On weak higher dimensional categories

by Claudio Hermida[1,2] and Michael Makkai[3,4]
Department of Mathematics and Statistics
McGill University[5]

and

John Power[6,7]
Department of Computer Science
University of Edinburgh[8]

Abstract

Inspired by the concept of opetopic set introduced in a recent paper by John C. Baez and James Dolan, we give a modified notion called multitopic set. The name reflects the fact that, whereas the Baez/Dolan concept is based on operads, the one in this paper is based on multicategories. The concept of multicategory used here is a mild generalization of the same-named notion introduced by Joachim Lambek in 1969. Opetopic sets and multitopic sets are both intended as vehicles for concepts of weak higher dimensional category. Baez and Dolan define weak n-categories as $(n+1)$-dimensional opetopic sets satisfying certain properties. The version intended here, multitopic n-category, is similarly related to multitopic sets. Multitopic n-categories are not described in the present paper; they are to follow in a sequel. The present paper gives complete details of the definitions and basic properties of the concepts involved with multitopic sets. The category of multitopes, analogs of opetopes of Baez and Dolan, is presented in full, and it is shown that the category of multitopic sets is equivalent to the category of set-valued functors on the category of multitopes.

---

1 Supported by FCAR Quebec, Canada

2 e-mail: hermida@triples.math.mcgill.ca

3 Supported by FCAR Quebec, Canada, and NSERC Canada

4 e-mail: makkai@triples.math.mcgill.ca

5 address: 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6

6 Supported by EPSRC grant GR/J84205: Frameworks for programming language semantics and logic

7 e-mail: ajp@dcs.ed.ac.uk

8 address: King's Buildings, Edinburgh EH9 3JZ, Scotland

Introduction

In [B/D2] and [B/D3], John C. Baez and James Dolan have introduced a concept of weak higher dimensional category. The present paper is inspired by the work of Baez and Dolan. It is the first of two papers in which a modification of the Baez/Dolan proposal is offered and described in detail.

There are other proposals for related concepts; see [Ba], [T].

The problem of the identification of the weak higher dimensional categories has been recognized for some time; see e.g. [S1], [S2], [S3], [S4] . The motivations for the Baez/Dolan work were described in [B/D1]. In [M2], the second author of this paper describes another motivation, one that relates higher dimensional categories to the foundations of mathematics. In [M2], a program for a new type-theoretical foundation, termed *structuralist*, is described in which there is a hierarchy of totalities of higher and higher dimensions, starting with sets. In this framework, sets are taken to be totalities with an equality predicate. However, no equality is assumed between elements of different sets, and, essentially as a consequence, no equality of sets is contemplated. Because of this, sets do not form a set, or even a set-like totality like a class. Instead, sets form a *category*, the category of sets; and the role of equality as principle of identity is taken over by *isomorphism*, a concept derived from the structure of category. When we say that equality of objects is not part of the structure of the category, we have in mind a notion of category that is not the same as the one we deal with on the basis of the standard set-theoretical foundation. The negative statement of the denial of equality can be given objective content only by specifying a suitably constrained language to be adopted as the formal language of the structuralist foundation. The work [M3] proposes First Order Logic with Dependent Sorts (FOLDS) as the basis for such a language.

Classically, categories form 2-categories; the latter concept can already be found in [M L]. The structuralist foundation involves the program of revising category-theoretical concepts in which *equality of objects* of a category is used by replacing that equality by *specified isomorphisms of objects*. As a matter of fact, it has been widely accepted among category theorists that equality of objects should be avoided; the tendency to replace equality of objects by isomorphism is a common one in category theory. Jean Benabou's notion [Be] of bicategory is an instance of this tendency. In the case of a 2-category, the 1-arrows from a fixed 0-cell to another 0-cell form a(n ordinary) category. Applying the "isomorphisms-for-equality" treatment to the part of the definition of 2-category which explicitly refers to equality of 1-arrows (e.g., the the associative law of composition of 1-arrows) results in the concept of bicategory. We do not simply require the existence of certain isomorphism-2-arrows, but introduce specified ones (*coherence isomorphisms*), and we attach them to the structure. Furthermore, certain natural *coherence conditions* are imposed on the coherence isomorphisms (the Mac Lane pentagon is an example; see [M L], p. 158, formulated for monoidal categories, that is, bicategories with a single 0-cell). It should be emphasized that the concept of bicategory was motivated in the first place by more mathematical considerations than the ones connected to the structuralist foundation. Bicategories have turned out to be extremely useful, and a great deal more flexible than 2-categories.

The paper [M1] deals with a more elementary instance of replacing equality of objects by isomorphism; the notion of (*saturated*) *anafunctor* is introduced, in which the value-object of a(n ana)functor at any given argument-object is determined (strictly) up to isomorphism. Anafunctors are "mathematically equivalent" to functors, but only at the cost of an application of the Axiom of Choice. The replacement of the composition-functors in the definition of a bicategory by anafunctors results in *anabicategories*, which are held, in [M1] and [M3], to be the right concept for totalities of categories, at least from the point of view of the structuralist foundation. Saturated anabicategories are equivalent to bicategories, again via Choice. Saturated anabicategories are equivalent in a canonical manner, without the use of Choice, to the Baez/Dolan weak 2-categories, and the *multitopic 2-categories* that the sequel to this paper will describe.

Besides being the first answer to a long-standing problem, the Baez/Dolan proposal has several remarkable features. The main one is a complete elimination of explicit lists of coherence structure and conditions. This feature is already fully apparent when one looks at the case $n=2$, a Baez/Dolan weak 2-category. It is related to a bicategory as a fibration is related to a pseudo-functor [G]. The coherence isomorphisms and conditions present in the definition of pseudo-functor are, in the corresponding fibration, eliminated in favor of a structure defined by a universal property, that of Cartesian arrows. Such an elimination of coherence takes place in a Baez/Dolan (B/D) weak n-category as well, for all n. For $n=2$, the composition of 1-cells is defined by a universal property, and accordingly, its result, the composite, is not a uniquely defined thing, but one which is determined up to isomorphism; recall that the last feature is present also in anabicategories. There are no coherence isomorphisms (such as the associativity isomorphism), no coherence conditions (such as Mac Lane's pentagon). The way this is achieved is similar to the case of fibrations inasmuch one adds *more entities* to the original (pseudo-functor, respectively, bicategory) to get the new structure (fibration, respectively, B/D weak 2-category). In the case of a fibration, the arrows between objects in *different fibers* of the total category are new with respect to the data of the pseudo-functor. In the case of the B/D weak 2-category, we have 2-cells whose domain is a composable string of 1-cells, of arbitrary finite lengths in fact, instead of just a single 1-cell. These "multi-arrows" are new entities with respect to the corresponding (ana)bicategory, and they are taken away when one passes from the B/D 2-category to the corresponding bicategory; of course, before being taken away, they are used to define the data for the bicategory.

Multitopic higher dimensional categories, as we will call the objects that we intend to introduce, will share the above general aspects of the Baez/Dolan weak higher dimensional categories.

Although the proposal to be explained here was directly inspired by the B/D proposal, its exposition will not make this fact clear. In fact, at the present time, we do not see the *precise* equivalence of the two proposals. A conspicuous difference is the absence here, and the presence in [B/D3], of actions of permutation groups. It is possible to introduce an "up to isomorphism" variant of the basic notion of *multicategory* used in this paper (more on this will follow soon); this higher-dimensional variant of "multicategory" (in which, for instance, isomorphisms between arrows in a multicategory would appear) seems more directly related to [B/D3] than what is found here.

On the other hand, even if there are close ties between the proposal of [B/D3] and that of this paper, their mathematical forms are entirely different. The [B/D3] concept is abstract and conceptual; ours here is concrete and geometric.

The above description concerning the 2-dimensional case already indicates the starting point of the approach of the present paper. We define a concept of *k-dimensional cell*, or k-cell, for all $k=0, 1, 2, 3, \ldots$, in an inductive way. For $k>0$, a k-cell has a *domain* and a *codomain*; the codomain is an $(k-1)$-cell, but the domain is a *pasting diagram* of $(k-1)$-cells. The inductive character of the definition lies in the definition of pasting diagrams. These are related to what go under the same name in the literature (see e.g. [P1], [P2]), but are greatly simplified by the fact that the codomains of cells is always a single cell. Despite the fact that the Baez/Dolan concept is not explained in terms of cells whose domains are pasting diagrams of lower cells, the crucial restriction to single-cell codomains also originates in [B/D3].

The present paper's approach is consciously geometrical. At the same time, great care is taken to express everything in algebraic terms. The main algebraic tool we use is the concept of *multicategory*, a modified form of the same-named notion introduced by Joachim Lambek in 1969; see [L1] and [L2]. It is worth remarking that one of the first uses Lambek made of multicategories was to proof-theory, for an algebraic formulation of Gentzen's proof-system for intuitionistic propositional logic.

Lambek's concept is closely related to monoidal categories. A multicategory may be said to be mathematically equivalent to a strictly associative monoidal category in which the monoid of the objects under the tensor-product is a free monoid (on the objects of the multicategory as generators). In a multicategory, we have objects and arrows; each arrow has a *source* which is a finite tuple of objects, and a *target*, a single object. The main distinguishing point about the notion of multicategory is that it is phrased in terms of a *composition*, a ternary operation, two of whose arguments are arrows, the third being the *place* where the target of one the arrows is to fit into the source of the other; of course, the result of composition is an arrow. From the point of view of the arrows, we have a system of binary compositions. Two of the laws are an associative law and a commutative law of composition as in the ordinary binary case, but suitably decorated with places.

We generalize Lambek's notion in two steps, one major and a minor. The major step is to make explicit and generalize the *amalgamation* that takes place in composition. When two arrows are composed, the source of the composite results by *amalgamating* the sources of the original arrows in a certain way. In the Lambek case, this amalgamation is the standard one of *inserting* the source of one of the arrows into the source of the other at the given place. In the generalized concept, the amalgamation is made arbitrary, subject to certain laws. It should be noted that for the precise statement of the laws of multicategory, one has to make an explicit reference to this amalgamation already in Lambek's case. Lambek does not make the amalgamation explicit, but there is an acknowledgement of the resulting incompleteness of the formulation in lines 12 and 11 from the bottom on p.222 of [L2].

It does not seem possible to relate the general concept of multicategory with that of monoidal category as closely as in the case of the Lambek multicategory. The new concept is "essentially geometric"; it has geometric instances (see below), but it does not seem to have "semantical" instances, apart from the standard Lambek case, which does have many "semantical" examples.

On the other hand, the generalized concept is a *mild* generalization. This is witnessed by the fact that the free multicategory in the Lambek sense on a set of objects and generating arrows is also the free multicategory on the same generating data in the generalized sense.

The main point of the new notion is that multicategories with non-standard amalgamation appear in nature. The *multicategory of function-replacement* derived from a free multicategory plays a central role in our work; it is needed for the definition of the domain, a $(k\text{-}1)$-pasting diagram, of a k-pasting diagram.

The first section of the paper is an extended informal introduction. After the next three sections on multicategories, on morphisms of multicategories, and free multicategories, respectively, section 5 gives the construction of the multicategory of function-replacement.

Section 6 uses the preceding machinery to put together the definition of *multitopic set*, the main notion arrived at in this paper. A *multitopic n-category,* the main object we want, will, in the sequel to this paper, be defined as an $(n+1)$-dimensional multitopic set *with additional properties*; no new data are needed. Baez and Dolan used *opetopic sets* instead; the name of their notion is derived from *operads*, the abstract algebraic concept at the basis of their work. Let us note that by a multitopic set. we mean what also could be called an ω-dimensional multitopic set; an n-dimensional one is in fact a truncated one.

Section 7 identifies a particular category, the category $\mathtt{Multitope}$ of *multitopes*, and identifies multitopic sets defined in the section 6 as set-valued functors on the category of multitopes. More precisely, we prove that $\mathtt{MSet}$, the naturally defined category of multitopic sets, is equivalent to the category of functors from $\mathtt{Multitope}$ to $\mathtt{Set}$. $\mathtt{Multitope}$ is related to the terminal object $\mathcal{T}$ of $\mathtt{MSet}$. The objects of $\mathtt{Multitope}$ are identical to the pasting diagrams of the multitopic set $\mathcal{T}$; on the other hand, the identification of the arrows of $\mathtt{Multitope}$ takes additional work. It should be emphasized that all the complexity involved in the definition of multitopic sets in general is already present in the definition of the terminal one , $\mathcal{T}$, despite the fact that this object is absolutely uniquely given.

The category $\mathtt{Multitope}$ and, for any $n=0,1,2,\ldots,$ its truncation $\mathtt{Multitope}[n]$ to include k-pasting diagrams of $\mathcal{T}$ for $k=0,\ldots,n$, are fundamental from the point of view taken in this paper. In [M3], a concept of L-equivalence, for variable signatures L for FOLDS, is introduced, and it is shown that, when used in conjunction with the ana-concepts of

[M1], L-equivalence becomes identified with categorical equivalence in many cases, for instance in the case of biequivalence for bicategories. `Multitope[`n`]` is the FOLDS-signature for multitopic n-sets. In view of the fact that multitopic n-categories are multitopic n-sets with additional properties *formulated in FOLDS*, `Multitope[`n`]` is the FOLDS-signature also for multitopic n-categories. Thus, now, even before we have given the further details of the definition of multitopic n-category, we have a notion of equivalence of multitopic n-categories. In Baez's and Dolan's work, we also find a notion of equivalence for weak n-categories. The comparison awaits further work.

The Appendix contains some details of proofs for sections 4 and 5.

It should be emphasized that this paper is only a part, in fact, just a beginning, of the work of establishing the concept of weak higher dimensional category. Even when we have the full definition (which is given by [B/D3], and promised, in a modified form, to be given by the sequel to this paper), the accompanying structures are still to be provided.

The second author thanks Marek Zawadowski for extended discussions, and valuable suggestions. He thanks members of the Department of Mathematics of the University of New Hampshire, especially Donovan van Osdol and Steven Shnider, and the BRICS group of the Department of Computer Science of Aarhus University, in particular Carsten Butz, Prakash Panangaden and Glynn Winskel, for their kind hospitality and their interest in the work presented here.

1. An informal description

1.1. n-graphs and multitopic sets.

In the classical, strict, concept of higher-dimensional category (HDC), an HDC $\mathbf{A}$ consists of k-*cells* in each of several *dimensions* k, where k ranges over a set $\{0,\ ...,\ n\}$ (n-*category*), or over all natural numbers (ω-*category*). Let us denote the class of all k-cells of $\mathbf{A}$ by C_k. For $k>0$, each k-cell a is "based on" two $(k-1)$-cells, the *domain* $\mathrm{d}a$ and *codomain* $\mathrm{c}a$ of a; when $b=\mathrm{d}a$, $c=\mathrm{c}a$, we write $a:b\longrightarrow c$; we have the assignments $\mathrm{d}_k=\mathrm{d}:C_k\longrightarrow C_{k-1}$, $\mathrm{c}_k=\mathrm{c}:C_k\longrightarrow C_{k-1}$ as part of the structure of the HDC $\mathbf{A}$. The part of the structure of $\mathbf{A}$ so far described is an n-*graph* in the case of "n-category", ω-*graph* in the case of "ω-category"; the data for an n-graph can be summarized in the diagram.

$$C_0 \underset{\mathrm{c}_1}{\overset{\mathrm{d}_1}{\leftleftarrows}} C_1 \underset{\mathrm{c}_2}{\overset{\mathrm{d}_2}{\leftleftarrows}} C_2 \ \ldots\ C_{n-1} \underset{\mathrm{c}_n}{\overset{\mathrm{d}_n}{\leftleftarrows}} C_n \qquad (1)$$

A feature of n-graphs, is *globularity*: for any $a\in C_k$, $k\geq 1$, $b=\mathrm{d}a$ and $c=\mathrm{c}a$ must be *parallel*, that is, either $k-1=0$, or else $\mathrm{d}b=\mathrm{d}c$, $\mathrm{c}b=\mathrm{c}c$:

$$e \underset{c}{\overset{b}{\rightrightarrows}} f$$

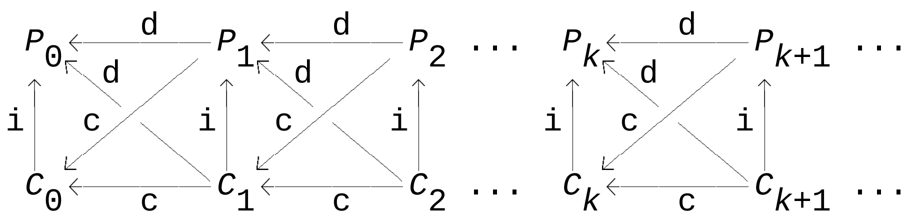
where $e=\mathrm{d}b=\mathrm{d}c$, $f=\mathrm{c}b=\mathrm{c}c$. Put another way,

$$\mathrm{dd}=\mathrm{dc}\ ,\ \mathrm{dc}=\mathrm{cc}\ , \qquad (2)$$

where d and c ambiguously denote any of the domain, respectively codomain maps $\mathrm{d}_k:C_k\longrightarrow C_{k-1}$, $\mathrm{c}_k:C_k\longrightarrow C_{k-1}$, with the restriction that the composites intended should be meaningful. n-graphs are defined by having data as in (1), the domain/codomain assignments satisfying globularity (2). An n-category (in the usual sense) has several additional operations of *composition*; see, e.g., [S2].

The notion of HDC of the present paper will retain the above general features, except for one thing: the *domain* of a cell is no longer a cell itself; rather, it is a *pasting diagram* (see below) of cells. Note the asymmetry: we only mentioned "domain", not "codomain"; codomains will remain single cells.

The role of n-graphs is taken up by (n-dimensional) *multitopic sets*; below, there will be an explanation for the choice of the name of the concept. The data for a multitopic set are summarized in the diagram



where C_k is the set of k-cells, P_k the set of k-dimensional pasting diagrams (k-pd's for short), each $\mathtt{i}$ is an inclusion map, and the $\mathtt{d}$ and $\mathtt{c}$ are domain and codomain maps. All meaningful instances of the globularity condition (2) will hold.

In the next subsection, we will explain the notion of pasting diagram; here, we note that they are not independent data governed by relations and properties; rather, they are defined explicitly in terms of cells. The most important point to keep in mind that there is an essential recursive character to the notion of multitopic set; this is because the notion of $(k+1)$-cell cannot be explained before we know what k-pd's are, and k-pd's, in turn, are defined in terms of k-cells.

The higher dimensional categories, *multitopic n-categories,* whose definition is the eventual goal of the present paper, are based on multitopic sets, just as n-categories are based on n-graphs. As a compensation for the increased complexity in multitopic sets in comparison to higher dimensional graphs, we have the fundamental fact that a multitopic n-category is an $(n+1)$-dimensional multitopic set *with additional properties* only; no additional data are required. (Note, however, the placing of the prefixes n and $n+1$ in this description.)

1.2. Pasting diagrams

The expression of "pasting diagram" refers to the idea of a *composable diagram*, one which, *if* a concept of composition of cells were available, would result in a single cell after all the meaningful compositions denoted in the diagram are performed. This is an approximate expression of an intuitive idea. It turns out that composability in higher dimensions is a difficult concept, and despite several contributions (e.g., [S1], [J], [P1], [P2]) it is not yet completely clarified. It is to be emphasized that the concept "composable diagram" is a geometric one in that it does not involve composition of cells in the algebraic sense. Composability is the *geometric precondition* of (iterated) composition.

An important point for this paper, inspired by the Baez/Dolan work, is the restriction of cells to the form $a:\alpha\to b$, where α is a pasting diagram (pd), but b is a single cell. The first consequence is that the notion of pd itself becomes simple, and abstractly manageable, in comparison with the (potential) more comprehensive concept that would allow both the domain and the codomain to be arbitrary pd's. The "Baez/Dolan restriction" (as we may call the above-mentioned restriction) is not a *necessary* feature of the intended notion of HDC; it is, rather, a simplifying idea; the thus simplified notion of pd turns out to be *sufficient* for carrying the intended structure of an HDC.

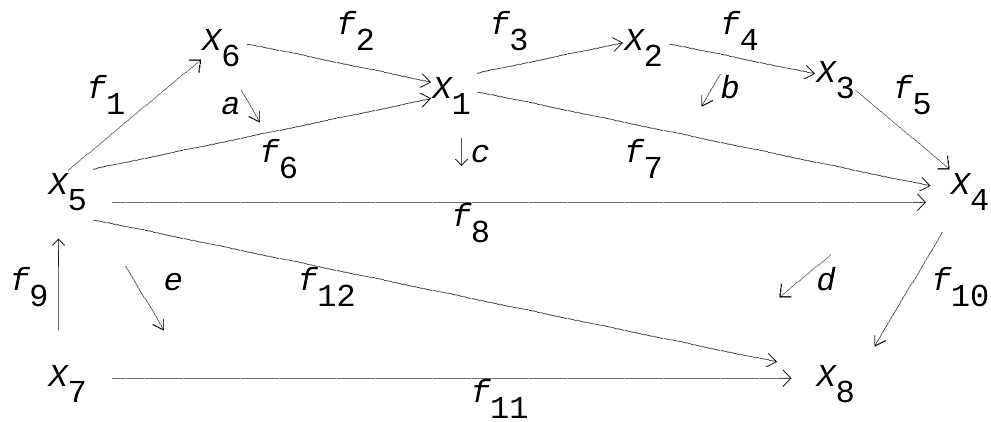
A 0-dimensional pd (or 0-pd) is just a 0-cell (object). A 1-dimensional pd (or 1-pd) is a composable string of 1-cells:

$$X_1 \xrightarrow{f_1} X_2 \xrightarrow{f_2} X_3 \longrightarrow \dots \xrightarrow{f_n} X_{n+1} \qquad (3)$$

where the X_i's are 0-cells, the f_i's are 1-cells. $n=0$ is allowed, in which case there are no arrows; but in this case, there is still an object, X_1, and we have the empty string of arrows starting and ending in X_1.

A 2-pd consists of 0-, 1- and 2-cells; each 2-cell in it is *from* a 1-pd, a string of one cells (possibly empty), *to* a single 1-cell; and the whole thing is *composable*. Here is an example of a 2-pd, which we denote by the single letter

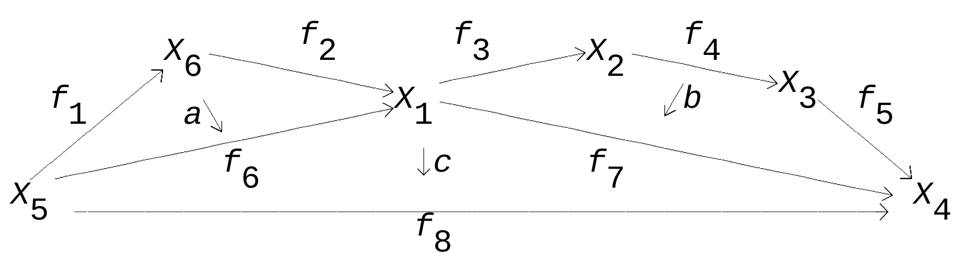
γ :



γ consists of the 0-cells X_i, $1 \leq i \leq 8$, 1-cells f_j, $1 \leq j \leq 11$ (numbered in no particular order), and the 2-cells a, b, c, d, e. The figure is supposed to make clear the domain/codomain relations among the cells and 1-pd's involved. Notice the constraint that each 2-cell targets a single 1-cell; in a 2-pd in a more general sense, both domains and codomains could be general 1-pd's. Perhaps it is superfluous to say that the 2-pd γ is the totality of the items listed; it is not the result of some kind of composition performed on those items. Of course, the relative position of its component 2-cells is part of the defining data of the 2-pd.

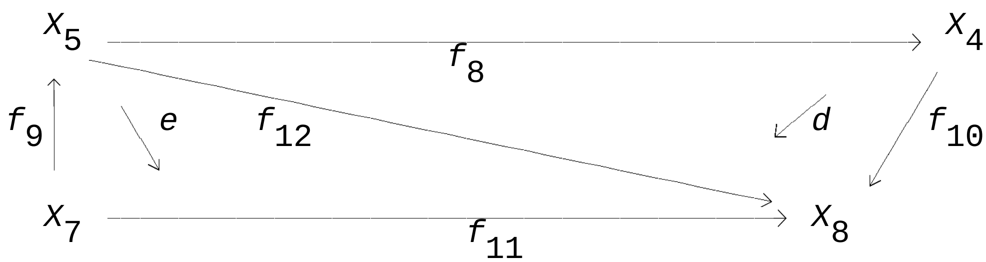
There are features of 2-pd's that become important elements of the general concept of a k-pd. The above 2-pd can be regarded as obtained by *composition*, in a new sense of "formal" composition, which applies to pd's rather than cells. This composition may also be called *grafting*. For instance, γ is obtained by grafting from the following two pd's α and β :

α :



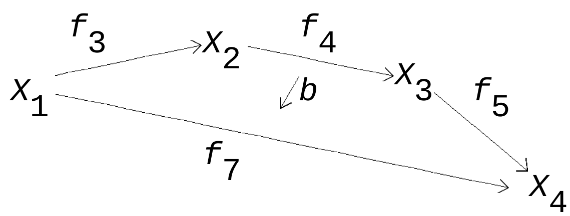
and

β :



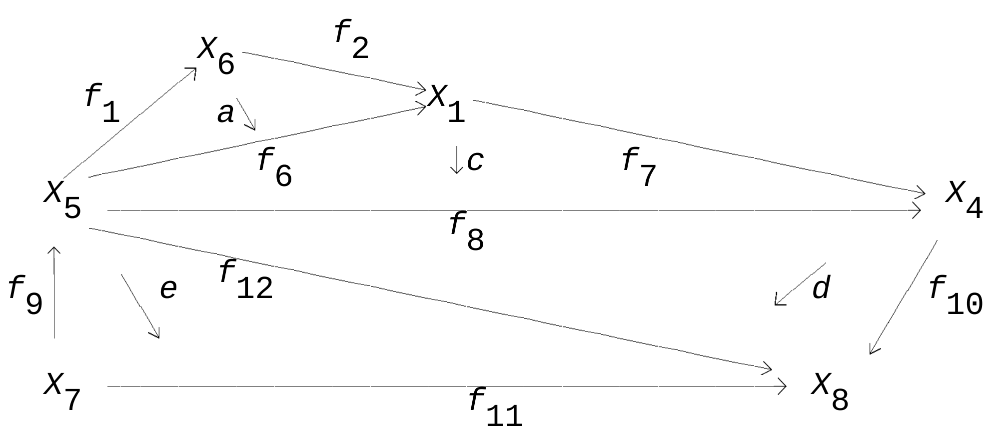
More precisely, we graft α into β *at* f_8 , and obtain the original γ . Of course, the same pd γ can also be obtained in several other ways as the result of grafting, e.g. by grafting δ into ε , where

δ :



(which is a pd consisting of a single 2-cell), and

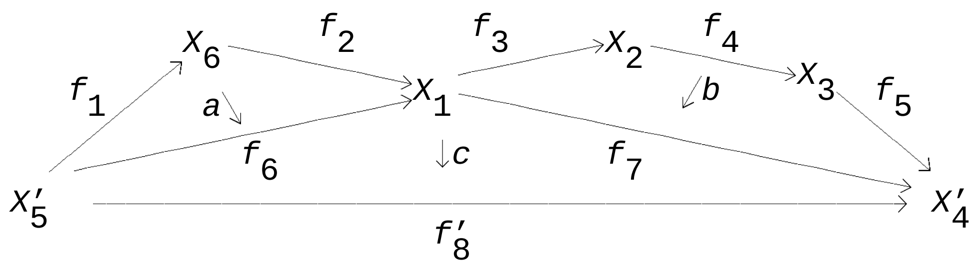
ε :



The grafting composition is a binary operation as far as the number of arguments that are pd's is concerned; but it also has a third argument, the *place* at which the grafting takes place. The two grafting compositions displayed are denoted as $\beta\circ_{f_8}\alpha=\gamma$, and $\varepsilon\circ_{f_7}\delta=\gamma$; read e.g. the first as α composed (grafted) into β at f_8 is γ .

Given β as above, and, say,

α' :



where the primed items may or may not be equal to the corresponding non-primed items in β , the composite $\beta\circ_{f_8}\alpha'$ is meaningful if and only if $f'_8=f_8$, and as a consequence, $X'_5=X_5$, $X'_4=X_4$. f'_8 is distinguished as the *target*-1-*cell* of α ; $\mathrm{t}(\alpha)\overset{=}{\mathrm{def}}f'_8$. For the given β , and an undetermined α' , the condition for $\beta\circ_{f_8}\alpha'$ to be well-defined is that $\mathrm{t}(\alpha')=f_8$.

It is perfectly possible that several items in the above pd's that are now denoted by different symbols are actually the same. For instance, it is possible that all the 0-cells are the same, and all the 1-cells are the same. If so, the 2-cells a, c, d, e could all be the same, although b cannot be the same as those since its shape is different: its domain pd is a length-3 1-pd, whereas the domains of the others are of length 2. Assuming, e.g., that all the said coincidences actually take place, the subscript f_8 in $\beta\circ_{f_8}\alpha$ cannot refer to the f_8 simply as a 1-cell; it has to refer to *the place of* f_8 ; we have $f_8=f_9=f_{10}$, and we can just as well compose α into β at the *two other places*, now denoted f_9 and f_{10} , and the results of these compositions are all very different, distinguished already by their shapes. This tells us that in the concept of pd there has to be an essential element that we may call *place*; in a 2-pd, there are *places* for 1-cells, each of which carries the "occurrence" of a particular 1-cell.

Note that it does not make sense to compose anything into α at f_{11} , or into γ at places other than $f_9, f_1, f_2, f_3, f_4, f_5, f_{10}$; the result would not be a "composable diagram".

The listed places of γ, the ones at which it is legitimate to compose something into γ, are the *source* places of γ; they are, together with the *target* place f_{11}, "outer places"; the "inner places" are the rest, f_6, f_7, f_8, f_{12}. $\mathtt{s}(\gamma)$ denotes the tuple $\langle f_{10}, f_5, f_4, f_3, f_2, f_1, f_9\rangle$, and it is called the *source* of γ; the reason for the order will be explained below. In the example, $\mathtt{s}(\gamma)$ is a function on the set $[1,6]=\{1,2,3,4,5,6\}$, and its values are $\mathtt{s}(\gamma)(1)=f_{10}$, *etc.* The source places themselves of γ are identified with the natural numbers $1,2,3,4,5,6$; the place 1 carries an occurrence of f_{10}, the place 2 one of f_5, etc. Writing $|\mathtt{s}(\gamma)|$ for the domain of the function $\mathtt{s}(\gamma)$, the source-places of γ are the elements of $|\mathtt{s}(\gamma)|$.

Similarly, $\mathtt{s}(\beta)=\langle f_{10}, f_8, f_1\rangle$. Since the place of f_8 in β is 2, we will write $\beta\circ_2\alpha$ for $\beta\circ_{f_8}\alpha$; we have $\gamma=\beta\circ_2\alpha$.

For general 2-pd's α and β,

(3) $\beta\circ_p\alpha$ makes sense if and only if $p\in|\mathtt{s}(\beta)|$ and $\mathtt{s}(\beta)(p)=\mathtt{t}(\alpha)$.

We have identified what we take to be the essential structure on pd's: the *placed composition* $\alpha\circ_p\beta$, a ternary operation as explained above.

1.3. Multicategories

The abstract concept of structure for the operation of placed composition is called *multicategory*. Multicategories were introduced by J. Lambek in 1969 [L1]; one of the uses he made of them was to define a multicategory of proofs in the Gentzen formal system for intuitionistic logic, where the placed composition corresponds to the Cut-rule. A *Lambek multicategory* $\boldsymbol{C}$ has a set $O=\mathrm{O}(\boldsymbol{C})$ of objects, and a set $A=\mathrm{A}(\boldsymbol{C})$ of arrows; each arrow α has a source $\mathtt{s}(\alpha)$ which is a finite tuple of objects, and a target $\mathtt{t}(\alpha)$ which is a single object ; when $\mathtt{s}(\alpha)=\vec{X}$, $\mathtt{t}(\alpha)=Y$, we write $\alpha:\vec{X}\longrightarrow Y$; $\boldsymbol{C}$ has, for each object X, an identity arrow $1_X:\langle X\rangle\to X$; and $\boldsymbol{C}$ has a placed composition as in (3) above. These data are to satisfy certain laws, the first of which regulates the source and the target of a composite, with the remaining laws being two identity laws, an associativity law, and a commutativity law. The definition will be given in section 2; the reader will notice that the definition in section 2 is, initially, something more general and more complicated than the one indicated here; later in that section, however, it is pointed out what exactly the Lambek concept is as a special case. Later in this introduction we will turn to the reasons why we need the more general concept of multicategory.

Thus, the 2-pd's (in a given HDC $\boldsymbol{A}$) form a Lambek multicategory (the 1-pd's also do, in fact, they form an ordinary category). More is true: the 2-pd's form a *free* multicategory, with objects the 1-cells, and generating arrows the 2-cells. Hence, all 2-pd's are generated by the 2-cells by using the operation of placed composition. This should be seen as an intuitively natural fact about pasting (composable) diagrams. (Let us remind ourselves that here we are in the business of *defining* what pasting diagrams are; the definition is constrained by intuitive ideas, which we are trying to make explicit.) Freeness is meant here in the sense of a strict universal property; it will be crucial later that the free Lambek multicategory maintains its universal property in the larger context of all (generalized) multicategories in the sense of Section 2.

For precise definitions concerning morphisms of multicategories, and free multicategories, see sections 3 and 4. Here we only give a brief idea.

Let O be a set of *objects*, L a set of *arrows*, with each $f\in L$ equipped with a source $\mathrm{s}(f)\in O^{*}$, and a target $\mathrm{t}(f)\in O$; data as described define a *language* $\mathcal{L}$. The terminology is natural, since $\mathcal{L}$ is exactly what is usually called a language (signature) for multi-sorted algebras; the elements of O are the sorts; the elements of L are the sorted operation symbols. The free multicategory, $\boldsymbol{C}=\mathcal{F}(\mathcal{L})$, on $\mathcal{L}$ is defined by the conditions that $\mathrm{O}(\boldsymbol{C})=O$, $L\subset\mathrm{A}(\boldsymbol{C})$, and any "interpretation" (a rather obvious notion) $\mathcal{L}\longrightarrow\boldsymbol{D}$ to any multicategory $\boldsymbol{D}$ can be uniquely extended to a morphism $\boldsymbol{C}\longrightarrow\boldsymbol{D}$. It turns out that the concrete description of $\mathcal{F}(\mathcal{L})$ is very simple. Its arrows are the *terms*, in the sense used in describing the syntax of first order logic, built up from sorted variables and the operation symbols of L , with the further simplification that we use only a single variable for each sort X , which variable therefore may just as well be identified with X itself.

Thus, we now have a *term-representation* of 2-pd's. Turning to the examples above, we have the following:

$$\gamma : e(d(f_{10}, c(b(f_5, f_4, f_3), a(f_2, f_1))), f_9)\ ,$$
$$\alpha : \quad c(b(f_5, f_4, f_3), a(f_2, f_1))\ ,$$
$$\beta : e(d(f_{10}, f_8), f_9)\ ,$$
$$\delta : \quad b(f_5, f_4, f_3)$$
$$\varepsilon : e(d(f_{10}, c(f_7, a(f_2, f_1))), f_9)$$

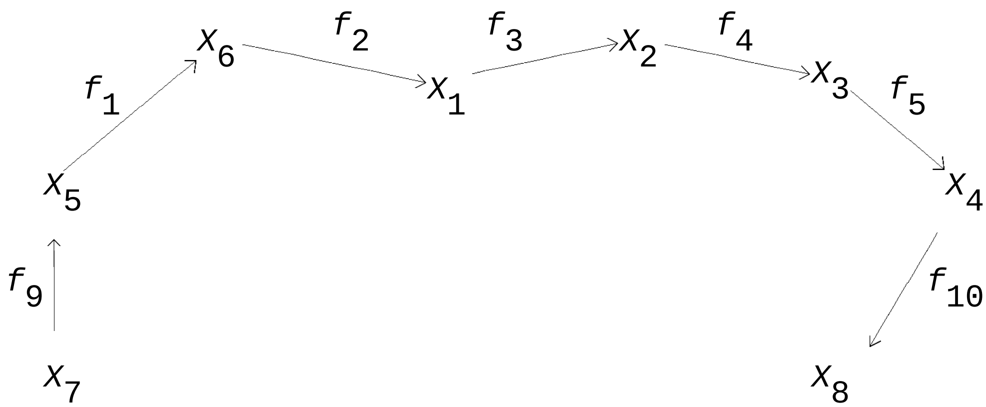
To understand these, consider the following. Any expression $x(y, z, \ldots)$ stands for a repeated composition; $x(y, z, \ldots) = \ldots(x \circ_1 y) \circ_{\bar{2}} z \ldots$; $\bar{2}$ is the place in $x \circ_1 y$ that "corresponds to" the place 2 in x. Each f_i stands for 1_{f_i}, the identity arrow $\langle f_i \rangle \xrightarrow{1_{f_i}} f_i$. Since $a(f_2, f_1)$ is a with identities composed into a, $a(f_2, f_1)$ equals a itself; we could write a in place of $a(f_2, f_1)$ above, except that in that case we would have not used the normal form which is intended by the term-representation. For $t_1 = b(f_5, f_4, f_3)$, $t_2 = a(f_2, f_1)$, the term $\alpha = c(t_1, t_2)$ is, really, the multicategory composite $(c \circ_1 t_1) \circ_{\bar{2}} t_2 = (c \circ_2 t_2) \circ_{\bar{1}} t_1$; the equality is the commutative law; $\bar{2} = 4$, $\bar{1} = 1$ (why?). We also see that placed composition corresponds to *substitution:* the fact that $\beta \circ_2 \alpha = \gamma$ is reflected in the fact that γ is the result of substituting α for f_8 in β.

The term-representation is a simple linear way of writing down 2-pd's; in fact, it will also be available for k-pd's for any k. However, note that in this notation, several elements that are clear in the geometric picture are suppressed. All 0-cells, and all but the input 1-cells are suppressed, although they can be be recovered by the information concerning the *targets* of the 2-cells involved.

Let us note that the 1-pd's also admit a term representation, since they also form a multicategory, which in fact is an *ordinary category*, since only unary arrows appear. The 1-pd in (1) is represented by the term $f_n(\ldots(f_2(f_1(X_1)))\ldots)$. The source-assignment to 2-cells above follows the left-to-right order in the term-representation; this is the reason why we used the "reverse" order for those sources above.

Let us move from dimension 2 to dimension 3 .

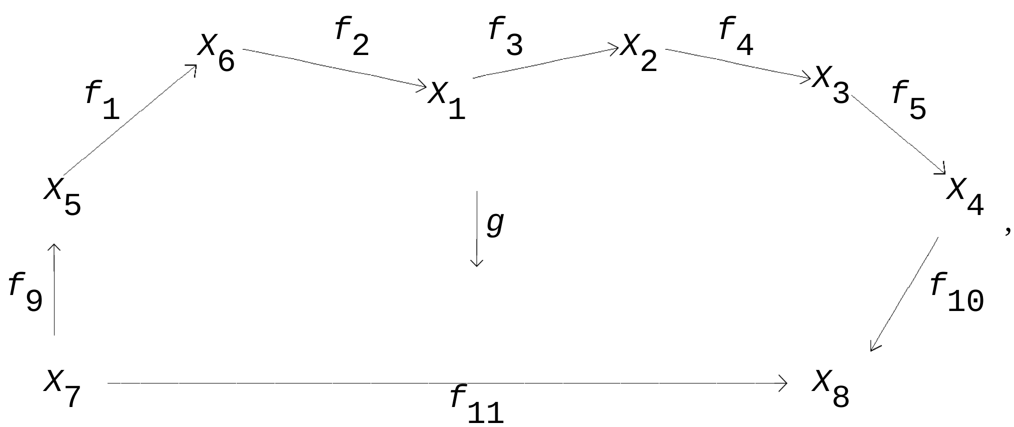
A 3-cell u is to have a 2-pd $\mathrm{d}u (= \mathrm{d}_3 u)$ as domain, and a 2-cell $\mathrm{c}u$ as codomain. Globularity requires that we should have $\mathrm{dd}u = \mathrm{dc}u$, $\mathrm{cd}u = \mathrm{cc}u$; however, we have not defined $\mathrm{d}\alpha$, $\mathrm{c}\alpha$ for 2-pd's α as yet, and we need them now for $\alpha=\mathrm{d}u$. The definition of the domain of a pd is a major issue in our enterprise; the codomain is easy. In the case of the example γ above, $\mathrm{d}\gamma$ is



that is,

$$X_7 \xrightarrow{f_9} X_5 \xrightarrow{f_1} X_6 \xrightarrow{f_2} X_1 \xrightarrow{f_3} X_2 \xrightarrow{f_4} X_3 \xrightarrow{f_5} X_4 \xrightarrow{f_{10}} X_8 \,;$$

this is the "upper part of the contour (boundary) of γ". $\mathrm{c}\gamma$ is the 1-cell $X_7 \xrightarrow{f_{11}} X_8$, the "lower" part of the contour of γ, the cell that "closes off" $\mathrm{d}\gamma$. Thus, a 3-cell u for which $\mathrm{d}u=\gamma$, with γ as in the example, looks necessarily like $u:\gamma\to g$, where g is a 2-cell of the following "shape":



which means that $\mathrm{d}g=\mathrm{d}\gamma$, $\mathrm{c}g=\mathrm{c}\gamma$. One cannot faithfully represent u in a 2-dimensional drawing; but u has a good 3-dimensional geometric representation; in this, the 2-pd γ is placed in the plane of the table, say; the 2-cell g is spanned out in a curved surface above the table, with its contour joining the contour of γ according to the the identifications inherent in the facts $\mathrm{d}g=\mathrm{d}\gamma$, $\mathrm{c}g=\mathrm{c}\gamma$; the 3-cell u "fills" the space between γ and g, "in the direction" from γ to g.

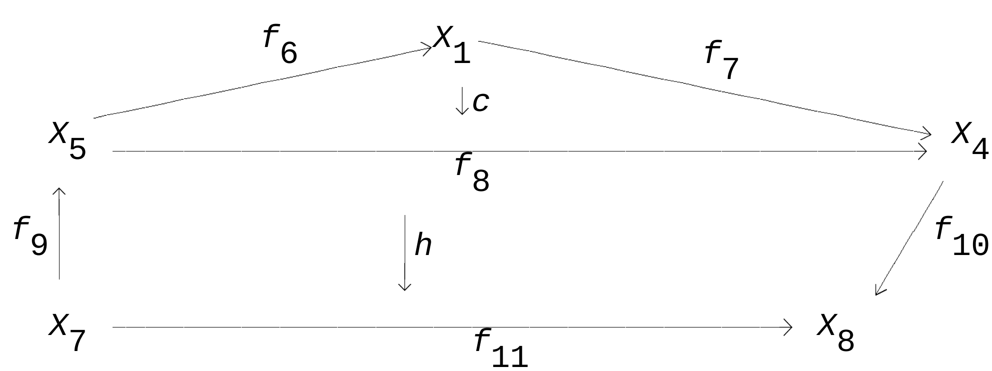
3-pd's will be construed as arrows in the free multicategory on the language whose objects (sorts) are the 2-cells, operation-symbols the 3-cells, and in which the sorting of the latter is given as follows. Every 3-cell u comes with $\mathrm{d}u$, a 2-pd; regard $\mathrm{d}u$ in the term-representation; look at all the operation-symbol occurrences in $\mathrm{d}u$, which are 2-cells; define $\mathrm{s}(u)$ to be the left-to-right tuple $\langle \mathrm{d}u \rangle$ of those occurrences; $\mathrm{s}(u)\in C_2^*$ as it should be. $\mathrm{t}(u)$ is defined to be $\mathrm{c}(u)$.

For instance, for $u:\gamma\to g$ considered above, $\mathrm{s}(u)=\langle e, d, c, b, a\rangle$.

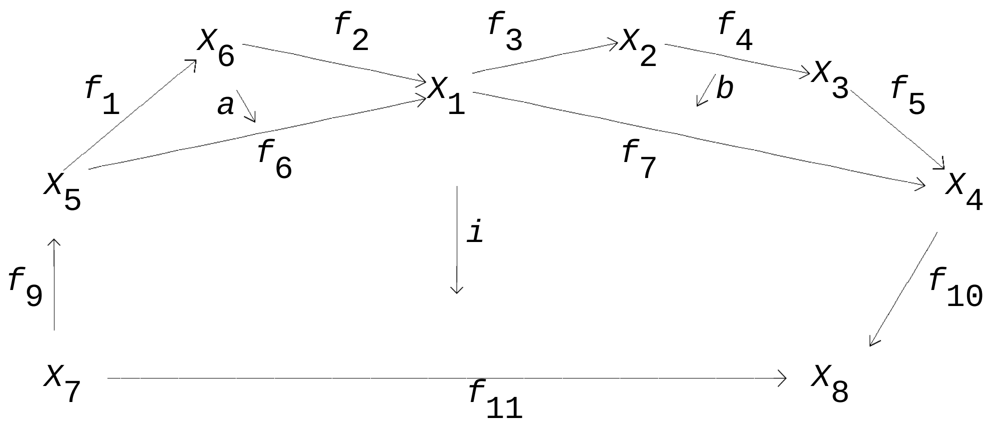
We will now describe a 3-pd φ which is *parallel* to the 3-cell u considered before. This involves the statement that $\mathrm{d}\varphi=\mathrm{d}u$, and therefore involves the determination of the domain $\mathrm{d}\varphi$ of a 3-pd φ. The systematic way of defining the domain of a pd is our main task.

Let us use the 2-pd's β and δ introduced above, as well as the following η and λ; we will use two new 2-cells, h and i:

$$\eta = h(f_{10}, c(f_7, f_6), f_9):$$



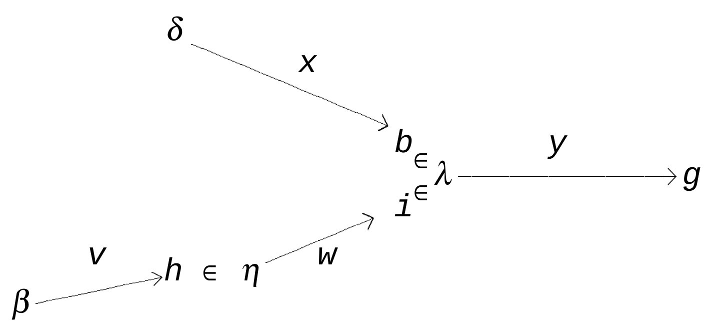
$$\lambda = i(f_{10}, b(f_5, f_6, f_3), a(f_2, f_1), f_9) :$$



Now we introduce the 3-cells:

$$\beta \xrightarrow{v} h\ ,\ \ \eta \xrightarrow{w} i\ ,\ \ \delta \xrightarrow{x} b\ ,\ \ \lambda \xrightarrow{y} g\ .$$

The first thing to check is that these are well-formed, that is, in each case the assigned domain (a 2-pd) and codomain (a 2-cell) are parallel; this is true. Now, notice that these four 3-cells "line up" as follows:



(3')

In fact, we have

$$s(v) = \langle e, d\rangle\ ,\ \ s(w)=\langle h, c\rangle\ ,\ \ s(x)=\langle b\rangle\ ,\ \ s(y)=\langle i, b, a\rangle\ ;$$

$$h=s(w)(1)\ ,\ \ i=s(y)(1)\ ,\ \ b=s(y)(2)\ ;$$

and

$$\varphi \underset{\text{def}}{=} y(w(v(e,d),c),x(b),a)$$

is well-defined as a 3-pd. Note that, to an even larger extent than before, what φ really is cannot be directly seen on its defining expression; only by taking into account the descriptions of all the ingredients, which themselves were defined in similar ways, can we grasp what φ is. The faithful geometric representation of the 3-pd φ is a 3-dimensional object, obtained by joining the 3-dimensional cells v, w, x, y ; the target 2-cell h of v is joined with the occurrence of h in η , similarly for i and b ; we get a spherical (simply connected) 3-dimensional object subdivided appropriately. The full entity φ involves four levels of ingredients: k-cells for all of $k=0, 1, 2, 3$. The 2-dimensional boundary of this object consists of the 2-pd γ as domain, and the 2-cell g as codomain; we have $\mathrm{d}\varphi=\gamma$, $\mathrm{c}\varphi=g$. The 2-cells h , i and one of the occurrences of b are "inner" 2-cells in φ , not denoted in the term representation. φ is indeed parallel to the 3-cell $u:\gamma\rightarrow g$; as a consequence, a 4-cell of the shape $\varphi\longrightarrow g$ is possible.

1.4. The domain of a pasting diagram

We turn to explaining how $\mathrm{d}\varphi$, and in general, the domain of an arbitrary pd, is determined algebraically.

As explained before for the cases $k=1$, 2 and 3 , we construe the set P_k of k-pd's of the HDC $\boldsymbol{A}$ as the arrows of a free multicategory $\boldsymbol{C}_k^0$ whose objects are the elements of C_{k-1} ($(k-1)$-cells), and whose generating arrows are the elements of C_k . (We use the superscript 0 since there will be a modified ("twisted") variant $\boldsymbol{C}_k$ which will be the final version.) The k-cells $a\in C_k$ come with a domain $\mathrm{d}a\in P_{k-1}$ and a codomain $\mathrm{c}a\in C_{k-1}$. For the determination of $\boldsymbol{C}_k^0$, we also need $\mathrm{s}a$ and $\mathrm{t}a$ for $a\in C_k$; as done above for low values of k , we put $\mathrm{s}a=\langle \mathrm{d}a\rangle$, and $\mathrm{t}a=\mathrm{c}a$.

Let $k\geq 1$ be arbitrary, and let $\alpha\in P_{k+1}$. For any $\gamma\in P_k$, we let $\langle\gamma\rangle$ denote the left-to-right list of function-symbol occurrences in γ . Thus, $\mathrm{s}\alpha$ is a tuple of elements of C_k , and $\mathrm{d}\alpha$ is to be defined in such a way that $\langle \mathrm{d}\alpha\rangle$ is also a tuple of elements of C_k .

The first fact on how $\mathrm{d}\alpha$ is defined is that $\mathrm{s}\alpha$ and $\langle \mathrm{d}\alpha \rangle$ are *almost* equal; one is obtained from the other by a permutation. That is, $|\mathrm{s}\alpha| = |\langle \mathrm{d}\alpha \rangle|$, and there is a permutation $\theta_{\alpha}=\theta\colon |\mathrm{s}\alpha| \xrightarrow{\cong} |\langle \mathrm{d}\alpha \rangle|$ such that

$$\begin{array}{ccc} |\mathrm{s}\alpha| & \xrightarrow{\quad\theta\quad} & |\langle \mathrm{d}\alpha \rangle| \\ & {\scriptstyle \mathrm{s}\alpha}\searrow \quad \circ \quad \swarrow{\scriptstyle \langle \mathrm{d}\alpha \rangle} & \\ & C_n & \end{array}$$

Note that, by what was said above, for α a single cell, $\mathrm{d}\alpha$ is already defined, and $\mathrm{s}\alpha=\langle \mathrm{d}\alpha \rangle$; for such α , θ_{α} can be taken to be the identity.

The second, and main, fact about the way $\mathrm{d}\alpha$ is defined is that there is an operation assigning a new "composite" $\gamma\square_{q}\delta$ to any $\gamma, \delta\in P_{k}$ and $q\in |\langle \gamma \rangle|$ satisfying certain conditions of compatibility (that we will see below in detail) such that

$$\mathrm{d}(\alpha\circ_{p}\beta) = (\mathrm{d}\alpha)\square_{\theta(p)}(\mathrm{d}\beta) \; ; \qquad (4)$$

that is, the domain of the grafting composite of two $(k+1)$-pd's is the $\square$-composite of the domains of the $(k+1)$-pd's. This, together with knowing what $\mathrm{d}\alpha$ is for single-cell pd's α determines the operation d .

Let us describe the operation $\square$. In fact, this can be done on an arbitrary free multicategory.

Start with $\boldsymbol{C}=\mathcal{F}(\mathcal{L})$, the free Lambek multicategory on the arbitrary language $\mathcal{L}$; we use the notation we had before; $O=\mathrm{O}(\mathcal{L})=\mathrm{O}(\boldsymbol{C})$ is the set of objects of $\boldsymbol{C}$; $A=\mathrm{A}(\boldsymbol{C})$ is the set of arrows of $\boldsymbol{C}$; we write $\mathrm{s}\alpha$ for $\mathrm{s}_{\boldsymbol{C}}(\alpha)$, $\mathrm{t}\alpha$ for $\mathrm{t}_{\boldsymbol{C}}(\alpha)$. For any $\alpha\in A$, we let $\langle\alpha\rangle$ denote the left-to-right list of function-symbol occurrences in α , as we did before. We let $\mathrm{T}(\alpha)\overset{=}{_{\mathrm{def}}}(\mathrm{s}\alpha,\mathrm{t}\alpha)$. Note that $\mathrm{T}(\alpha)=\mathrm{T}(\beta)$ means that α and β are "parallel in the multicategory $\boldsymbol{C}$ ".

We are going to define a partial operation

$$(\alpha, \beta, p)\longmapsto\alpha\square_{p}\beta \qquad (\alpha, \beta\in A, p\in |\langle\alpha\rangle| ;\; \alpha\square_{p}\beta\in A) .$$

defined whenever

$$T(\beta)=T(\langle\alpha\rangle(p)) \,. \qquad (5)$$

The intuitive idea behind the operation $\square$, called *function-replacement,* is that $\alpha\square_p\beta$ is the function obtained by evaluating, at the place p and only at that place, the function-variable $\langle\alpha\rangle(p)$ as the composite function β. The condition (5) says that β is of the same type as $\langle\alpha\rangle(p)$, and therefore, the said evaluation is possible.

Given $\alpha\in A$ and $p\in|\langle\alpha\rangle|$, let $f=\langle\alpha\rangle(p)\in L$. Then α can be written in the form

$$\alpha=\alpha'\circ_q f(\alpha_1, \ldots, \alpha_n) \,, \qquad (6)$$

where $\alpha', \alpha_1, \ldots, \alpha_n\in A$, and q is a suitable place $q\in s(\alpha')$. Note that if f occurs in more than one place in α, then this *decomposition at* f of α is not unique; however, we have in mind the *decomposition of* α *at the place* p, in which f "stands for the occurrence at p ". What these obscure words mean is intuitively clear, and will be made precise in section 5. The notation $f(\alpha_1, \ldots, \alpha_n)$ follows the term-representation explained above; it is, structurally, a repeated (or *simultaneous*, because of the presence of an appropriate commutative law) composition, as it was also indicated above.

Now, suppose, that, in addition, $\beta\in A$ such that (5). Let
We put

$$\alpha\square_p\beta \;\overset{=}{\text{def}}\; \alpha'\circ_q\beta(\alpha_1, \ldots, \alpha_n) \,. \qquad (7)$$

Here, $\beta(\alpha_1, \ldots, \alpha_n)=\beta(\alpha_1/1, \ldots, \alpha_n/n)$ is simultaneous composition. $T(\beta)=T(f)$ implies that $s(\beta)=s(f)$, and so $t(\alpha_i)=s(f)(i)=s(\beta)(i)$, which makes the term $\beta(\alpha_1, \ldots, \alpha_n)$ well-defined; but also, $T(\beta)=T(f)$ implies that $t(\beta)=t(f)$, which ensures that $t(\beta(\alpha_1, \ldots, \alpha_n))=t(\beta)=t(f)=s(\alpha')(q)$, and thus, the composition at q is well-defined.

Let us see how this works for the examples of 3-pd's ($k=2$) in the previous subsection. We are going to make the discussion easier to follow, by replacing the place-number p by the symbol which occurs at p in the given term; since the terms in the examples are *separated*, that is, have no repeated occurrences of symbols, this will not introduce ambiguity. Note that under this convention, with $f=(s\alpha)(p)=t\alpha$, (4) becomes

$$d(\alpha\circ_f\beta) = (d\alpha)\square_f(d\beta)\ ;$$

and the role of θ disappears (of course, for the general, non-separated case, the said simplification is not valid).

The 3-pd φ introduced in the previuos subsection can be written in the following two ways:

$$\varphi = (y\circ_i(w\circ_h v))\circ_b x = (y\circ_b x)\circ_i(w\circ_h v)$$

(compare (3')). Let us go through the definition of the domain of each of the constituent 3-pd's here.

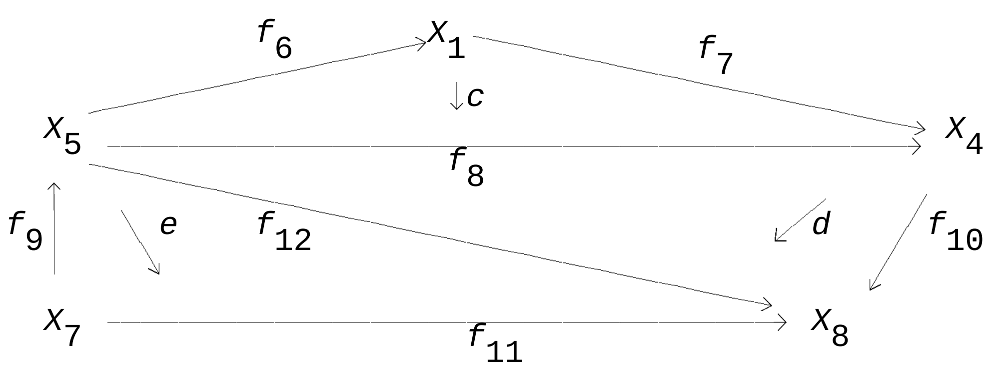
$$d(w\circ_h v) = (dw)\square_h(dv) = \eta\square_h\beta\ .$$

The decomposition of η at h has $\eta'=1_{f_{11}}$ (we are writing η' for what was α' in the general case (6)); that is, now α' can be ignored in (6) and (7). (7) gives

$$\xi \overset{\text{def}}{=} \eta\square_h\beta = \beta(f_{10}/f_{10}, c(f_7, f_6)/f_8, f_9) = e(d(f_{10}, c(f_7, f_6)), f_9)\ ;$$

that is,

ξ :



ξ is obtained by replacing h with β as it should.

Next,

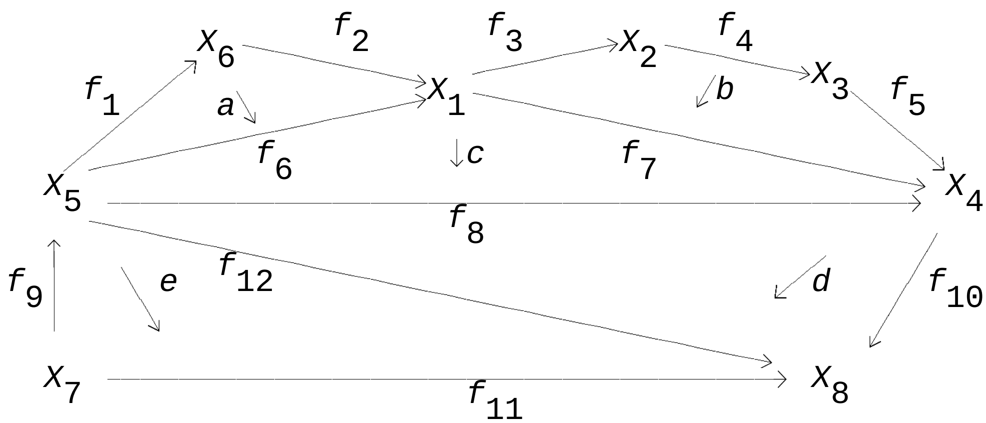
$$\mathrm{d}(y\circ_i(w\circ_h v)) = (\mathrm{d}y)\square\circ_i \mathrm{d}(w\circ_h v)) = \lambda\square_i\xi \ .$$

i is again the head-operation in λ, and so

$$\zeta \ \overline{\overline{\mathrm{def}}}\ \lambda\square_i\xi = \xi(f_{10}/f_{10}, b(f_5, f_6, f_3)/f_7, a(f_2, f_1)/f_6, f_9/f_9)$$
$$e(d(f_{10}, c(b(f_5, f_6, f_3), a(f_2, f_1)), f_9) \ ,$$

that is,

ζ :



Note that ζ is the result of replacing i by ξ in λ.

Finally,

$$\mathrm{d}\varphi = \mathrm{d}(y\circ_i(w\circ_h v))\square_b \mathrm{d}x = \zeta\square_b\delta = \zeta\square_b b = \zeta \ ;$$

note that $\delta{=}b$, and when b is replaced by b, nothing happens. Of course, $\zeta{=}\gamma$, for our initial γ, so this calculation confirms what we said "geometrically" about φ and γ.

Let us look at the other way of expressing φ. We have

$$\mathrm{d}(y\circ_b x) = \mathrm{d}y = \lambda \ ,$$

for the same reason as in the preceding case. $\mathrm{d}(w\circ_h v)=\xi$ was calculated above. Then

$$\mathrm{d}\varphi=\mathrm{d}(y\circ_b x)\square_i\mathrm{d}(w\circ_h v)=\lambda\square_i\xi=\zeta=\gamma\,,$$

as it should be the case.

In this subsection, we described the way the domain-function $\mathrm{d}:P_{k+1}\to P_k$ is actually calculated, and saw that, in some examples at least, it agrees with the geometric intuition. However, thereby the problem of definining d is far from resolved. For instance, it is not clear that, in general, (4) is a compatible way of determining $\mathrm{d}\gamma$ for $\gamma\in P_{k+1}$, ; usually, γ can be written in more than one way as $\gamma=\alpha\circ_p\beta$, and we must see that the corresponding right-hand side expressions for $\mathrm{d}\gamma$ give the same result. There are other problems too. E.g., we have to see that if in (4), the left side is well-defined, so is the right side. Also note that we have not made any reference yet to the fact that d and c on C_{k+1} are determined so that the globularity condition (2) is satisfied. It is worth noting that that condition refers, besides d on P_{k+1}, also to d as defined on P_k. This suggests that d_k *must be defined recursively in* k.

1.5. Generalizing multicategories

The operation $\square$ used in the last subsection looks like a multicategory operation. Let us start with $\boldsymbol{C}$, a free multicategory on $\mathcal{L}$ as we had in the last subsection for the purposes of defining the operation $\square$ on the arrows of $\boldsymbol{C}$; let's use the same accompanying notation. We are going to define $\boldsymbol{D}$, a new multicategory, albeit in a somewhat generalized sense with respect to what we had above. $\boldsymbol{D}$ is called the *multicategory of function-replacement*. The arrows of $\boldsymbol{D}$ are the same as those of $\boldsymbol{C}$: $\mathrm{A}(\boldsymbol{D})=\mathrm{A}(\boldsymbol{C})=A$. The idea is to consider each $\alpha\in A$ to be a function not of its variable-occurrences, but of its function-symbol occurrences. The objects of $\boldsymbol{D}$ are pairs $(\vec{X};Y)$ where $\vec{X}\in O^*$ is a tuple of objects of $\boldsymbol{C}$, and Y is a single object: $\mathrm{O}(\boldsymbol{D})=O^*\times O$. If $\langle\alpha\rangle=\langle f_1,\ldots,f_n\rangle$, then, by definition, $\mathrm{s}_{\boldsymbol{D}}(\alpha)=\mathrm{S}(\alpha)\overset{=}{\mathrm{def}}\langle\mathrm{T}f_1,\ldots,\mathrm{T}f_n\rangle$ and $\mathrm{t}_{\boldsymbol{D}}(\alpha)\overset{=}{\mathrm{def}}\mathrm{T}(\alpha)=(\mathrm{s}\alpha,\mathrm{t}\alpha)$. The operation $\circ_{\boldsymbol{D}}$ is defined to be the operation $\square$ explained in the previous subsection.

Before we say more on to what extent $\mathbf{D}$ is a multicategory, let us point out in what aspect it fails to be one.

Consider a language $\mathcal{L}$ in which we have sorts U, V, W, X, Y and function-symbols

$$f:\langle U, V\rangle \to W\ ,\ g:\langle X\rangle \to U\ ,\ h:\langle U, Y\rangle \to W\ ,\ i:\langle V\rangle \to Y\ . \tag{8}$$

Let $\beta=f(g(X), V)$, $\alpha=h(U, i(V))$, terms in $\mathrm{A}(\mathcal{L})$. We have $\alpha:\langle U, V\rangle \to W$, thus $\mathrm{T}(\alpha)=\mathrm{T}(f)$, and so $\beta\Box_1\alpha=\beta\Box_f\alpha$ is well-defined. Now, we have $\beta=1_W\circ_1 f(g(X), V)$ as the decomposition of β at 1 (at f), so

$$\beta\Box_1\alpha = 1_W\circ_1\alpha(g(X)/U, V/V) = \alpha(g(X)/U, V/V) = h(g(X), i(V))\ .$$

Also,

$$\langle\beta\rangle=\langle f, g\rangle\ ,\ \mathrm{S}(\beta)=\langle \mathrm{T}f, \mathrm{T}g\rangle=\langle(\langle U, V\rangle; W), (\langle X\rangle; U)\rangle\ ,$$
$$\langle\alpha\rangle=\langle h, i\rangle\ ,\ \mathrm{S}(\alpha)=\langle \mathrm{T}h, \mathrm{T}i\rangle=\langle(\langle U, Y\rangle; W), (\langle V\rangle; Y)\rangle\ ,$$
$$\langle\beta\Box_1\alpha\rangle=\langle h, g, i\rangle\ ,$$
$$\mathrm{S}(\beta\Box_1\alpha)=\langle \mathrm{T}h, \mathrm{T}g, \mathrm{T}i\rangle=\langle(\langle U, Y\rangle; W), (\langle X\rangle; U), (\langle V\rangle; Y)\rangle\ .$$

In a Lambek multicategory $\mathbf{E}$, if $\mathrm{s}_{\mathbf{E}}(\beta)=\langle b_1, \ldots, b_n\rangle$, $\mathrm{s}_{\mathbf{E}}(\alpha)=\langle a_1, \ldots, a_m\rangle$, then for $\beta\circ_p\alpha=\beta\circ_p^{(\mathbf{E})}\alpha$, we have

$$\mathrm{s}_{\mathbf{E}}(\beta\circ_p\alpha)=\langle b_1, \ldots, b_{p-1}, a_1, \ldots, a_m, b_{p+1}, \ldots, b_n\rangle\ ;$$

$\mathrm{s}_{\mathbf{E}}(\alpha)$ is inserted in the place of a_p ; this is what we mean by *standard amalgamation* of the sources. The operation $\Box$ is much like a multicategory composition, except for the standard amalgamation. If $\mathbf{D}$ had standard amalgamation, $\mathrm{S}(\beta\Box_1\alpha)$ would have to be the result of inserting $\langle \mathrm{T}h, \mathrm{T}i\rangle$ into $\langle \mathrm{T}f, \mathrm{T}g\rangle$ in the place of $\mathrm{T}f$, resulting in $\langle \mathrm{T}h, \mathrm{T}i, \mathrm{T}g\rangle$; but $\mathrm{S}(\beta\Box_1\alpha)$ is, rather, $\langle \mathrm{T}h, \mathrm{T}g, \mathrm{T}i\rangle\neq\langle \mathrm{T}h, \mathrm{T}i, \mathrm{T}g\rangle$.

We cannot hope that another simple "rule of amalgamation" applies, either. Suppose that in the above, $U=V$, but all other objects listed are distinct; so we have the previous example, still with non-standard amalgamation. But also, for $\beta' = f(U, g(X))$, $\beta'\Box_f\alpha=h(U, i(g(X)))$, and $\mathrm{S}(\beta'\Box_f\alpha)=\langle h, i, g\rangle\neq\mathrm{S}(\beta\Box_f\alpha)$, despite the fact that

$S(\beta')=S(\beta)$. That is, the source of a composite does not depend just on the sources (and targets) of the composed arrows, unlike in the ordinary, Lambek, multicategory.

There is a generalized notion of "multicategory" which allows for "non-standard" amalgamation. In this we have, as part of the structure, so-called *amalgamating maps* $\psi=\psi[\beta, \alpha, p]$, $\varphi=\varphi[\beta, \alpha, p]$:

$$s(\beta)\backslash p \xrightarrow{\ \psi\ } s(\beta\circ_p\alpha) \xleftarrow{\ \varphi\ } s(\alpha)$$

associated with any meaningful composition $(\beta, \alpha, p)\longmapsto\beta\circ_p\alpha$, which puts together the source of $\beta\circ_p\alpha$ in a specific, but *a priori* undetermined, way from the source of β (take away the symbol at place p) and the source of α . The notation abbreviates the following: ψ is a map from the set $|s(\beta)|-\{p\}$ to the set $|s(\beta\circ_p\alpha)|$ (where $|s|=\mathrm{dom}(s)$, and $s\backslash p=s\upharpoonright(|s|-\{p\})$) such that

$$\begin{array}{ccc} |s(\beta)|-\{p\} & \xrightarrow{\ \psi\ } & |s(\beta\circ_p\alpha)| \\ & s(\beta)\backslash p \searrow \quad \circ \quad \swarrow s(\beta\circ_p\alpha) & \\ & O & \end{array} ,$$

and similarly for φ . In the standard case, the amalgamating maps correspond to the fact that in $s(\beta\circ_p\alpha)$, " $s\alpha$ is inserted in $s\beta$ in the place p ". In the generalized concept, there are coherence conditions on the amalgamating maps, one for each of the four laws: the unit laws, the associative law, and the commutative law. The above-described structure $\boldsymbol{D}$ is a multicategory in the generalized sense (in comparing this part with the official definition of section 2, and the definition of $\boldsymbol{D}$ in section 5 , note that the concept being described here is a *1-level* multicategory as opposed the more general 2-level version given in those sections; we will comment on the reason for the 2-level version later in the introduction).

The reason for the general concept of multicategory and for the particular multicategory $\boldsymbol{D}$ is to provide a concept under which $d:P_{k+1}\longrightarrow P_k$ becomes a morphism of multicategories. A *morphism* $F:\boldsymbol{C}\rightarrow\boldsymbol{D}$ *of multicategories* maps objects to objects, arrows to arrows, but, instead of being compatible strictly with the source-assignments, it has a system of transition isomorphisms $\theta_\alpha: |s_{\boldsymbol{C}}(\alpha)| \xrightarrow{\ \cong\ } |s_{\boldsymbol{D}}(F\alpha)|$ ($\alpha\in A(\boldsymbol{C})$) such that

$$\begin{array}{ccc} |s_{\boldsymbol{C}}(\alpha)| & \xrightarrow{\theta_\alpha} & |s_{\boldsymbol{D}}(F\alpha)| \\ {\scriptstyle s_{\boldsymbol{C}}(\alpha)}\downarrow & \circ & \downarrow{\scriptstyle s_{\boldsymbol{D}}(F\alpha)} \\ \mathrm{O}(\boldsymbol{C}) & \xrightarrow[F]{} & \mathrm{O}(\boldsymbol{D}) \end{array} .$$

F is to preserve placed composition; in formulating this, the transition isomorphisms play a role: given that $\beta\circ_p\alpha$ is well-formed in $\boldsymbol{C}$, $F\beta\circ_q F\alpha$ for $q=\theta_\beta(p)$ is well-formed in $\boldsymbol{D}$; we require that $F(\beta\circ_p\alpha)=F\beta\circ_q F\alpha$. It is also required that the θ_α be compatible with the amalgamating maps.

There is a trade-off between amalgamating maps and transition isomorphisms. Given any morphism $F:\boldsymbol{C}\to\boldsymbol{D}$ of multicategories, there is a factorization of F,

$$\begin{array}{ccc} \boldsymbol{C} & \xrightarrow{\quad F \quad} & \boldsymbol{D} \\ {\scriptstyle \Phi}\searrow & \cong & \nearrow{\scriptstyle F'} \\ & \boldsymbol{C}' & \end{array} ,$$

in which Φ is an isomorphism, and in fact, it is an identity on both objects and arrows, and F' is *strict*, its transition isomorphisms are all identities. In other words, by changing the domain to an isomorphic copy, albeit with "twisted" amalgamating maps, it is possible to turn a morphism into a strict one.

1.6. Constructing higher dimensional cells

We are ready to summarize the construction of higher dimensional cells. Assuming that we have a set C_k of k-cells for $k=0, 1, ..., n$, and we have defined k-pd's for the same k's, with domain and codomain maps $\mathrm{d}:P_{k+1}\to P_k$, $\mathrm{c}:P_{k+1}\to C_k$, we introduce $(n+1)$-cells $a\in C_{n+1}$ by declaring each $\mathrm{d}a=\mathrm{d}_{n+1}(a)$ and $\mathrm{c}a=\mathrm{c}_{n+1}(a)$ to be a specific n-pd $\alpha=\mathrm{d}a$, resp. n-cell $b=\mathrm{c}a$ *such that* $\mathrm{d}\alpha=\mathrm{d}b$, $\mathrm{c}\alpha=\mathrm{c}b$, that is,

$$\mathrm{dd}a=\mathrm{dc}a\ ,\ \mathrm{cd}a=\mathrm{cc}a\ . \tag{9}$$

We let $\boldsymbol{D}_n$ be the multicategory of function-replacement based on $\boldsymbol{C}_n$, the free multicategory with arrows the n-pd's, and $\boldsymbol{C}^0_{n+1}$ be the free multicategory with standard amalgamation, and with objects the n-cells, and generating arrows the $(n+1)$-cells just declared; in other words, $\boldsymbol{C}^0_{n+1}=\mathcal{F}(\mathcal{L})$ where $\mathrm{O}(\mathcal{L})=C_n$, $\mathrm{L}(\mathcal{L})=C_{n+1}$, and in which $\mathrm{s}_{\mathcal{L}}(a)=\langle \mathrm{d}a\rangle$, $\mathrm{t}_{\mathcal{L}}(a)=\mathrm{c}a$ ($a\in C_{n+1}$). P_{n+1} is the set of arrows in $\boldsymbol{C}^0_{n+1}$. The main step in the definition is to define $\mathrm{d}^0=\mathrm{d}^0_{n+1}:\boldsymbol{C}^0_{n+1}\longrightarrow\boldsymbol{D}_n$ by the freeness of $\boldsymbol{C}^0_{n+1}$ as to extend the determination of d on C_{n+1}. For this, it is crucial that $\boldsymbol{C}^0_{n+1}$, although it is defined as a Lambek multicategory, it remains free on $\mathcal{L}$ in the larger category of all multicategories with possibly non-standard amalgamation. Finally, we alter $\boldsymbol{C}^0_{n+1}$ to the *isomorphic copy* $\boldsymbol{C}_{n+1}$ by "twisting" the amalgamation maps to ensure that $\mathrm{d}:\boldsymbol{C}_{n+1}\longrightarrow\boldsymbol{D}_n$ is strict. As a result, we get the *main formula* saying that

$$\mathrm{d}(\beta\circ_p\alpha) = (\mathrm{d}\beta)\square_p(\mathrm{d}\alpha) \qquad (10)$$

every time $\beta\circ_p\alpha$ is a meaningful composition in $\boldsymbol{C}_{n+1}$.

Let us see the effect of the above general procedure for some particular 3-cells and 3-pd's. In what follows, $U, V, W, \ldots$ denote 1-cells, $f, g, h, \ldots$ 2-cells, $u, v,$ 3-cells; Greek letters are used to denote pd's of various dimensions.

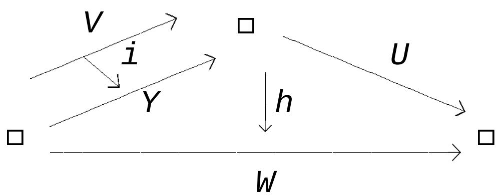
We adopt a single 0-cell that we indicate by $\square$; the 1-cells U, V, W, X, Y are all like $\square\longrightarrow\square$. The 2-cells f, g, h, i are as in (8). We add

$$k:\langle X, V\rangle\longrightarrow W\,,\quad \ell:\langle U, X\rangle\longrightarrow W\,.$$

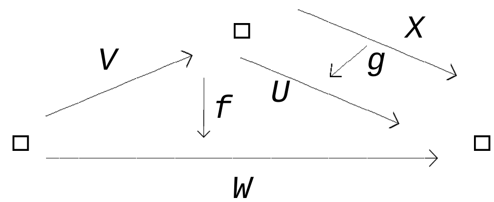
We are assuming that $U=V$, but all other 1-cells denoted differently are distinct.

Consider the 2-pd's $\alpha=h(U, i(U))$ and $\beta=f(g(X), U)$:

α :

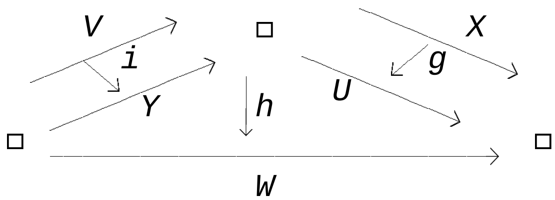


β :



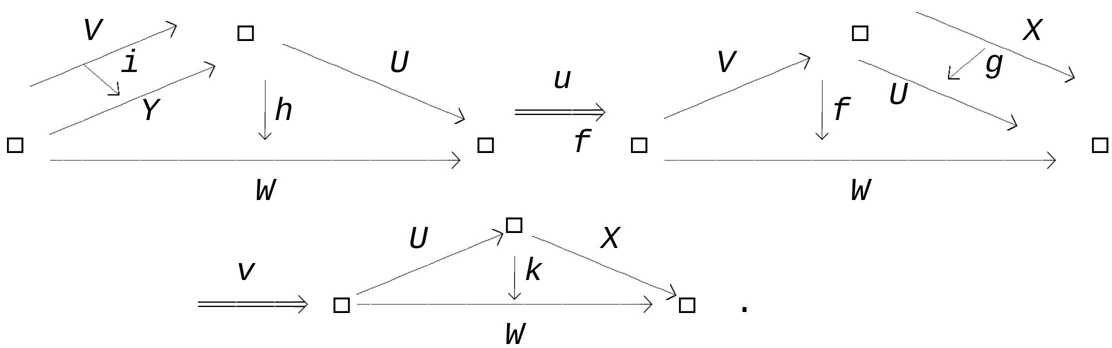
We have $\beta\square_1\alpha = h(g(X), i(U))$:

$\beta\square_1\alpha$:



We introduce the 3-cells u and v by declaring $\mathrm{d}u=\alpha$, $\mathrm{c}u=f$ and $\mathrm{d}v=\beta$, $\mathrm{c}v=k$; the globularity conditions (9) are satisfied. We let

$$\psi = v\circ_1 u = v(u(h, i), g) :$$



We have

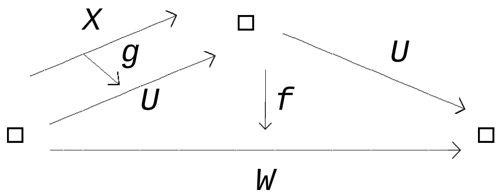
$$\mathrm{d}\psi = \mathrm{d}(v\circ_1 u) = (\mathrm{d}v)\square_1(\mathrm{d}u) = \beta\square_1\alpha ;$$

so

$$s_{\mathbf{C}_3}(\psi) = \langle d\psi \rangle = \langle h, g, i \rangle \ .$$

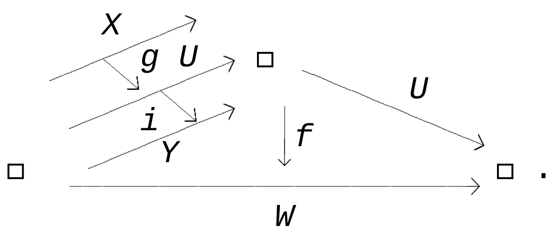
Now, look at

$$\beta' = f(U, g(X)) \ :$$



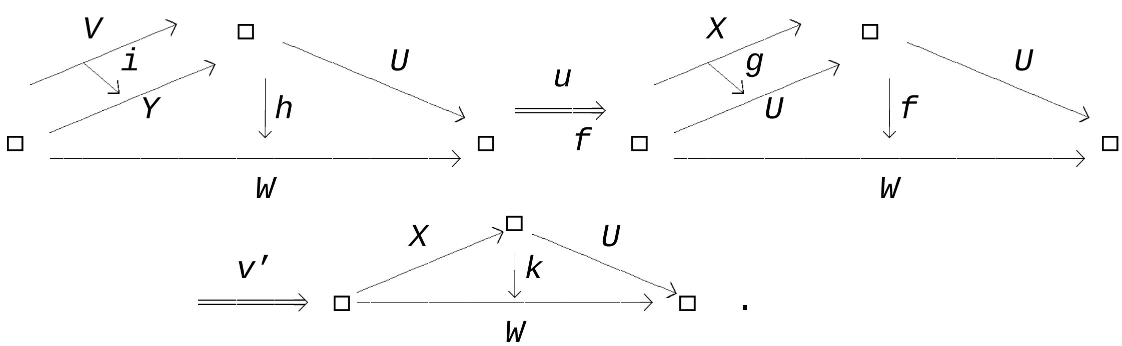
and

$$\beta' \square_1 \alpha = \ h(U, i(g(X))) \ :$$



We let $v' \in C_3$ with $dv' = \beta'$, $cv' = k$ ($d\beta' = dk$, $c\beta' = ck$ hold), and

$$\psi' = v' \circ_1 u = v'(u(h, i), g) \ :$$



We have

$$s_{\mathbf{C}_3}(\psi') = \langle d\psi' \rangle = \langle h, i\, g \rangle \ .$$

We just have to get used to the fact that

$$s_{\mathbf{C}_3}(v(u(h, i), g)) = \langle h, g, i \rangle ,$$

and

$$s_{\mathbf{C}_3}(v'(u(h, i), g)) = \langle h, i, g \rangle$$

at the same time. Of course, this does not look so surprising if we look at the *full* representations of the two 3-pd's $\psi = v(u(h, i), g)$ and $\psi' = v'(u(h, i), g)$, which are different "geometrically".

1.7 Introducing two levels of objects

Some remarks concerning the "2-leveled" version for the notion of multicategory, for whose definition we refer to section 2. This is introduced purely for technical convenience. The 2-leveled notion packs more structure into the multicategory $\mathbf{D}$ of function-replacement, structure that is already there "naturally". For instance, instead of having the source of α as $s_{\mathbf{D}}(\alpha) = \langle T(f_1), \ldots, T(f_n) \rangle$, we have it, in the 2-leveled version of $\mathbf{D}$, as $s_{\mathbf{D}}(\alpha) = \langle f_1, \ldots, f_n \rangle = \langle \alpha \rangle$. The effect is to *restrict the scope* of the composition operation $\square$; composition in the 2-leveled version remains the same as in the 1-leveled version, but it is defined for a subset of the domain of the 1-leveled composition. For $\gamma, \delta \in \mathbf{D}$, the composite $\delta \square_p \gamma$ is meaningful, in the 2-leveled version, if and only if $p \in |\langle \delta \rangle|$, and for $f = \langle \delta \rangle(p)$, we have $df = d\gamma$ and $cf = c\gamma$. This is in fact the case exactly when the function-replacement composite is the *meaningful geometrically*. Under the 1-leveled version, the multicategory $\mathbf{D}$ has composites that cannot be realized geometrically in Euclidean space.

The 2-leveled concept helps technically. An example is the equality $d\alpha = \langle \alpha \rangle$ holding for all $\alpha \in \mathbf{C}_{n+1}$. This is immediate if d is defined by the freeness of $\mathbf{C}_{n+1}$ with respect to the 2-leveled version of "multicategory"; it would require additional arguments if we used the 1-leveled version.

1.8. Final remarks

Obviously, for any fixed n , n-graphs are the objects of a category of the form $\mathtt{Set}^{\mathtt{g}_n}$; here, $\mathtt{g}_n$ is the category whose shape in given in (1). It turns out that n-dimensional multitopic sets, with a natural notion of morphism, also form a category of the form $\mathtt{Set}^{\mathtt{mt}_n}$. In this case, the description of the exponent category $\mathtt{mt}_n$, the category of n-dimensional *multitopes*, is less easy to describe. In fact, there is, apparently, no other way of describing $\mathtt{mt}_n$ than by the same recursive process that serves defining multitopic sets in general. The objects of $\mathtt{mt}_n$ are the same as the pasting diagrams (elements of the P_n-component) in the *terminal* n-dimensional multitopic set, the one that has exactly one cell in each possible type (domain/codomain pair; in fact, here "domain" suffices; this description is an oversimplification, and neglects an inherent recursion). The arrows of $\mathtt{mt}_n$ are more difficult to explain. The definition of the $\mathtt{mt}_n$ and the proof of their connection to multitopic sets in general are given in section 7.

The fact just stated is the justification for the name "multitopic set". It is a similar construction to "simplicial set" , with "simplices" in the background, and also to "opetopic set" of [B/D3], based on "opetopes", in which *operads*, the basic abstract concept for [B/D3], are referred to. We copied and modified "opetope" and "opetopic set" of [B/D3], bearing in mind multicategories as the basic abstract concept, replacing operads.

We note that "higher dimensional (or: n-dimensional) multicategory", a term that may seem at first to be the appropriate one for our concept of multitopic set, is in fact incorrect and misleading. "Higher dimensional multicategory" would rightly be expected to generalize "multicategory"; however, in our multitopic sets only special multicategories figure, namely, the free ones, and another particular kind, the multicategories of function replacement, closely tied to the free multicategories. For multitopic sets, *particular* multicategories are used as a tool to describe a specific geometric arrangement, that of cells of various dimensions fitting together in pasting diagrams. Of course, this is similar to the use of operads in [B/D3].

2. Multicategories

For $\ell\in\mathbb{N}=\{0, 1, 2, \ldots\}$, we write $[1, \ell]$ for the set $\{1, 2, \ldots, \ell\}$; $[1, \ell]=\emptyset$ when $\ell=0$.

Let O be a set. A *tuple* (string) of elements of O is a function of the form $\varphi:[1, \ell]\to O$, for some $\ell\in\mathbb{N}$. We write $|\varphi|$ for the set $[1, \ell]$, and $\ell h(\varphi)$ for ℓ. O^* is the set of all tuples of elements of O. $\bot\in O^*$ is the empty tuple ($|\bot|=\emptyset$). For $X\in O$, $\langle X\rangle\in O^*$ is the one-term tuple whose only term is X; $|\langle X\rangle|=[1, 1]=\{1\}$, $\langle X\rangle(1)=X$.

It will be convenient to work with the following category $O^{\#}$. Its objects are each a function s whose domain $|s|$ is a finite set (possibly empty) of positive integers, and whose range is a subset of O; $s:|s|\to O$. An arrow $s\to t$ is a function $f:|s|\to|t|$ such that

$$\begin{array}{ccc} |s| & \xrightarrow{\quad f \quad} & |t| \\ & {\scriptstyle s}\searrow \quad \circ \quad \swarrow{\scriptstyle t} & \\ & O & \end{array}$$

(the circle in a diagram denotes the assertion that the diagram commutes).

A *multicategory* $\boldsymbol{C}$ is given by data (i) to (vii) and conditions (viii) to (xi) as follows.

(i) A set $O=\mathrm{O}(\boldsymbol{C})$ of *upper level objects*, or simply, *objects*.

(ii) A set $\dot{O}=\dot{\mathrm{O}}(\boldsymbol{C})$ of *lower level objects*.

(iii) A map $O\longrightarrow\dot{O}:X\mapsto\dot{X}$, assigning a lower-level object $\dot{X}$ to every object X .

(iv) A set $A=\mathrm{A}(\boldsymbol{C})$ of *arrows*.

(v) To each arrow f a *source* $\mathrm{s}(f)=\mathrm{s}f=\mathrm{s}_{\mathbf{C}}(f)\in O^{*}$, and a *target* $\mathrm{t}(f)=\mathrm{t}_{\mathbf{C}}(f)\in\dot{O}$ is assigned; we write $\vec{X}\xrightarrow{f}A$ if $\mathrm{s}(f)=\vec{X}$, $\mathrm{t}(f)=A$; here, $\vec{X}\in O^{*}$, $A\in\dot{O}$.

(vi) Given $\mathrm{s}(f)\xrightarrow{f}\mathrm{t}(f)$, $\mathrm{s}(g)\xrightarrow{g}\mathrm{t}(g)$, and $p\in|\mathrm{s}(g)|$ such that $(\mathrm{s}(g)(p))^{\cdot}=\mathrm{t}(f)$, which situation we indicate by the notation

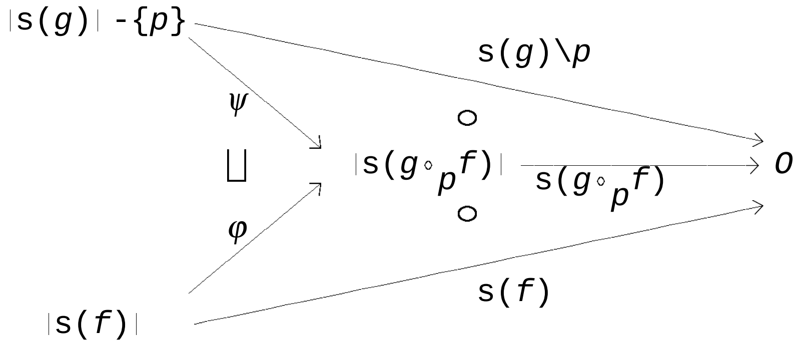
$$\mathrm{s}(f)\xrightarrow[p]{f}\mathrm{s}(g)\xrightarrow{g}\mathrm{t}(g)\ ,$$

a *composite* $g\circ_p f$ is defined; it is an arrow; we have $\mathrm{t}(g\circ_p f)=\mathrm{t}(g)$; furthermore, we have specified *amalgamating maps*

$$\psi=\psi[g,f,p]:\mathrm{s}(g)\backslash p\longrightarrow\mathrm{s}(g\circ_p f)\ ,$$

$$\varphi=\varphi[g,f,p]:\mathrm{s}(f)\longrightarrow\mathrm{s}(g\circ_p f)$$

(morphisms in $O^{\#}$), forming the coprojections of a coproduct in $O^{\#}$. ($\mathrm{s}(g)\backslash p$ means the restricted function $\mathrm{s}(g)\restriction(|\mathrm{s}(g)|-\{p\})$; also, for a subset $P\subset|\mathrm{s}(g)|$, we use the notation $\mathrm{s}(g)\backslash P$ in a similar sense.) In plain words, the set $|\mathrm{s}(g\circ_p f)|$ is the disjoint sum of the sets $|\mathrm{s}(g)|-\{p\}$ and $|\mathrm{s}(f)|$, with injections ψ and φ ; and these injections are morphisms of the functions ($O^{\#}$-objects) $\mathrm{s}(g)\backslash p$, $\mathrm{s}(f)$; that is, we have the commutative diagram



It follows that $|\mathrm{s}(g\circ_p f)|$ is given as $[1,\ell+m-1]$ where $|\mathrm{s}(f)|=[1,\ell]$, $|\mathrm{s}(g)|=[1,m]$; however, this fact leaves open multiple possibilities for the amalgamating maps φ and ψ. Let me emphasize that in general, $\psi=\psi[g, f, p]$, $\varphi=\varphi[g, f, p]$ depend in an essential way on all three arguments g, f, p; in particular, it is possible that $s(g')=s(g)$, $s(f')=s(f)$, but $s(g'\circ_p f')\neq s(g\circ_p f)$.

(vii) For each $Y\in O$, an identity map $\langle Y\rangle \xrightarrow{1_Y} \dot{Y}$.

For the data listed, we require the following laws to be obeyed.

(viii) (**unit law 1**) Whenever $g\in A$, $p\in |s(g)|$, and $Y=s(g)(p)$, which, in particular, implies $\langle Y\rangle \xrightarrow[p]{1_Y} s(g) \xrightarrow{g} t(g)$ (although the latter only says that $\dot{Y}=\dot{Y}_p$, which is weaker than what we are assuming now), we require that

$$g\circ_p 1_Y=g\ .$$

Moreover, we require that

$$\psi=\psi[g, 1_Y, p]=\mathrm{incl.}:(|s(g)|-\{p\})\to |s(g)|\ ,$$
$$\varphi=\varphi[g, 1_Y, p]=(1\mapsto p)$$

(which imply that $(\varphi, \psi\backslash p)$ is a coproduct pair in $O^{\#}$).

(ix) (**unit law 2**) Under the assumption that $s(f) \xrightarrow[1]{f} \langle Y\rangle \xrightarrow{1_Y} Y$ (that is, $t(f)=\dot{Y}$), we require that

$$1_Y\circ_1 f = f\ ,$$

and $\varphi=\varphi[1_Y, f, p]=\mathrm{id}$ (making $(\varphi, \psi\backslash 1=\perp)$ a coproduct pair in $O^{\#}$).

(x) (**associative law**) In the situation

$$s(f) \xrightarrow[p]{f} s(g) \xrightarrow[q]{g} s(h) \xrightarrow{h} t(h)\ ,$$

we require that

$$(h\circ_q g)\circ_{\bar{p}} f = h\circ_q (g\circ_p f)\ ;$$

here, $\bar{p}=\varphi[h, g, q](p)$. Let us refer to the four compositions by the numbers as in

$$\overset{3}{(h\circ_q} g)\overset{4}{\circ}_{\bar{p}} f\ ,\quad h\overset{2}{\circ}_q (g\overset{1}{\circ}_p f)\ .$$

Note that the compositions 1 and 3 are well-defined by the assumptions. 2 is meaningful since we have $t(g\circ_p f)=t(g)$. 4 is meaningful since, for $\varphi=\varphi[g, h, q]$,

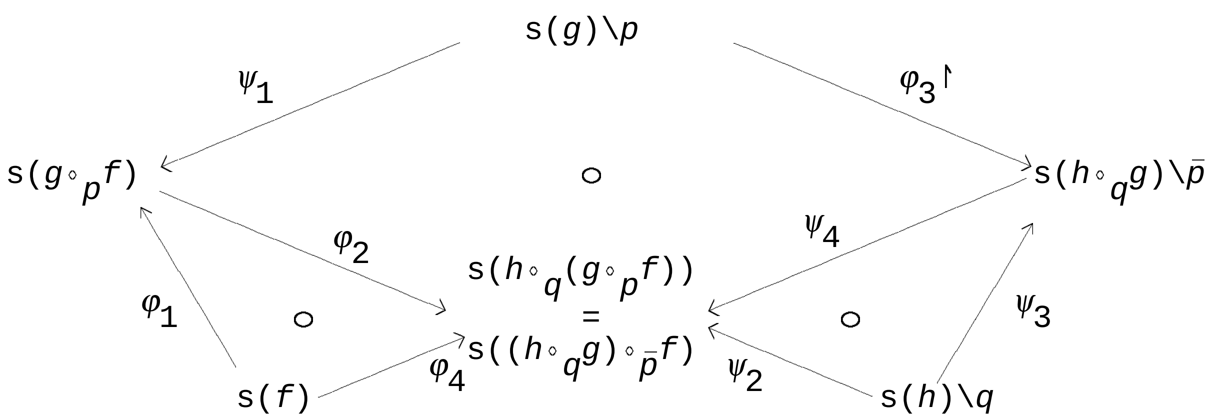
$$s(h\circ_q g)(\bar{p}) = s(h\circ_q g)(\varphi(p)) = s(g)(p) = t(f)\ ;$$

the second equality because we have $\varphi{:}s(g)\longrightarrow s(h\circ_q g)$ in $O^{\#}$.

We abbreviate

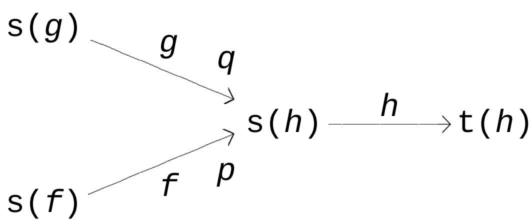
$$\varphi_1=\varphi[g, f, p]\ ,\ \psi_1=\psi[g, f, p]\ ,\ \varphi_2=\varphi[h, g, q]\ ,\ \psi_2=\psi[h, g, q]\ ,$$
$$\varphi_3=\varphi[h, g, q]\ ,\ \psi_3=\psi[h, g, q]\ ,\ \varphi_4=\varphi[h\circ_q g, f, p]\ ,\ \psi_4=\psi[h\circ_q g, f, p]\ .$$

We require that the following diagram be commutative:



Here, $\varphi_3\restriction$ is the restriction of φ_3 to the appropriate domain. Since (φ_3, ψ_3) is a coproduct, $\bar{p}=\varphi_3(p)\notin \mathrm{Im}(\psi_3)$, thus the use of ψ_3 in the diagram is legitimate.

(xi) (**commutative law**) In the situation

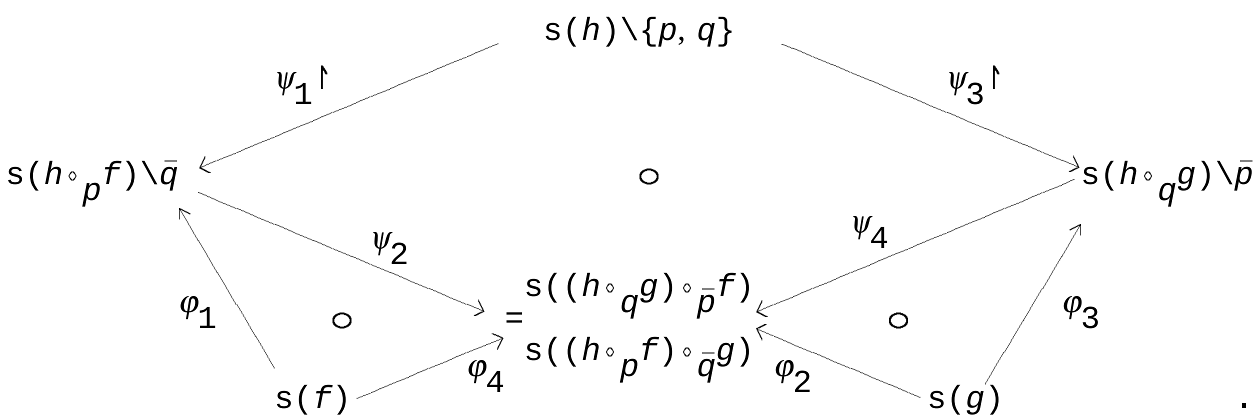


, $p\neq q$,

for $\bar{q}=\psi[h, f, p](q)=\psi_1(q)$ (since $p\neq q$, $q\in \mathrm{dom}(s(h)\backslash p)$, so $\psi_1(q)$ is defined), $\bar{p}=\psi[h, g, q](p)=\psi_3(p)$, we require

$$(h\overset{1}{\circ}_p f)\overset{2}{\circ}_{\bar{q}} g = (h\overset{3}{\circ}_q g)\overset{4}{\circ}_{\bar{p}} f \ .$$

The fact that the composites 2, 4 are well-defined is seen as in the previous case. With φ_i, ψ_i similarly as above, we require the commutativities as in



.

Since $\bar{q}\in \mathrm{Im}(\psi_1)$, we have $\bar{q}\notin \mathrm{Im}(\varphi_1)$, so the use of φ_1 is justified; similarly for φ_3. The map ψ_1 is injective; so, $\bar{q}\notin \mathrm{Im}(\psi_1\restriction(|s(h)|-\{p, q\}))$, and the use of the restricted $\psi_1\restriction$ is justified. Similarly for $\psi_3\restriction$

(**end** of definition of "multicategory").

The standard definition of multicategory (see [L1], [L2]) is the special case in which (i) $\dot{O}=O$, $\dot{X}=X$ for all $X\in O$ (a *one-level* multicategory, as opposed to the general *two-level* notion), and (ii) we make the *standard* choice for the amalgamating maps as explained now.

For $\vec{X}=\langle X_i\rangle_{i\in[1,\ell]}$, $\vec{Y}=\langle Y_j\rangle_{j\in[1,n]}$, both in O^*, and for $p\in[1,m]$, a particular index, $\vec{Y}\Box_p\vec{X}$ denotes the result of *inserting* $\vec{X}$ *into* $\vec{Y}$ *at the place* p; this in effect replaces Y_p by $\vec{X}$. This means that $\vec{Y}\Box_p\vec{X}=\vec{Z}=\langle Z_k\rangle_{k\in[1,n]}$, where $n=\ell+m-1$, $Z_k=Y_k$ when $1\le k<p$, $Z_k=X_{k-p+1}$ when $p\le k<p+\ell$, and $Z_k=Y_{k-\ell+1}$ when $p+\ell\le k\le n$. Define

$$\varphi=\varphi[\vec{Y},\vec{X},p]:|\vec{X}|\to|\vec{Z}|\ ,\ \psi=\psi[\vec{Y},\vec{X},p]:|\vec{Y}|-\{p\}\to|\vec{Z}|$$

by $\varphi(i)=p+i-1$; $\psi(j)=j$ when $1\le j<p$, and $\varphi(j)=p-j+1$ when $p<j\le n$; we have the coproduct diagram

$$\vec{X}\xrightarrow{\ \varphi\ }\vec{Z}\xleftarrow{\ \psi\ }\vec{Y}\backslash p$$

in $O^\#$. When the multicategory has the just specified connecting maps: for the composition $g\circ_p f$, $\varphi[g,f,p]=\varphi[\mathrm{s}(g),\mathrm{s}(f),p]$, $\psi[g,f,p]=\psi[\mathrm{s}(g),\mathrm{s}(f),p]$, we talk about a multicategory *with standard amalgamation*. In particular, the source of $g\circ_p f$ depends, in the standard case, only on the sources of the factors, and the place p; not necessarily so in the general case.

Note that in the standard case, the commutativities required for associativity and commutativity ((ix) and (x)) are automatic.

There is a further remark to be made about the commutative diagrams in the laws of associativity and commutativity, to the effect that they are, to a large extent, automatically true. Referring to the notation in (x), suppose that the functions $\mathrm{s}(f)$, $\mathrm{s}(g)$, $\mathrm{s}(h)$ are one-to-one (non-repeating tuples), and their ranges are pairwise disjoint. I claim that, as a consequence of the preceding conditions, the commutativities required in (x) are now true. First of all, since $s\overset{=}{\mathrm{def}}\mathrm{s}(h\circ_q(g\circ_p f))$ is a coproduct (in $O^\#$) of non-repeating (generalized) tuples, it is itself non-repeating. But then for any $t\in O^{\#}$, there can be at most one morphism $t\to s$ in $O^{\#}$. This implies each of the three commutativities in (x). The same can be said about (xi). We will exploit this fact in section 5.

As a consequence of the definition, in any multicategory, we have a concept of *simultaneous composition*. Assume $g\in A$, $p_i\in|\mathrm{s}(g)|$ for $i=1,\ldots,m$, $p_i\neq p_j$ when $i\neq j$. Assume that $f_i\in A$ for $i=1,\ldots,m$, such that $\mathrm{t}(f_i)=(\mathrm{s}(g)(i))$ for all $i\in[1,m]$. Then we define

$$h \overset{=}{\scriptstyle\mathrm{def}} g(f_1/p_1, f_2/p_2, \ldots, f_m/p_m)\,, \qquad (1)$$

and with $P=\{p_1,\ldots,p_m\}$, the amalgamating functions

$$\psi:\mathrm{s}(g)\backslash P\longrightarrow\mathrm{s}(h)\,, \qquad \psi=\psi[g,\langle f_j\rangle_{j\in[1,m]},\langle p_j\rangle_{j\in[1,m]}]$$

and

$$\varphi_i:\mathrm{s}(f_i)\longrightarrow\mathrm{s}(h)\,, \qquad \varphi_i=\varphi_i[g,\langle f_j\rangle_{j\in[1,m]},\langle p_j\rangle_{j\in[1,m]}]$$

such that $\mathrm{s}(h)$ is the coproduct of the $O^{\#}$-objects $\mathrm{s}(g)\backslash P$, $\mathrm{s}(f_i)$ $(i\in[1,m])$ via the coprojections ψ, φ_i $(i\in[1,m])$. The definition is by recursion on m. When $m=0$, $h=g$, we have $\psi=\mathrm{id}_{\mathrm{s}(g)}$. Suppose $m\geq1$, and assume that $\bar{h}=g(f_1/p_1, f_2/p_2, \ldots, f_{m-1}/p_{m-1})$ has been defined, with corresponding amalgamating functions

$$\bar{\psi}:\mathrm{s}(g)\backslash P^-\longrightarrow\mathrm{s}(\bar{h}) \quad (P^-=\{p_i: i\in[1,m-1]\}),$$

$$\bar{\varphi}_i:\mathrm{s}(f_i)\longrightarrow\mathrm{s}(\bar{h}) \quad (i\in[1,m-1])\,.$$

We put

$$h=\bar{h}\circ_{\bar{p}_m} f_m\,,$$

where $\bar{p}_m=\bar{\psi}(p_m)$, and, with $\check{\varphi}=\varphi[\bar{h},f_m,\bar{p}_m]$, $\check{\psi}=\psi[\bar{h},f_m,\bar{p}_m]$, we define the amalgamating functions for (1) as $\psi=\check{\psi}\circ(\bar{\psi}\restriction P)$, $\varphi_i=\check{\psi}\circ\bar{\varphi}_i$ $(i\in[1,m-1])$ and $\varphi_m=\check{\varphi}$.

In the simultaneous composition, the order of the composed-in factors is immaterial. Precisely speaking, we have the *generalized commutative law*, which says the following:

$$g(f_1/p_1, f_2/p_2, \ldots, f_m/p_m) = g(\hat{f}_1/\hat{p}_1, \hat{f}_2/\hat{p}_2, \ldots, \hat{f}_m/\hat{p}_m)$$

provided for a permutation $\sigma:[1,m]\xrightarrow{\cong}[1,m]$, we have $\hat{p}_i=p_{\sigma i}$ and $\hat{f}_i=f_{\sigma i}$ $(i\in[1,m])$; moreover,

$$\hat{\psi}=\psi,\quad \hat{\varphi}_i=\varphi_{\sigma i}$$

where, of course, we have used the obvious notation for the corresponding amalgamating functions, that is,

$$\hat{\psi}=\psi[g, \langle\hat{f}_j\rangle_{j\in[1,m]}, \langle\hat{p}_j\rangle_{j\in[1,m]}],$$
$$\hat{\varphi}_i=\varphi_i[g, \langle\hat{f}_j\rangle_{j\in[1,m]}, \langle\hat{p}_j\rangle_{j\in[1,m]}].$$

For the case $m=2$, the generalized commutative law is identical to the original form of the commutative law (including the commutativity of the corresponding diagram). The general case be proved by using the commutative law alone, by representing the arbitrary permutation σ as a product of transpositions each of which exchanges two elements standing next to each other in the "previous" permutation.

Therefore, the best way of looking at simultaneous composition is that we have an arrow g, a set $P\subset|s(g)|$, and a function $p\mapsto f_p:P\to A$, such that $t(f_p)=s(g)(p)$ $(p\in P)$, giving rise to the composite $h=g(\langle f_p/p\rangle_{p\in P})$, and to the amalgamating maps $\psi=\psi[g,\langle f_p\rangle_{p\in P}]:s(g)\backslash P\to s(h)$, $\varphi_p=\varphi_p[g,\langle f_p\rangle_{p\in P}]:s(f_p)\to s(h)$ $(p\in P)$. In fact, we can define

$$h=g(\langle f_p/p\rangle_{p\in P}) = g(\bar{f}_1/p_1, \bar{f}_2/p_2, \ldots, \bar{f}_m/p_m)$$

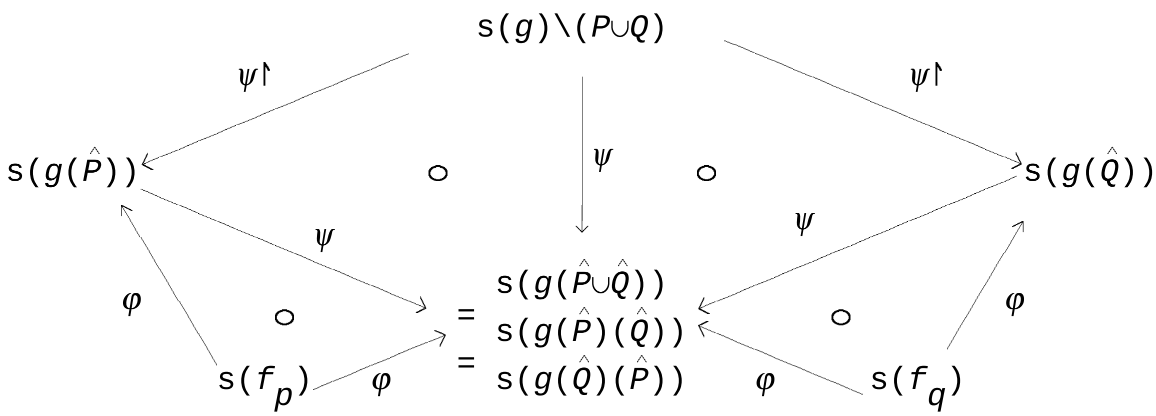
for an arbitrary repetition-free enumeration $\langle p_j\rangle_{j\in[1,m]}$ of P, and for $\bar{f}_i = f_{p_i}$; of course,

$$\psi[g, \langle f_p \rangle_{p\in P}] = \psi[g, \langle \bar{f}_j \rangle_{j\in[1,m]}, \langle p_j \rangle_{j\in[1,m]}],$$

and

$$\varphi_{p_i}[g, \langle f_p \rangle_{p\in P}] = \varphi_i[g, \langle \bar{f}_j \rangle_{j\in[1,m]}, \langle p_j \rangle_{j\in[1,m]}].$$

Suppose $P, Q \subset |s(g)|$, $P \cap Q = \emptyset$; write $\hat{P}$ and $\hat{Q}$ for the assignments $\hat{P} = \langle f_p/p \rangle_{p\in P}$, $\hat{Q} = \langle f_q/q \rangle_{q\in Q}$. Suppose both $g(\hat{P}) = g(\langle f_p/p \rangle_{p\in P})$, $g(\hat{Q}) = g(\langle f_q/q \rangle_{q\in Q})$ are well-defined. We can consider $g(\hat{P}\cup\hat{Q}) = g(\langle f_r/r \rangle_{r\in P\cup Q})$, and we have $g(\hat{P}\cup\hat{Q}) = g(\hat{P})(\hat{Q}) = g(\hat{Q})(\hat{P})$, with the following diagram commuting:



.

Here, the further specification of the maps is self-explanatory. It should be mentioned that the map $\varphi : s(f_p) \longrightarrow s(g(\hat{P}\cup\hat{Q})) = s(g(\hat{Q})(\hat{P}))$ has two meanings, which coincide: $\varphi = \varphi[g, \hat{P}\cup\hat{Q}, p] = \varphi[g(\hat{Q}), \hat{P}, p]$; similarly for q in place for p.

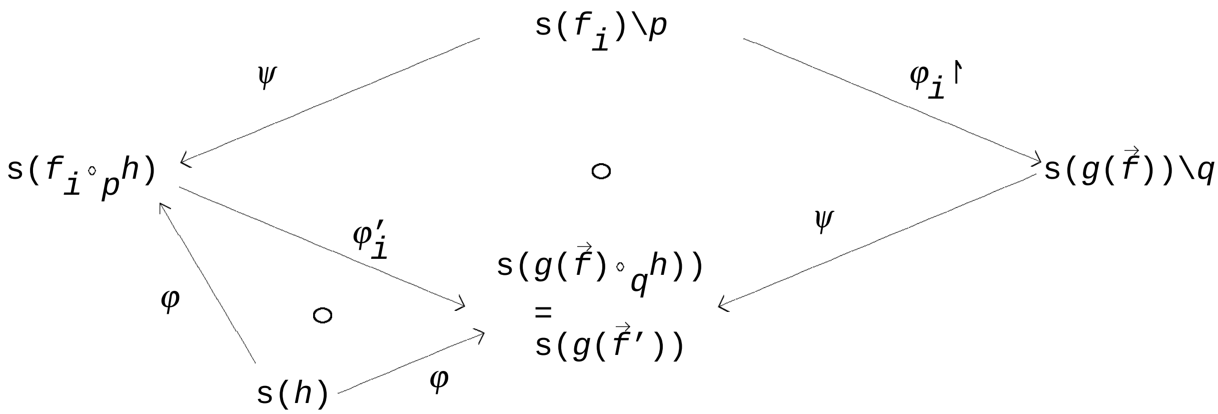
We write $g(f_1, f_2, \dots, f_m)$, or $g(\langle f_i \rangle_{i=1}^m)$, for $g(f_1/1, f_2/2, \dots, f_m/m)$. The notation $g(f_1, f_2, \dots, f_m)$ (in which there is no notation of the place where each f_i is being composed into g) will never be used unless *all* places of g are involved (that is, $m = \ell h(s(g))$), and f_i is composed into g at the place i. Now, $P = [1, m]$: the ψ-map for $g(f_1, f_2, \dots, f_m)$ is empty: its domain is the empty set $|s(g)| - P = \emptyset$; $s(g(f_1, f_2, \dots, f_m))$ is the coproduct of the $s(f_i)$ via the maps

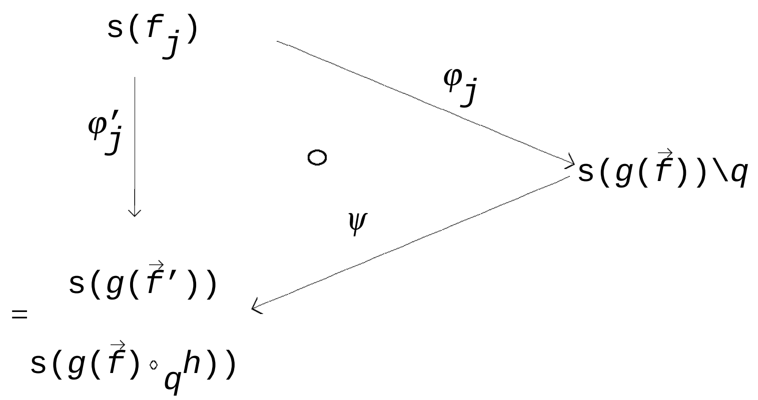
$$\varphi_i = \varphi_i[g, \langle f_i \rangle_{i\in[1,m]}] : s(f_i) \to s(g(f_1, f_2, \dots, f_m)) \quad (i \in [1, m]).$$

Let us formulate a version of the transitive law, involving a simultaneus composition. Using the notation of the previous paragraph, let $i \in [1, m]$, $p \in |s(f_i)|$, and suppose $f_i \circ_p h$ is well-defined. Then for $q = \varphi_i(p)$, we have

$$g(f_1, f_2, \ldots, f_m) \circ_q h = g(f_1, \ldots, f_{i-1}, f_i \circ_p h, f_{i+1}, \ldots, f_m) \ .$$

The coherence commutativities in this case are:





we have used the abbreviations $g(\vec{f}) = g(f_1, f_2, \ldots, f_m)$,
$g(\vec{f}') = g(f_1, \ldots, f_{i-1}, f_i \circ_p h, f_{i+1}, \ldots, f_m)$; $j \neq i$.

3. Morphisms of multicategories

Given multicategories $\boldsymbol{C}$, $\boldsymbol{D}$, a morphism $F:\boldsymbol{C}\longrightarrow\boldsymbol{D}$ is given by data and conditions as follows.

(i) Maps $F:\mathrm{O}(\boldsymbol{C})\to\mathrm{O}(\boldsymbol{D})$, $\dot{F}:\dot{\mathrm{O}}(\boldsymbol{C})\to\dot{\mathrm{O}}(\boldsymbol{D})$ on objects such that

$$\begin{array}{ccc} \mathrm{O}(\boldsymbol{C}) & \xrightarrow{(\ )^{\cdot}} & \dot{\mathrm{O}}(\boldsymbol{C}) \\ {\scriptstyle F}\downarrow & \circ & \downarrow{\scriptstyle \dot{F}} \\ \mathrm{O}(\boldsymbol{D}) & \xrightarrow[(\ )^{\cdot}]{} & \dot{\mathrm{O}}(\boldsymbol{D}) \end{array} \quad .$$

(ii) A map $F:\mathrm{A}(\boldsymbol{C})\longrightarrow\mathrm{A}(\boldsymbol{D})$ on arrows; $\mathrm{t}_{\boldsymbol{D}}(Ff)=\dot{F}(\mathrm{t}_{\boldsymbol{C}}(f))$ is required for all $f\in\mathrm{A}(\boldsymbol{C})$.

(iii) For any $f\in\mathrm{A}(\boldsymbol{C})$, a *transition* bijection

$\theta_f:|\mathrm{s}_{\boldsymbol{C}}(f)|\xrightarrow{\cong}|\mathrm{s}_{\boldsymbol{D}}(Ff)|$ such that

$$\begin{array}{ccc} |\mathrm{s}_{\boldsymbol{C}}(f)| & \xrightarrow{\theta_f} & |\mathrm{s}_{\boldsymbol{D}}(Ff)| \\ {\scriptstyle \mathrm{s}_{\boldsymbol{C}}(f)}\downarrow & \circ & \downarrow{\scriptstyle \mathrm{s}_{\boldsymbol{D}}(Ff)} \\ \mathrm{O}(\boldsymbol{C}) & \xrightarrow[F]{} & \mathrm{O}(\boldsymbol{D}) \end{array} \quad . \qquad (1)$$

Note that this is the same as to say that $\theta_f:F\circ\mathrm{s}_{\boldsymbol{C}}(f)\xrightarrow{\cong}\mathrm{s}_{\boldsymbol{D}}(Ff)$ in $O(\boldsymbol{D})^{\#}$.

(iv) F preserves identities: $F(\langle Y\rangle\xrightarrow{1_Y}\dot{Y})=\langle FY\rangle\xrightarrow{1_{FY}}(FY)^{\cdot}$.

(v) F preserves composition. Given $f,g\in\mathrm{A}(\boldsymbol{C})$, $p\in|\mathrm{s}_{\boldsymbol{C}}(g)|$, $\mathrm{t}_{\boldsymbol{C}}(f)=(\mathrm{s}_{\boldsymbol{C}}(g)(p))^{\cdot}$ (so that $g\circ_p f$ is well-defined), for $\hat{p}=\theta_g(p)\in|\mathrm{s}_{\boldsymbol{D}}(Fg)|$ we have that

$$t_{\mathbf{D}}(Ff) \underset{\uparrow\,(ii)}{=} \dot{F}(t_{\mathbf{C}}(f)) = \dot{F}(s_{\mathbf{C}}(g)(p)^{\cdot}) \underset{\uparrow\,(i)}{=} F(s_{\mathbf{C}}(g)(p)^{\cdot} \underset{\uparrow\,(iii)}{=} s_{\mathbf{D}}(Fg)(\hat{p})^{\cdot} \ ,$$

thus $(Fg)\circ_{\hat{p}}(Ff)$ is well-defined. We require that

$$F(g\circ_p f) = (Fg)\circ_{\hat{p}}(Ff) \ ;$$

moreover,

$$\begin{array}{ccc}
|s_{\mathbf{C}}(g)|-\{p\} & \xrightarrow{\ \theta_g\upharpoonright\ } & |s_{\mathbf{D}}(Fg)|-\{\hat{p}\} \\
\psi[f,g,p]\downarrow & \circ & \downarrow\psi[Ff,Fg,\hat{p}] \\
|s_{\mathbf{C}}(g\circ_p f)| & \xrightarrow{\ \theta_{g\circ_p f}\ } & |s_{\mathbf{D}}(F(g\circ_p f))| = |s_{\mathbf{D}}((Fg)\circ_{\hat{p}}(Ff))| \\
\varphi[f,g,p]\uparrow & \circ & \uparrow\varphi[Ff,Fg,\hat{p}] \\
|s_{\mathbf{C}}(f)| & \xrightarrow[\ \theta_f\ ]{} & |s_{\mathbf{D}}(Ff)|
\end{array} \ .$$

There is a composition of morphisms of multicategories. Given $\mathbf{C}\xrightarrow{F}\mathbf{D}\xrightarrow{G}\mathbf{E}$, for $H=G\circ F:\mathbf{C}\longrightarrow\mathbf{E}$, we have H is the usual composite as far as the effect on objects and arrows is concerned, and $\theta_f\colon |s_{\mathbf{C}}(f)| \longrightarrow |s_{\mathbf{E}}(Hf)|$ is given as the composite $|s_{\mathbf{C}}(f)| \xrightarrow{\theta_f^{(F)}} |s_{\mathbf{D}}(Ff)| \xrightarrow{\theta_{Ff}^{(G)}} |s_{\mathbf{E}}(GFf)|$. It is fairly clear that H is so defined is indeed a morphism of multicategories. We also have the obvious *identity morphism* $\mathrm{Id}_{\mathbf{C}}:\mathbf{C}\to\mathbf{C}$.

The said items form the category Multicat of (small) multicategories and their morphisms.

Let us emphasize that *multicategories* are treated here as 0-dimensional objects, that is, objects of a 1-dimensional, ordinary, category, in contrast to the fact that *categories* are usually treated as 1-dimensional objects in a 2-dimensional category. This fact is the key specific feature of our approach. There are *isomorphisms* of multicategories, but there are, at least for us, *no equivalences* of them.

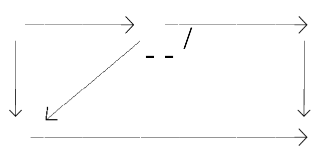
Let us note that every morphism $F:\mathbf{C}\to\mathbf{D}$ can be factored, in a unique manner, in the form

$$\begin{array}{ccccc} \mathbf{C} & & \xrightarrow{\quad F \quad} & & \mathbf{D} \\ & {\scriptstyle \Phi\ \cong}\searrow & & \nearrow {\scriptstyle F'} & \\ & & \mathbf{C}' & & \end{array}$$

so that the isomorphism Φ is the identity on objects and arrows, and F' is *strict*, that is, all its transition maps are identities. To define $\mathbf{C}'$, we put $O(\mathbf{C}')=O(\mathbf{C})=O$, $A(\mathbf{C}')=A(\mathbf{C})=A$, and, for any $f\in A$, $t_{\mathbf{C}'}(f)=t_{\mathbf{C}}(f)=t(f)$. For any $f\in A$, $s_{\mathbf{C}'}(f)$ is defined by $|s_{\mathbf{C}'}(f)|=|s_{\mathbf{C}}(f)|$, and the commutative diagram

$$\begin{array}{ccc} |s_{\mathbf{C}}(f)| & \xrightarrow{\quad \theta_f \quad} & |s_{\mathbf{C}'}(f)| \\ & \circ & \\ {\scriptstyle s_{\mathbf{C}}(f)}\searrow & & \swarrow{\scriptstyle s_{\mathbf{C}'}(f)} \\ & O & \end{array}$$

(this diagram is obtained by decomposing (1) in the form

)

using the transition map θ_f for F. Given $f, g\in A$, $\hat{p}\in |s_{\mathbf{C}'}(g)|$ such that $t(f)=s_{\mathbf{C}'}(f)(\hat{p})$, we put $g\circ^{\mathbf{C}'}_{\hat{p}} f = g\circ^{\mathbf{C}}_{p} f$ with $p=\theta_f^{-1}(\hat{p})$. To define the amalgamating functions $\varphi'=\varphi_{\mathbf{C}'}[f, g, \hat{p}]$, $\psi'=\psi_{\mathbf{C}'}[f, g, \hat{p}]$, we use the commutative squares in

$$\begin{array}{ccc} |s_{\mathbf{C}}(g)|-\{p\} & \xrightarrow{\quad \theta_g\restriction \quad} & |s_{\mathbf{C}'}(g)|-\{\hat{p}\} \\ {\scriptstyle\psi}\downarrow & \circ & \downarrow{\scriptstyle\psi'} \\ |s_{\mathbf{C}}(g\circ_p f)| & \xrightarrow{\quad \theta_{g\circ_p f} \quad} & |s_{\mathbf{C}'}(g\circ_{\hat{p}} f))| \\ {\scriptstyle\varphi}\uparrow & \circ & \uparrow{\scriptstyle\varphi'} \\ |s_{\mathbf{C}}(f)| & \xrightarrow[\quad \theta_f \quad]{} & |s_{\mathbf{C}'}(f)| \end{array} .$$

where φ and ψ are the amalgamating functions given with $\boldsymbol{C}$. It is immediate that $\boldsymbol{C}'$ is well-defined. The transition maps for Φ are the given θ_f; the effect of F' on objects and arrows is that of F.

4. The free multicategory

We need only the free multicategory in the case there is only one level of objects; therefore, we restrict the definition to this case. However, note that the free one-level multicategory will be free with respect to the general, two-level, variety.

Suppose O is a set (of "objects"), L is a set (of "generating arrows"), and for each $f\in L$, we are given $\mathrm{s}(f)\in O^*$, and $\mathrm{t}(f)\in O$. Such data determine a *language* $\mathcal{L}$; we may write $O=\mathrm{O}(\mathcal{L})$, $L=\mathrm{L}(\mathcal{L})$, $\mathrm{s}=\mathrm{s}_{\mathcal{L}}$, $\mathrm{t}=\mathrm{t}_{\mathcal{L}}$. The free multicategory $\mathcal{F}(\mathcal{L})=\boldsymbol{C}$ on the given language is defined by the universal property as follows. We have that $\mathrm{O}(\mathcal{L})=\mathrm{O}(\boldsymbol{C})$, $\mathrm{L}(\mathcal{L})\subset\mathrm{A}(\boldsymbol{C})$, $\mathrm{s}_{\boldsymbol{C}}$, $\mathrm{t}_{\boldsymbol{C}}$ extend the given maps $\mathrm{s}_{\mathcal{L}}$ and $\mathrm{t}_{\mathcal{L}}$; and every time $\boldsymbol{D}$ is a multicategory, and we are given $F(X)\in\mathrm{O}(\boldsymbol{D})$, $F(f)\in\mathrm{O}(\boldsymbol{D})$ for $X\in O$, $f\in L$ such that $\mathrm{t}_{\boldsymbol{D}}(F(f))=(F(\mathrm{t}_{\mathcal{L}}(f)))$, and we are also given $\theta_f\colon |\mathrm{s}_{\mathcal{L}}(f)| \xrightarrow{\cong} |\mathrm{s}_{\boldsymbol{D}}(Ff)|$ such that

$$\begin{array}{ccc}
|\mathrm{s}_{\mathcal{L}}(f)| & \xrightarrow{\ \theta_f\ } & |\mathrm{s}_{\boldsymbol{D}}(Ff)| \\
{\scriptstyle \mathrm{s}_{\mathcal{L}}(f)}\downarrow & \circ & \downarrow{\scriptstyle \mathrm{s}_{\boldsymbol{D}}(Ff)} \\
O(\mathcal{L}) & \xrightarrow[\ F\ ]{} & O(\boldsymbol{D})
\end{array}$$

(when θ_f is the identity, $\mathrm{s}_{\boldsymbol{D}}(Ff)=F\circ\mathrm{s}_{\mathcal{L}}(f)$) for all $f\in\mathrm{L}(\mathcal{L})$, there is a unique morphism $F\colon\boldsymbol{C}\to\boldsymbol{D}$ of multicategories extending the given data F and $\theta_{(\ )}$. The uniqueness, up to isomorphism, of $\mathcal{F}(\mathcal{L})$ is clear; its existence could be proved routinely by the Adjoint Functor Theorem (or, Initial Object Theorem; see [M L]). Instead, we will find a direct description and proof of existence for $\mathcal{F}(\mathcal{L})$.

We first formulate a characterization.

(1) Let $\mathcal{L}$ be a language as above. Suppose $\boldsymbol{C}$ is a 1-level multicategory with $\mathrm{O}(\boldsymbol{C})=\mathrm{O}(\mathcal{L})$, $\mathrm{L}(\mathcal{L})\subset\mathrm{A}(\boldsymbol{C})$, and $s_{\boldsymbol{C}}$, $t_{\boldsymbol{C}}$ extend $\mathrm{s}_{\mathcal{L}}$ and $\mathrm{t}_{\mathcal{L}}$, respectively. Then $\boldsymbol{C}$ is free on $\mathcal{L}$ **if and only if** the following condition (2) holds:

(2) ***(Unique Readability)*** for every $\alpha \in \mathrm{A}(\boldsymbol{C})$,

either **(a)** $\alpha = 1_X$ for some $X \in \mathrm{O}(\mathcal{L})$,
or **(b)** $\alpha = f(\alpha_1, \alpha_2, \ldots, \alpha_m)$ for some $f \in \mathrm{L}(\mathcal{L})$,
$m = \ell\mathrm{h}(\mathrm{s}(f))$, $\alpha_i \in \mathrm{A}(\boldsymbol{C})$ such that $\mathrm{t}(\alpha_i) = \mathrm{s}(f)(i)$ $(i \in [1, m])$

($f(\alpha_1, \alpha_2, \ldots, \alpha_m)$ refers to simultaneous composition; see the end of section 1.);

and furthermore,

exactly one of **(a), (b)** is the case; in case **(a)**, X is uniquely determined by α , and in case **(b)**, the items f , α_i are uniquely determined by α .

Note that there are no additional conditions put on the amalgamating functions.

Note that in case **(b)**, for each i , $\ell\mathrm{h}(\mathrm{s}(\alpha_i)) < \ell\mathrm{h}(\mathrm{s}(\alpha))$, which fact implies that under (2), $\mathrm{A}(\boldsymbol{C})$ is generated by $\mathrm{L}(\mathcal{L})$ in the obvious sense: $\mathrm{A}(\boldsymbol{C})$ is the least set $\mathcal{X}$ containing each 1_X $(X \in \mathrm{O}(\mathcal{L}))$ and such that if $f \in L(\mathcal{L})$, $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathcal{X}$ and $f(\alpha_1, \alpha_2, \ldots, \alpha_m)$ is well-defined, then $f(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathcal{X}$. In fact, if the condition (2) holds, we may apply *structural induction*, respectively, *structural recursion*, to prove that a property holds for all arrows of $\boldsymbol{C}$, respectively, to define a function, say Φ , whose domain is $\mathrm{A}(\boldsymbol{C})$. In the latter case, we should have the definition of the function Φ at arguments 1_X , $X \in \mathrm{O}(\boldsymbol{C})$, and a way that determines the value of Φ at any argument of the form $f(\alpha_1, \alpha_2, \ldots, \alpha_m)$ $(f \in \mathrm{L}(\mathcal{L}))$ from the following data: $f, \alpha_1, \alpha_2, \ldots, \alpha_m$ and $\Phi(\alpha_1), \Phi(\alpha_2), \ldots, \Phi(\alpha_m)$; unique readability ensures that thereby Φ is uniquely determined.

The proof of the **if** part consists in verifying the universal property of $\boldsymbol{C}$ under the condition (2). Let us use the notation in the statement of the universal property. The effect of F on the arrows α of $\boldsymbol{C}$, including the connecting maps θ_α , is defined by structural recursion on α . Of course, the amalgamating maps for the composition in $\boldsymbol{C}$ and those for the composition in $\boldsymbol{D}$ are used in this definition. The details are put into the Appendix.

(Note that the **if** part of (1) is an important piece in the justification of the generalized notion of multicategory introduced in this paper; the **if** part of (1) shows that the generalized notion is, after all, not so far from the standard concept of multicategory; in fact, the **if** part of (1) shows that, in a sense, the generalized notion is the algebraic essence of the standard notion.)

Next, for a given language $\mathcal{L}$, we exhibit a particular multicategory $\mathcal{F}(\mathcal{L})$ *with standard amalgamation* satisfying the condition (2). Then, by the **if** part already shown, $\mathcal{F}(\mathcal{L})$ is free on $\mathcal{L}$; and since any multicategory free on $\mathcal{L}$ is isomorphic to $\mathcal{L}$, and the condition (2) is clearly invariant under isomorphism, the "only if" part will follow.

$\boldsymbol{C}=\mathcal{F}(\mathcal{L})$, in the specific sense now to be adopted, is defined to have objects $\mathrm{O}(\boldsymbol{C})=\mathrm{O}(\mathcal{L})$. The arrows are defined inductively as follows:

(i) each $X\in S$ is an *arrow*; $\mathrm{s}_{\boldsymbol{C}}(X)=\langle X\rangle$, $\mathrm{t}_{\boldsymbol{C}}(f)=X$.

(ii) whenever $f\in L$ with $\ell\mathrm{h}(f)=n$, and, for each $i\in[1,n]$, α_i is an *arrow* such that $\mathrm{t}_{\boldsymbol{C}}(\alpha_i)=\mathrm{s}(f)(i)$, then

$$\alpha \overset{=}{\mathrm{def}} f(\langle\alpha_i\rangle_{i\in[1,m]}) = f(\alpha_1,\alpha_2,\ldots,\alpha_m) \tag{3}$$

is an *arrow*, and $\mathrm{s}_{\boldsymbol{C}}(\alpha)$ is the *concatenation* $\mathrm{s}_{\boldsymbol{C}}(\alpha_1)\hat{}\,\mathrm{s}_{\boldsymbol{C}}(\alpha_2)\hat{}\ldots\hat{}\,\mathrm{s}_{\boldsymbol{C}}(\alpha_{\mathrm{m}})$; that is, with $n=\ell\mathrm{h}(\mathrm{s}_{\boldsymbol{C}}(\alpha))$, $n_i=\ell\mathrm{h}(\mathrm{s}_{\boldsymbol{C}}(\alpha_i))$, we have $n=\sum_{i=1}^{n} n_i$, and for any $j\in[1,n]$, with $i\in[1,m]$ determined such that $j\in[(\sum_{h<i}n_h)+1,\ \sum_{h\le i}n_h])$, we have $\mathrm{s}_{\boldsymbol{C}}(\alpha)(j)=\mathrm{s}_{\boldsymbol{C}}(\alpha_i)(j-\sum_{h<i}n_h)$.

In (3), $f(\alpha_1,\alpha_2,\ldots,\alpha_m)$ means something determined from f and the α_i so that, conversely, f and the α_i can be recovered from it. Thus, $f(\alpha_1,\alpha_2,\ldots,\alpha_m)$ may be the concatenation of the strings $\langle f\rangle$, α_1, α_2, $\ldots$, α_n. This is all right if and only if the sets $\mathrm{O}(\mathcal{L})$ and $\mathrm{L}(\mathcal{L})$ are disjoint. In the general case, a set-theoretical construct such as $\langle 1, f, \alpha_1, \alpha_2, \ldots, \alpha_m\rangle$ (function on $[1, m+2]$ with values as listed) can be taken for $f(\alpha_1,\alpha_2,\ldots,\alpha_m)$; now, for clause (i), we take $\langle 0, X\rangle$ to be the arrow, rather than plain X.

The notation in (3) is in agreement with the notation for simultaneous composition, as will become clear when we have defined composition in $\mathcal{F}(\mathcal{L})$.

The composition in $\mathcal{F}(\mathcal{L})$ is defined by *substitution*. Given

$$s(\beta) \xrightarrow[p]{\beta} s(\alpha) \xrightarrow{\alpha} t(\alpha) \,,$$

$\alpha \circ_p \beta$ is defined as $\alpha(\beta/p)$, the result of *substituting* β in α for $t(\beta)=s(\alpha)(p)$ at the place p. The value of the expression $\alpha(\beta/p)$ is defined by recursion on the complexity of α. If $\alpha=X\in O$, and thus $s(\alpha)(p)=X$, then $\alpha(\beta/p)=\beta$. If $\alpha=f(\langle\alpha_i\rangle_{i\in[1,m]})$, then, using the notation adopted under (ii), for a specific $j\in[1,m]$, we $p=\sum_{h<j} n_j+q$ with $q\in[1,n_j]$; and we put $\alpha(\beta/p)=f(\langle\hat{\alpha}_i\rangle_{i\in[1,m]})$, where $\hat{\alpha}_i=\alpha_i$ when $i\in[1,m]-\{j\}$, and $\hat{\alpha}_j=\alpha_j(b/q)$. It is left to the reader to verify that in this way we have defined a multicategory with standard amalgamation.

5. A 2-level multicategory with non-standard amalgamation

Let $\mathcal{L}$ be a language, and $\boldsymbol{C}$ a (not necessarily standard, but 1-level) multicategory free over $\mathcal{L}$ (see section 4). For $\alpha\in A=\mathrm{A}(\boldsymbol{C})$, we define $\langle\alpha\rangle$ to be "the tuple of occurrences of operation symbols in α, listed from the left to the right". For a formal definition, we use unique redability (4.(1)), which enables us to employ a recursion. For $X\in O$, $\langle 1_X\rangle=\bot$, the empty tuple. For $f\in L(\mathcal{L})$, $m=\ell\mathrm{h}(\mathrm{s}(f))$, $\alpha_i\in A$, $n_i=\ell\mathrm{h}(\langle\alpha_i\rangle)$, $\alpha=f(\alpha_1, \alpha_2, ..., \alpha_m)\in A$, we put $\ell\mathrm{h}(\langle\alpha\rangle)=1+\sum_{i=1}^{m} n_i$, $\langle\alpha\rangle(1)=f$, and for $i\in[1, m]$, $k\in[1, n_i]$, $j=1+\sum_{h<i} n_h+k$, we define $\langle\alpha\rangle(j)=\langle\alpha_i\rangle(k)$. This is the same as saying that

$$\langle f(\alpha_1, \alpha_2, ..., \alpha_m)\rangle = \langle f\rangle^\wedge\langle\alpha_1\rangle^\wedge\langle\alpha_2\rangle^\wedge ...^\wedge\langle\alpha_m\rangle \,, \tag{1}$$

where we used the well-known operation of concatenation of tuples; $\langle f\rangle$ means the one-term tuple whose only term is f; of course, it is also the same as $\langle\alpha\rangle$ in the sense being defined now, for $\alpha=f$.

Let us fix a multicategory $\boldsymbol{E}$, not necessarily 1-level, or with standard amalgamation. A *free multicategory over* $\boldsymbol{E}$ is a system $(\mathcal{L},\boldsymbol{C},\mathrm{d}:\boldsymbol{C}\to\boldsymbol{E})$ where $\mathcal{L}$ is a language, $\boldsymbol{C}$ is a (1-level) multicategory free over $\mathcal{L}$ such that $(\mathrm{O}(\mathcal{L})=)\mathrm{O}(\boldsymbol{C})=\mathrm{O}(\boldsymbol{E})$, the morphism $\mathrm{d}:\boldsymbol{C}\to\boldsymbol{E}$ of multicategories is the identity on upper-level objects, and it is strict. A *morphism* $H:(\mathcal{L}_1,\boldsymbol{C}_1,\mathrm{d}_1)\to(\mathcal{L}_2,\boldsymbol{C}_2,\mathrm{d}_2)$ of free multicategories over $\boldsymbol{E}$ is a mapping $H:\mathrm{A}(\mathcal{L}_1)\to\mathrm{A}(\mathcal{L}_2)$ such that for any $f\in\mathrm{A}(\mathcal{L}_1)$, we have $\mathrm{s}_{\mathcal{L}_2}(H(f))=\mathrm{s}_{\mathcal{L}_1}(f)$, $\mathrm{t}_{\mathcal{L}_2}(H(f))=\mathrm{t}_{\mathcal{L}_1}(f)$, and for the induced strict morphism $H:\boldsymbol{C}_1\to\boldsymbol{C}_2$ (which is the identity on objects) we have the commutative diagram

$$\begin{array}{ccc} \boldsymbol{C}_1 & \xrightarrow{\;H\;} & \boldsymbol{C}_2 \\ & \circ & \\ {\scriptstyle \mathrm{d}_1}\searrow & & \swarrow{\scriptstyle \mathrm{d}_2} \\ & \boldsymbol{E} & \end{array} \quad .$$

We also say that H is a *morphism of languages*, and write $H:\mathcal{L}_1\to\mathcal{L}_2$ (note that by a "morphism of languages" one might *a priori* mean something more general; we do not need the more general concept).

Theorem. There is a uniquely determined assignment of a multicategory $\mathbf{D}=\mathbf{D}[\mathcal{L},\mathbf{C},\mathrm{d}]$ to any free multicategory $(\mathcal{L},\mathbf{C},\mathrm{d})$ over $\mathbf{E}$ such that conditions (i) to (vi) hold.

(i) $\mathrm{O}(\mathbf{D})=\mathrm{L}(\mathcal{L})$, $\dot{\mathrm{O}}(\mathbf{D})=\mathrm{A}(\mathbf{E})\times\mathrm{O}(\mathbf{E})$, $\mathrm{A}(\mathbf{D})=\mathrm{A}(\mathbf{C})$.

(ii) Using the abbreviation $\mathrm{T}(\alpha)=(\mathrm{d}(\alpha),\mathrm{t}_{\mathbf{C}}(\alpha))$ ($\alpha\in A=\mathrm{A}(\mathbf{D})=\mathrm{A}(\mathbf{C})$) the mapping $\mathrm{O}(\mathbf{D})\longrightarrow\dot{\mathrm{O}}(\mathbf{D})$ is $s\mapsto\langle\mathrm{T}(s(i))\rangle_{i\in|s|}$.

(iii) For $\alpha\in A$, we have $\mathrm{t}_{\mathbf{D}}(\alpha)=\mathrm{T}(\alpha)$ and $\mathrm{s}_{\mathbf{D}}(\alpha)=\langle\alpha\rangle$.

(iv) For $f\in\mathrm{L}=\mathrm{O}(\mathbf{D})$, $1_f^{(\mathbf{D})}=f$.

(v) Let us write $\square$ for $\circ^{(\mathbf{D})}$, and $\circ$ for $\circ^{(\mathbf{C})}$. Whenever $\alpha\square_p\beta$ is well-defined, we have that

$$\alpha=\alpha'\circ_q f(\alpha_1,\ldots,\alpha_n) \tag{2}$$

and

$$\alpha\square_p\beta \overset{\mathrm{def}}{=} \alpha'\circ_q\beta(\alpha_1,\ldots,\alpha_n) \tag{3}$$

for $f=\langle\alpha\rangle(p)$, and for some $\alpha',\alpha_1,\ldots,\alpha_n\in A$ and $q\in|\mathrm{s}_{\mathbf{C}}(\alpha)|$ (we are referring here to simultaneous composition in $\mathbf{C}$, discussed in the last section).

(vi) Whenever $H:(\mathcal{L}_1,\mathbf{C}_1,d_1)\to(\mathcal{L}_2,\mathbf{C}_2,d_2)$ is a morphism of free multicategories over $\mathbf{E}$, $\mathbf{D}_j=\mathbf{D}[\mathcal{L}_j,\mathbf{C}_j,d_j]$, the mappings

$$H:\mathrm{O}(\mathbf{D}_1)=\mathrm{L}(\mathcal{L}_1)\longrightarrow\mathrm{O}(\mathbf{D}_2)=\mathrm{L}(\mathcal{L}_2)$$
$$\mathrm{id}:\dot{\mathrm{O}}(\mathbf{D}_1)=\mathrm{A}(\mathbf{E})\times\mathrm{O}(\mathbf{E})\longrightarrow\dot{\mathrm{O}}(\mathbf{D}_2)=\mathrm{A}(\mathbf{E})\times\mathrm{O}(\mathbf{E})$$
$$H:\mathrm{A}(\mathbf{D}_1)=\mathrm{A}(\mathbf{C}_1)\longrightarrow\mathrm{A}(\mathbf{D}_2)=\mathrm{A}(\mathbf{C}_2)$$

constitute a strict morphism $H:\mathbf{D}_1\longrightarrow\mathbf{D}_2$ of multicategories.

The multicategory $\mathbf{D}=\mathbf{D}[\mathcal{L},\mathbf{C},\mathrm{d}]$ is called the *multicategory of function-replacement associated with* $(\mathcal{L},\mathbf{C},\mathrm{d})$. The name derives from the main clause, (v). This clause *tries to* *say* that

" $\alpha\square_p\beta$ results by replacing the "function-symbol" $f=\langle\alpha\rangle(p)$ at the place p in α by the arrow β ";

however, it actually says less, namely that

" $\alpha\square_p\beta$ results by replacing the "function-symbol" $f=\langle\alpha\rangle(p)$ at *some* place in α by the arrow β ".

Note the difficulty of saying the first of these two statements mathematically; this difficulty comes from the fact that f may occur at more than one place in α. The theorem avoids specifying the *particular* decomposition (2) that "belongs to" the place p , and still manages to give the complete definition of the concept of the multicategory of function replacement. The price we pay is that we do not have the definition of $\mathbf{D}[\mathcal{L},\mathbf{C},\mathrm{d}]$ for any *particular* $(\mathcal{L},\mathbf{C},\mathrm{d})$ spelled out in detail; rather, we have the complete definition of the *global* assignment $(\mathcal{L},\mathbf{C},\mathrm{d})\longmapsto\mathbf{D}[\mathcal{L},\mathbf{C},\mathrm{d}]$.

The rest of this section is devoted to the proof of the theorem; certain technical details will be relegated to the Appendix.

We prove the theorem in two stages. In the first, we fix $(\mathcal{L},\mathbf{C},\mathrm{d})$, a free multicategory over $\mathbf{E}$, and define the operations for $\mathbf{D}=\mathbf{D}[\mathcal{L},\mathbf{C},\mathrm{d}]$ partially, for certain combinations of arguments only, ones that we will call "separated". The second stage will involve the use of morphisms of free multicategories over $\mathbf{E}$ to complete the definition.

With the fixed $(\mathcal{L},\mathbf{C},\mathrm{d})$, we have $O=\mathrm{O}(\mathcal{L})=\mathrm{O}(\mathbf{C})$, $L=\mathrm{L}(\mathcal{L})$, $A=\mathrm{A}(\mathbf{C})$. α,β,γ denote elements of A , f, g elements of L .

The construction of $\mathbf{D}$ takes place in $\mathbf{C}$. The role $\mathbf{E}$ and d have in the construction is summarized in the following lemma:

(4) Lemma.

(i) $\mathrm{d}(\alpha)=\mathrm{d}(\beta)$ implies that $\mathrm{s}(\alpha)=\mathrm{s}(\beta)$.

(ii) $d(\alpha\circ_r\beta) = (d\alpha)\circ_r^{(\mathbf{E})}(d\beta)$,

and more generally

$$d(\alpha(\alpha_1,...,\alpha_n))=(d\alpha)(d\alpha_1,...,d\alpha_n))$$

whenever $\alpha\circ_r\beta$, $\alpha(\alpha_1, ..., \alpha_n)$ are well-defined; here, we refer to simultaneous composition, in $\mathbf{C}$ on the left and in $\mathbf{E}$ on the right.

(iii) Suppose that $d(\alpha_1)=d(\alpha_2)$, $d(\beta_1)=d(\beta_2)$ and $c(\beta_1)=c(\beta_2)$. Suppose $\alpha_1\circ_r\beta_1$ is well-defined. Then also,

(a) $\alpha_2\circ_r\beta_2$ is well-defined;

(b) $d(\alpha_1\circ_r\beta_1) = d(\alpha_2\circ_r\beta_2)$;

as a consequence,

(c) $s(\alpha_1\circ_r\beta_1) = s(\alpha_2\circ_r\beta_2)$;

and we have, for the amalgamating functions for $\mathbf{C}$, that

(d) $\varphi[\alpha_1, \beta_1, r]=\varphi[\alpha_2, \beta_2, r]$, $\psi[\alpha_1, \beta_1, r]=\psi[\alpha_2, \beta_2, r]$.

This is essentially immediate from the fact that $d:\mathbf{C}\to\mathbf{E}$ is a strict morphism which is the identity on objects; here are some details.

For (i): since $d:\mathbf{C}\to\mathbf{E}$ is a strict morphism which is the identity on objects, $s_{\mathbf{E}}(d(\alpha))=s_{\mathbf{C}}(\alpha)=s(\alpha)$. Therefore, $d(\alpha)=d(\beta)$ implies $s(\alpha)=s(\beta)$.

For (ii): this is a consequence of the fact that d is a strict morphism $d:\mathbf{C}\to\mathbf{E}$ of multicategories.

For (iii):

Remember that $\alpha\circ_r\beta$ is well-defined iff $r\in|s(\alpha)|$ and $s(\alpha)(r)=c(\beta)$. Thus, (a) is clear.

Writing α,β for α_i, β_i , for either $i=1$ or $i=2$, we have $d(\alpha\circ_r\beta)=d(\alpha)\circ_r^{\mathbf{E}}d(\beta)$, implying (b). By the strictness of d , $\varphi[\alpha,\beta,r]=\varphi^{\mathbf{E}}[d(\alpha),d(\beta),r]$, and similarly for ψ . The equalities $d(\alpha_1)=d(\alpha_2)$, $d(\beta_1)=d(\beta_2)$ now clearly imply (d).

Let us write $\|\alpha\|$ for the range of the function $\langle\alpha\rangle$, the set of function-symbols occurring in α . The definition of $\langle\alpha\rangle$ gives that

$$\|f(\alpha_1, \ldots, \alpha_n)\| = \{f\}\cup\|\alpha_1\|\cup\ldots\cup\|\alpha_n\|$$

and by induction on α, we see that

$$\|\alpha\circ_q\beta\| = \|\alpha\| \cup \|\beta\| . \tag{5}$$

Note the obvious fact that, for $\alpha, \beta\in A$, the existence of at least one map $\langle\beta\rangle \longrightarrow \langle\alpha\rangle$ in $L^\#$ is equivalent to the condition $\|\beta\|\subset\|\alpha\|$.

Let $\alpha\in A$ and let $f\in\|\alpha\|$. Any representation of α in the form of (2), with suitable α', etc., is called a *decomposition of* α *at* f.

(6) Lemma. Assume $f\in\|\alpha\|$. There is at least one decomposition of α at f.

Proof: see Appendix.

α is *separated* if $\langle\alpha\rangle$ is a repetition-free tuple: the function $\langle\alpha\rangle: |\langle\alpha\rangle| \longrightarrow L$ is one-to-one. A system $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ of terms is *separated* if each α_i is separated, and the ranges $\|\alpha_i\|$ are pairwise disjoint sets. Note the obvious fact that if α is separated, β is any term, then there is at most one arrow $\langle\beta\rangle \to \langle\alpha\rangle$ in $L^\#$. For $f\in L$, the well-defined term $f(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is separated iff the system $(f, \alpha_1, \alpha_2, \ldots, \alpha_n)$ is separated. By induction on α, we see that $\alpha\circ_q\beta$ is separated iff (α, β) is a separated system; in particular, if $\alpha\circ_q\beta$ is separated, then $\|\alpha\|\cap\|\beta\|=\emptyset$.

The following lemma is intuitively obvious.

(7) Lemma. For a separated term α, the decomposition at $f\in\|\alpha\|$ (see (2)) is unique.

Proof: see Appendix.

Let $\alpha, \beta \in A$, and $f \in L$. We declare $\alpha \square_f \beta$ to be *well-defined* if and only if α is separated, $f \in \|\alpha\|$, and $\mathrm{T}(f)=\mathrm{T}(\beta)$ (for T, see (ii) of the theorem); if so, $\alpha \square_f \beta$ is given by the expression (3), that is,

$$\alpha \square_f \beta = \alpha' \circ_q \beta(\alpha_1, \ldots, \alpha_n) ,$$

where we refer to (2), the (unique) decomposition of α at f (by (6) and (7)). Note that, instead of a "place" p, we now have a function-symbol f in the subscript position. We still have to see that the expression defining $\alpha \square_f \beta$ is well-defined.

Note that $\mathrm{T}(f)=\mathrm{T}(\beta)$ implies that $\mathrm{s}(f)=\mathrm{s}(\beta)$ and $\mathrm{t}(f)=\mathrm{t}(\beta)$. The simultaneous composition $\beta(\alpha_1, \ldots, \alpha_n)$ is well-defined since $(\mathrm{s}(\beta)(i))=(\mathrm{s}(f)(i))=\mathrm{t}(\alpha_i)$, the second equality from the fact that $f(\alpha_1, \ldots, \alpha_n)$ is well-defined. The composition at q is well-defined since

$$(\mathrm{s}(\alpha')(q)) = \mathrm{t}(f) = \mathrm{t}(\beta) = \mathrm{t}(\beta(\alpha_1, \ldots, \alpha_n)) ;$$

the first equality holding since (2) is well-defined.

The first thing we check is that

$$\mathrm{T}(\alpha \square_f \beta)=\mathrm{T}(\alpha) . \tag{8}$$

provided $\alpha \square_f \beta$ is well-defined. Applying d to the expressions (2) and (3), and applying (4) repeatedly, we get that

$$\mathrm{d}(\alpha) = \mathrm{d}(\alpha') \circ_q \mathrm{d}(f)(\mathrm{d}(\alpha_1), \ldots, \mathrm{d}(\alpha_n)) ,$$

and

$$\mathrm{d}(\alpha \square_f \beta) = \mathrm{d}(\alpha') \circ_q \mathrm{d}(\beta)(\mathrm{d}(\alpha_1), \ldots, \mathrm{d}(\alpha_n)) ;$$

on the right-hand side, we have simultaneous composition in $\boldsymbol{E}$. The equality $\mathrm{d}(f)=\mathrm{d}(\beta)$ ensures that $\mathrm{d}(\alpha \square_f \beta)=\mathrm{d}(\alpha)$. We also have that $\mathrm{c}(\alpha \square_f \beta)=\mathrm{c}(\alpha')=\mathrm{c}(\alpha)$. By (4), it follows that (8) holds.

The definition of $\alpha\Box_f\beta$ gives immediately that we have

$$\|\alpha\Box_f\beta\| = (\|\alpha\| - \{f\}) \cup \|\beta\| \; . \qquad (9)$$

Next, we claim

(10) Lemma. Assuming the pair (α, β) is separated and $f \in \|\alpha\circ_r\beta\|$,

$$(\alpha\circ_r\beta)\Box_f\gamma = \begin{cases} (\alpha\Box_f\gamma)\circ_r\beta & \text{if } f\in\|\alpha\| \\ \alpha\circ_r(\beta\Box_f\gamma) & \text{if } f\in\|\beta\| \end{cases}$$

(Note that the same place r appears on the two sides. The upper right-hand occurrence of the composition $\circ_r$ at r is meaningful, since by (8) and (4), $s(\alpha\Box_f\gamma)=s(\alpha)$.)

Proof: see Appendix.

Note that, together with the equality $f\Box_f\beta=\beta$ (provided $f\Box_f\beta$ is well-defined), (10) determines the value of $\alpha\Box_f\gamma$ in all cases, since the generating arrows $f\in L$, together with the identities generate A . Of course, (10) cannot be used to define $\Box$, directly at least, since terms can, in general, be written in the form $\alpha\circ_r\beta$ in more than one way.

The definition of $\alpha\Box_f\beta$ in terms of $\circ_{\mathbf{c}}$, and what we know about separatedness and $\circ_{\mathbf{c}}$, makes it clear that if the pair (α, β) is separated, then so is the term $\alpha\Box_f\beta$ (a little less would in fact suffice).

We are ready to state and prove the associative and commutative laws for the separated case; the proof uses (10).

(11) Lemma. Assume that the triple (α, β, γ) is separated, $f\in\|\alpha\|$, $g\in\|\alpha\Box_f\beta\|=(\|\alpha\|-\{f\})\cup\|\beta\|$, $T(\beta)=T(f)$ and $T(\gamma)=T(g)$. Then

$$(\alpha\Box_f\beta)\Box_g\gamma = \begin{cases}(\alpha\Box_g\gamma)\Box_f\beta & \text{if } g\in\|\alpha\| \\ \alpha\Box_f(\beta\Box_g\gamma) & \text{if } g\in\|\beta\|\end{cases} .$$

(Note that the assumption implies that the left-hand side is well-defined. Note also that the right-hand expressions are well-defined too; if $g\in\|\beta\|$, we have, by (8), $\mathrm{T}(\beta\Box_g\gamma)=\mathrm{T}(\beta)=\mathrm{T}(f)$, making $\alpha\Box_f(\beta\Box_g\gamma)$ well-defined.)

Proof: see Appendix.

Assume again that (α, β) is separated, and $f\in\|\alpha\|$, so $\alpha\Box_f\beta$ is well-defined. Let $p\in|\langle\alpha\rangle|$ such that $\langle\alpha\rangle(p)=f$. As we have noted, the separatedness of $\alpha\Box_f\beta$ ensures there is at most one morphism $\langle\alpha\rangle\backslash p\longrightarrow\langle\alpha\Box_f\beta\rangle$, and at most one $\langle\beta\rangle\longrightarrow\langle\alpha\Box_f\beta\rangle$. But also, since $\|\alpha\Box_f\beta\|=(\|\alpha\|-\{f\})\cup\|\beta\|$, there are such morphisms

$$\langle\alpha\rangle\backslash p\xrightarrow{\ \psi\ }\langle\alpha\Box_f\beta\rangle\xleftarrow{\ \varphi\ }\langle\beta\rangle , \qquad (15)$$
$$\psi=\psi_\Box(\alpha, \beta, f) ,\ \varphi=\varphi_\Box(\alpha, \beta, f) ;$$

we have defined the amalgamating maps for the composition $\circ_{\boldsymbol{D}}=\Box$, partially, for the "separated case". Finally, we note that, provided the triple (α, β, γ) is separated, each one of the diagrams made up of amalgamating maps for the composition $\circ_{\boldsymbol{D}}=\Box$, associated with either the associative law $\delta\overset{=}{\mathrm{def}}(\alpha\Box_f\beta)\Box_g\gamma=\alpha\Box_f(\beta\Box_g\gamma)$ in 1.(xi) or the commutative law $\delta\overset{=}{\mathrm{def}}(\alpha\Box_f\beta)\Box_g\gamma=(\alpha\Box_f\beta)\Box_g\gamma$ in 1.(xii) as the case may be, is automatically commutative, by the separatedness of the term δ , which implies that into $\langle\delta\rangle$ from any other object of $L^{\#}$ there is at most one morphism.

This completes the work of establishing the multicategory structure in the restricted sense of applying to the "sufficiently separated" arguments. We now enter the second stage of the proof of the theorem.

Let $F:(\hat{\mathcal{L}},\hat{\boldsymbol{C}},\hat{\mathrm{d}})\longrightarrow(\mathcal{L},\boldsymbol{C},\mathrm{d})$ be a morphism of free multicategories over $\boldsymbol{E}$. We say that F is *ample*, or that $(\hat{\mathcal{L}},\hat{\boldsymbol{C}},\hat{\mathrm{d}})$ is an *ample expansion of* $(\mathcal{L},\boldsymbol{C},\mathrm{d})$ *via* F , if for each $f\in L$ there are infinitely many distinct $\hat{f}\in\hat{L}$ such that $F(\hat{f})=f$.

It is almost obvious that we have

(12) Lemma. Any free multicategory $(\mathcal{L},\boldsymbol{C},\mathrm{d})$ over $\boldsymbol{E}$ has ample expansions.

Proof: see Appendix.

Let $F:\hat{\boldsymbol{C}}\rightarrow\boldsymbol{C}$ be an ample expansion . Since F is strict, for any $\beta\in\hat{A}=\mathrm{A}(\hat{\boldsymbol{C}})$, $\mathrm{s}_{\hat{\boldsymbol{C}}}(\beta)=\mathrm{s}_{\boldsymbol{C}}(F(\beta))$. We will write $\mathrm{s}(\beta)$, $\mathrm{t}(\beta)$ for $\mathrm{s}_{\hat{\boldsymbol{C}}}(\beta)$, $\mathrm{t}_{\hat{\boldsymbol{C}}}(\beta)$, respectively, just like in $\boldsymbol{C}$. We have, for any $\hat{\alpha}\in\hat{A}$, that

$$\langle F(\hat{\alpha})\rangle=F\circ\langle\hat{\alpha}\rangle\ :$$

$$\begin{array}{ccc} \langle F(\hat{\alpha})\rangle| = |\langle\hat{\alpha}\rangle| & \xrightarrow{\ \langle\hat{\alpha}\rangle\ } & \hat{L} \\ & {\scriptstyle\langle F(\hat{\alpha})\rangle}\searrow \quad \circ \quad \swarrow{\scriptstyle F} & \\ & L & \end{array} \tag{13}$$

Of course, "separated" terms in $\hat{\boldsymbol{C}}$ are meant as they were in $\boldsymbol{C}$.

Here is another "obvious" lemma.

(14) Lemma. $F:\hat{\boldsymbol{C}}\longrightarrow\boldsymbol{C}$ is surjective: For any $\alpha\in L$, there is at least one separated $\beta\in\hat{L}$ such that $F(\beta)=\alpha$. In fact, for any $\alpha\in L$, and any finite set I , there is at least one separated $\beta\in\hat{L}$ such that $F(\beta)=\alpha$ and $\|\beta\|\cap I=\emptyset$.

Proof: see Appendix.

We use the notations $\hat{\mathrm{c}}(\hat{\alpha})=\mathrm{t}_{\hat{\boldsymbol{C}}}(\hat{\alpha})$, $\hat{\mathrm{T}}(\hat{\alpha})=(\hat{\mathrm{d}}(\hat{\alpha}),\hat{\mathrm{c}}(\hat{\alpha}))$ for $\hat{\alpha}\in\hat{A}$. Since $\hat{\mathrm{d}}(\hat{\alpha})=\mathrm{d}(F\hat{\alpha})$, $\hat{\mathrm{c}}(\hat{\alpha})=\mathrm{c}(F\hat{\alpha})$, we have that $\hat{\mathrm{T}}(\hat{\alpha})=\mathrm{T}(\alpha)$ for $\alpha=F(\hat{\alpha})$.

Now, we can define $\alpha\Box_p\beta$ for any $\alpha,\ \beta\in A=\mathrm{A}(\boldsymbol{C})$, and $p\in|\langle\alpha\rangle|$ such that, for $f=\langle\alpha\rangle(p)$, $\mathrm{T}(\beta)=\mathrm{T}(f)$. Let $\hat{\alpha}\in\hat{A}$ be separated such that $F(\hat{\alpha})=\alpha$ by (14); next, let $\hat{\beta}\in\hat{A}$ be separated such that $F(\hat{\beta})=\beta$ and $\|\beta\|\cap\|\alpha\|=\emptyset$ by (14) again. Let $\hat{f}=\langle\hat{\alpha}\rangle(p)$. By (13), $F(\hat{f})=f$; we have $\hat{\mathrm{T}}(\hat{\beta})=\hat{\mathrm{T}}(\hat{f})$, and $\hat{\alpha}\Box_p\hat{\beta}=\hat{\alpha}\Box_{\hat{f}}\hat{\beta}$ is well-defined. We define

$$\alpha\Box_p\beta \overset{\text{def}}{=} F(\hat{\alpha}\Box_{\hat{f}}\hat{\beta}).$$

Further, we define the amalgamating maps

$$\langle\alpha\rangle\backslash p \xrightarrow{\ \psi\ } \langle\alpha\Box_p\beta\rangle \xleftarrow{\ \varphi\ } \langle\beta\rangle, \qquad (15)$$
$$\psi=\psi_\Box(\alpha,\ \beta,\ p),\ \ \varphi=\varphi_\Box(\alpha,\ \beta,\ p)$$

as the F-images of the maps

$$\langle\hat{\alpha}\rangle\backslash p \xrightarrow{\ \hat{\psi}\ } \langle\hat{\alpha}\Box_{\hat{f}}\hat{\beta}\rangle \xleftarrow{\ \hat{\varphi}\ } \langle\hat{\beta}\rangle,$$
$$\hat{\psi}=\psi_\Box(\hat{\alpha},\ \hat{\beta},\ \hat{f}),\ \ \varphi=\varphi_\Box(\hat{\alpha},\ \hat{\beta},\ \hat{f}).$$

As maps of sets, ψ and φ are the same as $\hat{\psi}$ and $\hat{\varphi}$, respectively; by (13) , e.g., $|\langle\alpha\rangle\backslash p| = |\langle\hat{\alpha}\rangle\backslash p|$, $|\langle\alpha\Box_p\beta\rangle| = |\langle\hat{\alpha}\Box_{\hat{f}}\hat{\beta}\rangle|$, thus, we can define $\psi\colon |\langle\alpha\rangle\backslash p| \longrightarrow |\langle\alpha\Box_p\beta\rangle|$ as $\psi=\hat{\psi}$; similarly, $\varphi=\hat{\varphi}$; then by (13) again, φ and ψ are maps as in (15). It is also obvious that $(\varphi,\ \psi)$ are the coprojections of a coproduct in $L^{\#}$.

Let us show that this definition is legitimate: that is, the result does not depend on the choice of the ample expansion $F:\hat{\boldsymbol{C}}\to\boldsymbol{C}$, and the choice of $\hat{\alpha},\hat{\beta}$.

Let $\alpha,\beta\in\mathrm{A}(\boldsymbol{C})$, $p\in|\langle\alpha\rangle|$, $f=\langle\alpha\rangle(p)$. Let $\hat{\boldsymbol{C}},\hat{\alpha},\hat{\beta}$ and $\hat{f}$ be as above. Assume $G:\check{\boldsymbol{C}}=\mathcal{F}(\check{\mathcal{L}})\to\boldsymbol{C}$ is another ample expansion, $(\check{\alpha},\ \check{\beta})$ a separated pair of terms in $\check{A}=\mathrm{A}(\check{\boldsymbol{C}})$

such that $G(\tilde{\alpha})=\alpha$, $G(\tilde{\beta})=\beta$; let $\tilde{f}=\langle\tilde{\alpha}\rangle(p)$; we want to show that

$$F(\hat{\alpha}\square_{\hat{f}}\hat{\beta})=G(\tilde{\alpha}\square_{\tilde{f}}\tilde{\beta}) \ , \tag{16}$$

and that the amalgamating maps φ,ψ also come out to be the same when we use the new data.

We claim that there is a morphism $H:\hat{\mathcal{L}}\to\tilde{\mathcal{L}}$ of languages (see above) such that

$$\begin{array}{ccc} \hat{\mathcal{L}} & \xrightarrow{\quad H \quad} & \tilde{\mathcal{L}} \\ & F \searrow \quad \circ \quad \swarrow G & \\ & \mathcal{L} & \end{array} \tag{17}$$

and such that $H\circ\langle\hat{\alpha}\rangle=\langle\tilde{\alpha}\rangle$, $H\circ\langle\hat{\beta}\rangle=\langle\tilde{\beta}\rangle$ for the particular α and β given to us.. The latter two conditions determine the effect of H on the subset $\|\hat{\alpha}\|\dot{\cup}\|\hat{\beta}\|$ of $\hat{L}$; these conditions are possible to fulfill since $\langle\hat{\alpha}\rangle$, $\langle\hat{\beta}\rangle$ are injective functions, and $\|\hat{\alpha}\|\cap\|\hat{\beta}\|=\emptyset$. The restriction of H to $\|\hat{\alpha}\|\dot{\cup}\|\hat{\beta}\|$ so determined will satisfy what it has for (17) to hold. On the rest of the set $\hat{L}$, H can be defined arbitrarily, except for being subject to (17); the ampleness of G ensures that for every $\hat{g}\in\hat{L}$ there is $\tilde{g}\in\tilde{L}$ such that $G(\tilde{g})=F(\hat{g})$; we may put $H(\hat{g})=\tilde{g}$.

H gives rise to a morphism $H:\hat{\mathbf{C}}\longrightarrow\tilde{\mathbf{C}}$ for which the transition isomorphisms $\theta_{\hat{g}}$ $(\hat{g}\in\hat{L})$ are all identities. H is the identity on objects. It follows from the strictness of F and G, and (17) that H is strict.

We claim that $H(\hat{\alpha})=\tilde{\alpha}$, $H(\hat{\beta})=\tilde{\beta}$. Note that for $\tilde{\alpha}_1=H(\hat{\alpha})$, $\tilde{\beta}_1=H(\hat{\beta})$, we have $G(\tilde{\alpha}_1)=G(\tilde{\alpha})$ and $G(\tilde{\beta}_1)=G(\tilde{\beta})$; also, $\langle\tilde{\alpha}_1\rangle=\langle\tilde{\alpha}\rangle=H\circ\langle\hat{\alpha}\rangle$, $\langle\tilde{\beta}_1\rangle=\langle\tilde{\beta}\rangle$. The assertion then follows from the following observation: if $\tilde{\gamma},\tilde{\delta}\in\tilde{L}$, $G(\tilde{\gamma})=G(\tilde{\delta})$, $\langle\tilde{\gamma}\rangle=\langle\tilde{\delta}\rangle$ then $\tilde{\gamma}=\tilde{\delta}$; this is proved by an induction on the length of $G(\tilde{\gamma})=G(\tilde{\delta})$.

The definition of $\hat{\alpha}\square_{\hat{f}}\hat{\beta}$ via the structure of $\hat{\boldsymbol{C}}$, and that of $\tilde{\hat{\alpha}}\square_{\tilde{\hat{f}}}\tilde{\hat{\beta}}$ via the structure of $\tilde{\hat{\boldsymbol{C}}}$ tell us that $H(\hat{\alpha}\square_{\hat{f}}\hat{\beta})=\tilde{\hat{\alpha}}\square_{\tilde{\hat{f}}}\tilde{\hat{\beta}}$ holds as a consequence of $H(\hat{\alpha})=\tilde{\hat{\alpha}}$, $H(\hat{\beta})=\tilde{\hat{\beta}}$, the facts that H is a strict morphism, and that H is the identity on objects. For this, one notes, in the first place, that the decomposition of $\hat{\alpha}$ at $\hat{f}$ is carried by H into the decomposition of $\tilde{\hat{\alpha}}$ at $\tilde{\hat{f}}$. Now, (16) follows by (17). The assertion on the amalgamating maps is also clear.

The condition (v) is clearly met by the construction.

We now prove assertion (vi) of the Theorem, even though we have not yet proved that $\boldsymbol{D}[\mathcal{L},\boldsymbol{C},\mathrm{d}]$ is a multicategory. Assume the data for (vi) as shown. The claim is that in this case we have

$$H(1_f^{(\boldsymbol{D}_1)}) = 1_{H(f)}^{(\boldsymbol{D}_2)} \tag{18}$$

$$H(\alpha\square_p^1\beta) = (H\alpha)\square_p^2(H\beta) \tag{19}$$

and

$$\varphi_{\square^1}[\alpha, \beta, p] = \varphi_{\square^1}[H\alpha, H\beta, p]\ ,\ \ \psi_{\square^1}[\alpha, \beta, p] = \psi_{\square^1}[H\alpha, H\beta, p]\ , \tag{20}$$

every time $\alpha\square_p^1\beta$ is well-defined ($\square^i$ is the composition in $\boldsymbol{D}_i$).

(18) is immediate. The proofs of (19) and (20) are also easy: having set up $F:\hat{\boldsymbol{C}}\to\boldsymbol{C}_1$, $\hat{\alpha}\overset{F}{\longmapsto}\alpha$, $\hat{\beta}\overset{F}{\longmapsto}\beta$ as is needed for the definition of $\alpha\square_p^1\beta$ as $\alpha\square_p^1\beta=F(\hat{\alpha}\square_{\hat{f}}^1\hat{\beta})$, we have a valid set-up $HF:\hat{\boldsymbol{C}}\to\boldsymbol{C}_2$, $\hat{\alpha}\overset{HF}{\longmapsto}H\alpha$, $\hat{\beta}\overset{HF}{\longmapsto}H\beta$ for the disambiguated definition of $(H\alpha)\square_p^2(H\beta)$ as $(H\alpha)\square_p^2(H\beta)=(HF)(\hat{\alpha}\square_{\hat{f}}^1\hat{\beta})=(H(F(\hat{\alpha}\square_{\hat{f}}^1\hat{\beta}))$, which shows (19); (20) is similarly seen.

The relation

$$\mathrm{T}(\alpha\square_p\beta) = \mathrm{T}(\alpha)$$

follows instantly from the definition of $\alpha\square_p\beta$, the variant (8) of the same law established in the "separated case", and applied in $\hat{\mathbf{C}}$, and by the fact $\hat{\mathrm{T}}(\hat{\alpha})=\mathrm{T}(\alpha)$ when $F(\hat{\alpha})=\alpha$.

The rest of the required laws for $\square$ are also easily established. We treat the associative law.

Suppose $\alpha, \beta, \gamma\in A$, $\alpha\square_p\beta$ and $\beta\square_q\gamma$ are well-defined, and $\bar{q}=\varphi_\square[\alpha, \beta, p](q)$, to see that $(\alpha\square_p\beta)\square_{\bar{q}}\gamma=\alpha\square_p(\beta\square_q\gamma)$. Choose an ample $F:(\hat{\mathcal{L}},\hat{\mathbf{C}},\hat{\mathrm{d}})\to(\mathcal{L},\mathbf{C},\mathrm{d})$ and by (14), choose arrows $\hat{\alpha},\hat{\beta},\hat{\gamma}\in\mathrm{A}(\hat{\mathbf{C}})$ such that $\hat{\alpha}$ is separated, $F(\hat{\alpha})=\alpha$, $\hat{\beta}$ is separated, $\|\hat{\beta}\|\cap\|\hat{\alpha}\|=\emptyset$, $F(\hat{\beta})=\beta$, $\hat{\gamma}$ is separated, $\|\hat{\gamma}\|\cap(\|\hat{\alpha}\|\cup\|\hat{\beta}\|)=\emptyset$, and $F(\hat{\gamma})=\gamma$. Then, with $\hat{f}=\langle\hat{\alpha}\rangle(p)$, $\hat{g}=\langle\hat{\beta}\rangle(p)$, we have $(\hat{\alpha}\square_p\hat{\beta})\square_{\bar{q}}\hat{\gamma}=(\hat{\alpha}\square_{\hat{f}}\hat{\beta})\square_g\hat{\gamma}$, $\hat{\alpha}\square_p(\hat{\beta}\square_q\hat{\gamma})=\hat{\alpha}\square_{\hat{f}}(\hat{\beta}\square_g\hat{\gamma})$, and $(\hat{\alpha}\square_{\hat{f}}\hat{\beta})\square_g\hat{\gamma}=\hat{\alpha}\square_{\hat{f}}(\hat{\beta}\square_g\hat{\gamma})$ as (α, β, γ) is a separated triple, and in the separated case, we know that the associative law holds (see (11)). By (vi) (already proved), $F((\hat{\alpha}\square_p\hat{\beta})\square_{\bar{q}}\hat{\gamma}) = (\alpha\square_p\beta)\square_{\bar{q}}\gamma$ and $F(\hat{\alpha}\square_p(\hat{\beta}\square_q\hat{\gamma}))=\alpha\square_p(\beta\square_q\gamma)$. It follows that $(\alpha\square_p\beta)\square_{\bar{q}}\gamma = \alpha\square_p(\beta\square_q\gamma)$.

The commutative law and the commutative diagrams involving the amalgamating maps are established similarly.

We have left the treatment of the identity arrows for the end.

For $f\in L=\mathrm{O}(\mathbf{D})$, $1_f^{(\mathbf{D})}$ is defined to be $f\in A$ itself; since $\mathrm{s}_{\mathbf{D}}(f)=\langle f\rangle$, and $\mathrm{t}_{\mathbf{D}}(f)=\dot{f}=\mathrm{T}(f)$, the source and the target of $1_f^{(\mathbf{D})}$ are as they should be.

To see that the first unit law holds, let us assume first that $\beta\in A$ is separated, $p\in|\langle\beta\rangle|$, $f=\langle\beta\rangle(p)$, to see that $\beta\square_p f=\beta$ and for $\varphi=\varphi_\square[\beta, f, p]$, $\psi=\psi_\square[\beta, f, p]$, we have $\varphi(1)=p$ and $\psi(i)=i$ for $i\in[1, lh(\beta)]-\{p\}$. By condition (v) in the Theorem (which we have proved), it is clear that $\beta\square_p f=\beta$. The separatedness of β ensures that φ and ψ cannot be anything else but the ones described. Turning to the general case, let $F:\hat{\mathbf{C}}\to\mathbf{C}$ be an ample expansion, and $\hat{\beta}\in\hat{A}$ such that $F(\hat{\beta})=\beta$. Let $\hat{f}=\langle\hat{\beta}\rangle(p)$. Then, by what we just saw, $\hat{\beta}\Box_p\hat{f}=\hat{\beta}$ and for $\hat{\varphi}=\varphi[\hat{\beta}, \hat{f}, p]$, $\hat{\psi}=\psi[\hat{\beta}, \hat{f}, p]$, we have $\hat{\varphi}(1)=p$ and $\hat{\psi}(i)=i$ for $i\in[1, \ell h(\hat{\beta})]-\{p\}$. Since $\Box$ is preserved by F (condition (vi), already proved), the desired result follows.

The other unit law is similarly seen.

The uniqueness assertion in the Theorem is clear from what we have gone through.

6. Multitopic sets

An *multitopic set* S , by definition, consists of data (i) to (iii), subject to conditions (iv) to (viii):

(i) a sequence $\langle C_k \rangle_{k\in\mathbb{N}}$ of sets [to indicate dependence on S , we may write $C_k(S)$ for C_k , and similarly for the other ingredients to follow],

(ii) sequences $\langle \boldsymbol{C}_k \rangle_{k\in\mathbb{N}}$, $\langle \boldsymbol{D}_k \rangle_{k\in\mathbb{N}}$ of multicategories,

and

(iii) morphisms $\mathrm{d}_{k+1}:\boldsymbol{C}_{k+1}\longrightarrow\boldsymbol{D}_k$ $(k\in\mathbb{N})$ of multicategories

such that

(iv) $\boldsymbol{C}_0$ has only identity arrows, and $\mathrm{O}(\boldsymbol{C}_0)=C_0$;

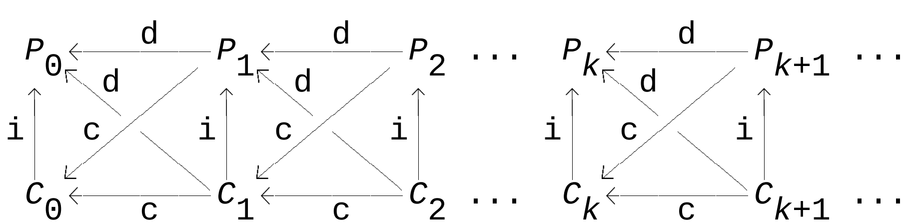
(v) for $k\geq 1$, $\boldsymbol{C}_k$ is free on a language $\mathcal{L}_k$ for which $\mathrm{O}(\mathcal{L}_k)=C_{k-1}$, $\mathrm{L}(\mathcal{L}_k)=C_k$;

(vi) $\boldsymbol{D}_0=\boldsymbol{C}_0$;

(vii) $\mathrm{O}(\boldsymbol{D}_k)=C_k=\mathrm{O}(\boldsymbol{C}_{k+1})$, and $\mathrm{d}_{k+1}:\boldsymbol{C}_{k+1}\longrightarrow\boldsymbol{D}_k$ is a strict morphism which is the identity on upper level objects;

(viii) for $k\geq 1$, $\boldsymbol{D}_k$ is the multicategory of function-replacement associated with $\boldsymbol{C}_k\xrightarrow{d_k}\boldsymbol{D}_{k-1}$ (see the previous section).

The multitopic set S gives rise to the following diagram of sets and functions:



Here, C_k is the set of k-*cells*, and it is given in (i) in the data for S . $P_k=A(\boldsymbol{C}_k)=A(\boldsymbol{D}_k)$; its elements are called the k-*pasting diagrams* of S . We have omitted subscripts from the maps; each should be understood with the same subscript as its domain; e.g., $\mathrm{d}_{k+1}:P_{k+1}\rightarrow P_k$, which is the effect on arrows of the morphism $\mathrm{d}_{k+1}:\boldsymbol{C}_{k+1}\rightarrow\boldsymbol{D}_k$.

$\mathrm{c}_{k+1}=\mathrm{t}_{\boldsymbol{C}_{k+1}}$.

In general, $\mathrm{d}(a)$ is the *domain*, $\mathrm{c}(a)$ is the *codomain* of a whatever a is to which the d or c in question applies. In particular, we talk about domain and codomain of both cells and pasting diagrams.

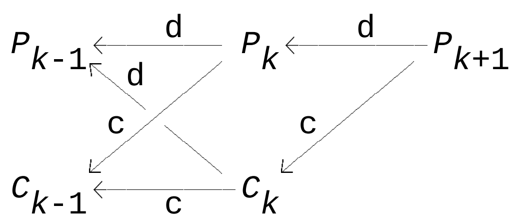
$i_k:C_k\rightarrow P_k$ is the inclusion of the generating arrows into the free $\boldsymbol{C}_k=\mathcal{F}(\mathcal{L}_k)$. d_k and c_k with domain C_k are the composites $\mathrm{d}_k\mathrm{i}_k$, $\mathrm{c}_k\mathrm{i}_k$, respectively.

We have

$$\mathrm{dd}=\mathrm{dc}\ ,\ \ \mathrm{cd}=\mathrm{cc}\ ;$$

in more detail,

$$\mathrm{d}_k\mathrm{d}_{k+1}=\mathrm{d}_k\mathrm{c}_{k+1}\ ,\ \mathrm{c}_k\mathrm{d}_{k+1}=\mathrm{c}_k\mathrm{c}_{k+1}\ : \qquad (1)$$



for all $k\in\mathbb{N}-\{0\}$. This is the familiar "globular" aspect of higher-dimensional category theory: it says that the domain and the codomain of a cell of dimension $k+1$ greater than 1 are *parallel* to each other, that is, they agree as far as their domains and codomains are concerned. Note however that domains and codomains here are very different things; the domains are pasting diagrams, the codomains are individual cells.

To see (1), let us abbreviate d_{k+1} and c_{k+1} by d and c , respectively. d on upper level objects is the identity id_{C_k} . Recall that, by the definition of $\boldsymbol{D}_k$ as in (viii) , we have that $\dot{a}$, the lower level object corresponding to the upper level object $a\in C_k$, is $\dot{a}=\mathrm{T}(a)=(\mathrm{d}_k(a),\mathrm{c}_k(a))$. Remember also that $\mathrm{t}_{\boldsymbol{D}_k}(\beta)=\mathrm{T}(\beta)=(\mathrm{d}_k(\beta),\mathrm{c}_k(\beta))$ ($\beta\in P_k$).

On lower level objects, the effect $\dot{\mathrm{d}}$ of d is forced by the required commutativity of

$$\begin{array}{ccc}
\mathrm{O}(\boldsymbol{C}_{k+1}) & \xrightarrow[\mathrm{id}]{\mathrm{d}} & \mathrm{O}(\boldsymbol{D}_{k}) \\
\mathrm{id}\Big\downarrow & \circ & \Big\downarrow\mathrm{T} \\
\dot{\mathrm{O}}(\boldsymbol{C}_{k+1}) & \xrightarrow[\dot{\mathrm{d}}]{} & \dot{\mathrm{O}}(\boldsymbol{D}_{k})
\end{array} ,$$

to be $\dot{\mathrm{d}}(a)=\mathrm{T}(a)$. Since $\mathrm{d}:\boldsymbol{C}_{k+1}\longrightarrow\boldsymbol{D}_{k}$ preserves "target", for any $\alpha\in P_{k+1}$, $\mathrm{t}_{\boldsymbol{D}_{k}}(\mathrm{d}\alpha))=\dot{\mathrm{d}}(\mathrm{t}_{\boldsymbol{C}_{k+1}}(\alpha))$, that is $\mathrm{T}(\mathrm{d}\alpha)=\mathrm{T}(\mathrm{c}\alpha)$, which is (1).

Note that $\mathrm{O}(\boldsymbol{D}_{k})=\mathrm{L}(\mathcal{L}_{k})=C_{k}$, $\dot{\mathrm{O}}(\boldsymbol{D}_{k})=\mathrm{A}(\boldsymbol{D}_{k-1})\times\mathrm{O}(\boldsymbol{C}_{k})=P_{k-1}\times C_{k-1}$, and we have $\mathrm{A}(\boldsymbol{D}_{k})=\mathrm{A}(\boldsymbol{C}_{k})$.

The fact that $\mathrm{d}_{k+1}:\boldsymbol{C}_{k+1}\longrightarrow\boldsymbol{D}_{k}$ is a strict morphism which is the identity on objects implies that $\mathrm{s}_{\boldsymbol{C}_{k+1}}(\alpha)=\mathrm{s}_{\boldsymbol{D}_{k}}(\alpha)=\langle\alpha\rangle$, where $\langle\alpha\rangle$ is defined as the left-to-right tuple of function-symbol occurrences in α ; see the previous section.

The fundamental equality is

$$\mathrm{d}(\alpha\circ_{p}\beta) = (\mathrm{d}\alpha)\square_{p}(\mathrm{d}\beta) \qquad\qquad (p\in|\langle\alpha\rangle| \ , \ \mathrm{c}\beta=\langle\alpha\rangle(p)),$$

signifying part of the fact that $\mathrm{d}=\mathrm{d}_{k+1}$ is a strict morphism of multicategories; here $\alpha, \beta\in C_{k+1}$, $\circ$ is the composition in $\boldsymbol{C}_{k+1}$, $\square$ is the composition in $\boldsymbol{D}_{k}$.

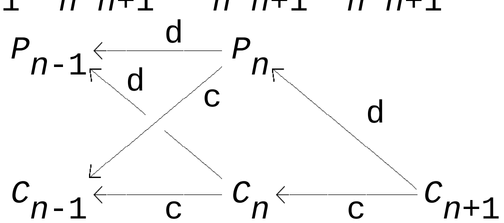
It is possible to build an multitopic set recursively. An n-*truncated multitopic set* is given by data $\langle C_{k}\rangle_{k\in[0, n]}$, $\langle\boldsymbol{C}_{k}\rangle_{k\in[0, n]}$, $\langle\boldsymbol{D}_{k}\rangle_{k\in[0, n-1]}$ as above in (i), (ii), (iii), but with the index k ranging over the integers 0 to n inclusive in the first two sequences, and up to $n-1$ in the last; the conditions (iv) to (viii) are assumed. Every (full) multitopic set S gives rise to its n-*truncation* $S\restriction n$, an n-truncated multitopic set; also, if $n\geq m$, an n-truncated one, S , gives rise to $S\restriction m$, an m-truncated one. On the other hand, if for each n , there is given S_{n} , an n-truncated multitopic set , and $S_{n+1}\restriction n=S_{n}$ for all n , then there is a unique multitopic set S for which $S\upharpoonright n=S_n$ for all n .

With C_0 an arbitrary set (of 0-cells), we let $\mathbf{C}_0=\mathcal{F}(\mathcal{L}_0)$ be the (free) multicategory whose objects are the 0-cells, and whose only arrows are identities ($\mathrm{O}(\mathcal{L}_0)=\mathrm{C}_0$, $\mathrm{L}(\mathcal{L}_0)=\emptyset$). We have $P_0=\mathrm{A}(C_0)\cong C_0$. This is all there is to a 0-truncated multitopic set.

After having determined C_0 , we let C_1 , the set of 1-cells, be any set, and we let $\mathrm{d}_1:C_1\to P_0=C_0$, $\mathrm{c}_1:C_1\to C_0$ be arbitrary functions (the domain and codomain assignments for 1-cells). We let $\mathbf{D}_0=\mathbf{C}_0$. The language $\mathcal{L}_1$ has $\mathrm{O}(\mathcal{L}_1)=C_0$, and $\mathrm{L}(\mathcal{L}_1)=C_1$; $\mathrm{s}_{\mathcal{L}_1}(f)=\langle \mathrm{d}_1(f)\rangle\in C_0^*$ (singleton tuple) , $\mathrm{t}_{\mathcal{L}_1}(f)=\mathrm{c}_1(f)\in C_0$. There is nothing to say about the amalgamating functions for $\mathbf{C}_1$.

Given an

n-truncated multitopic set S , with notation used above,
an arbitrary set C_{n+1} (of $(n+1)$-cells);
functions $\mathrm{c}_{n+1}:C_{n+1}\to C_n$, $\mathrm{d}_{n+1}:C_{n+1}\to P_n$ such that
$\mathrm{d}_n\mathrm{d}_{n+1}=\mathrm{d}_n\mathrm{c}_{n+1}$, $\mathrm{c}_n\mathrm{d}_{n+1}=\mathrm{c}_n\mathrm{c}_{n+1}$:



,

we have a uniquely determined $(n+1)$-truncated multitopic set $\hat{S}$ which extends the given data. To define $\hat{S}$, we first let $\mathbf{D}_n$ be the multicategory of function-replacement associated with $d_n:\mathbf{C}_n\to\mathbf{D}_{n-1}$.

Next, we define the language $\mathcal{L}_{n+1}$ to have $\mathrm{O}(\mathcal{L}_{n+1})=C_n$, $\mathrm{L}(\mathcal{L}_{n+1})=C_{n+1}$. For $f\in C_{n+1}$, $\mathrm{s}_{\mathcal{L}_{n+1}}(f) \overset{=}{def} \langle \mathrm{d}_{n+1}(f)\rangle\in C_n^*$; We let $t_{\mathcal{L}_{n+1}}(f) \overset{=}{def} \mathrm{c}_{n+1}(f)\in C_n$.

We define $\mathbf{C}_{n+1}^0$ (not yet $\mathbf{C}_{n+1}$) as the free multicategory on $\mathcal{L}_{n+1}$ with standard amalgamation.

To define $d^{\Theta}=d^{\Theta}_{n+1}:\mathbf{C}^{\Theta}_{n+1}\longrightarrow\mathbf{D}_n$, a morphism of multicategories, we use the freeness of $\mathbf{C}^{\Theta}_{n+1}$. d^{Θ} is defined on upper level objects as the identity:
$d^{\Theta}:O(\mathbf{C}^{\Theta}_{n+1})=C_n\longrightarrow O(\mathbf{D}_n)=C_n$ is id_{C_n}. On lower level objects, the effect of d^{Θ} is forced by the required commutativity of

$$\begin{array}{ccc}
O(\mathbf{C}^{\Theta}_{n+1}) & \xrightarrow[id]{d^{\Theta}} & O(\mathbf{D}_n) \\
{\scriptstyle id}\downarrow & \circ & \downarrow{\scriptstyle T} \\
\dot{O}(\mathbf{C}^{\Theta}_{n+1}) & \xrightarrow[\dot{d}^{\Theta}]{} & \dot{O}(\mathbf{D}_n)
\end{array} \quad ;$$

to be $\dot{d}^{\Theta}=T$, we put, for $a\in\dot{O}(\mathbf{C}^{\Theta}_{n+1})=C_n$, $\dot{d}^{\Theta}(a)=T(a)=(d_n(a),c_n(a))$.

On a generating arrow $f\in L(\mathcal{L}_{n+1})=C_{n+1}$, we put $d^{\Theta}(f)=d_{n+1}(f)$. We need to have that $t_{\mathbf{D}_n}(d^{\Theta}(f))=\dot{d}^{\Theta}(t_{\mathcal{L}_{n+1}}(f))$; but this means $T(d_{n+1}(f))=T(c_{n+1}(f))$, which reduces to $d_n(d_{n+1}(f))=d_n(c_{n+1}(f))$ and $c_n(d_{n+1}(f))=c_n(c_{n+1}(f))$, which are true by the assumptions we have made on d_{n+1}, c_{n+1}.

For $f\in C_{n+1}$, $s_{\mathcal{L}_{n+1}}(f)=\langle d_{n+1}(f)\rangle$, and $s_{\mathbf{D}_n}(d^{\Theta}(f))=\langle d_{n+1}(f)\rangle$; also, the effect of d^{Θ} on the upper level objects is the identity; therefore, it is legitimate to define the transition isomorphism $\theta_f:|s_{\mathcal{L}_{n+1}}(f)|\xrightarrow{\cong}|s_{\mathbf{D}_n}(d^{\Theta}(f))|$ (see section 3.) to be the identity.

The freeness of $\mathbf{C}^{\Theta}_{n+1}$ on $\mathcal{L}_{n+1}$ ensures the existence and uniqueness of $d^{\Theta}:\mathbf{C}^{\Theta}_{n+1}\longrightarrow\mathbf{D}_n$, a morphism of multicategories, extending the determination of d^{Θ} given on $\mathcal{L}_{n+1}$. In particular, d^{Θ} is the identity on the upper level objects. However, d^{Θ} is not, in general, a strict morphism, since $\mathbf{D}_n$ may have nonstandard amalgamation. We factor d^{Θ} as in

$$\begin{array}{ccc} \mathbf{C}^{0}_{n+1} & \xrightarrow{\quad d^{0} \quad} & \mathbf{D}_n \\ & {\scriptstyle \Phi} \searrow \quad \cong \quad \circ \quad \nearrow {\scriptstyle d_{n+1}} & \\ & \mathbf{C}_{n+1} & \end{array}$$

so that Φ is an isomorphism which acts as the identity on objects and arrows (but may be non-strict), and d_{n+1} is strict (see the end of section 2.). This is the definition of the desired $d_{n+1}:\mathbf{C}_{n+1}\longrightarrow\mathbf{D}_n$. Since $\mathbf{C}_{n+1}\cong\mathbf{C}^{0}_{n+1}$, by an isomorphism which is the identity on objects and arrows, $\mathbf{C}_{n+1}$ is also free on $\mathcal{L}_{n+1}$, with possibly nonstandard amalgamation. We have completed the definition of $\hat{S}$.

Putting $S_{n+1}=\hat{S}_n$, we produce a sequence of truncated multitopic sets S_n which togerther define a full multitopic set S . The definition of the S_n and S is by a non-deterministic recursion, with the data on the n-cells and their domains and codomains being parameters that are to some extent arbitrary.

7. The category of multitopes

Suppose $S=(C_n, \mathbf{C}_n, \mathbf{D}_n, d_n)_{n\in\mathbb{N}}$, $\bar{S}=(\bar{C}_n, \bar{\mathbf{C}}_n, \bar{\mathbf{D}}_n, \bar{d}_n)_{n\in\mathbb{N}}$ are multitopic sets. A *morphism* $\Phi:S\to\bar{S}$ *of multitopic sets* consists of maps $\Phi_n:C_n\to\bar{C}_n$ such that

Φ_n and Φ_{n+1} combine to induce a, necessarily unique, strict morphism of multicategories $\mathbf{C}_{n+1}\longrightarrow\bar{\mathbf{C}}_{n+1}$;

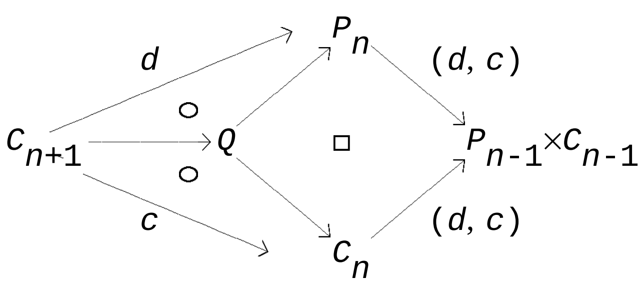
and

the Φ_n are compatible with the d's: $\Phi_{n-1}\circ\mathrm{d}_n=\mathrm{d}_n\circ\Phi_n$.

Because of the definition of $\mathbf{D}_n$ in terms of $\mathbf{C}_n$ and the d's , it follows that the Φ_n induce a strict morphism $\mathbf{D}_n\longrightarrow\bar{\mathbf{D}}_n$.

Under this definition of morphism, we have a category $\mathtt{MSet}$ of multitopic sets, with obvious composition and identities.

We are particularly interested in the terminal multitopic set $\mathcal{T}$. This is obtained if we stipulate that for each n , there be exactly one n-cell of any possible type; that is, exactly one 0-cell altogether; exactly one 1-cell altogether; and for each $n\geq 1$, for each pair $(\beta, b)\in P_n\times C_n$ such that $\mathrm{d}\beta=\mathrm{d}b$, $\mathrm{c}\beta=\mathrm{c}b$, there be exactly one $a\in C_{n+1}$ such that $da=\beta$, $ca=b$; in other words that for each $n\geq 1$, the mapping $C_{n+1}\longrightarrow Q$ in the diagram



induced by the pullback $Q=P_n\times_{P_{n-1}\times C_{n-1}}C_n$, be an isomorphism. It is easy to see, by going through the recursive buildup of any given multitopic set S as given in the previous section, that if $\mathcal{T}$ is such as described, then there is a unique map $S\to\mathcal{T}$: $\mathcal{T}$ is terminal.

On the other hand, the *existence* of $\mathcal{T}$ as described is also assured by the same recursive procedure. Note that in this case, the non-deterministic element of the recursion, the arbitrary choice of the $(C_{n+1}, d_{n+1}, c_{n+1})$, is eliminated by the condition given.

Let us use the notations $\mathbf{C}_n=\boldsymbol{C}_n(\mathcal{T})$, $\mathbf{D}_n=\boldsymbol{D}_n(\mathcal{T})$, $\mathrm{C}_n=C_n(\mathcal{T})$, $\mathrm{P}_n=P_n(\mathcal{T})$; we continue to use d, c and i without further specifying tags both for $\mathcal{T}$ and for other multitopic sets as they might come up. It will be convenient to have a new element $*$, and set $\mathrm{P}_{-1}=\{*\}$.

Given an arbitrary multitopic set S, we use the notation $\Phi:S\to\mathcal{T}$ for the terminal map, as well as for its various components. For any entity a in S, $\Phi(a)$ is its *type*.

The first remark is that $\mathrm{d}:\mathrm{C}_n\xrightarrow{\cong}\mathrm{P}_{n-1}$ ($n\geq 0$). This is true when $n=0$ and $n=1$. Let $n\geq 2$. We know that

$$\begin{array}{ccc} \mathrm{C}_n & \xrightarrow{\cong} & \mathrm{P}_{n-1}\times_{\mathrm{P}_{n-2}\times\mathrm{C}_{n-2}}\mathrm{C}_{n-1} \\ a & \longmapsto & (\mathrm{d}a, \mathrm{c}a) \end{array}.$$

Therefore, it suffices to show that

$$\begin{array}{ccc} \mathrm{P}_{n-1}\times_{\mathrm{P}_{n-2}\times\mathrm{C}_{n-2}}\mathrm{C}_{n-1} & \xrightarrow{\cong} & \mathrm{P}_{n-1} \\ (\beta, b) & \longmapsto & \beta \end{array}.$$

But for (β, b) as shown,

$$\mathrm{d}b=\mathrm{d}\beta\ ,\ \mathrm{c}b=\mathrm{c}\beta\ ; \tag{1}$$

and, by the definition of $\mathcal{T}$, with any given $\beta\in\mathrm{P}_{n-1}$, there is exactly one $b\in\mathrm{C}_{n-1}$ satisfying (1).

What the last-shown fact signifies is that the sets C_n may be dropped, the sets P_n may be used for them as well. However, to avoid confusion, we continue to use the C_n. When

$\rho\in \mathrm{P}_{n-1}$, $\bar{\rho}$ is $\mathrm{d}^{-1}(\rho)\in \mathrm{C}_n$.

Our aim here is to show that there is a specific category $\mathtt{Multitope}$, th*e category of multitopes,* such that the multitopic sets are the same as set-valued functors on $\mathtt{Multitope}$: $\mathtt{MSet}\simeq\mathtt{Set}^{\mathtt{Multitope}}$. The objects of $\mathtt{Multitope}$ are the elements of the P_n for all $n\in\mathbb{N}\cup\{-1\}$ (we write $\mathrm{P}_{-1}=\{*\}$); $\mathrm{Ob}(\mathtt{Multitope})=\dot{\bigcup}_{n\in\mathbb{N}\dot{\cup}\{*\}}\mathrm{P}_n$.

In what follows, S denotes an (arbitrary) multitopic set, with the notation for its ingredients we used before; $\Phi:S\to\mathcal{T}$, $\Phi(a)$ is the *type* of a.

Given any $\rho\in\mathrm{P}_{n-1}$, including the possibility $\rho=*\in\mathrm{P}_{-1}$, $C_n(\rho)$ is the set of n-cells of type $\bar{\rho}$: $C_n(\rho)=\{a\in C_n:\Phi(a)=\bar{\rho}\}$ for $n\geq 1$, and $C_0(*)=C_0$ for $n=-1$. Similarly, for $\rho\in\mathrm{P}_n$, let $P_n(\rho)=(\alpha\in P_n:\Phi(\alpha)=\rho\}$.

For emphasis, let us write $[\rho]$, or even $[\rho]_n$ (the subscript is n, not $n-1$, since $[\rho]_n$ is the "sort" for n-cells, not $(n-1)$-cells), for the object of $\mathtt{Multitope}$ corresponding to $\rho\in\mathrm{P}_{n-1}$ ($n\geq 0$). For the $\mathtt{Set}$-valued functor $\bar{S}:\mathtt{Multitope}\longrightarrow\mathtt{Set}$ corresponding to S, we will have $\bar{S}([\rho]_n)=C_n(\rho)$; $\bar{S}(*)=C_0$.

To identify the arrows of $\mathtt{Multitope}$, we have to do more work.

For $n\geq 1$, let us call $\alpha\in P_n$ *proper* if $\alpha\neq 1_b^{(C_n)}$ for any $b\in C_{n-1}$, i.e., if $|\langle\alpha\rangle|\neq\emptyset$; *improper* otherwise. For $n=0$, all $\alpha\in P_0=C_0$ are *proper*. We use "proper", "improper" for elements of P_n in a similar way.

Note the following fact. To know a *proper* pasting diagram $\alpha\in P_n$, it suffices to know its type $\rho=\Phi(\alpha)$ and the n-cells filling its places; in other words,

(1') supposing that $\alpha,\beta\in P_n(\rho)$ have the same *proper* type ρ, and for all $p\in|\langle\alpha\rangle|=|\langle\beta\rangle|=|\langle\rho\rangle|$, we have $\langle\alpha\rangle(p)=\langle\beta\rangle(p)$, then $\alpha=\beta$.

This is intuitively clear, and seen very easily by an induction on $\ell h(\langle\rho\rangle)$. Note, however, that the assignment $p\mapsto\langle\alpha\rangle(p)$ satisfies conditions, due to "links" within α; therefore, to determine α we need, beyond its type ρ, a suitably "linked" filling out its places with n-cells. Next, we give the description of these links.

Let $n\geq 1$, and $\rho\in P_n$. Recall that $s_{\mathbf{D}_n}(\rho)=\langle\rho\rangle$ and $s_{\mathbf{C}_n}(\rho)=\langle d\rho\rangle$; $p\in|\langle\rho\rangle|$ is a place where a "function-symbol", $\langle\rho\rangle(p)$, occurs in ρ; $r\in|\langle d\rho\rangle|$ is a place where a "variable", $\langle d\rho\rangle(r)$, occurs in ρ. If $\rho=f(\rho_1,\ldots\rho_m)$, $\langle\rho\rangle$ is the concatenation of $\langle f\rangle$ and the $\langle\rho_i\rangle$; we have the injections $\mu_i=\mu_i[f,\rho_1,\ldots\rho_m]:\langle\rho_i\rangle\longrightarrow\langle\rho\rangle$; $\mu_i(j)=1+\sum_{h<i}\ell h(\langle\rho_h\rangle)+j$: $\langle\rho\rangle$ is the coproduct of $\langle f\rangle$ and the $\langle\rho_i\rangle$ via the coprojections $\nu:|\langle f\rangle|\longrightarrow\rho:1\mapsto 1$ and the μ_i.

$\mathrm{Link}_1(\rho)$ is a set, specified below, of certain triples (p,q,s), the so-called 1-*links* of ρ, such that $p,q\in|\langle\rho\rangle|$, and for

$$\dot{p}=\langle\rho\rangle(p)\ ,\quad \dot{q}=\langle\rho\rangle(q)\ , \tag{2}$$

we have, in particular, $s\in|\langle d\dot{p}\rangle|$, and

$$\langle d\dot{p}\rangle(s)\ =\ c\dot{q}\ . \tag{3}$$

Intuitively, $(p,q,s)\in\mathrm{Link}_1(\rho)$ means that the the function-symbol occurrence of $\langle\rho\rangle(q)$ at q in ρ plugs directly into the occurrence of $\langle\rho\rangle(p)$ at p in ρ at the place s of the function-symbol $\langle\rho\rangle(p)$. To see an exmaple, let $X,Y\in C_{n-1}$, $a,b,c\in C_n$, $a:\langle Y,Y\rangle\to X$, $b:\langle X,Y\rangle\to Y$, $c:\langle X,Y,Y\rangle\to Y$ in the multicategory $\mathbf{C}_n$, and let

$$\rho\ =\ a(c(X,b(X,Y)),Y),b(a(Y,Y),Y))\ \in\ P_n\,.$$

Then $\langle\rho\rangle=\langle a,c,b,b,a\rangle$, and

$$\mathrm{Link}_1(\rho)=\{(1,2,1),(2,3,2),(1,4,2),(4,5,1)\}\ ;$$

for instance, $(4,5,1)\in\mathrm{Link}_1(\rho)$ since the second occurrence of a, which is at place 5 in ρ, plugs directly into the second occurrence of b, which is at place 4 in ρ, and this takes place at the 1st place of the function-symbol b.

Here is the formal, recursive, definition. When ρ is improper, $\mathrm{Link}_1(\rho)=\emptyset$. Let $\rho=f(\rho_1,\ldots\rho_m)$. We define

$$\mathrm{Link}_1(\rho)=\bigcup_{i=1}^{m}\{(\mu_i(p),\mu_i(q),s):\ (p,q,s)\in\mathrm{Link}_1(\rho_i)\}$$
$$\cup\ \{(1,\mu_i(1),i):\ i\in[1,m],\ |\langle\ \rho_i\rangle|\neq\emptyset\}\ .$$

The terms of the first union are there because all the 1-links in the ρ_i give rise to 1-links in ρ via the maps μ_i; the final term says that the head-occurrence (if any) of a function-symbol in ρ_i, which occurs at $\mu_i(1)$ in ρ, is plugged directly into the head-occurrence of f in ρ, at the argument-place i of f.

(3) (under (2)) is seen immediately for $(p,q,s)\in\mathrm{Link}_1(\rho)$.

Given ρ, let $\alpha\in P_n(\rho)$. For $p\in|\langle\alpha\rangle|=|\langle\rho\rangle|$, let us write $\hat{p}$ for $\langle\alpha\rangle(p)$; then $\dot{p}=\langle\rho\rangle(p)=\Phi(\hat{p})$; we have $\hat{p}\in C_n(\mathrm{d}\dot{p})$. Given $(p,q,s)\in\mathrm{Link}_1(\rho)$, we have

$$\langle\mathrm{d}\hat{p}\rangle(s)=\mathrm{c}\hat{q}\ ; \tag{3'}$$

this is a consequence of (3). We claim that, conversely,

(4) given $\rho\in\mathrm{P}_n$ *proper,* and for each $p\in|\langle\rho\rangle|$, a cell $a_p\in C_n(\mathrm{d}\dot{p})$ such that for every $(p,q,s)\in\mathrm{Link}_1(\rho)$,

$$\langle\mathrm{d}a_p\rangle(s)=\mathrm{c}a_q\quad(\in C_{n-1}),$$

then there is a unique $\alpha\in P_n(\rho)$ such that $\hat{p}\overset{=}{_{\mathrm{def}}}\langle\alpha\rangle(p)=a_p$ ($p\in|\langle\rho\rangle|$).

The proof is a relatively straightforward induction.

Next, $\mathrm{Link}_2(\rho)$ will denote a set of certain triples (p, r, s), the 2-*links* in ρ, such that, amomg others, $p\in|\langle\rho\rangle|$, $r\in|\langle d\rho\rangle|$ and, for $\dot{p}=\langle\rho\rangle(p)$, $s\in|\langle d\dot{p}\rangle|$,

$$\langle d\dot{p}\rangle(s) = \langle d\rho\rangle(r) . \qquad (5)$$

Intuitively, the 2-links (p, r, s) are those for which the variable-occurrence in ρ at r plugs directly into the function-symbol occurrence in ρ at p, at the place s of the function-symbol $\langle\rho\rangle(p)$. For instance, let us consider the ρ taken as an example above, and assume that as far as ρ and its subterms are concerned, the multicategory $\mathbf{C}_n$ has standard amalgamation; in particular, $s_{\mathbf{C}_n}(\rho)=\langle X, X, Y, Y, Y, Y, Y\rangle$, and in fact each i of the seven places 1 to 7 refers to the ith occurrence from the left of an element of C_{n-1} in ρ. (This assumption is, of course, not automatically true.) In this case, we have

$$\mathrm{Link}_2(\rho) = \{((2,1,1),(3,2,1),(3,3,2),(1,4,2),(5,5,1),(5,6,2),(4,7,2)\}$$

Formally, we define $\mathrm{Link}_2(\rho)$ as follows. When ρ is improper, $\mathrm{Link}_2(\rho)=\emptyset$. Let $\rho=f(\rho_1, \ldots, \rho_m)$. We have the amalgamating functions $\varphi_i:s_{\mathbf{C}_n}(\rho_i)\longrightarrow s_{\mathbf{C}_n}(\rho)$, that is, $\varphi_i:\langle d\rho_i\rangle\longrightarrow\langle d\rho\rangle$;

$$\mathrm{Link}_2(\rho) = \bigcup_{i=1}^{m}\{(\mu_i(p), \varphi_i(r), s):\ (p, r, s)\in\mathrm{Link}_2(\rho_i)\} \cup \{(1, \mu_i(1), i):\ i\in[1, m],\ |\langle\rho_i\rangle|=\emptyset\} .$$

(5) is immediately seen.

It is easy to see that

(5') if ρ is proper, then for every $r\in|\langle d\rho\rangle|$, there is a unique $(p, r, s)\in\mathrm{Link}_2(\rho)$ with second component the given r.

Supposing that $\alpha\in P_n(\rho)$, $p\in|\langle\rho\rangle|$, $\hat{p}=\langle\rho\rangle(p)=\Phi(\dot{p})$, and $(p, r, s)\in\mathrm{Link}_2(\rho)$, we have

$$\langle d\hat{p}\rangle(s) = \langle d\alpha\rangle(r)\,, \tag{6}$$

as a consequence of (5).

We are ready to define the category $\mathtt{Multitope}$. As we said, its objects are $[\rho]_n$, one for each $n\in\mathbb{N}$, $\rho\in P_{n-1}$. Next, we give *generating arrows* for $\mathtt{Multitope}$. They are

Multitope[1] $$[\rho]_{n+1} \xrightarrow{d_{\rho,\,p}} [d\dot{p}]_n\,,$$

one for each $n\geq -1$, $\rho\in P_n$ and $d\dot{p}=d(\langle\rho\rangle(p))\in P_{n-1}$;

Multitope[2] $$[\rho]_{n+1} \xrightarrow{c_{\rho}} [d\rho]_n\,,$$

one for each $n\geq -1$ and $\rho\in P_n$.

Finally, we give *defining relations* that the generating arrows are to satisfy. These come in four groups $\mathtt{Multitope[3]}$ to $\mathtt{Multitope[6]}$.

Multitope[3]:

$$\begin{array}{ccccc} & & [\rho]_{n+1} & & \\ & \swarrow^{d_{\rho,\,p}} & & \searrow^{d_{\rho,\,q}} & \\ [d\dot{p}]_n & & \circ & & [d\dot{q}]_n \\ & \searrow_{d_{d\dot{p},\,s}} & & \swarrow_{c_{d\dot{q}}} & \\ & & [\theta]_{n-1} & & \end{array} \tag{7}$$

one for each $n\geq 0$, $\rho\in P_n$, and $(p, q, s)\in \mathrm{Link}_1(\rho)$. Note that the codomain of $d_{d\dot{p},\,s}$ is $[\theta_1]_{n-1}$ for $\theta_1= d(\langle d\dot{p}\rangle(s))$, and the codomain of $c_{d\dot{q}}$ is $[\theta_2]_{n-1}$ for $\theta_2=d(d\dot{q})$; by the fact that $d(d\dot{q})=d(c\dot{q})$, and by (3), we have $\theta_1=\theta_2$, and so, the diagram (7) makes sense.

Multitope[4]:

$$\begin{array}{ccccc} & & [\rho]_{n+1} & & \\ & \swarrow d_{\rho,\, p} & & c_{\rho} \searrow & \\ [\mathrm{d}\dot{p}]_{n} & & \circ & & [\dot{\mathrm{d}}\rho]_{n}, \\ & d_{\mathrm{d}\dot{p},\, s} \searrow & & \swarrow d_{\mathrm{d}\rho,\, r} & \\ & & [\tau]_{n-1} & & \end{array} \tag{8}$$

one for each $n\geq 1$, $\rho\in P_n$, and $(p, r, s)\in \mathrm{Link}_2(\rho)$. Note that the codomain of $d_{\mathrm{d}\dot{p},\, s}$ is $[\tau_1]_{n-1}$ for $\tau_1=\mathrm{d}(\langle \mathrm{d}\dot{p}\rangle(s))$, and that of $d_{\mathrm{d}\rho . r}$ is $[\tau_2]_{n-1}$ for $\tau_2=\mathrm{d}(\langle \mathrm{d}\rho\rangle(r))$, and by (5), $\tau_1=\tau_2$, thus, (8) is meaningful.

Multitope[5] :

$$\begin{array}{ccccc} & & [\rho]_{n+1} & & \\ & \swarrow d_{\rho,\, 1} & & c_{\rho} \searrow & \\ [\mathrm{d}\dot{1}]_{n} & & \circ & & [\dot{\mathrm{d}}\rho]_{n}, \\ & c_{\mathrm{d}\dot{1}} \searrow & & \swarrow c_{\mathrm{d}\rho} & \\ & & [\xi]_{n-1} & & \end{array} \tag{9}$$

one for each $n\geq 0$, *proper* $\rho\in P_n$. Note that $1\in |\langle\rho\rangle|$; $\dot{1}=\langle p\rangle(1)$. The codomain of $c_{\mathrm{d}\dot{1}}$ is $\mathrm{dd}\dot{1}$, that of $c_{\mathrm{d}\rho}$ is $\mathrm{dd}\rho$. We have $c\dot{1}=c\rho$ as a general, and obvious, rule for all proper ρ. Therefore $\mathrm{dd}\dot{1}=\mathrm{dc}\dot{1}=\mathrm{dc}\rho=\mathrm{dd}\rho$, thus, (9) is meaningful.

Multitope[6] :

$$\begin{array}{ccc} & [1_g]_{n+1} & \\ & \downarrow c_{1_g} & \\ & [d\rho]_n=[ig]_n & \\ d_{ig,1}\downarrow & & \downarrow c_{ig} \\ & [dg]_{n-1} & \end{array}$$

commutes: $d_{ig,1}\circ c_{1_g}=c_{ig}\circ c_{1_g}$,

whenever $g\in C_{n-1}$. Note that $1_g=1_g^{(\mathbf{C}_n)}\in P_n$, $ig=i_{n-1}(g)\in P_{n-1}$, $|\langle ig\rangle|=\{1\}$, $\langle ig\rangle(1)=g$, $dig=dg$.

The category $\mathtt{Multitope}$ is the one whose arrows are generated by the generating arrows, under the identification of arrows forced by the defining relations; briefly, the category whose presentation was given above.

Given any multitopic set S , we may define $\bar{S}:\mathtt{Multitope}\longrightarrow\mathtt{Set}$ as follows. We put $\bar{S}([\rho]_n)=C_n(\rho)$. For the generating arrow in $\mathtt{Multitope[1]}$, we note that $\bar{S}([d\dot{p}]_{n-1})=C_{n-1}(d\dot{p})$, and if $a\in\bar{S}([\rho]_n)=C_n(\rho)$, then $\alpha\overset{=}{\mathrm{def}}da\in P_{n-1}(\rho)$, and for $\hat{p}=\langle d\alpha\rangle(p)$, $\Phi(\hat{p})=\dot{p}$, and thus $\hat{p}\in C_{n-1}(d\dot{p})$; this means that we can define

$$\bar{S}(d_{\rho,p}):C_n(\rho)\longrightarrow C_{n-1}(d\dot{p})\ :\ a\longmapsto\hat{p}=\langle da\rangle(p)\ .$$

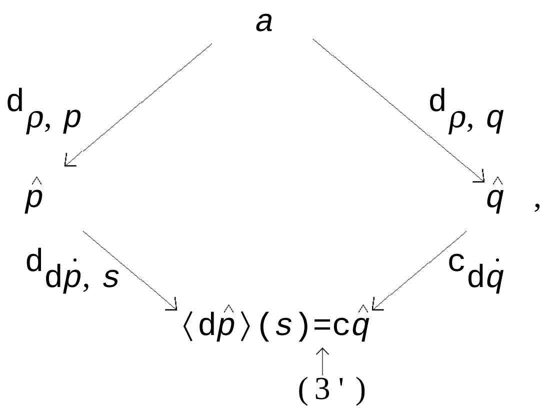
As for $\mathtt{Multitope[2]}$,

$$\bar{S}(c_\rho):C_n(\rho)\longrightarrow C_{n-1}(d\rho)\ :\ a\longmapsto ca\ ;$$

note that since $d\Phi(ca)=dc\Phi(a)=dd\Phi(a)=d\Phi(da)=dd\rho$, we have $\Phi(ca)=d\rho$, and so $ca\in C_{n-1}(d\rho)$.

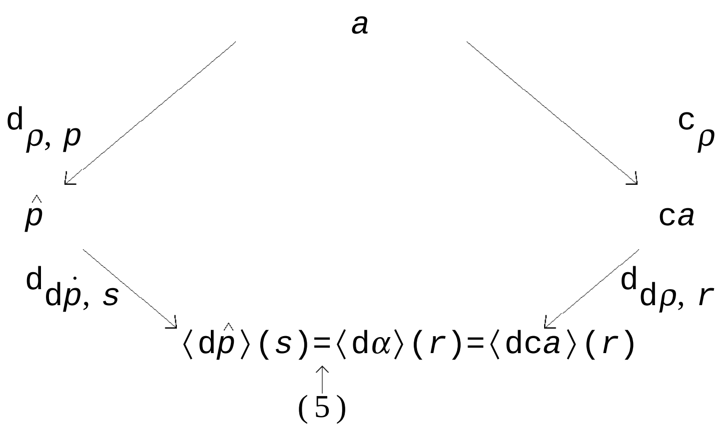
The diagrams which are the images under $\bar{S}$ of the ones in `Multitope[3]` to `Multitope[6]` commute:

`Multitope[3]`:



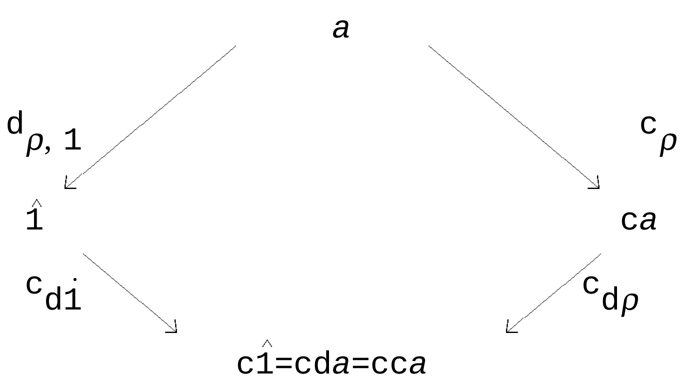
, (10)

`Multitope[4]`:



, (11)

`Multitope[5]`:



(12)

Multitope[6]:

$$\begin{array}{ccc} & a & \\ & \downarrow & \\ & \mathrm{c}a & \\ \swarrow & & \searrow \\ \langle \mathrm{dc}a\rangle(1) & = & \mathrm{cc}a \end{array} \quad ; \tag{13}$$

the reason for the last equality is that now $\mathrm{d}a=\alpha=1_f$ for some $f\in C_{n-1}$, and so $\mathrm{dd}a=\mathrm{i}f\in P_{n-1}$, $\mathrm{cc}a=\mathrm{cd}a=f$, and so $\langle \mathrm{dc}a\rangle(1)=\langle \mathrm{dd}a\rangle(1)=f=\mathrm{cc}a$.

Therefore, the $\bar{S}:\mathtt{Multitope}\longrightarrow\mathtt{Set}$ is well-defined.

It is clear that any arrow $\Psi:S_1\to S_2$ in MSet gives rise to a natural transformation $\bar{\Psi}:\bar{S}_1\to\bar{S}_2$, whose components $C^1_n(\rho)\longrightarrow C^2_n(\rho)$ are the restrictions of $\Psi:C^1_n\longrightarrow C^2_n$. We have a functor $\Sigma:\mathtt{MSet}\longrightarrow\mathtt{Set}^{\mathtt{Multitope}}:S\mapsto\bar{S},\ \Psi\mapsto\bar{\Psi}$.

I omit the (easy) proof of the fact that Σ is full and faithful.

Let us show that Σ is surjective on objects. Let $T\in\mathtt{Set}^{\mathtt{Op}}$, to construct $S\in\mathtt{MSet}$ for which $\bar{S}=T$; we are going to use the standard notation for S; $C_k=C_k(S)$, etc.

We define $C_0 \overset{\text{def}}{=} T([*]_0)$.

Let $n\geq 0$. Suppose we have constructed the n-truncation (see the previous section) $S\restriction n$ of S, so that

$$\overline{(S\restriction n)} = T\restriction n\ ; \tag{14}$$

here we are using a self-explanatory notation. $S\restriction n$ involves k-cells and k-pasting diagrams for all $k=-1, 0, \ldots, n$. $\mathtt{Multitope}\restriction n$ is the category whose objects are the $[\rho]_k$ for $\rho\in\mathrm{P}_k$, and for the same range of k, and whose arrows the arrows generated by the generators and relations in Multitope[1] to Multitope[6], with the value of n appearing in Multitope[1] to Multitope[6] restricted to the range from 0 the present n (inclusive). $T\restriction n$ is the $\mathtt{Set}$-valued functor on $\mathtt{Multitope}\restriction n$ which is the restriction of T. For the canonical functor $\Xi:\mathtt{Multitope}\restriction n\longrightarrow\mathtt{Multitope}$, $T\restriction n$ is $T\circ\Xi$. (Incidentally, Ξ is full and faithful.)

Next, we define

(15) C_{n+1} and $\mathrm{d}_{n+1}:C_{n+1}\longrightarrow P_n$, $\mathrm{c}_{n+1}:C_{n+1}\longrightarrow C_n$ such that when $n\geq 1$, also $\mathrm{d}_n\mathrm{d}_{n+1}=\mathrm{d}_n\mathrm{c}_{n+1}$, $\mathrm{c}_n\mathrm{d}_{n+1}=\mathrm{c}_n\mathrm{c}_{n+1}$.

We put $C_{n+1}\overset{\mathrm{def}}{=}\dot{\bigcup}_{\rho\in\mathrm{P}_n}T([\rho]_{n+1})$.

Let $\rho\in\mathrm{P}_n$, $a\in T([\rho]_{n+1})\subset C_{n+1}$.

We let $\mathrm{c}_{n+1}(a)=\mathrm{c}a\overset{\mathrm{def}}{=}T(\mathrm{c}_\rho)(a)$, for the arrow c_ρ from $\mathtt{Multitope[2]}$. Since $\mathrm{c}a\in T([\mathrm{d}\rho]_n)$, and (9), we have $\mathrm{c}a\in C_n(\mathrm{d}\rho)$.
To define $?:\alpha\overset{\mathrm{def}}{=}\mathrm{d}_{n+1}(a)=\mathrm{d}a\in P_n$, we distinguish two cases. First, assume ρ is improper; $\rho=1_g^{(\mathbf{C}_n)}$ for some $g\in\mathrm{C}_{n-1}$. The type of α has to be $\Phi\alpha=\rho$; thus, $\alpha=1_c^{(\mathbf{C}_n)}$ for a suitable $c\in C_{n-1}$ for which $\Phi(c)=g$; moreover, $c=\mathrm{c}\alpha=\mathrm{cd}a=\mathrm{cc}a$, and $\mathrm{c}a\in C_n(\mathrm{d}\rho)$ were determined above. So, let us define $\mathrm{d}a\overset{\mathrm{def}}{=}1_c^{(\mathbf{C}_n)}$ for $c\overset{\mathrm{def}}{=}\mathrm{cc}a$.

For $n\geq 1$, we need $\mathrm{dd}a=\mathrm{dc}a$, $\mathrm{cd}a=\mathrm{cc}a$.

We have $\mathrm{d}\rho=\mathrm{d}(1_g)=\mathrm{i}g$; for $b=\mathrm{c}a$, we have $b\in C_n(\mathrm{i}g)$; so, $\mathrm{d}b=\mathrm{i}f$ for some $f\in C_{n-1}$; $f=\langle\mathrm{d}b\rangle(1)$. Applying $\mathtt{Multitope[6]}$ to T (and remembering (9)), we get that $\langle\mathrm{dc}a\rangle(1)=\mathrm{cc}a$: read (13) as to imply the equality stated in it. This means that $\mathrm{dc}a=\mathrm{icc}a$. Also, $\mathrm{dd}a=\mathrm{d}(1_c)=ic=\mathrm{icc}a$. Therefore, $\mathrm{dd}a=\mathrm{dc}a$ is established.

Since $\mathrm{d}a=1_c$, $\mathrm{cd}a=c$; $\mathrm{cc}a=c$ by the definition of c; $\mathrm{cd}a=\mathrm{cc}a$ is established.

Second, assume ρ is proper. We now apply (4). We let $a_p\overset{\mathrm{def}}{=}T(\mathrm{d}_{\rho,\,p})(a)$ $(p\in|\langle\rho\rangle|)$. The fact that T respects the diagram $\mathtt{Multitope[3]}$ (compare (10)) gives that the condition in (4) is satisfied. Therefore, we have $\alpha\in P_n(\rho)$ such that

$\hat{p}=\langle\alpha\rangle(p)=T(\mathrm{d}_{\rho,\,p})(a)$ for all $p\in|\langle\rho\rangle|$. We let $\mathrm{d}a \overset{=}{\mathrm{def}} \alpha$.

To see that $\mathrm{cd}a=\mathrm{cc}a$, it suffices to invoke the fact that T "satisfies" Multitope[5]; see (12).

We have $\mathrm{dd}a$, $\mathrm{dc}a\in P_{n-1}(\mathrm{d}\rho)$. To show the equality $\mathrm{dd}a=\mathrm{dc}a$, we distinguish two cases: $\mathrm{d}\rho$ is proper (Case 1), $\mathrm{d}\rho$ is improper (Case 2).

Case 1. Let $\alpha=\mathrm{d}a$, $\beta=\mathrm{dc}a$. We have that $\beta\in \mathrm{P}_{n-1}(\mathrm{d}\rho)$, and β satisfies the condition $\langle \mathrm{d}\hat{p}\rangle(s)=\langle\beta\rangle(r)$ for all $(p,r,s)\in\mathrm{Link}_2(\rho)$, by Multitope[4] applied to T. By (6), and (5'), thus $\langle\beta\rangle(r)=\langle\mathrm{d}\alpha\rangle(r)$ for all $r\in|\langle\mathrm{d}\rho\rangle|$, which means, by (1'), that $\beta=\mathrm{d}\alpha$; this is what we wanted.

Case 2. Now, $n-1\geq 1$. Any (improper) $\beta\in P_{n-1}(\mathrm{d}\rho)$ is determined by $\mathrm{c}\beta$; if $\beta_1,\beta_2\in P_{n-1}(\mathrm{d}\rho)$, and $\mathrm{c}\beta_1=\mathrm{c}\beta_2$, then $\beta_1=\beta_2$. For $\beta_1=\mathrm{dd}a$, $\beta_2=\mathrm{dc}a$,

$$\mathrm{c}\beta_1=\mathrm{cdd}a \underset{1}{\overset{}{=}} \mathrm{ccd}a \underset{2}{=} \mathrm{ccc}a \underset{3}{=} \mathrm{cdc}a,$$

where 1 and 3 hold by the law " $cd=cc$ " holding in $S\restriction n$, 2 by the fact that $\mathrm{cd}a=\mathrm{cc}a$.

We have completed (15).

By what we have done in the previous section, we now have $\hat{S}=S\restriction(n+1)$, the $(n+1)$-truncation of the desired S; by recursion, we have the full S.

The construction of S clearly ensures that $\bar{S}=T$, for the effect on both objects and arrows of Multitope.

This completes the proof that $\mathrm{MSet}\simeq\mathrm{Set}^{\mathrm{Multitope}}$.

We develop a notation for multitopes.

In what follows, we work in $\mathcal{T}$, the terminal multitopic set. To have a notation for cells and pd's, we make the following decisions:

(i) We make each of the bijections $d_n:C_n \xrightarrow{\cong} P_{n-1}$ ($n\geq 0$) equal the identity map.

(ii) For elements of the free multicategory $\mathcal{F}(\mathcal{L})$ we may use Polish notation. That is, instead of writing $f(\alpha_1, ..., \alpha_n)$ ($f\in L(\mathcal{L})$, $\alpha_i\in A(\mathcal{F}(\mathcal{L}))$), we may write $f\alpha_1...\alpha_n$. In particular, if $n=0$, then the term is just f. (It is known that there is no loss of unique readability when we so omit all parentheses and commas. In a complex expression, the key to decoding is to use the arity of each operation symbol. In our notation, the arity will be contained in the notation of the operation symbol itself.) Alternatively, we may use the original parenthesis/comma notation as well.

(iii) The ingredients of the notation of members of $A(\mathcal{F}(\mathcal{L}))$ are the elements of $O(\mathcal{L})$ and those of $L(\mathcal{L})$. Thus, the ingredients for the elements of $P_n=A(\mathcal{F}(\mathcal{L}_n))$ ($n\geq 1$) are the elements of $C_{n-1}=P_{n-2}$, and those of $C_n=P_{n-1}$. When $\beta\in C_{n-1}=P_{n-2}$ is used in P_n, we put β in brackets $[\ ]$; when $\alpha\in C_n=P_{n-1}$ is used in P_n, we put α in curly brackets $\{\ \}$. When $n=0$, only the second part applies.

(iv) We denote the single element of $C_0=P_{-1}$ by $*$ (as we already did above).

The above fixes the notation, with a choice of using Polish notation, or the parenthesis/comma notation. The Polish normal form for the generating arrow $\beta\in C_n=L(\mathcal{L}_n)=P_{n-1}$ $(n\geq 1)$ as it appears in $P_n=A(\boldsymbol{C}_n)$ is $\{\beta\}[\gamma_1]...[\gamma_\ell]$ where $\langle\beta\rangle=\langle\gamma_1,...,\gamma_\ell\rangle\in P_{n-2}$. As an abbreviation, we will write $\{\beta\}$ for $\{\beta\}[\gamma_1]...[\gamma_\ell]$. On the other hand, for $\gamma\in P_{n-2}$, we have $[\gamma]\in P_n$ in proper notation.

The single element of $C_1=P_0$ is $\{*\}$. The elements of $C_2=P_1$ are the expressions $\{\{*\}\}\{\{*\}\}...\{\{*\}\}[*]$, with zero or more parts of the form $\{\{*\}\}$. If $\alpha=\{\{*\}\}\{\{*\}\}...\{\{*\}\}[*]\in P_1=C_2=L(\mathcal{L}_2)\subset A(\boldsymbol{C}_2)$ has ℓ occurrences of $\{\{*\}\}$, then $s_{\mathcal{L}_2}(\alpha)=s_{\boldsymbol{C}_2}(\alpha)=\langle d\alpha\rangle=\langle\alpha\rangle=\langle\{*\},\{*\},...,\{*\}\rangle\in C_1^*$ (ℓ occurrences), and $t_{\mathcal{L}_2}(\alpha)=t_{\boldsymbol{C}_2}(\alpha)=\{*\}\in C_1$. Consequently, an inductive definition of $C_3=P_2$ is as follows:

(a) $[\{*\}]\in P_2$;

(b) if $\{\{*\}\}\{\{*\}\}...\{\{*\}\}[*]\in P_1$ has ℓ occurrences of $\{*\}$, and $\alpha_1,...,\alpha_\ell\in P_2$, then $\{\{\{*\}\}\{\{*\}\}...\{\{*\}\}[*]\}\alpha_1,...,\alpha_\ell\in P_2$.

Note that, for $\ell=0$, (b) gives $\{[*]\}$ as an example of a member of P_2; of course, it is different from the one in (a).

To have an example, consider the 2-pd γ considered in section 1; $\gamma\in P_2(S)$ in some multitopic set S that accommodates the cells involved. The type of γ, $\Phi(\gamma)\in P_2$, is as follows. Let us write upper bar for "type", $\bar{x}$ for $\Phi(x)$. Then $\bar{a}=\bar{c}=\bar{d}=\bar{e}=\{\{*\}\}\{\{*\}\}[*]=\varphi\in C_2=P_1$, and $\bar{b}=\{\{*\}\}\{\{*\}\}\{\{*\}\}[*]=\eta\in C_2=P_1$. The type of all of the f_i is $\{*\}\in C_1=P_0$. Thus, abbreviating $\#=[\{*\}]$, $\hat{\varphi}=\{\varphi\}$, $\hat{\eta}=\{\eta\}$, we have

$$\bar{\gamma} = \hat{\varphi}(\hat{\varphi}(\#, \hat{\varphi}(\hat{\eta}(\#, \#, \#), \hat{\varphi}(\#, \#))), \#)$$

or, in Polish notation,

$$\bar{\gamma} = \hat{\varphi}\hat{\varphi}\#\hat{\varphi}\hat{\eta}\#\#\#\hat{\varphi}\#\#\#$$

or, without abbreviations,

$$\begin{aligned}\bar{\gamma} = \{\{\{*\}\}\{\{*\}\}[*]\}\{\{\{*\}\}\{\{*\}\}[*]\}[\{*\}]\{\{\{*\}\}\{\{*\}\}[*]\}\\ \{\{\{*\}\}\{\{*\}\}\{\{*\}\}[*]\}[\{*\}][\{*\}][\{*\}]\\ \{\{\{*\}\}\{\{*\}\}[*]\}[\{*\}][\{*\}][\{*\}] .\end{aligned}$$

This is not meant as a particularly intuitive representation; it is a systematic notation well-suited for mechanical manipulation.

The inductive definition of P_n and $c(\alpha)$ for $\alpha\in P_n$ is this:

(a) for any $\gamma\in P_{n-2}$, $[\gamma]\in P_n$; and $c[\gamma]=\{\gamma\}\in P_{n-1}$;

(b) whenever $\beta\in P_{n-1}$, $\langle\beta\rangle=\langle\gamma_1,\dots,\gamma_\ell\rangle\in P^*_{n-2}$, $\alpha_i\in P_n$, $c\alpha_i=\gamma_i$, we have $\{\beta\}\alpha_1\dots\alpha_\ell\in P_n$ and $c(\{\beta\}\alpha_1\dots\alpha_\ell)=d\beta$.

For d on P_n, we have

$$d(\alpha_1 \circ_p \alpha_2) = (d\alpha_1)\square_p(d\alpha_2)$$

whenever $p\in |\langle d\alpha_1\rangle|$. Also, $d(\{\beta\})=\beta$ and $d([\gamma])=\{\gamma\}$ for $\beta\in P_{n-1}$, $\gamma\in P_{n-2}$.

We look at the types of the 3-cells considered in subsection 1.6 of section 1. Let's use φ, $\hat{\varphi}$, # as above; $\sigma=\{\{*\}\}[*]$, $\hat{\sigma}=\{\sigma\}$. We have $\bar{f}=\bar{h}=\varphi$, $\bar{i}=\bar{g}=\sigma$, $\bar{\alpha}=\hat{\varphi}\#\hat{\sigma}\#$, $\bar{\beta}=\hat{\varphi}\hat{\sigma}\#\#$; also, $\overline{\beta\square_1\alpha}=\bar{\beta}\square_1\bar{\alpha}=\hat{\varphi}\hat{\sigma}\#\hat{\sigma}\#$; $\bar{\alpha},\bar{\beta},\bar{\beta}\square_1\bar{\alpha}\in P_2$. For the 3-cells u, v, we have $\bar{u}=\bar{\alpha}$, $\bar{v}=\bar{\beta}$, both elements of $C_3=P_2$. The 3-pd $\psi=v\circ_1 u=v(u(h,i),g)$ (from 1.6 too),

$$\bar{\psi} = \{\bar{\beta}\}\{\bar{\alpha}\}[\varphi][\sigma][\sigma] = \{\hat{\varphi}\hat{\sigma}\#\#\}\{\hat{\varphi}\#\hat{\sigma}\#\}[\varphi][\sigma][\sigma] ;$$

$$d\bar{\psi} = \bar{\beta}\square_1\bar{\alpha} = \hat{\varphi}\hat{\sigma}\#\hat{\sigma}\# \ .$$

Also, $\bar{\beta}'=\hat{\varphi}\#\hat{\sigma}\#=\bar{\alpha}$, $\bar{v}'=\bar{\beta}'$, $\psi'=v'\circ_1 u=v'(u(h,i),g)$, and

$$\bar{\psi}' = \{\bar{\alpha}\}\{\bar{\alpha}\}[\varphi][\sigma][\sigma] = \{\hat{\varphi}\#\hat{\sigma}\#\}\{\hat{\varphi}\#\hat{\sigma}\#\}[\varphi][\sigma][\sigma] ;$$

$$d\bar{\psi} = \bar{\beta}'\square_1\bar{\alpha} = \hat{\varphi}\#\hat{\sigma}\hat{\sigma}\# \ .$$

Appendix

We prove the **if** direction of 3.(1). We use the notation introduced in and before the statement of 3.(1). On the basis of the data $\mathcal{L}$, $\boldsymbol{C}$, F defined on $\mathcal{L}$, and $\langle\theta_f\rangle_{f\in L}$, satisfying the conditions in the definition of the free multicategory, we have to define $F:\boldsymbol{C}\to\boldsymbol{D}$, including $\langle\theta_f\rangle_{f\in A}$.

We write s, t, O, L, A for $\mathrm{s}_{\boldsymbol{C}}=\mathrm{s}_{\mathcal{L}}$, $\mathrm{t}_{\boldsymbol{C}}=\mathrm{t}_{\mathcal{L}}$, $\mathrm{O}(\boldsymbol{C})=\mathrm{O}(\mathcal{L})$, $\mathrm{L}(\mathcal{L})$, $\mathrm{A}(\boldsymbol{C})$, respectively. Let's write $\bar{O}$ for $\mathrm{O}(\boldsymbol{D})$, $\bar{A}$ for $\mathrm{A}(\boldsymbol{D})$, $\bar{X}$ for $F(X)$ ($X\in O$), $\bar{f}$ for $F(f)$ ($f\in A$). We will write $\mathrm{s}(\beta)$, $\mathrm{t}(\beta)$ for $\mathrm{s}_{\boldsymbol{D}}(\beta)$, $\mathrm{t}_{\boldsymbol{D}}(\beta)$, respectively, when $\beta\in\bar{A}$.

We define $\bar{\alpha}=F(\alpha)$ by structural recursion.

If $\alpha=1_X:\langle X\rangle\to\dot{X}$, we let $\bar{1}_X:\langle\bar{X}\rangle\to\dot{\bar{X}}$.

Assume that $f\in L$, $|\mathrm{s}(f)|=[1,n]$, $\alpha_1,\dots,\alpha_n\in A$, $\mathrm{t}(\alpha_i)=(\mathrm{s}(f)(i))^{\cdot}$; thus, $\alpha=f(\vec{\alpha})=f(\alpha_1,\dots,\alpha_n)$ is well-defined. Consider

$$\theta_f:F\circ\mathrm{s}(f)\xrightarrow{\cong}\mathrm{s}(\bar{f}), \tag{1}$$

and write $\bar{i}$ for $\theta_f(i)$, and $\tilde{j}$ for $\theta_f^{-1}(j)$ ($i,j\in[1,n]$). Assume that $\bar{\alpha}_i$ is defined for all $i\in[1,n]$, and $\mathrm{t}(\bar{\alpha}_i)=(\overline{\mathrm{t}(\alpha_i)})^{\cdot}$ (induction hypothesis). We let

$$\bar{\alpha}=\bar{f}(\bar{\alpha}_{\tilde{1}},\bar{\alpha}_{\tilde{2}},\dots,\bar{\alpha}_{\tilde{n}})=\bar{f}(\bar{\alpha}_{\tilde{1}}/1,\bar{\alpha}_{\tilde{2}}/2,\dots,\bar{\alpha}_{\tilde{n}}/n).$$

We observe that this is well-defined:

$$\mathrm{t}(\bar{\alpha}_{\tilde{j}})=(\overline{\mathrm{t}(\alpha_{\tilde{j}})})^{\cdot}=(\overline{\mathrm{s}(f)(\tilde{j})})^{\cdot}\underset{\substack{\uparrow\\ *}}{=}\mathrm{s}(\bar{f})(\theta_f(\tilde{j}))^{\cdot}=\mathrm{s}(\bar{f})(j)^{\cdot};$$

here, the equality marked $*$ is the content of (1). Also,

$t(\bar{\alpha})=t(\bar{f})=(\overline{t(f)})^{\cdot}=(\overline{t(\alpha)})^{\cdot}$. This completes the definition of the mapping $\alpha\mapsto\bar{\alpha}=F(\alpha)$.

Next, we define $\theta_{\alpha}:F\circ s(\alpha)\xrightarrow{\cong}s(\bar{\alpha})$. For $\alpha=1_X$, there is no choice. Consider $\alpha=f(\vec{\alpha})$ as above. Let $\vec{\bar{\alpha}}=\langle\bar{\alpha}_1, ..., \bar{\alpha}_n\rangle$, and

$$\begin{aligned} \varphi_i &= \varphi_i^{(\mathbf{C})}[f, \vec{\alpha}]:s(\alpha_i)\xrightarrow{\cong}s(\alpha) ,\\ \bar{\varphi}_j &= \varphi_j^{(\mathbf{D})}[\bar{f}, \vec{\bar{\alpha}}]:s(\bar{\alpha}_j)\xrightarrow{\cong}s(\bar{\alpha}) , \end{aligned}$$

arrows in $O^{\#}$, $\bar{O}^{\#}$, respectively. We know that the morphisms $\bar{\varphi}_j$ are coprojections making $s(\bar{\alpha})$ a coproduct of the $s(\bar{\alpha}_j)$ in $\bar{O}^{\#}$. Therefore, there is a uniquely determined morphism $\theta_{\alpha}:F\circ s(\alpha)\longrightarrow s(\bar{\alpha})$ for which the following diagram in $\bar{O}^{\#}$ commutes for each $j\in[1, n]$:

$$\begin{array}{ccc} F\circ s(\alpha) & \xrightarrow{\quad\theta_{\alpha}\quad} & s(\bar{\alpha}) \\ {\scriptstyle\varphi_{\tilde{j}}}\Big\uparrow & \circ & \Big\uparrow{\scriptstyle\bar{\varphi}_j} \\ F\circ s(\alpha_{\tilde{j}}) & \xrightarrow[\quad\theta_{\alpha_{\tilde{j}}}\quad]{} & s(\bar{\alpha}_{\tilde{j}}) \end{array} \quad ; \tag{2}$$

it is also clear that θ_{α} is an isomorphism. This completes the definition of the θ_{α} .

We need to verify the requisite properties of $F:\mathbf{C}\to\mathbf{D}$ defined by the above specifications. First, we show

$$F(\alpha\circ_p\beta)=\overline{\alpha\circ_p\beta}=\bar{\alpha}\circ_{\bar{p}}\bar{\beta} ; \tag{3}$$

here, $\bar{p}=\theta_{\alpha}(p)$.

We employ a structural induction on α. When $\alpha=1_X$, the assertion is obviously true. Let $\alpha=f(\vec{\alpha})$, with all the accompanying notation used above. Assume that $\alpha\circ_p\beta$ is well-defined. There are uniquely determined $i\in[1,n]$ and $q\in|s(\alpha_i)|$ for which $p=\varphi_i(q)$; the associative law, in the form applying to a simultaneous composition, gives that

$$\alpha\circ_p\beta = f(\alpha_1, ..., \alpha_{i-1}, \alpha_i\circ_q\beta, \alpha_{i+1}, ..., \alpha_n) . \tag{4}$$

Let $j=\bar{i}$; $\tilde{j}=i$. According to the definition of $\gamma\longmapsto\bar{\gamma}$, we have

$$\overline{\alpha\circ_p\beta}=\bar{f}(\bar{\alpha}_{\tilde{1}}/1, ..., \bar{\alpha}_{\widetilde{(j-1)}}/(j-1), \overline{\alpha_i\circ_q\beta}/j, \bar{\alpha}_{\widetilde{(j+1)}}/j+1, ..., \bar{\alpha}_{\tilde{n}}/n) \tag{5}$$

Let $\bar{q}=\theta_{\alpha_i}(q)$; by the induction hypothesis,

$$\overline{\alpha_i\circ_q\beta} = \bar{\alpha}_i\circ_{\bar{q}}\bar{\beta} . \tag{6}$$

Consider the diagram (2), and chase the element q in the lower left corner. We obtain that $\bar{\varphi}_j(\bar{q})=\bar{p}$. Since we have

$$\bar{\alpha} = \bar{f}(\bar{\alpha}_{\tilde{1}}/1, \bar{\alpha}_{\tilde{2}}/2, ..., \bar{\alpha}_{\tilde{n}}/n) ,$$

we get, by associativity in $\boldsymbol{D}$, that

$$\bar{\alpha}\circ_{\bar{p}}\bar{\beta}=\bar{f}(\bar{\alpha}_{\tilde{1}}/1, ..., \bar{\alpha}_{\widetilde{(j-1)}}/(j-1), \bar{\alpha}_i\circ_{\bar{q}}\bar{\beta}/j, \bar{\alpha}_{\widetilde{(j+1)}}/j+1, ..., \bar{\alpha}_{\tilde{n}}/n) \tag{7}$$

which, after a comparison with (5) and (6), gives (3).

It remains to show that the θ_α satisfy the coherence condition in 2.(v).

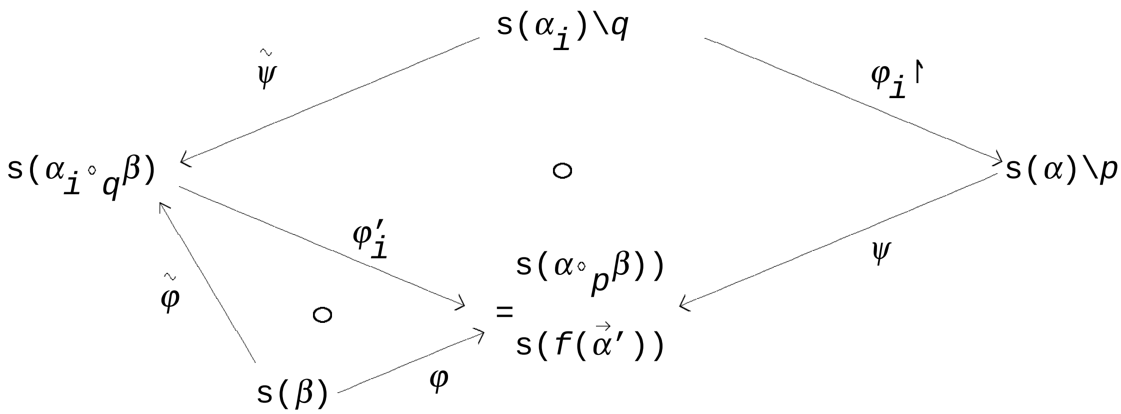
We consider $\alpha = f(\vec{\alpha})$ and $\alpha \circ_p \beta$ as before; we write $f(\vec{\alpha}')$ to abbreviate the right-hand side of (4), $\bar{f}(\vec{\bar{\alpha}}')$ for the right-hand side of (5) (equivalently, (7)). We want to prove the commutativities in

$$\begin{array}{ccc}
(F \circ s(\alpha))\backslash p & \xrightarrow{\theta_\alpha \restriction} & s(\bar{\alpha})\backslash\bar{p} \\
\psi \big\downarrow & \circ ? & \big\downarrow \bar{\psi} \\
F \circ s(\alpha \circ_p \beta) & \xrightarrow{\theta_{\alpha \circ_p \beta}} & s(\overline{\alpha \circ_p \beta}) = s(\bar{\alpha} \circ_{\bar{p}} \bar{\beta}) \\
\varphi \big\uparrow & \circ ? & \big\uparrow \bar{\varphi} \\
F \circ s(\beta) & \xrightarrow[\theta_\beta]{} & s(\bar{\beta})
\end{array} \quad . \tag{8}$$

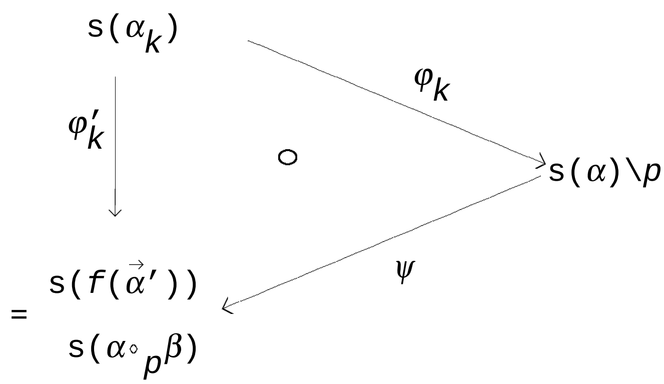
What we have to go on are the following facts. First, as the induction hypothesis, the commutativities

$$\begin{array}{ccc}
(F \circ s(\alpha_i))\backslash q & \xrightarrow{\theta_{\alpha_i} \restriction} & s(\bar{\alpha}_i)\backslash\bar{q} \\
\tilde{\psi} \big\downarrow & \circ & \big\downarrow \hat{\psi} \\
F \circ s(\alpha_i \circ_p \beta) & \xrightarrow{\theta_{\alpha_i \circ_p \beta}} & s(\overline{\alpha_i \circ_q \beta}) = s(\bar{\alpha}_i \circ_{\bar{q}} \bar{\beta}) \\
\tilde{\varphi} \big\uparrow & \circ & \big\uparrow \hat{\varphi} \\
F \circ s(\beta\ ) & \xrightarrow[\theta_\beta]{} & s(\bar{\beta})
\end{array} \tag{9}$$

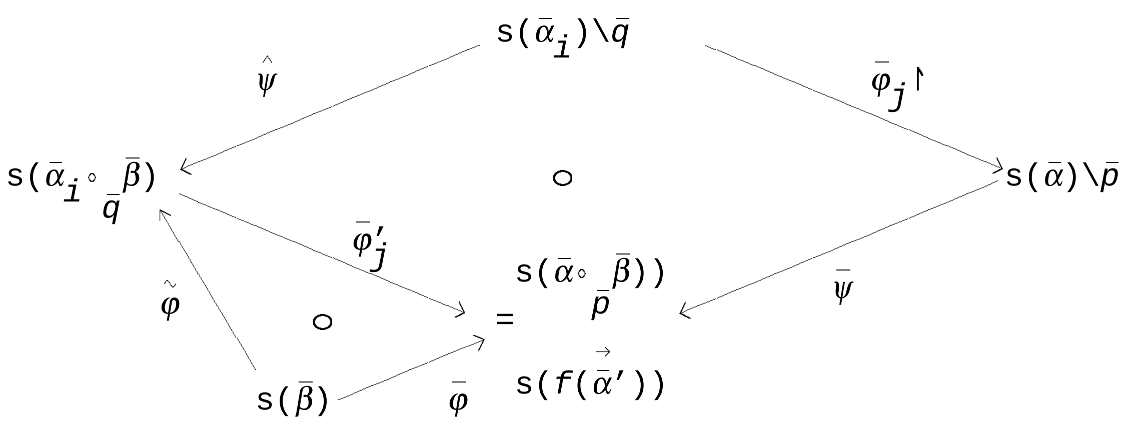
where we used the notations $\tilde{\psi}, \tilde{\varphi}, \hat{\psi}, \hat{\varphi}$ in the obvious senses. Second, we have the coherence conditions associated with associativity, both in $\boldsymbol{C}$ and $\boldsymbol{D}$, to wit



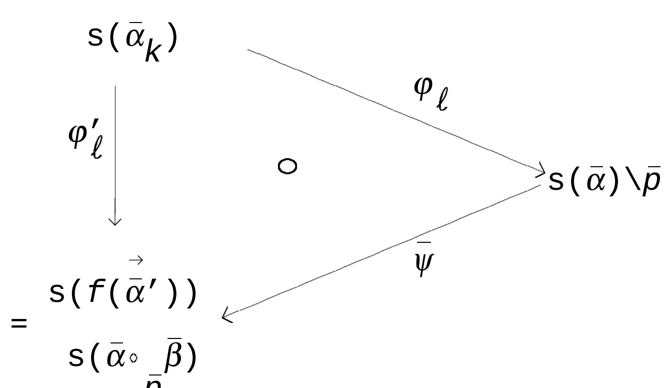
(10)



$(k \neq i)$ (11)



(12)



$(k \neq i,\ \ell = \theta_f(k))$ (13)

Further, we have the definition of θ_α in terms of the diagrams (2), that is

$$\begin{array}{ccc} F\circ s(\alpha) & \xrightarrow{\theta_\alpha} & s(\bar\alpha) \\ \varphi_i\Big\uparrow & \circ & \Big\uparrow\bar\varphi_j \\ F\circ s(\alpha_i) & \xrightarrow[\theta_{\alpha_i}]{} & s(\bar\alpha_i) \end{array} \qquad \begin{array}{ccc} F\circ s(\alpha) & \xrightarrow{\theta_\alpha} & s(\bar\alpha) \\ \varphi_k\Big\uparrow & \circ & \Big\uparrow\bar\varphi_\ell \\ F\circ s(\alpha_k) & \xrightarrow[\theta_{\alpha_k}]{} & s(\bar\alpha_k) \end{array} \qquad (14)$$

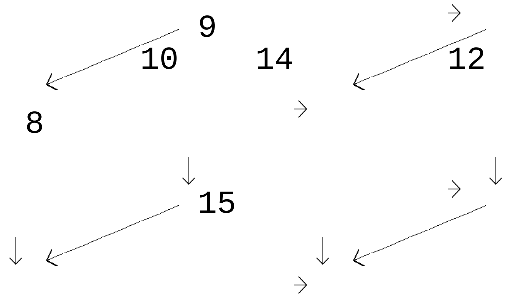
and the definition of $\theta_{\alpha\circ_p\beta}=\theta_{f(\vec{\alpha}')}$, that is

$$\begin{array}{ccc} F\circ s(\alpha\circ_p\beta) & \xrightarrow{\theta_\alpha} & s(\bar\alpha\circ_{\bar p}\bar\beta) \\ \varphi'_i\Big\uparrow & \circ & \Big\uparrow\bar\varphi'_j \\ F\circ s(\alpha_i\circ_q\beta) & \xrightarrow[\theta_{\alpha_i}]{} & s(\bar\alpha_i\circ_{\bar q}\bar\beta) \end{array} \qquad \begin{array}{ccc} F\circ s(\alpha\circ_p\beta) & \xrightarrow{\theta_\alpha} & s(\bar\alpha\circ_{\bar p}\bar\beta) \\ \varphi'_k\Big\uparrow & \circ & \Big\uparrow\bar\varphi'_\ell \\ F\circ s(\alpha_k) & \xrightarrow[\theta_{\alpha_k}]{} & s(\bar\alpha_k) \end{array} \qquad (15)$$

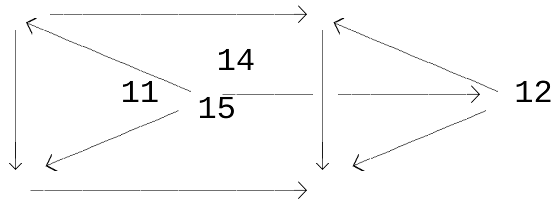
Let us remark that in each of these diagram, we actually have functions on sets, and as functions of sets, e.g. $\varphi_i:s(\alpha_i)\longrightarrow s(\alpha)$ and $\varphi_i:F\circ s(\alpha_i)\longrightarrow F\circ s(\alpha)$ are the same, namely $\varphi_i:|s(\alpha_i)|\longrightarrow|s(\alpha)|$.

Consider the upper commutativity in (8). $s(\alpha)\backslash p$ is the coproduct of $s(\alpha_i)\backslash q$ and $s(\alpha_k)$ $(k\neq i)$ via the coprojections $\varphi_i\restriction$, φ_k; therefore, the required commutativity is proved if we can show that it holds when we precompose it with the said coprojections.

We observe (not without a certain amount of experimentation) that using some of the diagrams above, we can fit together the cube



in which the front face is (8) upper (the one we want to see commute), the left face is (10) upper, the right face is (12) upper, the bottom face is (15) left, the top face is (14) left, and the back face is (9) upper. All of these, except the front face, commute. It follows that the front face commutes when precomposed by the left upper edge, which is $\varphi_i\upharpoonright$; this is the first thing we want. As far as precomposing with φ_k $(k\neq i)$ is concerned, the back face of the cube collapses, and we get



here again, the front face is (8) upper, the left face is (11), the right face is (13), the bottom face is (15) right, the top face is (14) right; thus, we again have the desired conclusion. As we said, this shows the commutativity of (8) upper.

The proof for (8) lower is left to the reader.

This completes the proof of the existence of the morphism $F:\mathbf{C}\to\mathbf{D}$; the uniqueness of F is clear from what we have seen.

Next, the proofs of some lemmas in section 4 are provided.

Proof of 4.(6) Lemma:

We employ induction according to 3.(2).

If $\alpha=1_X$, then $|\langle\alpha\rangle|$ is empty; the assertion is vacuously true.

Let $\alpha=g(\beta_1, ..., \beta_m)$, with $g\in L$; $m_i=\ell h(s(\beta_i))$. There are two cases: $p=1$ (Case 1), $p\neq1$ (Case 2).

In Case 1, we have $f=\langle\alpha\rangle(1)=g$ and $n=m$; we put $\alpha'=1_X$ where $X=t(g)$, and $\alpha_i=\beta_i$ ($i\in[1, m]$); (1) is clear.

In Case 2, by (1'), there is $i\in[1, m]$ such that $p\in|\langle\beta_i\rangle|$. By assumption, we have

$$\beta_i=\gamma\circ_r f(\alpha_1, ...\alpha_n)$$

for suitable $\gamma, r, \alpha_1, ...\alpha_n$. Then for

$$\delta \overset{=}{\text{def}} g(\beta_1, ..\beta_{i-1}, 1_Y, \beta_{i+1}, ..., \beta_m),$$
$$\varepsilon \overset{=}{\text{def}} f(\alpha_1, ...\alpha_n),$$
$$s \overset{=}{\text{def}} \varphi_i[g; \beta_1, ..\beta_{i-1}, 1_Y, \beta_{i+1}, ..., \beta_m](1)\in|s(\delta)|,$$
$$\alpha' \overset{=}{\text{def}} \delta\circ_s\gamma = g(\beta_1, ..\beta_{i-1}, \gamma, \beta_{i+1}, ..., \beta_m)$$

[α' is well-defined since $t(\gamma)=t(\beta)$],

$$q \overset{=}{\text{def}} \varphi[\delta, \gamma, r](r),$$

we have 4.(2):

$$\alpha'\circ_q f(\alpha_1, ...\alpha_n) = (\delta\circ_s\gamma)\circ_q\varepsilon = \delta\circ_s(\gamma\circ_q\varepsilon) = \delta\circ_s\beta_i = \alpha .$$

This completes the proof.

Proof of 4.(7) Lemma:

For the proof, let us first make a general statement. Given an identity

$$g(\beta_1, \ldots, \beta_m) = \eta \circ_s \xi$$

with $g \in L$, $\beta_1, \ldots, \beta_m, \eta, \xi \in A$, it follows that

either $\eta = 1_X$ and $g(\beta_1, \ldots, \beta_m) = \xi$,
or for unique $i \in [1, m]$, $\hat{\beta}$ and $t \in |s(\hat{\beta})|$, we have

$$\beta_i = \hat{\beta} \circ_t \xi \text{ and } \eta = g(\beta_1, \ldots, \beta_{i-1}, \hat{\beta}, \beta_{i+1}, \ldots, \beta_m),$$

and for $\varphi_i = \varphi_i[g; \beta_1, \ldots, \beta_{i-1}, \hat{\beta}, \beta_{i+1}, \ldots, \beta_m]$, we have $s = \varphi_i(t)$.

This is proved by an induction on η, according to 3.(2); no separatedness is involved.

Let $f \in L$, $n = \ell h(s(f))$, $\vec{\beta} = (\beta_1, \ldots, \beta_n)$, $\vec{\delta} = (\delta_1, \ldots, \delta_n)$, $f(\vec{\beta})$, $f(\vec{\delta})$ well-defined. Assume

$$\varepsilon = \alpha \circ_p f(\vec{\beta}) = \gamma \circ_q f(\hat{\delta})$$

and ε is separated, to show that $\alpha = \gamma$, $\vec{\beta} = \vec{\delta}$ and $p = q$. We do an induction on α.

Let first $\alpha = 1_X$; $X = t(f)$ necessarily. Then $\varepsilon = f(\vec{\beta})$. If $\gamma = 1_Y$, then $Y = X = t(f)$, and $\alpha = \gamma$. Otherwise, $\|\gamma\| \neq \emptyset$, and $\|\gamma\| \cap \|f(\hat{\delta})\| = \emptyset$, and so $f \notin \|\gamma\|$; but $\gamma = g(\vec{\gamma})$ for a suitable $g \in L$, $\vec{\gamma} \in A^*$, and so $\gamma \circ_q f(\hat{\delta}) = g(\vec{\eta})$ for a suitable $\vec{\eta} \in A^*$; $\varepsilon = f(\vec{\beta}) = \gamma \circ_q f(\hat{\delta}) = g(\vec{\eta})$ is impossible; contradiction.

Next, $\alpha = g(\vec{\alpha})$, $g \in L$, $\vec{\alpha} = (\alpha_1, \ldots, \alpha_m) \in L^*$. Then for a unique $i \in [1, m]$, for $\varphi_i = \varphi_i[g, \vec{\alpha}]$, and for a unique $r \in |s(\alpha_i)|$, we have $p = \varphi_i(r)$, and for

$\tilde{\alpha}=\alpha_i \circ_r f(\vec{\beta})$, we have

$$\varepsilon=\alpha \circ_p f(\vec{\beta})=g(\alpha_1, \ldots \alpha_{i-1}, \tilde{\alpha}, \alpha_{i+1}, \ldots, \alpha_m) .$$

Since ε is separated, we have that $g \neq f$.

$$\varepsilon=\gamma \circ_q f(\hat{\delta}) = g(\alpha_1, \ldots \alpha_{i-1}, \tilde{\alpha}, \alpha_{i+1}, \ldots, \alpha_m)$$

implies that

either $\gamma=1_X$ and $\varepsilon=f(\hat{\delta})$, a case excluded by $f \neq g$;

or for some $j \in [1, m]-\{i\}$, $\hat{\alpha}$ and $s \in |s(\hat{\alpha})|$, we have $\alpha_j=\hat{\alpha} \circ_s f(\hat{\delta})$, a case again excluded since $\|\alpha_j\| \cap \|\tilde{\alpha}\|=\emptyset$ by the separatedness of ε ;

or for some $\hat{\alpha}$ and $s \in |s(\hat{\alpha})|$, we have $\tilde{\alpha}=\hat{\alpha} \circ_s f(\hat{\delta})$ and

$$\gamma=g(\alpha_1, \ldots \alpha_{i-1}, \hat{\alpha}, \alpha_{i+1}, \ldots, \alpha_m) ,$$

and for $\hat{\varphi}_i=\varphi_i[g; \alpha_1, \ldots \alpha_{i-1}, \hat{\alpha}, \alpha_{i+1}, \ldots, \alpha_m]$, we have $\hat{\varphi}_i(s)=q$.

In this case, we have $\tilde{\alpha}=\alpha_i \circ_r f(\vec{\beta})=\hat{\alpha} \circ_s f(\hat{\delta})$. By the induction hypothesis applied to α_i , we get $\alpha_i=\hat{\alpha}$, $r=s$ and $\vec{\beta}=\vec{\delta}$. But then

$$\begin{aligned}\hat{\varphi}_i=\varphi_i[g; \alpha_1, \ldots \alpha_{i-1}, \hat{\alpha}, \alpha_{i+1}, \ldots, \alpha_m]\\ =\varphi_i[g; \alpha_1, \ldots \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_m]=\varphi_i ,\end{aligned}$$

$$p = \varphi_i(r)= \hat{\varphi}_i(s) = q ,$$

and

$$\alpha=g(\alpha_1, \ldots \alpha_{i-1}, \alpha_i, \alpha_{i+1}, \ldots, \alpha_m)=g(\alpha_1, \ldots \alpha_{i-1}, \hat{\alpha}, \alpha_{i+1}, \ldots, \alpha_m)=\gamma ;$$

and the proof is complete.

Proof of 4.(10) Lemma:

Assume first that $f\in\|\alpha\|$. Write α in the form 4.(2), with $\vec{\alpha}$ abbreviating $\alpha_1, \ldots, \alpha_n$:

$$\alpha = \alpha' \circ_q f(\vec{\alpha}) ,$$

and

$$\alpha\square_f\gamma = \alpha' \circ_q \gamma(\vec{\alpha}) \quad .$$

We have that $\mathrm{s}(\alpha)$ is the coproduct of $\mathrm{s}(\alpha')$ and the $\mathrm{s}(\alpha_i)$ $(i\in[1, n])$, via the amalgamating maps

$$\psi \underset{\mathrm{def}}{=} \psi[\alpha', f(\vec{\alpha}), q]:\mathrm{s}(\alpha')\to\mathrm{s}(\alpha) ,$$
$$\varphi \underset{\mathrm{def}}{=} \varphi[\alpha', f(\vec{\alpha}), q]:\mathrm{s}(f(\vec{\alpha}))\to\mathrm{s}(\alpha) .$$

Therefore, we have the two mutually exclusive cases $r\in\mathrm{Im}(\psi)$ (Case 1) , and $r\in\mathrm{Im}(\varphi)$ (Case 2). Leaving aside the easier Case 1, we take up Case 2. $\mathrm{s}(f(\vec{\alpha}))$ is the coproduct of the $\mathrm{s}(\alpha_i)$ $(i\in[1, n])$, via the maps

$$\varphi_i=\varphi_i[f, \vec{\alpha}]:\mathrm{s}(\alpha_i)\longrightarrow\mathrm{s}(f(\vec{\alpha})) .$$

We have $r = \varphi(s)$ for a uniquely determined s ; and $s=\varphi_i(t)$ for uniquely determined i and t . Two applications of the associative law (one of the original form, the other of the form related to simultaneous composition) give that

$$\alpha\circ_r\beta = (\alpha' \circ_q f(\vec{\alpha}))\circ_r\beta = \alpha' \circ_q (f(\vec{\alpha})\circ_s\beta) =$$
$$\alpha' \circ_q f(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i \circ_t \beta, \alpha_{i+1}, \ldots, \alpha_n) \quad .$$

Therefore,

$$(\alpha \circ_r \beta)\Box_f \gamma = \alpha' \circ_q \gamma(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i \circ_t \beta, \alpha_{i+1}, \ldots, \alpha_n) \ . \tag{1}$$

On the other hand,

$$\alpha \Box_f \gamma = \alpha' \circ_q \gamma(\vec{\alpha}) \ .$$

Comparing $d(\alpha' \circ_q \gamma(\vec{\alpha}))$ and $d(\alpha' \circ_q f(\vec{\alpha}))$, by 4.(4) we see that they are equal, because $d(\gamma) = d(f)$; also, $\varphi[\alpha', \gamma(\vec{\alpha}), q] = \varphi[\alpha', f(\vec{\alpha}), q] = \varphi$, and similarly for the ψ's. Therefore,

$$(\alpha \Box_f \gamma) \circ_r \beta = (\alpha' \circ_q \gamma(\vec{\alpha})) \circ_r \beta = \alpha' \circ_q (\gamma(\vec{\alpha}) \circ_s \beta) \ ,$$

with the same s as the one determined above. For similar reasons, we have, for the same i and t as above,

$$\gamma(\vec{\alpha}) \circ_s \beta = \gamma(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i \circ_t \beta, \alpha_{i+1}, \ldots, \alpha_n) \ ,$$

and so,

$$(\alpha \Box_f \gamma) \circ_r \beta = \alpha' \circ_q \gamma(\alpha_1, \ldots, \alpha_{i-1}, \alpha_i \circ_t \beta, \alpha_{i+1}, \ldots, \alpha_n) \tag{2}$$

(1) and (2) confirm the desired equality $(\alpha \circ_r \beta)\Box_f \gamma = (\alpha \Box_f \gamma) \circ_r \beta$.

We will not consider Case 1, neither the second case in the lemma, $f \in \|\beta\|$, and take 4.(10) as having been proved.

Proof of 4.(11) Lemma:

In the possession of 4.(10), we can perform an induction on α, in the sense of showing the equalities for $\alpha=1_X$ and $\alpha\in L$, and assuming them for δ and ε in place of α, with all other parameters unconstrained, we prove it for $\alpha=\delta\circ_r\varepsilon$. Since A is the least set containing all 1_X, L and closed under the operations $(\delta,\varepsilon)\longmapsto\delta\circ_r\varepsilon$, this is a valid procedure.

The basis case $\alpha=1_X$ is vacuous. Let $\alpha=h\in L$. Then $g\in\|\alpha\|$ would mean $g=h$, and since $f\in\|\alpha\|$, $f=h=g$; however, $f\notin\|\alpha\square_f\beta\|$ and $g\in\|\alpha\square_f\beta\|$, contradiction. It remains to consider the case $g\in\|\beta\|$. Since $\alpha\square_f\beta$ is well-defined, $h=f$, and $\alpha\square_f\beta=\beta$. Both $(\alpha\square_f\beta)\square_g\gamma$ and $\alpha\square_f(\beta\square_g\gamma)$ are equal to $\beta\square_g\gamma$.

For the induction step, we let $\alpha=\delta\circ_r\varepsilon$, and distinguish, because of $\|\alpha\|=\|\delta\|\dot\cup\|\varepsilon\|$, the following six mutually exclusive and jointly exhaustive cases:

[1]:	$f\in\|\delta\|$, $g\in\|\delta\|$;
[2]:	$f\in\|\delta\|$, $g\in\|\varepsilon\|$;
[3]:	$f\in\|\varepsilon\|$, $g\in\|\delta\|$;
[4]:	$f\in\|\varepsilon\|$, $g\in\|\varepsilon\|$;
[5]:	$f\in\|\delta\|$, $g\in\|\beta\|$;
[6]:	$f\in\|\varepsilon\|$, $g\in\|\beta\|$.

[1] : we want to show: $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma\overset{?}{=}((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta;$$

we have:

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma=((\delta\square_f\beta)\circ_r\varepsilon)\square_g\gamma=((\delta\square_f\beta)\square_g\gamma)\circ_r\varepsilon$$
$$\underset{*}{=}((\delta\square_g\gamma)\square_f\beta)\circ_r\varepsilon=((\delta\square_g\gamma)\circ_r\varepsilon)\square_f\beta=((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta;$$

all equalities except the one marked $(*)$ which is the induction hypothesis for δ, are valid instances of 4.(10), sometimes used "backwards".

[2] : $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma \overset{?}{=} ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ :$$

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma = ((\delta\square_f\beta)\circ_r\varepsilon)\square_g\gamma = (\delta\square_f\beta)\circ_r(\varepsilon\square_g\gamma)$$
$$= (\delta\circ_r(\varepsilon\square_g\gamma))\square_f\beta = ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ ;$$

in this case, the induction hypothesis is not used.

[3]: $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma \overset{?}{=} ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ :$$

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma = (\delta\circ_r(\varepsilon\square_f\beta))\square_g\gamma = (\delta\square_g\gamma)\circ_r(\varepsilon\square_f\beta)$$
$$= ((\delta\square_g\gamma)\circ_r\varepsilon)\square_f\beta = ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ .$$

[4]: $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma \overset{?}{=} ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ :$$

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma = (\delta\circ_r(\varepsilon\square_f\beta))\square_g\gamma = \delta\circ_r((\varepsilon\square_f\beta)\square_g\gamma)$$
$$\underset{*}{=} \delta\circ_r((\varepsilon\square_g\gamma)\square_f\beta) = (\delta\circ_r(\varepsilon\square_g\gamma))\square_f\beta = ((\delta\circ_r\varepsilon)\square_g\gamma)\square_f\beta \ .$$

[5]: $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma \overset{?}{=} (\delta\circ_r\varepsilon)\square_f(\beta\square_g\gamma) \ :$$

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma = ((\delta\square_f\beta)\circ_r\varepsilon)\square_g\gamma = ((\delta\square_f\beta)\square_g\gamma)\circ_r\varepsilon$$

$$\overset{*}{=} (\delta\square_f(\beta\square_g\gamma))\circ_r\varepsilon = (\delta\circ_r\varepsilon)\square_f(\beta\square_g\gamma) \ .$$

[6]: $$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma \overset{?}{=} (\delta\circ_r\varepsilon)\square_f(\beta\square_g\gamma) \ :$$

$$((\delta\circ_r\varepsilon)\square_f\beta)\square_g\gamma = (\delta\circ_r(\varepsilon\square_f\beta))\square_g\gamma = \delta\circ_r((\varepsilon\square_f\beta)\square_g\gamma)$$

$$\overset{*}{=} \delta\circ_r(\varepsilon\square_f(\beta\square_g\gamma)) = (\delta\circ_r\varepsilon)\square_f(\beta\square_g\gamma) \ .$$

This completes the proof of 4.(11).

Proof of 4.(12) Lemma:

Assume $\mathcal{L}_1$, $\mathcal{L}_2$ are languages with $\mathrm{O}(\mathcal{L}_1)=\mathrm{O}(\mathcal{L}_2)$. Temporarily, we will mean by an morphism $H:\mathcal{L}_1\to\mathcal{L}_2$ of languages a mapping $H:\mathrm{A}(\mathcal{L}_1)\to\mathrm{A}(\mathcal{L}_2)$ such that $\mathrm{s}_{\mathcal{L}_2}(H(f))=\mathrm{s}_{\mathcal{L}_1}(f)$ and $\mathrm{t}_{\mathcal{L}_2}(H(f))=\mathrm{t}_{\mathcal{L}_1}(f)$ for all $f\in\mathrm{L}(\mathcal{L}_1)$ (thus, we do not consider any action on the objects themselves now). Note that an "ample expansion" $F:\hat{\mathbf{C}}\to\mathbf{C}$ restricts to a morphism $\hat{\mathcal{L}}\to\mathcal{L}$.

We first choose a language $\hat{\mathcal{L}}$ with a morphism $F:\hat{\mathcal{L}}\to\mathcal{L}$ (as described) such that for every $f\in L$ there are infinitely many $\hat{f}\in\hat{L}=\mathrm{L}(\hat{\mathcal{L}})$ with $F(\hat{f})=f$. Take $\tilde{\mathbf{C}}=\mathcal{F}(\hat{\mathcal{L}})$ to be the free multicategory on $\hat{\mathcal{L}}$, say, with standard amalgamation. Using $\theta_g=\mathrm{id}_{\mathrm{s}(g)}$ for all $g\in\hat{L}$, we have a uniquely determined morphism $F_0:\mathcal{F}(\hat{\mathcal{L}})\longrightarrow\mathbf{C}$ of multicategories for which $F_0(g)=F(g)$ as given to begin with, for each $g\in\hat{L}$, and whose transition isomorphisms, at each $g \in \hat{L}$, are identities. "Twist" $\mathcal{F}(\hat{\mathcal{L}})$ by using the transition isomorphisms of F_0; that is, use the factorization of F_0 into an isomorphism Φ which is the identity on objects and arrows, and a strict morphism as in 2.:

$$\begin{array}{ccc} \mathcal{F}(\hat{\mathcal{L}}) & \xrightarrow{\quad F_0 \quad} & \boldsymbol{C} \\ & {\scriptstyle\Phi}\searrow \quad \circ \quad \nearrow{\scriptstyle F} & \\ & \hat{\boldsymbol{C}} & \end{array}$$

$\hat{\boldsymbol{C}}$ is free as well, since, by the characterization 3.(1), being free is invariant under twisting (or, because $\hat{\boldsymbol{C}} \cong \tilde{\hat{\boldsymbol{C}}}$). $F : \hat{\boldsymbol{C}} \to \boldsymbol{C}$ is the desired ample expansion. We define $\hat{\mathrm{d}} : \hat{\boldsymbol{C}} \to \boldsymbol{E}$ as the composite $\hat{\mathrm{d}} = \mathrm{d} \circ F$; $\hat{\mathrm{d}}$ is a strict morphism which is the identity on objects;

Proof of 4.(13) Lemma:

The proof is by induction on α.

If $\alpha = 1_X$, $\beta = 1_X$ is the necessary choice. Let $\alpha = f(\alpha_1, \ldots \alpha_n)$. We apply the induction hypothesis successively to $\alpha_1, \ldots, \alpha_n$; there is separated $\beta_1 \in \hat{A}$ with $F(\beta_1) = \alpha_1$ such that $I \cap \|\beta_1\| \neq \emptyset$; there is $\beta_2 \in \hat{A}$ with $F(\beta_2) = \alpha_2$ and $\|\beta_2\| \cap (I \cup \|\beta_1\|) = \emptyset$; $\ldots$; there is $\beta_n \in \hat{A}$ with $F(\beta_n) = \alpha_n$ and $\|\beta_n\| \cap (I \cup \|\beta_1\| \cup \|\beta_2\| \cup \ldots \cup \|\beta_{n-1}\|) = \emptyset$. Now, let $J = I \cup \|\beta_1\| \cup \|\beta_2\| \cup \ldots \cup \|\beta_n\|$. By assumption, there is $g \in \hat{L}$ such that $g \notin \psi$, and $F(g) = F(f)$. In particular, $\mathrm{s}(g) = \mathrm{s}(f)$. Thus, $\beta = g(\beta_1, \ldots, \beta_n) \in \hat{A}$ is well-defined, and since the sets $\{g\}$, $\|\beta_1\|$, $\|\beta_2\|$, $\ldots$, $\|\beta_n\|$ are pairwise disjoint, β is separated. Also, the construction ensures that $\|\beta\| \cap I = \emptyset$. Since F is a morphism, $F(\beta) = F(g)(F(\beta_1), \ldots, F(\beta_n)) = \alpha$. The proof is complete.

References

[B/D1] J. Baez and J. Dolan, Higher-dimensional algebra and topological quantum field theory, *Journal of Mathematical Physics* **36** (1995), 6073-6105.

[B/D2] J. Baez and J. Dolan, letter to R. Street, Nov. 30, 1995, corrected version as of Dec. 3, 1995; available at http://math.ucr.edu/home/baez/ .

[B/D3] J. Baez and J. Dolan, Higher dimensional algebra III: n-categories and the algebra of opetopes. Preprint, February 5, 1997.

[Ba] M. Batanin, On the definition of weak ω-category. *Macquarie Mathematics Reports* no. **96/207**.

[Be] J. Benabou, Introduction to bicategories. In: Lecture Notes in Mathematics, vol **47**, Springer Verlag, 1967; 1-77.

[G/P/S] R. Gordon, A. J. Power and R. Street, Coherence for tricategories. *Memoirs of the American Mathematical Society*, vol. 117 (1995), no.**558**.

[G] A. Grothendieck, Categories fibrees et descente. In: *Revetements Etales et Groupe Fondamental*. Lecture Notes in Mathematics **224**, Springer-Verlag, Berlin (1970); Expose VI, pp. 145-194.

[J] M Johnson, *Pasting Diagrams in n-Categories with Applications to Coherence Theorems and Categories of Paths.* PhD Thesis, University of Sydney, October 1987.

[L1] J. Lambek, Deductive systems and categories II. In: Lecture Notes in Mathematics, no. **86**, Springer-Verlag (1969); pp. 76-122.

[L2] J. Lambek, Multicategories revisited. In: *Categories in Computer Science and Logic,* Proceedings, Boulder, 1987. Contemporary Mathematics 92, Americal Mathematical Society (1989); pp. 217-240.

[M L] S. Mac Lane, *Categories for the Working Mathematician*. Springer Verlag, 1971.

[M1] M. Makkai, Avoiding the axiom of choice in general category theory. *Journal of Pure and Applied Algebra* **108** (1996), 109-173.

[M2] M. Makkai, Towards a categorical foundation of mathematics. To appear in the *Proceedings of the Haifa Logic Colloquium, August 1995*.

[M3] M. Makkai, *First Order Logic with Dependent Sorts, with Applications to Category Theory*. Research monograph. Accepted for publication in Lecture Notes in Logic, Springer Verlag. Available electronically.

[M4] M. Makkai, On structuralism in mathematics. To appear in a memorial volume dedicated to the memory of John Macnamara. Edited by R. Jackendoff, Brandeis University.

[P1] A. J. Power, A 2-categorical pasting theorem. *Journal of Algebra* **129** (1990), 439-445.

[P2] A. J. Power, An n-categorical pasting theorem. In: *Category Theory,* Proceedings, Como 1990, Springer Lecture Notes in Mathematics **1488**, Springer-Verlag, Berlin (1991), 326-358.

[S1] R. Street, The algebra of oriented simplexes. *Journal of Pure and Applied Algebra* **49** (1987), 283-335.

[S2] R. Street, Parity complexes. *Cahiers Topologie Geom. Differentielle Categoriques* **35** (1991), 315-343.

[S3] R. Street, Higher categories, strings, cubes and simplex equations. *Applied Categorical Structures* **3** (1995), 29-77.

[S4] R. Street, Categorical Structures. In: *Handbook of Algebra,* vol.1. Edited by M. Hazewinkel. Elsevier Science B. V. (1996); pp. 529-577.

[T] Z. Tamsamani, *Sur des notions de n-categorie et n-groupoide non-strict via des ensembles multi-simpliciaux.* These, Laboratoire de Topologie et Geometrie, Universite Paul Sabatier, Toulouse, France.