $\mathscr{E}$ because it was defined as the coimage of a. We can take $c = id$, so that $f_1 = b$. Now we have

$$af_0 = id,$$

$$bf_0 = f_1 f_0 = f.$$

The arrow f_0 is thus a 2-cell between the maps $\langle id, f \rangle$ and $r = \langle a, b \rangle$. By Corollary 3.4, f_0 is an iso. The map r is isomorphic to $\langle id, f \rangle$. $\quad\square$

Finally, let us prove the main lemma, crucial in the first part of the preceeding proof.

Lemma 5.4. *Let $s = \langle c, d \rangle : S \to C \times D$ be an $\mathscr{M}$-relation. Suppose that c and d are $\mathscr{E}$-arrows, and that $\ker(c) \subseteq \ker(d)$. Then s is a map.*

Proof. Once again, translate the assumption $\ker(c) \subseteq \ker(d)$ into $dk_0 = dk_1$, where $\langle k_0, k_1 \rangle$ is the kernel of c. From this equation, and the fact that d, k_0 and k_1 are all in $\mathscr{E}$, it follows that

$$s \otimes s^o = \mathrm{im}\langle dk_0, dk_1 \rangle = \mathrm{im}\langle dk_0, dk_0 \rangle = \mathrm{im}(\delta_C dk_0) = \mathrm{im}(\delta_C) = \iota_C.$$
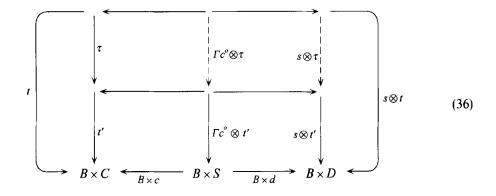
These equalities, of course, hold for a suitable choice of images. With this choice, the counit $\varepsilon : s \otimes s^o \to \iota_C$, induced by the arrow dk_0 from $\langle dk_0, dk_1 \rangle$ to δ_C, is represented by an identity. Furthermore, the 2-cell $\varepsilon \otimes s : s \otimes s^o \otimes s \to s$ can be chosen as identity too.

We shall construct a 2-cell $\eta : \iota_A \to s^o \otimes s$, such that $s \otimes \eta : s \to s \otimes s^o \otimes s$ is an identity. The adjunction equation

$$(\varepsilon \otimes s)(s \otimes \eta) = id$$

then holds trivially. Hence the result.

To analyse the way in which $s \otimes \eta$ comes about, we shall use a picture of the composition $\otimes$ based on Lemma 4.1(c). Let us first outline this picture abstractly. For relations $t, t' : B \nrightarrow C$, a 2-cell $\tau : t \to t'$, and a relation $s : C \nrightarrow D$, where $s = \langle c, d \rangle$, the 2-cell $s \otimes \tau : s \otimes t \to s \otimes t'$ is constructed in two steps, recalling that $s = \Gamma d \otimes \Gamma c^o$.

$$
\begin{array}{c}
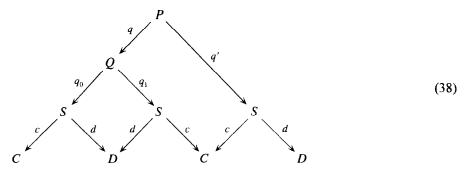\text{(diagram with } \tau,\ \Gamma c^{o}\otimes\tau,\ s\otimes\tau,\ t,\ \Gamma c^{o}\otimes t,\ s\otimes t,\ t',\ \Gamma c^{o}\otimes t',\ s\otimes t',\\[4pt]
B\times C \xleftarrow{\ B\times c\ } B\times S \xrightarrow{\ B\times d\ } B\times D)
\end{array}
\tag{36}
$$

The squares on the left-hand side are defined as pullbacks; the squares on the right-hand side are $\mathscr{E}\mathscr{M}$-decompositions. Namely, composing with a relation in the form Γc^{o}, for any arrow c from $\mathscr{C}$, boils down to pulling back along $B\times c$. On the other hand, a composite with a relation in the form Γd boils down to an image: given any $u : B \looparrowright S$ and $d : S \to D$, one easily calculates that $\Gamma d \otimes u = \mathrm{im}((B\times d)\circ u)$.

Now we apply the above scheme to the actual data.

$$
\begin{array}{c}
C \xleftarrow{\ c\ } S \\
\text{(I), (II), (III), (IV), (V), (VI), (VII) with arrows}\\
\tilde{e},\ \check{e},\ \hat{e},\ p,\ \dot{e},\ q,\ \eta,\ e,\ h,\ \ddot{e},\ s''\otimes s,\ \Gamma c^{o}\otimes s''\otimes s,\ \\
\Gamma c^{o}\otimes\eta,\ s\otimes\eta,\ s\otimes s''\otimes s = s,\ \iota,\ s,\\
C\times C \xleftarrow{\ C\times c\ } C\times S \xrightarrow{\ C\times d\ } S\times D
\end{array}
\tag{37}
$$

The objects P and Q, and the arrow $q : P \to Q$, are defined on the following pullback diagram:

$$
\begin{array}{c}
P \\
q\swarrow \qquad \searrow q' \\
Q \qquad\qquad\qquad \\
q_{0}\swarrow \quad \searrow q_{1} \qquad\qquad\\
S \qquad\quad S \qquad\quad S \\
c\swarrow\ \searrow d \quad d\swarrow\ \searrow c \quad c\swarrow\ \searrow d \\
C \qquad D \qquad C \qquad D
\end{array}
\tag{38}
$$

Evidently, P can be viewed as the limit of the zig-zag at the bottom, consisting of three copies of $s: S \to C \times D$. Three copies of $id: S \to S$ form a cone to this zig-zag. Let $p: S \to P$ be the limit factorisation of this cone. p is thus the common right inverse of $q_0 q$, $q_1 q$ and q'.

On the other hand, $s^o \otimes s$ in the $\mathcal{M}$-image of $\langle cq_0, cq_1 \rangle$. Let the arrow e on (37) be the coimage of this pair.

With the data defined like this, faces (I) and (II) on (37) commute when joined together:

$$(s^o \otimes s)\,eqp = \langle cq_0, cq_1 \rangle\, qp$$

$$= \langle cq_0 qp, cq_1 qp \rangle$$

$$= \langle c, c \rangle$$

$$= \delta c$$

$$= \iota \tilde{e} c. \tag{39}$$

The unit $\eta: \iota_C \to s^o \otimes s$ is determined as the arrow which makes (I) and (II) commute separately. Its existence and uniqueness are guaranteed by the fact that $\tilde{e}c \in \mathcal{E}$ and $s^o \otimes s \in \mathcal{M}$ are orthogonal to each other.

We proceed as on (36) and define (III) and (IV) as pullbacks, (VI) and (VII) as $\mathcal{E}\mathcal{M}$-factorisations.

Since it is obtained by pulling back $\delta_C = \iota_C \tilde{e}$ along $C \times c$, the arrow $(\Gamma c^o \otimes s^o \otimes s)(\Gamma c^o \otimes \eta)\hat{e}$ must be (isomorphic and can be chosen as) equal to $\langle id, c \rangle$. Postcomposed with $C \times d$, this arrow becomes s. The 2-cell $\check{e}\hat{e}$ is thus an endomorphism of s. Therefore, it must be in $\mathcal{M}$. On the other hand, as a composition of $\mathcal{E}$-arrows, it is obviously in $\mathcal{E}$. So $\check{e}\hat{e}$ is iso. Choose the images as to get $\check{e}\hat{e} = id_S$. Hence

$$s \otimes \eta = \ddot{e}(\Gamma c^o \otimes \eta)\hat{e}. \tag{40}$$

On the other hand, the equation

$$h(\Gamma c^o \otimes \eta)\hat{e} = \eta \tilde{e} c = eqp \tag{41}$$

and the fact that (III) is a pullback imply that (V) must commute. Eq. (40) now becomes

$$s \otimes \eta = \ddot{e}\dot{e}p. \tag{42}$$

We claim that $\ddot{e}\dot{e} = q_0 q$. If this is true, we are done, since p is by definition a right inverse of $q_0 q$, so that (42) yields

$$s \otimes \eta = \ddot{e}\dot{e}p = q_0 qp = id. \tag{43}$$

The following calculation actually completes the proof:

$$\ddot{s}\ddot{e}\dot{e} \overset{(1)}{=} (C \times d)(\varGamma c^o \otimes s^o \otimes s)\dot{e}$$

$$\overset{(2)}{=} \langle cq_0 q, dq' \rangle$$

$$\overset{(3)}{=} \langle cq_0 q, dq_1 q \rangle$$

$$\overset{(4)}{=} \langle cq_0 q, dq_0 q \rangle$$

$$= \langle c, d \rangle q_0 q$$

$$= sq_0 q. \tag{44}$$

Step (1) is just the commutativity of (VII). At step (2), we are using the equation

$$(\varGamma c^o \otimes s^o \otimes s)\dot{e} = \langle cq_0 q, q' \rangle.$$

To derive this one, recall that $(s^o \otimes s)e = \langle cq_0, cq_1 \rangle$ holds by definition, and that (III) and (IV) are pullbacks. So $(\varGamma c^o \otimes s^o \otimes s)\dot{e}$ is a pullback of $\langle cq_0, cq_1 \rangle$ along $C \times c$. But diagram (38) shows that this pullback yields $\langle cq_0 q, q' \rangle$.

At step (3) of (44), we are using the assumption $\ker(c) \subseteq \ker(d)$, which implies

$$cq' = cq_1 q \quad \Rightarrow \quad dq' = dq_1 q.$$

The equation on the left-hand side is true by the definition of q and q' in (38). Finally, the equation $dq_0 = dq_1$, used on step (4), is true by the definition of q_0 and q_1 in (38).

The desired equation $\ddot{e}\dot{e} = q_0 q$ follows from (44), using the uniqueness of decomposition. Calculation (43) is thus justified, and lemma is proved. $\square$

6. Final remarks

If needed, more precise characterisations could be extracted from the preceding proofs. For instance, one could prove that the situation when G is just full (rather than full and faithful, as in Theorem 5.1) is characterised by the requirement that each $\mathscr{E}$-arrow is a *weak* coequalizer of its kernel pair. The same scheme of the proof could be followed, with inessential but tedious complications at points where we used the separation properties. Instead of applying Lemmas 4.9 and 5.2(b) to "externalize" the equality of arrows, one would have to manipulate the "internal equality" of arrows, expressed in the form $\mathrm{im}\langle f, f' \rangle$.

Notice further that the condition $\mathscr{E} \subseteq \mathsf{Epi}^*$ obviously implies $\mathscr{M} \supseteq \mathsf{Mon}$. If $\mathscr{C}$ is a regular category, with all strong epis regular, the last condition is even equivalent to $\mathscr{E} \subseteq \mathsf{Epi}^*$, and thus characterises the function comprehension. In general, though, this may not be the case.

Going back to remarks preceding Proposition 4.6, let us mention what would be the meaning of Theorem 5.1 in the context of "sheaf theory" relative to a factorisation. If the category $\mathscr{C}$ with a stable factorisation $(\mathscr{E}, \mathscr{M})$ is regarded as a "site", the objects of $\mathscr{C}$ correspond to "representable presheaves". If being Cauchy complete means being a "sheaf", then $\mathscr{C}$ is function comprehensive if and only if the "topology" $(\mathscr{E}, \mathscr{M})$ is "subcanonical": all "representables" must be "sheaves".

Understood in this way, Theorem 5.1 appears familiar [2, 6.7, Proposition 4]: an ordinary site is indeed subcanonical if and only if the covers are (jointly) regular epi. Corollary 4.7 also comes up as a generalisation of a well-known fact that representables are separated with respect to a topology if and only if the covers are (jointly) epi.

The ways in which geometric logic of sites can be extended in a genuinely categorical (non-posetal) setting of strong constructivism will be systematically explored in the sequel to this paper [21].

Acknowledgements

References

[1] M. Barr, P.A. Grillet and D.H. van Osdol, Exact categories and categories of sheaves, Lecture Notes in Mathematics, Vol. 236 (Springer, Berlin, 1971).

[2] M. Barr and C. Wells, Toposes, Triples and Theories, Grundlehren Matematischen Wissenschaften, Vol. 278 (Springer, Berlin, 1985).

[3] J. Bénabou, Introduction to bicategories, in: Reports of the Midwest Category Seminar I, Lecture Notes in Mathematics, Vol. 47 (Springer, Berlin, 1967) 1–77.

[4] R. Betti, A. Carboni, R. Street and R. Walters, Variation through enrichment, J. Pure Appl. Algebra 29 (1983) 109–127.

[5] A. Carboni, P.J. Freyd and A. Scedrov, A categorical approach to realizability and polymorphic types, in: Proc. 3rd A.C.M. Workshop on Mathematical Foundations of Programming Language Semantics, Lecture Notes in Computer Science (Springer, Berlin, 1988) 23–42.

[6] A. Carboni, S. Kasangian and R. Street, Bicategories of spans and relations, J. Pure Appl. Algebra 33 (1984) 259–267.

[7] A. Carboni, G.M. Kelly and R. Wood, A 2-categorical approach to change of base and geometric morphisms I, Cahiers Topologie Géom. Différentielle Catégoriques 32 (1991) 47–95.

[8] A. Carboni and R.F.C. Walters, Cartesian bicategories I, J. Pure Appl. Algebra 49 (1987) 11–32.

[9] L. Coppey, Algèbres des Décompositions et Précatégories, Thèse de Doctorat d'Etat; Diagrammes 3 (1980).

[10] P. Freyd and G.M. Kelly, Categories of continuous functors, I, J. Pure Appl. Algebra 2 (1972) 169–191; erratum in J. Pure Appl. Algebra 4 (1974) 121.

[11] P.J. Freyd and A. Scedrov, Categories, Allegories, North-Holland Mathematical Library, Vol. 39 (North-Holland, Amsterdam, 1990).

[12] G.M. Kelly, Basic concepts of the enriched category theory, London Mathematical Society Lecture Note Series, Vol. 64 (Cambridge University Press, Cambridge, 1982).

[13] G.M. Kelly, A note on relations relative to a factorization system, in: A. Carboni et al., eds., Category Theory, Como 1990, Lecture Notes in Mathematics, Vol. 1488 (Springer, Berlin, 1991) 249–262.

[14] A. Klein, Relations in categories, Illinois J. Math. 14 (1970) 536–550.

[15] M. Korostenski and W. Tholen, Factorisation systems as Eilenberg–Moore algebras, J. Pure Appl. Algebra 85 (1993) 57–72.
[16] F.W. Lawvere, Metric spaces, generalized logic, and closed categories, Rend. Sem. Mat. Fis. Milano 43 (1973) 135–166.
[17] G. Longo and E. Moggi, Constructive natural deduction and its "modest" interpretation, Math. Structures Comput. Sci. 1 (1991) 215–254.
[18] J. Meisen, On bicategories of relations and pullback spans, Comm. Algebra 1 (1974) 377–401.
[19] D. Pavlović, Constructions and predicates, in: Category Theory and Computer Science, Lecture Notes in Computer Science, Vol. 530 (1991) 173–196.
[20] D. Pavlović, On completeness and cocompleteness in and around small categories, Ann. Pure Appl. Logic (1995), to appear.
[21] D. Pavlović, Maps II: chasing proofs in categorical logic, submitted.
[22] G. Richter, Mal'cev conditions for categories, in: Categorical Topology, Proc. Conf. Toledo, Ohio 1983 (Heldeman, 1984) 453–469.
[23] R. Street, Cauchy characterisation of enriched categories, Rend. Sem. Mat. Fis. Milano 51 (1981) 217–233.
[24] R. Street, Conspectus of variable categories, J. Pure Appl. Algebra 21 (1981) 307–338.
[25] A.S. Troelstra and D. van Dalen, Constructivism in Mathematics. An Introduction, Studies in Logic and Foundations of Mathematics, Vols. 121, 123 (North-Holland, Amsterdam 1988).
[26] D.R. Verity, Internal categories and change of base, Thesis, University of Cambridge, 1992.
[27] R.F.C. Walters, Sheaves and Cauchy-complete categories, Cahiers Topologie Géom. Différentielle (1981) 283–286.
[28] R.F.C. Walters, Sheaves on sites as Cauchy-complete categories, J. Pure Appl. Algebra 24 (1982) 95–102.
[29] A.N. Whitehead and B. Russell, Principia Mathematica, Vol. I (Cambridge University Press, Cambridge, 1910).
[30] N. Yoneda, On Ext and exact sequences, J. Fac. Sci. Univ. Tokyo, 7 (1954) 193–227.