Who am I, and why do research?

I've been a teacher at JAC in the mathematics department since the fall of 1979; in addition to teaching, I have also had a "parallel career" doing mathematical research, specifically work in logic and theoretical computer science, being funded by various agencies such as FCAR, NSERC, and currently, FRQNT in its Programme de Recherche pour les Enseignants de Collège. Over the years, I've published many research papers in scholarly journals, been an editor of journals and special collections of papers, and generally participated in the usual activities of an active scholar in mathematics. Most years I have had release time for my research, either funded by Québec agencies or by the college, and have been fortunate to have worked in an environment at John Abbott which has been remarkably supportive of pure academic research, as illustrated by today's activity.

I think such research, although often not directly related to our courses, nonetheless has a positive impact on our activities in the classroom. In my own case, I have found it possible to directly introduce material from my research into the college classroom, to the benefit of both my students and my colleagues, as may be seen from some examples. My work on monoidal categories has direct connections with mathematical linguistics, and I have included this application of monoidal category theory in my *Principles of Mathematics and Logic* course for Liberal Arts students at John Abbott. My course text even includes some excerpts from a couple of my research papers. Insights from my current work on algebraic approaches to differential calculus have influenced the way I teach college-level calculus (especially Calculus III). In general, my classroom material is influenced by my research in many ways, from subtle presentational matters to explicit descriptions of current mathematical research, which help students see that mathematics is a living, growing, even organic, subject, and not merely a dead repetition of meaningless formulas and rules. Many students have responded with enthusiasm to these additions to my classroom repertoire.

But perhaps most importantly, doing new, fresh, work can help keep one's own attitude to the subject new and fresh, and can give to one's class presentation an enthusiasm which helps grab students' attention and can even give them the impression that what's going on may actually be exciting(!). It also can help with one's interactions with colleagues; I know I've learned much from conversations with fellow teachers about matters sometimes closely, sometimes more distantly, related to my research, and always related to our course material.

In this poster, I'll try to give some idea of my current research project, although I'll try to keep technical material to a reasonable level. Feel free to skip over anything you don't understand! — there won't be an exam!.

Introduction (not too technical!)

(This intro should be understandable by anyone who has studied Calculus III; I make no such promise about the rest of this poster!)

What is the essential (algebraic) structure of undergraduate-level differential calculus? Specifically, what structure is needed for basic differentiation? There have been several answers in the literature; in collaboration with Richard Blute (University of Ottawa) and Robin Cockett (University of Calgary), I have developed (and continue to work on) answers to this question based on some intuitions from categorical logic, specifically some categorical models of linear logic.

Why would this be of interest? Well, there are structures occurring elsewhere in mathematics and in theoretical computer science that share many of the properties of differentiation, even though they are not at all what one studies in college calculus courses. In some cases, these structures are related to very practical matters, such as the efficiency of some computational algorithm (such as a computer programme), in other cases they are merely something which seems to observers to be a beautiful pattern, about which it would be nice to know more. In either case, there is strong motivation to study these structures, and if one can understand them better by seeing their connection—hitherto unnoticed, perhaps—to ordinary differentiation, then one is that much further to understanding what is behind the pattern.

Some examples to motivate the mathematics

Let's consider some simple examples. First, *really* simple(!): Take a function $f(x) = x^2$ (this goes from R to R). Usually, we say that the derivative is $f'(x) = 2x$: but what this really means is that for any input a, the slope is $2a$, *i.e.* the *slope function* takes an input u to output $2au$. So what we really have is a function which produces a linear function as output. All we actually will need is that this function preserves addition: we call such a function additive.

Let's generalize this a tiny bit by allowing multiple-value inputs and outputs. Suppose we have a smooth map (*i.e.* one with all necessary derivatives) $g\colon R^3 \to R$, such as $g(x,y,z) = xyz$. We recall the Jacobian of this map (which for such maps is essentially the derivative) is $[yz, xz, xy]$; this may be regarded as a smooth operator J which, given input (x,y,z) assigns a linear operator, $J(x,y,z)$, represented by the matrix $[yz, xz, xy]$, which in turn, given any input $[u,v,w]$ assigns an output

$$J(x,y,z)(u,v,w) = [yz, xz, xy] \cdot [u,v,w] = yzu + xzv + xyw$$

We can avoid the notion of a "function-valued function" by simply going to the last step, so that what we really have, from the initial smooth function g, is a smooth map $D_\times(g)\colon R^3 \times R^3 \to R$, which is additive in the first variable (in the first triple of coordinates):

$$D_\times(g)\colon ((u,v,w),(x,y,z)) \mapsto [yz, xz, xy] \cdot [u,v,w] = yzu + xzv + xyw$$

This suggests that the basic structure we need is the following. First two types of function: smooth and additive. Then a "derivative" operation, which, given a smooth function, produces another smooth function, which is additive in the first variable. In symbols:

$$\frac{f\colon X \to Y}{D_\times[f]\colon X \times X \to Y}$$

Introduction (technical summary)

The technical definition resulting from these ideas is the following:

Definition:

A **Cartesian differential category** is a Cartesian left additive category which has a **Cartesian differential operator** $D_\times$ on the maps of the category,

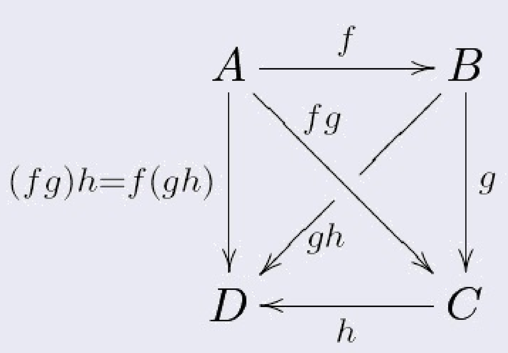
$$\frac{f\colon X \to Y}{D_\times[f]\colon X \times X \to Y}$$

which must satisfy certain equations.

Intuition?

To the non-expert, that definition may seem rather obscure. Without going into detail, maybe a few pointers can help direct one to an intuitive idea of what is intended; of course, the technical details are necessary to "get" the entire picture, and particularly its utility, but I hope the following remarks may be useful to some readers.

Category theory is in essence the abstract theory of functions and function composition, and a category may be viewed as consisting of objects and maps (or functions) between them. In general, these "objects" and "functions" may be very different from what one usually means by these terms, but for the examples we have in mind, that intuition isn't really very bad. One has "identity maps", and one can compose maps where appropriate; there are "obvious" equations that must hold (such as $f(gh) = (fg)h$).



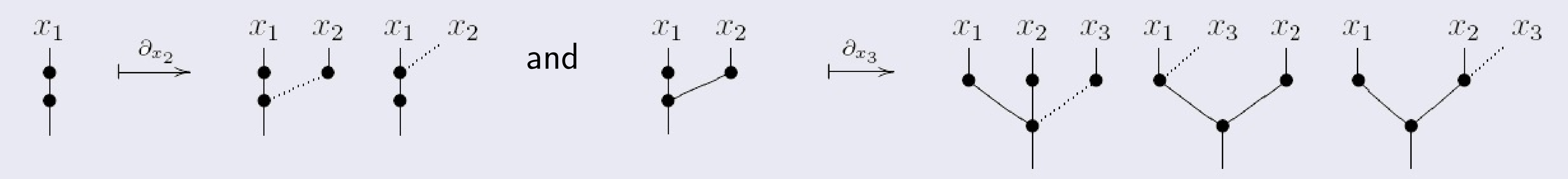
A "Cartesian left additive category" has more structure: there are "Cartesian products" $A \times B$ (as one has with sets), and one can add maps $f + g$ (as in ordinary algebra), again satisfying appropriate equations. So a Cartesian differential category is a structure where one can form products, add maps, and where one has a derivative operation. The axioms are just basic properties of differentiation, such as the condition "additive in the first variable", and the chain rule.

More intuition?

The next bit is rather technical, but I'll try to give some "feeling" for what it will say. It's a simple Cal I exercise to get a chain rule formula for second order (and higher) derivatives, and in a similar way, we can do that for our Cartesian differential categories. These the chain rules tell us how to compose (higher order) derivatives. So we can put together a "family of categories", whose composition is given by the (higher order) chain rules (this is called a fibration), which represents all the higher order categorical structure of differentiation. This family is the essence of differentiation, and is a very natural structure (even if it is rather complicated to describe in detail!). The Cartesian differential category we started with can be regained from this family (as the "bottom level").

The construction we arrived at (which we named Faà after a 19th century Italian priest named Francesco Faà di Bruno, beatified by Pope John Paul II apparently for his charitable work teaching young women mathematics, and who formulated the higher order chain rule for "ordinary, Calculus I, calculus") was, as was Faà di Bruno's original, a quite complicated piece of combinatorics. As neither of us (my coauthor Robin Cockett or myself) was experienced as a combinatorialist, we found it simpler to ignore Faà di Bruno's formula, and work from scratch. And we found it useful to create some "pictorial" structures to codify the categorical calculations, in particular, some "labelled trees", as seen below. And we had a notion of the "derivative" of such a tree, which would be a bag of trees created from the original by adding a new label appropriately in as many ways as possible, as illustrated below for two trees of height 2 (widths 1, 2).

Exercise: see if you can calculate the "derivative" ∂_{x_4} of one of the width 3 trees, and also see if you can work out how many such trees there are of a given height and width. (You didn't *really* think there'd be no homework, did you??)



Characterization of Cartesian differential categories

Theorem The category of coalgebras for the comonad Faà is equivalent to the category of Cartesian differential categories and Cartesian differential functors (Cartesian left additive functors which preserve the differential combinators).

Remark: Although I have not defined all the notions in the Theorem, in fact it gives a complete characterization of Cartesian differential categories as a natural byproduct of the structure of the (higher order) chain rule. In effect, this says the structure we have defined is as tightly connected with the essence of differentiation as one could hope.

As a personal remark, this result was a complete surprise to us both as we gradually realized what we had discovered, a very pleasant surprise indeed.

An alternate viewpoint

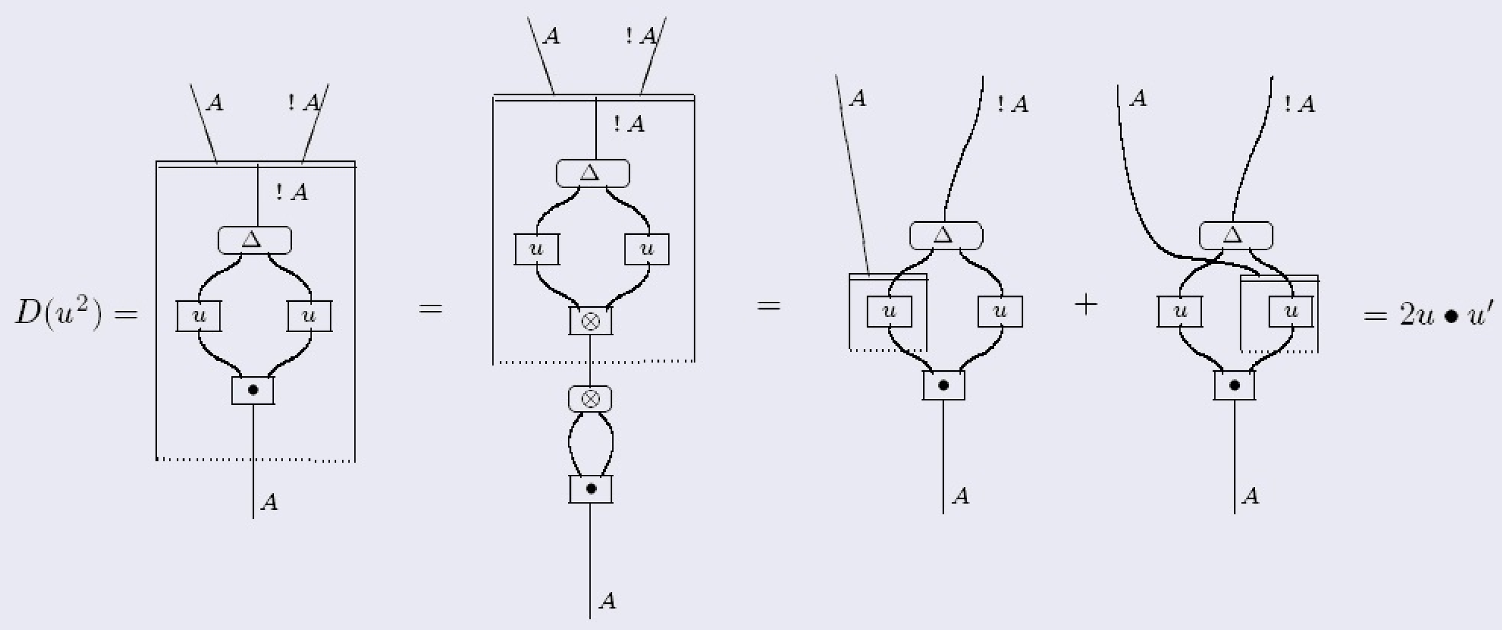
In fact, Cartesian differential categories were the second notion we developed to describe the categorical structure of differentiation. A couple of years earlier, we defined (tensor) differential categories in terms of two categories, one consisting of smooth maps (these are the maps to be differentiated), and the other consisting of linear maps (these are the derivatives), with a functor (a map of categories) which links them (in addition to the differential operator).

The link between these two notions is this: for any (tensor) differential storage category, its "coKleisli category" is a Cartesian differential category. Our current research is focussed on clarifying the relationship between the two notions of differential categories,specifically how to construct a tensor differential category from a Cartesian differential category. (Apologies for not explaining these technical notions; but those interested can read the original papers, found on my website. A talk describing partial results towards linking the two structures more closely may be found there also.)

Diagrammatic notation

One of the technical (and rather attractive) devices we have used often over the years is to represent maps in structured categories as graphs (we call them circuits). In the present case, as an illustration (and sorry, without a proper explanation of how the notation works*), here is the calculation of the derivative of a function u^2 in a differential category equipped with a commutative multiplication operator $\bullet\colon A \otimes A \to A$, so u^2 means $u \bullet u$.

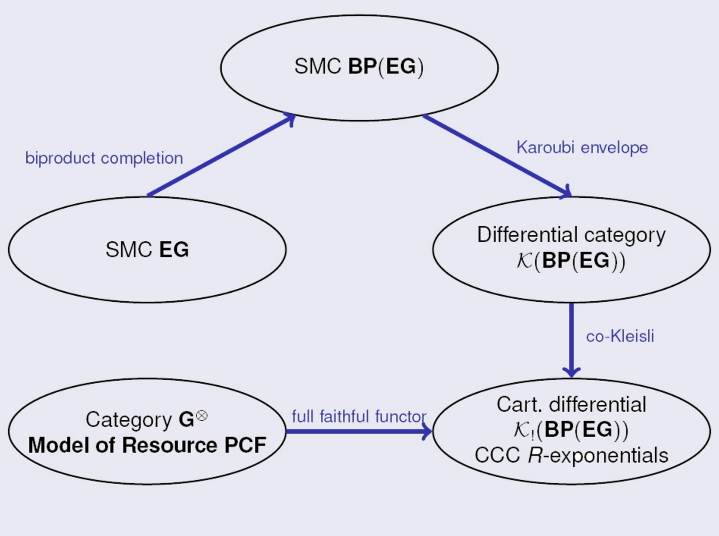
Then, using the appropriate rewrites for a differential combinator, we obtain



*Roughly, the boxes with the double line at top, dotted line at bottom, are applications of the derivative, the other boxes are maps or functions, labelled as shown, and the "wires" are objects, also labelled as shown.

Impact on the work of others

These two categorical notions of differential structure have found use in the work of other mathematicians, logicians, and computer scientists in the development of other structures involving differentiation. One interesting use in theoretical computer science is differential λ calculus and differential combinatory algebras. These (and related) structures are intended to provide theoretical tools for analysing things like the use of resources and the complexity of computations and programs, in other words, to measure the effectiveness of programming paradigms. Some examples: Laird, Manzonetto, and McCusker have modelled resource sensitive calculi (such as Resource PCF) in differential categories, which they use to unify Ehrhard & Regnier's differential λ calculus and Tranquilli's resource calculus. Paolo Tranquilli has used differential categories as the semantics of differential linear logic in his study of rewriting for this logic. Giulio Manzonetto has been looking at differential model theory for resource λ-calculus, using both of our notions of differential categories.



In this illustration, taken from a talk by Giulio Manzonetto, the construction of an embedding of a model of resource PCF into a Cartesian differential category is shown, starting from a category of games, G, constructing a related subcategory $G^{\otimes}$ (which is a Cartesian closed differential category), and then a category of "exhausting games", EG, which is a symmetric closed monoidal category.

In all these cases, having differential categories simplifies the semantics and provides a clean environment for the research being done.

Further information?

You can find several papers I've cowritten, and slides for talks I've given, on these matters at my webpage: `http://www.math.mcgill.ca/rags/`. All the technical details casually ignored in this poster may be found there.

In addition, there is a special session on this topic at meeting organized by the Association of Symbolic Logic at the University of Waterloo, 8-11 May 2013. `http://www.math.uwaterloo.ca/~asl2013/`