Bicategories in algebra and linguistics.[0)]

J. Lambek, McGill University

To Saunders Mac Lane on his 90th birthday.

Reporting on applications of bicategories to algebra and linguistics led me to take a new look at multicategories and polycategories: to replace free monoids by free categories and to introduce a new notation for Gentzen's cuts. This makes it clear that the equations holding in a multi- or polycategory are just those of the 2-category which contains it. Thus, a polycategory is almost the same as a 2-category whose underlying 1-category is freely generated by a graph, except that the class of 2-cells need not be closed under composition, but only under planar cuts.

0. Summary of contents

In Section 1 we point out that multicategories, slightly generalized, will do for bicategories what they originally did for monoidal categories, i.e. bicategories with one object. At the same time we introduce a new notation for Gentzen's "cut", to present it as a special case of composition in a 2-category.

In Section 2 we look at adjunctions in 2-categories and bicategories, with the aim of studying those bicategories in which each 1-cell has both a left and a right adjoint, namely compact noncommutative $*$-autonomous categories with several 0-cells.

In Section 3 we give a short exposition of some applications of bicategories to linguistics that were developed by Claudia Casadio and the present author. These touch on three deductive systems: the syntactic calculus, classical bilinear logic and compact bilinear logic.

In Section 4 we take a new look at polycategories, which are to classical bilinear logic as multicategories are to the syntactic calculus. Equations in a polycategory are explained by viewing the latter as contained in a 2-category.

In Section 5 we show that polycategories in the new sense will do for the linear bicategories of Cockett, Seely and Koslowski what multicategories can do for Bénabou's original bicategories.

In Section 6 we study adjoints in polycategories and show that, in a polycategory with residual quotients and "zero" 1-cells, every 1-cell has both a left and a right adjoint.

1. Multicategories recalled.

A good part of my book "Lectures on rings and modules" [1966] was devoted to the residuated bicategory of bimodules, although, at the time, I did not know what a bicategory was. Later I learned from Bénabou [1967] that a bicategory resembles a 2-category in having 0-cells (in my case rings $R, S, \cdots$), 1-cells (bimodules ${}_RA_S : S \to R$) and 2-cells (bimodule homomorphisms $f : {}_RA_S \to {}_RA'_S$), except that composition of 1-cells (the tensor product ${}_RA_S \otimes {}_SB_T$) satisfies the usual identity laws ($R \otimes A \cong A \cong A \otimes S$) and associative law only up to coherent isomorphism. All these properties of the tensor product of bimodules may be derived from Bourbaki's [1948] universal property, which stipulates a bilinear mapping $\mathbf{m}_{AB} : AB \to A \otimes B$ such that, for each bilinear mapping $f : AB \to C$ into a bimodule ${}_RC_T$, there is a unique homomorphism $g : A \otimes B \to C$ such that $g \circ \mathbf{m}_{AB} = f$.

Influenced by an early collaboration with George Findlay, I was particularly interested in the fact that the bicategory of bimodules was *residuated*, there being canonical isomorphisms

$$\text{Hom }(A \otimes B, C) \cong \text{ Hom }(A, C/B) \cong \text{ Hom }(B, A\backslash C),$$

where

$$\begin{aligned} {}_R(C/B)_S &= \text{ Hom}_S(B, C), \\ {}_S(A\backslash C)_T &= \text{ Hom}_R(A, C). \end{aligned}$$

To explain the universal properties of $\otimes$ (*tensor*), $/$ (*over*) and $\backslash$ (*under*) in general bicategories, I introduced the concept of a multicategory [L 1969[1)], 1989]. A multicategory consisted of multilinear maps

$$f : A_1 \cdots A_m \to B,$$

which might be viewed as context-free rules in grammar (I have reversed the usual arrow to reflect the hearer's point of view), as deductions in logic (called "sequents" by Gentzen, though here without his structural rules: interchange, contraction and weakening) or as multisorted operations in algebra (where one might write $fa_1 \cdots a_m \in B$, the a_i being variables or indeterminates of type A_i).

Originally, I had assumed that the left side of a multilinear map lives in the free monoid generated by a set, but one may as well let it live in the free category generated by a graph, so that f becomes an arrow from $S \xleftarrow{A_1} \cdots \xleftarrow{A_m} T$ to $S \xleftarrow{B} T$. The step of replacing a monoid by a category, while obvious in the bimodule example, was taken in linguistics by Brame [1984, 1985, 1987] (who may not, however, agree with my [1999a] interpretation of his ideas). It does not make sense in any logical system which admits the interchange rule.

Among the multilinear maps are the identities $1_A : A \to A$, as in a category, but composition is replaced by a more restricted notion, called "cut" by Gentzen:

$$\frac{f : \Lambda \to A \quad g : \Gamma A \Delta \to C}{g \circ \Gamma f \Delta : \Gamma \Lambda \Delta \to C},$$

where capital Greek letters denote *strings* in the free monoid or *chains* in the free category, say

$$\Lambda = R \xleftarrow{L_1} \cdots \xleftarrow{L_k} S.$$

(At one time, I had written $g \,\&\, f$ for this cut, gaining simplicity at the cost of sacrificing information.)

There are the expected identity and associative laws for the cut, but also a kind of commutative law: if $f : \Lambda \to A$, $g : \Lambda' \to A'$ and $h : \Gamma A \Delta A' \Theta \to C$ then

$$(h \circ \Gamma f \Delta A' \Theta) \circ \Gamma \Lambda \Delta g \Theta = (h \circ \Gamma A \Delta g \Theta) \circ \Gamma f \Delta \Lambda' \Theta.$$

This and similar equations will become clear if we think of the multicategory as being contained in a 2-category (see Section 4 below). It should be pointed out that Γ and Δ here do not denote terms of the multicategory, but merely serve to remove the ambiguity of the notation $g \& f$ by indicating where f is substituted into g. However, in the 2-category they may be interpreted as horizontal compositions of 1-cells.

In earlier papers [L 1989, 1993a] I found it useful to pass to an internal language, where the 1-cells are thought of as sorts and variables of each sort are admitted. Thus, to each operation $f : A_1 \cdots A_m \to B$ there is associated a term $f a_1 \cdots a_m$ of sort B, where the a_i are terms, e.g. variables, of sort A_i. Then $1_A : A \to A$ gives rise to the term $1_A x = x$, where x is a variable of sort A. If $\vec{x}, \vec{c}$ and $\vec{d}$ are appropriate strings of variables,

$$g \circ \Gamma f \Delta : \Gamma \Lambda \Delta \to C$$

gives rise to the term

$$(g \circ \Gamma f \Delta)(\vec{c}\vec{x}\vec{d}) = g \vec{c} f \vec{x} \vec{d},$$

which results by substituting $f\vec{x}$ for x in $g\vec{c}x\vec{d}$. The associative, commutative and identity laws can now be proved, provided we identify two operations f and f' whenever $f\vec{x} = f'\vec{x}$ is provable in the language. The variables may be called "indeterminates" in algebra and "assumptions" in logic.

Bourbaki's universal property for the tensor product stipulates a multilinear map $\mathbf{m}_{AB} : AB \to A \otimes B$ such that, for every $f : \Gamma AB\Delta \to C$, there exists a unique $g : \Gamma A \otimes B \Delta \to C$ such that $g \circ \Gamma \mathbf{m}_{AB} \Delta = f$, equivalently that

$$g\vec{c}\mathbf{m}_{AB}ab\vec{d} = f\vec{c}ab\vec{d}.$$

Similar universal properties may be given for the operations "over" and "under". Thus, one stipulates a multilinear map $\mathbf{e}_{CB} : (C/B)B \to C$ such that, for every $f : \Lambda B \to C$ there exists a unique $g : \Lambda \to C/B$ such that $\mathbf{e}_{CB} \circ gB = f$, equivalently that

$$\mathbf{e}_{CB} g\vec{l}b = f\vec{l}b.$$

It is now easy to prove that the operations $\otimes$, $/$ and $\backslash$ are bifunctors and that $\otimes$ is associative up to coherent isomorphism. In particular, one can derive Mac Lane's [1967] famous pentagonal condition. See e.g. [loc. cit.], where it was unnecessarily assumed that the multicategory has only one 0-cell.

Following Gentzen, one may reformulate the rules for tensor, over and under as introduction rules on the left and on the right:

$$\frac{\Gamma AB\Delta \to C}{\Gamma A\otimes B\Delta \to C}\quad,\quad \frac{\Gamma \to A \quad \Delta \to B}{\Gamma\Delta \to A\otimes B},$$

$$\frac{\Gamma C\Delta \to \Delta \quad \Lambda \to B}{\Gamma C/B\Lambda\Delta \to D}\quad,\quad \frac{\Lambda B \to C}{\Lambda \to C/B}.$$

The rules for \ are obtained from those for / by taking the mirror image on each side of the arrow. All these rules are subject to appropriate equations.

While these introduction rules incorporate some cuts, no further cuts are necessary. A categorical version of Gentzen's cut elimination theorem asserts the following:

PROPOSITION 1.1. Given a multicategory $\mathcal{M}$, one may construct the free residuated tensored multicategory $F(\mathcal{M})$ generated by $\mathcal{M}$ without using any cuts or identities, except those in $\mathcal{M}$.

For example, $1_{A\otimes B}$ and $1_{C/B}$ may be constructed as follows:

$$\cfrac{\cfrac{A \to A \quad B \to B}{AB \to A\otimes B}}{A\otimes B \to A\otimes B} \qquad\qquad \cfrac{\cfrac{C \to C \quad B \to B}{(C/B)B \to C}}{C/B \to C/B}$$

where we introduce $\otimes$ first on the right and then on the left, but we introduce / first on the left and then on the right. Similar rules hold for the cartesian product and coproduct. For proofs see [loc. cit.].

2. Adjoints in bicategories.

I have lately become interested in adjoints in 2-categories and bicategories. For the usual definition of a 2-category see [Mac Lane 1971]. For the present purpose, a 2-category may be described as having 0-cells (objects), 1-cells (arrows) and 2-cells (transformations). The first two constitute a category and the last act as arrows between 1-cells subject to a *vertical composition*

$$\frac{f : A \to B \quad g : B \to C}{g\circ f :\ A \to C}$$

and identity arrows $1_A : A \to A$ rendering the 1-cells $S \to R$ objects of a category. The 2-cells can be composed with 1-cells and behave like natural transformations in the familiar 2-category of categories: given

$$T \quad \begin{matrix} \xrightarrow{A} \\ \Downarrow f \\ \xrightarrow[B]{} \end{matrix} \quad S \quad \begin{matrix} \xrightarrow{F} \\ \Downarrow t \\ \xrightarrow[G]{} \end{matrix} \quad R$$

one has the commutative diagram

$$\begin{array}{ccc} FA & \xrightarrow{Ff} & FB \\ {\scriptstyle tA}\downarrow & & \downarrow{\scriptstyle tB} \\ GA & \xrightarrow[Gf]{} & GB \end{array}$$

which may be interpreted mnemonically as describing the *naturality* of t. Its diagonal defines the *horizontal* composition

$$tf = tB \circ Ff = Gf \circ tA.$$

Moreover, 1-cells *distribute* over composition of 2-cells:

$$F(g \circ f) = Fg \circ Ff, \quad (g \circ f)H = gH \circ fH.$$

In the same spirit,

$$F1_A = 1_{FA} = 1_F A.$$

The exchange property [Mac Lane 1971]

$$ug \circ tf = (u \circ t)(g \circ f)$$

may be deduced. Conversely, one presentation of 2-categories may be deduced from the traditional one [loc.cit.], by defining

$$tA = t1_A, \quad Ff = 1_F f.$$

An *adjunction* between 1-cells $F : R \to S$ and $U : S \to R$ in a 2-category is given by transformations $\eta : 1_R \to UF$ and $\varepsilon : FU \to 1_S$ such that

$$U\varepsilon \circ \eta U = 1_U, \quad \varepsilon F \circ F\eta = 1_F.$$

PROPOSITION 2.1 Adjoints in a 2-category are unique up to isomorphism.

While this is well-known, I have never seen a proof and shall produce one here, as it is a little tricky and because the same proof will also serve for the analogous result for polycategories in Section 6 below.

Proof: Suppose, for example, that U has another left adjoint F' given by $\eta' : 1_R \to UF'$ and $\varepsilon' : F'U \to 1_S$. We claim that the composite transformations

$$\varphi = \varepsilon F' \circ F\eta', \quad \psi = \varepsilon' F \circ F'\eta$$

are inverse to one another. Here, for example is a proof that $\psi \circ \varphi = 1_F$:

$$
\begin{aligned}
\psi \circ \varphi &= \psi \circ \varepsilon F' \circ F\eta' \\
&= \varepsilon F \circ FU\psi \circ F\eta' \quad \text{by naturality of } \varepsilon \\
&= \varepsilon F \circ FU\varepsilon' F \circ FUF'\eta \circ F\eta' \quad \text{by distributivity of } FU \\
&= \varepsilon F \circ FU\varepsilon' F \circ F(UF'\eta \circ \eta') \quad \text{by distributivity of } F \\
&= \varepsilon F \circ FU\varepsilon' F \circ F(\eta' UF \circ \eta) \quad \text{by naturality of } \eta' \\
&= \varepsilon F \circ F(U\varepsilon' \circ \eta' U)F \circ F\eta \quad \text{by distributivity of } F \\
&= \varepsilon F \circ F\eta \quad \text{since } U\varepsilon' \circ \eta' U = 1_U \\
&= 1_F.
\end{aligned}
$$

The notion of adjunction has been generalized to bicategories, e.g. by Kelly [1972] and Street and Walters [1978]. When is $F : R \to S$ left adjoint to $U : S \to R$? We require 2-cells $I_R \to U \otimes F$ and $F \otimes U \to I_S$ such that the composite 2-cell $U \to I_R \otimes U \to (U \otimes F) \otimes U \to U \otimes (F \otimes U) \to U \otimes I_S \to U$ is the identity on U and similarly for the analogous 2-cell $F \to F$. For example, we may ask: when does a bimodule ${}_RU_S$ have a left adjoint ${}_SF_R$? (Note that I_R is the bimodule ${}_RR_R$.) I was surprised to find the answer in the exercises to Section 4.1 in my 1966 book: ${}_RU_S$ has a left adjoint ${}_SF_R$ if and only if U_S is finitely generated and projective, and then $F = S/U$.

I have become particularly interested in bicategories in which each 1-cell has both a left and a right adjoint. Such bicategories are called *compact*, following Kelly [1972][2)]. For expository purposes, let me now confine attention to 2-categories with one object in which all 2-cells are just partial orders.

A *pregroup* is a partially ordered monoid with two operations $(-)^\ell$ and $(-)^r$ satisfying

$$a^\ell a \to 1 \to aa^\ell, \quad aa^r \to 1 \to a^r a$$

for each element a, the arrow denoting the partial order. In the discrete case, when the arrow denotes equality, a pregroup is just a group. More generally, in the cyclic[3)] case, when $a^\ell = a^r$, a pregroup is just a partially ordered group. My favorite example [L1994, 1995b] of a non-cyclic pregroup is the monoid of unbounded monotone mappings $f : \mathbf{Z} \to \mathbf{Z}$ under composition with elementwise order. Then adjoints may be defined thus:

$$
\begin{aligned}
f^r(y) &= \max\{x \in \mathbf{Z} | f(x) \le y\}, \\
f^\ell(y) &= \min\{x \in \mathbf{Z} | y \le f(x)\}.
\end{aligned}
$$

For example, let $f(x) = 2x$, then

$$f^r(y) = [y/2], \quad f^\ell(y) = [(y+1)/2].$$

We note that

$$f^{rr}(y) = 2y + 1 \neq f(y);$$

but, of course, $f^{r\ell} = f = f^{\ell r}$. Other examples of pregroups are provided by all natural languages, as we shall illustrate with English in the next section.

3. Linguistic applications.

We shall look at three possible applications of bicategories to linguistics, see [Casadio and Lambek t.a.]. For expository purposes, we first confine attention to partially ordered monoids:

(1) *Residuated monoids*, namely partially ordered monoids with two operations / and \ such that

$$a \cdot b \to c \ \text{ iff } \ a \to c/b \ \text{ iff } \ b \to a \backslash c.$$

(2) *Grishin algebras*, namely residuated monoids with a dualizing element 0 such that

$$0/(a\backslash 0) = a = (0/a)\backslash 0.$$

It is convenient to write

$$0/a = a^{\ell}, \quad a\backslash 0 = a^{r}$$

and one can introduce a second associative operation + by defining

$$a + b = (b^{\ell} \cdot a^{\ell})^{r} = (b^{r} \cdot a^{r})^{\ell}$$

such that

$$a + 0 = a = 0 + a.$$

One can then prove Grishin's [1983] mixed associative laws:

$$c(a + b) \to ca + b, \quad (a + b)c \to a + bc.$$

(3) *Pregroups* as defined above. These may also be viewed as *compact* Grishin algebras, in which

$$a + b = a \cdot b, \quad 0 = 1.$$

In the corresponding deductive systems, the arrow is not restricted to be a partial order, but equality of arrows is usually ignored. The above partially ordered monoids will then give rise to the following deductive systems:

(1) The *syntactic calculus*, introduced in [L1958] to study sentence structure.

(2) *Classical bilinear logic* [Abrusci 1991, L1993b], which was pioneered by Claudia Casadio [1997] for grammatical investigations.

(3) *Compact bilinear logic*, recently proposed by me [L 1999c] for linguistic applications.

Each of these deductive systems becomes a monoidal category (that is, a bicategory with one object), once attention is paid to equality between arrows:

(1) A residuated monoidal category.

(2) A noncommutative $*$-autonomouns category.

(3) A compact noncommutative $*$-autonomous category.

As already pointed out, one may remove the restriction to one object. In Linguistics, this step has already been taken by Brame [1984, 1985, 1987].

The idea common to the linguistic applications of these three systems is this: one assigns to each word, say of English, one or more *syntactic types*, namely elements of the free residuated monoid, Grishin algebra or pregroup generated by a partially ordered set of basic types and then calculates the type or types of any string of words. We shall illustrate this idea by looking at a single English sentence:

whom had she kissed–?

The dash at the end represents Chomsky's trace and is introduced for comparison only.

In (1), the words of this sentence are assigned the types

$$(\mathbf{q}'/(\mathbf{q}/\mathbf{o}))((\mathbf{q}_1/\mathbf{p}_2)/\pi_3)\pi_3(\mathbf{p}_2/\mathbf{o}) \to \mathbf{q}'.$$

The basic types employed here are:

$$\begin{aligned}
\mathbf{q}' &= \text{ question,}\\
\mathbf{q} &= \text{ yes-or-no question,}\\
\mathbf{q}_1 &= \text{ yes-or-no question in the present tense,}\\
\mathbf{o} &= \text{ object,}\\
\mathbf{p}_2 &= \text{ past participle,}\\
\pi_3 &= \text{ third person singular pronoun.}
\end{aligned}$$

The partial order on the set of basic types required for this example stipulates $\mathbf{q}_1 \to \mathbf{q} \to \mathbf{q}'$.

Although the method advocated here received a belated acceptance by a small group of linguists, I came to reject it myself for various reasons, one being the following. When a person hears the words *whom has*, she may calculate the type of this short string to be $(\mathbf{q}'/(\mathbf{p}_2/\mathbf{o}))/\pi_3$, but the formal proof of this, carried out in the syntactic calculus, is fairly long and, when put to paper, may occupy a quarter page. I had a strong feeling that this kind of calculation could not reflect the pychological reality of how people analyze speech.

In (2), the successive types for the same sentence are

$$(\mathbf{q}' + \mathbf{o}^{\ell\ell}\mathbf{q}^{\ell})(\mathbf{q}_1 + \mathbf{p}_2^{\ell} + \pi_3^{\ell})\pi_3(\mathbf{p}_2 + \mathbf{o}^{\ell}) \to \mathbf{q}'.$$

Here the type of *whom has* is easily calculated algebraically to be $\mathbf{q}' + \mathbf{o}^{\ell\ell}\mathbf{p}_2^{\ell} + \pi_3^{\ell}$; the calculation makes repeated use of the mixed associative laws.

In (3), the distinction between $+$ and $\cdot$ disappears and the successive types of the given sentence are simply

$$(\mathbf{q}'\mathbf{o}^{\ell\ell}\mathbf{q}^{\ell})(\mathbf{q}_1\mathbf{p}_2^{\ell}\pi_3^{\ell})\pi_3(\mathbf{p}_2\mathbf{o}^{\ell}) \to \mathbf{q}',$$

where the underlining indicates the cancellations

$$\mathbf{q}^{\ell}\mathbf{q}_1 \to \mathbf{q}^{\ell}\mathbf{q} \to 1, \quad \pi_3^{\ell}\pi_3 \to 1, \quad \mathbf{p}_2^{\ell}\mathbf{p}_2 \to 1, \quad \mathbf{o}^{\ell\ell}\mathbf{o}^{\ell} \to 1.$$

Here the type of the initial segment *whom has* is immediately seen to be $\mathbf{q}'\mathbf{o}^{\ell\ell}\mathbf{p}_2^{\ell}\pi_3^{\ell}$.

The reader should not be misled by this single example that only / and $(-)^{\ell}$ are useful in grammar and not \ and $(-)^{r}$. For example, in the sentence

she kissed me

the transitive verb has types

$$\pi_3\backslash(\mathbf{s}_2/\mathbf{o}), \quad \pi_3^{r} + \mathbf{s}_2 + \mathbf{o}^{\ell}, \quad \pi_3^{r}\mathbf{s}_2\mathbf{o}^{\ell}$$

in the three systems respectively, where

$$\mathbf{s}_2 = \text{ statement in the past tense.}$$

I believe that a good approximation to English grammar can be obtained by working with the *free* pregroup generated by a partially ordered set of basic syntactic types. For a better approximation, however, freeness must be abandoned. For example, it is difficult to justify the well-formedness of

people she knows like pizza

by the methods outlined above. The problem here is that there is no place for the type of the missing pronoun *whom*. To get around this, one may have to admit grammatical rules not listed in the dictionary, in other words, one may have to work with a pregroup which is not freely generated by its basic types.

4. A new look at polycategories.

If we adopt the slogan: don't ignore equality between deductions (also known in linguistics as derivations, productions or rewrite rules), a *production grammar* (also known as a semi-Thue system or rewrite system) is just a 2-category whose underlying 1-category is the free monoid generated by a set. Some greater generality is achieved if we allow the free category generated by a graph instead. In fact, for context-free grammars this generality has been advocated by Brame [loc.cit.].

We recall that a multicategory is essentially a context-free grammar dealing with deductions of the form

$$f : A_1 \cdots A_m \to B,$$

where juxtaposition on the left represents the tensor product, attention being paid to equality between deductions. In the presence of Gentzen's structural rules (interchange, contraction and weakening), these deductions are Gentzen's sequents for intuitionistic logic and the tensor product is just conjunction.

Gentzen also devised a deductive system for dealing with classical logic. Its sequents have the form

$$f : A_1 \cdots A_m \to B_1 \cdots B_n,$$

where juxtaposition on the left stands for conjunction and juxtaposition on the right stands for disjunction, although, in place of juxtaposition he had used commas on both sides. One may wonder why he did not use a comma on the left and a semicolon, say, on the right, to suggest the two different interpretations? We shall take advantage of his daring notation to embed polycategories, our categorical version of his system (in the absence of his structural rules) into 2-categories.

Polycategories may be regarded as underlying the grammar of Claudia Casadio [1997], where juxtaposition on the left and on the right of a deduction represent the tensor and the cotensor, its De Morgan dual, respectively. Compact polycategories, in which the tensor and cotensor are identified, are then essentially production grammars, hence 2- categories whose underlying 1-category is freely generated. Polycategories are like production grammars, except that composition of 1-cells is restricted to cuts.

Cuts in polycategories have the form

$$\frac{f : \Lambda \to \Gamma A \Delta \quad g : \Phi A \Psi \to \Theta}{\Gamma g \Delta \circ \Phi f \Psi : \Phi \Lambda \Psi \to \Gamma \Theta \Delta}$$

subject to the restriction that Γ or Φ is empty and Δ or Ψ is empty. Thus there are four cases:

Case 1. Φ and Ψ are empty and the conclusion is $\Gamma g \Delta \circ f : \Lambda \to \Gamma \Theta \Delta$.

Case 2. Φ and Δ are empty and the conclusion is $\Gamma g \circ f \Psi : \Lambda \Psi \to \Gamma \Theta$.

Case 3. Γ and Δ are empty and the conclusion is $g \circ \Phi f \Psi : \Phi \Lambda \Psi \to \Theta$.

Case 4. Ψ and Γ are empty and the conclusion is $g \Delta \circ \Phi f : \Phi \Lambda \to \Theta \Delta$.

The four cases may be illustrated by the following planar diagrams respectively:

$$\begin{array}{cccc} \Lambda & \Lambda & \Lambda & \Lambda \\ \overline{\Gamma A\Delta} & \overline{\Gamma A}\underline{\Psi} & \underline{\quad\Phi}\overline{A}\Psi & \underline{\quad\Phi}\overline{A\Delta} \\ \overline{\Theta} & \Theta & \Theta & \Theta \end{array}$$

where the 2-cells f and g are represented by horizontal lines.

In case the reader is still skeptical, here is a formal definition to convince her that a polycategory can be embedded into a 2-category with additional 2-cells.

Definition 4.1. A *polycategory over a graph*[4)] has 0-cells, 1-cells, 2-cells and equations between 2-cells:

- its 0-cells are the nodes of the graph;
- its 1-cells are the arrows of the free category generated by the graph;
- its 2-cells are certain arrows between 1-cells, Γ and Δ, assuming that Γ and Δ have the same source and target;
- among the 2-cells are all the identity arrows $1_A : A \to A$, where A is an arrow of the graph;
- the set of 2-cells is closed under the four kinds of cuts listed above;
- its equations are precisely those which hold in the 2-category obtained by allowing all identity 2-cells and arbitrary composition of 2-cells, Provided we interpret a cut with premisses $f : \Lambda \to \Gamma A\Delta$ and $g : \Phi A\Psi \to \Theta$ as the composition of $\Gamma g\Delta$ and $\Phi f\Psi$, as suggested by the above notation.

It was in order to get a grip on the possible equations between deductions that I had suggested the idea of a polycategory in [L1969]. However, I did not take the trouble to spell out exactly what equations had to hold. This was done by Szabo [1975], although he allowed too many cases of the cut for the substructural system of bilinear logic studied here. Polycategories were also investigated by Velinov [1988], who considered many variations, even the compact case. A detailed set of equations that meet my approval were presented by Cockett and Seely [1992], who were using polycategories to introduce the tensor and cotensor into what they then called "weakly distributive categories". In fact, they obtained an equivalence between the category of polycategories and the category of weakly distributive categories.

I believe that the present method of inferring all equations between deductions from those valid in 2-categories is new. (We shall ignore here another approach I have been exploring, which replaces the operations that had proved useful in multicategories by binary relations.) We shall look at a few examples of such equations. (A list of five such equations will be found in Cockett and Seely [1992]. I have not checked whether these five equations imply all the equations that can be inferred from those of 2-categories.)

EXAMPLE 1.

$$\frac{\Lambda \xrightarrow{f} \Gamma A \quad A\Delta \xrightarrow{g} \Phi B\Psi \quad B \xrightarrow{h} \Theta}{\Lambda\Delta \to \Gamma\Phi\Theta\Psi}$$

There are two ways of deriving the conclusion, depending on whether we first compose g with f or h with g. These are represented by the two sides of the equation

$$\Gamma\Phi h\Psi \circ (\Gamma g \circ f\Delta) = \Gamma(\Phi h\Psi \circ g) \circ f\Delta,$$

which is justified by associativity and distributivity in a 2-category. Note that the intermediate term

$$\Gamma(\Phi h\Psi \circ g) \circ f\Delta$$

does not live in the polycategory, but in the embedding 2-category.

EXAMPLE 2.

$$\frac{\Lambda \xrightarrow{f} \Gamma A\Delta \quad A \xrightarrow{g} \Phi B\Psi \quad B \xrightarrow{h} \Theta}{\Lambda \to \Gamma\Phi\Theta\Psi\Delta}$$

Here we have the equation

$$\Gamma\Phi h\Psi\Delta \circ (\Gamma g\Delta \circ f) = \Gamma(\Phi h\Psi \circ g)\Delta \circ f,$$

which is also justified by associativity and distributivity.

EXAMPLE 3.

$$\frac{\Phi \xrightarrow{f} A \quad \Psi \xrightarrow{g} B \quad \Gamma A\Delta B\Lambda \xrightarrow{h} \Theta}{\Gamma\Phi\Delta\Psi\Lambda \to \Theta}$$

Here the equation

$$(h \circ \Gamma A\Delta g\Lambda) \circ \Gamma f\Delta\Psi\Lambda = (h \circ \Gamma f\Delta B\Lambda) \circ \Gamma\Phi\Delta g\Lambda$$

may be reduced by associativity and distributivity to showing that

$$A\Delta g \circ f\Delta\Psi = f\Delta B \circ \Phi\Delta g,$$

which follows from naturality of f.

These examples should support the claim that the equations holding in a polycategory are precisely those which hold in the 2-category which contains it. In retrospect, the same is true for a multicategory. Perhaps a polycategory should have been called a "sesqui-category"! The algebraic derivations of the equations in the three examples above become redundant if one relies instead on the planar diagrams which illustrate how the conclusion is obtained:

$$\begin{array}{c} \Lambda \\ \hline \Gamma A\Delta \\ \Phi\underline{B}\Psi \\ \Theta \end{array} \qquad \begin{array}{c} \Lambda \\ \hline \Gamma A\Delta \\ \Phi\underline{\mathrm{B}}\Psi \\ \Theta \end{array} \qquad \begin{array}{c} \underline{\Phi} \quad \underline{\Psi} \\ \Gamma A\Delta B\Lambda \\ \hline \Theta \end{array}$$

where the horizontal lines represent the deductions f, g and h: two deductions are identified if they give rise to the same diagram.

A final example will illustrate the behaviour of the identity arrow.

EXAMPLE 4.

$$\frac{\Lambda \xrightarrow{f} \Gamma A\Delta \quad A \xrightarrow{1_A} A}{\Lambda \xrightarrow{f} \Gamma A\Delta}$$

Here we have

$$\Gamma 1_A \Delta \circ f = 1_{\Gamma A \Delta} \circ f = f.$$

5. Polycategories and linear bicategories.

After composing the first draft of this paper, I was presented with a copy of the article by Cockett, Koslowski and Seely [2000], in which they developed the notion of "linear bicategory" and studied "linear adjoints", a generalization of adjoints in the original bicategories of Bénabou.

One purpose of multicategories had been to introduce the tensor product $\otimes$ and the corresponding identity 1-cells I_R into a bicategory so that their properties can be proved instead of having to be postulated. Polycategories will do the same for linear bicategories in helping to introduce also the cotensor $\oplus$ and the corresponding zero 1-cells O_R. This program had in fact been carried out by Cockett and Seely [1992], although they had presented polycategories more directly than here.

We recapitulate the definitions of these operations in a present style polycategory:

$\otimes$ is given by $\mathbf{m}_{AB} : AB \to A \otimes B$ such that, for each $f : \Gamma AB\Delta \to \Theta$, there exists a unique $g : \Gamma A \otimes B\Delta \to \Theta$ such that $g \circ \Gamma \mathbf{m}_{AB}\Delta = f$.

I_R is given by $\mathbf{i}_R : 1_R \to I_R$ such that, for each $f : \Gamma\Delta \to \Theta$, there exists a unique $g : \Gamma I_R \Delta \to \Theta$ such that $g \circ \Gamma \mathbf{i}_R \Delta = f$. Here 1_R denotes the identity arrow $R \to R$ in a 2-category, that is, the empty chain between Γ and Δ in $\xleftarrow{\Gamma} R \xrightarrow{\Delta}$.

$\oplus$ is given by $\mathbf{n}_{AB} : A \oplus B \to AB$ such that, for each $f : \Theta \to \Gamma AB\Delta$, there exists a unique $g : \Theta \to \Gamma A \oplus B\Delta$ such that $\Gamma \mathbf{n}_{AB}\Delta \circ g = f$.

O_R is given by $\mathbf{j}_R : O_R \to 1_R$ such that, for each $f : \Theta \to \Gamma\Delta$, there exists a unique $g : \Theta \to \Gamma O_R \Delta$ such that $\Gamma \mathbf{j}_R \Delta \circ g = f$.

A *residuated polycategory* has residual quotients / and \, the first of which is introduced as follows:

/ is given by $\mathbf{e}_{AB} : (A/B)B \to A$ such that, for each $f : \Gamma B \to \Delta A$, there exists a unique $g : \Gamma \to \Delta A/B$ such that $\Delta \mathbf{e}_{AB} \circ gB = f$.

For \ one takes the mirror image of each side of the arrow.

One may also consider residual differences $\dot{-}$ (*less*) and $\dot{-}$ (*from*). For a discussion of these see [L1993b].

Gentzen style introduction rules for $\otimes$ and I_R take the following form, while those for $\oplus$ and O_R may be obtained by reversing the arrows:

$$\frac{\Gamma AB\Delta \to \Theta}{\Gamma A \otimes B\Delta \to \Theta} \quad , \quad \frac{\Gamma \to \Phi A \quad \Delta \to B\Psi}{\Gamma\Delta \to \Phi A \otimes B\Psi},$$

$$\frac{\Gamma\Delta \to \Theta}{\Gamma I_R \Delta \to \Theta} \quad , \quad 1_R \to I_R.$$

The introduction rules for / have the form

$$\frac{\Gamma A\Delta \to \Theta \quad \Lambda \to B}{\Gamma A/B\Lambda\Delta \to \Theta} \qquad \frac{\Gamma B \to \Delta A}{\Gamma \to \Delta A/B},$$

while those for $\backslash$ are obtained by taking the mirror image on each side of the arrow.

Here, for example, is how one may construct the arrow

$$(A \oplus B) \otimes C \to A \oplus (B \otimes C),$$

representing one of Grishin's mixed associative laws:

$$\cfrac{\cfrac{\cfrac{\cfrac{A \to A \quad B \to B}{A \oplus B \to AB \quad C \to C}}{(A \oplus B)C \to A(B \otimes C)}}{A \oplus B)C \to A \oplus (B \otimes C)}}{(A \oplus B) \otimes C \to A \oplus (B \otimes C)}$$

One introduces first $\oplus$ on the left and then $\otimes$ on the right, next $\oplus$ on the right and then $\otimes$ on the left,

At first sight, it looks as though the system comprising all these operations should enjoy the cut elimination property.[5)] Indeed, here are two cut-free proofs:

$$\cfrac{\cfrac{\cfrac{\cfrac{A \to A}{A \to AO}}{1 \to A(O/A)}}{I \to A(O/A)}}{I \to A \oplus (O/A)} \qquad\qquad \cfrac{\cfrac{O \to O \quad A \to A}{(O/A)A \to O}}{(O/A) \otimes A \to O}$$

where 1 denotes the empty string or chain and subscripts on 1, I and O have been omitted. Similarly one shows

$$I \to (B\backslash O) \oplus B \quad , \quad B \otimes (B\backslash O) \to O,$$

and one obtains

$$B\backslash O \to I \otimes (B\backslash O) \to (A \oplus (O/A)) \otimes (B\backslash O) \to A \oplus ((O/A) \otimes (B\backslash O)).$$

Taking $B = O/A$, so that

$$(O/A) \otimes (B\backslash O) \to O,$$

one thus obtains a deduction

$$(O/A)\backslash O \to A.$$

Evidently, this can have no cut-free proof.

6. Adjoints in polycategories.

The linear adjoints of Cockett, Koslowski and Seely may be traced back to polycategories. In fact, the definition of adjoints for 2-categories given in Section 2 remains valid for polycategories, once one realizes that the compositions $U\varepsilon \circ \eta U$ and $\varepsilon F \circ F\eta$ are cuts, illustrating cases 2 and 4 of Section 2:

$$\frac{\eta : 1_R \to UF \quad \varepsilon : FU \to 1_S}{U\varepsilon \circ \eta U : U \to U} \quad , \quad \frac{\eta : 1_R \to UF \quad \varepsilon : FU \to 1_S}{\varepsilon F \circ F\eta : F \to F}.$$

The proof of Proposition 2.1 also remains valid for polycategories, so we conclude:

PROPOSITION 6.1. In a polycategory any existing adjoints are unique up to isomorphism.

When can we infer that adjoints exist?

PROPOSITION 6.2. In a residuated polycategory with zero 1-cells, every 1-cell $A : S \to R$ has both a left and a right adjoint:

$$A^\ell = O_S/A \quad , \quad A^r = A\backslash O_R.$$

Proof: To show the existence of left adjoints, for example, we have to define

$$\varepsilon_A : (O_S/A)A \to 1_S \quad , \quad \eta_A : 1_R \to A(O_S/A)$$

and verify that

$$A\varepsilon_A \circ \eta_A A = 1_A \quad , \quad \varepsilon_A(O_S/A) \circ (O_S/A)\eta_A = 1_{O/A}.$$

We define $\varepsilon_A = \mathbf{j} \circ \mathbf{e}_{OA}$ by the cut

$$\frac{\mathbf{e}_{OA} : (O/A)A \to O \quad \mathbf{j} : O \to 1}{\varepsilon_A : (O/A)A \to 1}$$

and $\eta_A : 1 \to A(O/A)$ as the unique $g : 1 \to A(O/A)$ such that

$$A\mathbf{e}_{OA} \circ gA = f,$$

where $f : A \to AO$ is the unique arrow such that

$$A\mathbf{j} \circ f = 1_A :$$

$$\frac{A \xrightarrow{f} AO \quad O \xrightarrow{\mathbf{j}} 1}{A \to A} \quad , \quad \frac{1 \xrightarrow{g} A(O/A) \quad (O/A)A \xrightarrow{\mathbf{e}_{OA}} O}{A \to OA}.$$

Then

$$A\varepsilon_A \circ \eta_A A = A\mathbf{j} \circ A\mathbf{e}_{OA} \circ gA = A\mathbf{j} \circ f = 1_A.$$

To show the other equation to be proved, we recall that 1_{A^ℓ} is the unique $h : A^\ell \to A^\ell$ such that

$$\mathbf{e}_{OA} \circ hA = \mathbf{e}_{OA}.$$

Hence this equation has to be verified when

$$h = \varepsilon_A A^\ell \circ A^\ell \eta_A.$$

Indeed

$$\begin{aligned} \mathbf{e}_{OA} \circ hA &= \mathbf{e}_{OA} \circ \varepsilon_A A^\ell A \circ A^\ell \eta_A A \\ &= \varepsilon_A O \circ A^\ell A \mathbf{e}_{OA} \circ A^\ell \eta_A A \quad \text{by naturality of } \varepsilon_A \\ &= \mathbf{j}O \circ \mathbf{e}_{OA} O \circ A^\ell f \quad \text{by definition of } \varepsilon_a \text{ and } \eta_A \end{aligned}$$

Now, by lemma 6.3 below, we may replace $\mathbf{j}O$ by $O\mathbf{j}$, hence this

$$\begin{aligned}
&= \mathbf{e}_{OA} \circ A^{\ell}A\mathbf{j} \circ A^{\ell}f \quad \text{by naturality of } \mathbf{e}_{OA} \\
&= \mathbf{e}_{OA} \circ A^{\ell}(1_A) \quad \text{by definition of } f \\
&= \mathbf{e}_{OA} \circ 1_{A^{\ell}A} = \mathbf{e}_{OA}.
\end{aligned}$$

COROLLARY 6.4. In a residuated polycategory with zero 1-cells, the zeros are dualizing: for any 1-cell $A : S \to R$,

$$(O_S/A)\backslash O_R \cong A \cong O_S/(A\backslash O_R).$$

Proof: If A^{ℓ} is left adjoint to A, then both A and $A^{\ell r}$ are right adjoints of A^{ℓ}, hence $A^{\ell r} \cong A$ by Proposition 6.1. Similarly $A^{r\ell} \cong A$.

LEMMA 6.3. In a polycategory with zero 1-cells, $\mathbf{j}O = O\mathbf{j}$.

Proof: By the universal property of $\mathbf{j} : O \to 1$, any 2-cell $f : OO \to 1$ gives rise to a unique $g : OO \to O$ such that $\mathbf{j} \circ g = f$. Now take $f = \mathbf{j} \circ O\mathbf{j}$, then $g = O\mathbf{j}$. But, by naturality of $\mathbf{j}, \mathbf{j} \circ O\mathbf{j} = \mathbf{j} \circ \mathbf{j}O$, hence $g = \mathbf{j}O$.

7. Postscript.

This article is an elaboration of a talk at the 1999 category conference in Coimbra. Its major aim was to explain my idea of what the equations of a polycategory should be. I had introduced this concept in 1969 without spelling out these equations. In the mean time, attempts to produce such equations axiomatically were made by several authors, though not in agreement with one another. While the axioms provided by Cockett and Seely are “sound”, “completeness” with respect to the present treatment remains to be shown: the equations of a polycategory should be those that ensure its embedding into a 2-category to be faithful.

REFERENCES

V.M. Abrusci (1991), *Phase semantics and sequent calculus for pure noncommutative classical linear propositional logic*, J. Symbolic Logic **56**, 1403-1451.

M. Barr (1971), *$*$-Autonomous categories*, Springer LNM **752**.

....., (1995), *Non-symmetric $*$-autonomous categories*, Theoretical Computer Science **139**, 115-130.

J. Bénabou (1967), *Introduction to bicategories*, Springer LNM **47**, 1- 77.

N. Bourbaki (1948), *Algèbre multilinéaire*, Hermann, Paris.

M. Brame (1984, 1985, 1987), *Recursive categorical syntax and morphology* I, II, III, Linguistic Analysis **14**, 265-287; **15**, 137-176; **17**, 147-185.

C. Casadio (1997), *Unbounded dependencies in non-commutative logic*, in: Proc. Conference Formal Grammars, ESSLLI, Aix en Provence.

..... (2001), *Non-commutative linear logic in linguistics*, Grammars **3/4**, 1-19.

C. Casadio and J. Lambek (2002), *A tale of four grammars*, Studia Logica, **71**, 315-329.

J.R.B. Cockett and R.A.G. Seely (1997), *Weakly distributive categories*, J. Pure & Applied Algebra **114**, 133-173.

J.R.B. Cockett, J. Koslowski and R.A.G. Seely (2000), *Introduction to linear bicategories*, Math. Structures in Computer Science, **10**, 165-203.

V.N. Grishin (1983), *On a generalization of the Ajdukiewicz-Lambek system*, in: *Studies in non-commutative logics and formal systems*, Nauka, Moscow, 315- 343. English translation in: V.M. Abrusci and C. Casadio (eds), *New Perspectives in Logic and Formal Linguistics*, Bulzoni Editore, Roma 2002, 9-27.

G.M. Kelly (1972), *Many variable functorial calculus* I, Springer LNM **281**, 66-105.

S.C. Kleene (1952), *Introduction to metamathematics*, Van Nostrand, New York.

J. Lambek (1958), *The mathematics of sentence structure*, Amer. Math. Monthly **65**, 154-169.

..... (1968, 1976, 1986), *Lectures on rings and modules*, Blaisdell, Waltham Mass.; Ginn, New York, N.Y.; Chelsea, New York, N.Y.

..... (1969), *Deductive systems and categories* II, Springer LNM **87**, 76- 122.

..... (1989), *Multicategories revisited*, Contemporary Math. **92**, 217- 239.

..... (1993a), *Logic without structural rules*, in: K. Došen and P. Schroeder-Heister (eds.), *Substructural Logics*, Studies in Logic and Computation **2**, Oxford Science Publications, 179-206.

..... (1993b), *From categorial grammar to bilinear logic*, ibid., 207-237.

..... (1994), *Some Galois connections in elementary number theory*, J. Number Theory **47**, 371-377.

..... (1995a), *Bilinear logic in algebra and linguistics*, in: J.-Y. Girard et al. (eds), *Advances in Linear Logic*, London Math. Soc. Lecture Notes Series **222**, Cambridge University Press.

..... (1995b), *Some lattice models of bilinear logic*, Algebra Universalis **34**, 541-550.

..... (1999a), *Deductive systems and categories in linguistics*, in: H.J. Ohlbach and U. Reyle (eds), *Logic, language and reasoning*, Kluwer Academic Publishers, Dordrecht, 279-294.

..... (1999b), *Bilinear logic and Grishin algebras*, in: E. Orlowska (ed.), *Logic at work*, Essays dedicated to the memory of Helena Rasiowa, Physica-Verlag, Heidelberg, New York, 604-612.

..... (1999c), *Type grammars revisited*, in: A. Lecomte, F. Lamarche and G. Perrier (eds), *Logical Aspects of Computational Linguistics*, Springer LNAI **1582**, 1-27.

F.W. Lawvere (1973), *Metric spaces, generalized logic, and closed categories*, Rend. Sen. Mat. E Fis. Milano **43**, 135-166.

S. Mac Lane (1963), *Natural associativity and commutativity*, Rice University Studies **49**, 28-46.

..... (1971), *Categories for the working mathematician*, Springer-Verlag, New York, N.Y.

K.I. Rosenthal (1994), $*$-*autonomous categories of bimodules*, J. Pure & Applied Algebra **97**, 188-201.

R. Street and R.F.C. Walters (1978), *Yoneda structures on 2-categories*, J. Algebra **50**, 350-379.

M.E. Szabo (1975), *Polycategories*, Communications in Algebra **3**, 663-698.

Y. Velinov (1988), *An algebraic structure for derivations in rewriting systems*, Theoretical Computer Science *57*, 205-224.

D.N. Yetter (1990), *Quantales and (non-commutative) linear logic*, J. Symbolic Logic **55**, 41-64.

ENDNOTES

[0] This research was supported by NSERC and SSHRC.

[1] The cited paper contains some mistakes, but these do not affect the concept of a multicategory.

[2] Kelly used the word "compact" for symmetric monoidal categories in which each object has a right adjoint. These are the compact $*$-autonomous categories of Barr [1979], and the concept was generalized to bicategories by Street and Walters [1978].

[3] The word "cyclic" in this context is due to Yetter [1990].

[4] By a *graph* is here understood what graph theorists call an "oriented multigraph".

[5] I had mistakenly thought so in [L1993b], although a counterexample had been produced by Abrusci [1991].