Two categorical approaches to differentiation (R.A.G. Seely)

In the past decade, we (coauthors Rick Blute, Robin Cockett and I) have formulated two different abstract categorical approaches to differential calculus, based on the structure of linear logic (an idea of Ehrhard and Regnier). The basic idea has two types of maps ("analytic" or "smooth", and "linear"), a comonad S (a "coalgebra modality"), somewhat like the **!** of linear logic, and a differentiation operator. In our first approach (**monoidal differential categories**), the coKleisli category (the category of cofree coalgebras) of S consists of smooth maps, and differentiation operates on coKleisli maps to smoothly produce linear maps. Our second approach (**Cartesian differential categories**) reversed this orientation, directly characterizing the smooth maps and situating the linear maps as a subcategory. If S is a "storage modality", meaning essentially that the "exponential isomorphisms" from linear logic ($S(X \times Y) \simeq S(X) \otimes S(Y)$ and $S(1) \simeq S(\top)$) hold, we get a tight connection between these approaches in the Cartesian (monoidal) closed cases: the linear maps of a Cartesian closed differential storage category form a monoidal closed differential storage category, and the coKleisli category of a monoidal closed differential storage category is a Cartesian closed differential storage category. Two technical aides in proving these results are the development of a graphical calculus as well as a term calculus for the maps of these categories. With the term calculus, one can construct arguments using a language similar to that of ordinary undergraduate calculus.

Two categorical approaches to differentiation

R.A.G. Seely

McGill University & John Abbott College

Joint with Rick Blute & Robin Cockett

http://www.math.mcgill.ca/rags/

Preludium

2006 Monoidal differential categories—an additive monoidal category of “linear” maps, a (suitable) comonad whose coKleisli maps are “smooth”, and a differential combinator. (This gave a “categorical reconstruction” of Ehrhard & Regnier’s work)

2009 Cartesian differential categories—a left additive Cartesian category with a differential operator, and subcategories of “linear” maps. CDCs are the coalgebras of a “higher order chain rule fibration” comonad Faà [2011].

Preludium

- The coKliesli category of a monoidal differential (storage) category is a Cartesian differential category [2009].
- The linear maps of a Cartesian differential (storage) category form a monoidal differential category [2015].

Wished: Any Cartesian differential category is, or may be full&faithfully embedded into, the coKliesli category of a (suitable) monoidal differential category. (Partial answer: conditions sufficient to ensure "is" in [2015].)

Wished: Any monoidal differential category may be represented as the linear maps of a (suitable) Cartesian differential category. (Partial answer: this is true in the presence of "storage" [2015].)

Coalgebra modality

We need some of the structure of the ! comonad from linear logic, specifically, a coalgebra modality: namely a comonad $\langle S\colon \mathbb{X} \longrightarrow \mathbb{X},\ \delta\colon S \longrightarrow S^2,\ \epsilon\colon S \longrightarrow 1\rangle$ on an additive (commutative monoid enriched) symmetric monoidal category $\mathbb{X}$, so each object $S(X)$ is equipped with a natural coalgebra structure

$$\top \xleftarrow{e} S(X) \xrightarrow{\Delta} S(X) \otimes S(X)$$

satisfying the "obvious" coherence conditions:

Coalgebra modality: the conditions

$\langle S(X), \Delta, e \rangle$ is a comonoid (or coalgebra),

$$\begin{tikzcd} & S(X) \arrow[dl, equal] \arrow[d, "\Delta"] \arrow[dr, equal] & \\ S(X) & S(X) \otimes S(X) \arrow[l, "1\otimes e"] \arrow[r, "e\otimes 1"'] & S(X) \end{tikzcd}
\qquad
\begin{tikzcd} S(X) \arrow[r, "\Delta"] \arrow[d, "\Delta"'] & S(X) \otimes S(X) \arrow[d, "\Delta\otimes 1"] \\ S(X) \otimes S(X) \arrow[r, "1\otimes\Delta"'] & S(X) \otimes S(X) \otimes S(X) \end{tikzcd}$$

and δ is a comonoid (or coalgebra) homomorphism.

$$\begin{tikzcd} S(X) \arrow[rr, "\delta"] \arrow[dr, "e"'] & & S^2(X) \arrow[dl, "e"] \\ & \top & \end{tikzcd}
\qquad
\begin{tikzcd} S(X) \arrow[r, "\delta"] \arrow[d, "\Delta"'] & S^2(X) \arrow[d, "\Delta"] \\ S(X) \otimes S(X) \arrow[r, "\delta\otimes\delta"'] & S^2(X) \otimes S^2(X) \end{tikzcd}$$

Note that we are not assuming that S *etc.* are monoidal.

Monoidal differential categories—the intuition

The basic intuition comes from linear logic, where a "linear" map $!A \longrightarrow B$ corresponds to a "smooth" map $A \longrightarrow B$ (and so categorically, if ! is a comonad, "smooth" maps are just maps in the coKleisli category).

For a differential setting, consider a simple example from multivariable calculus: $f(x, y, z) = \langle x^2 + xyz, z^3 - xy \rangle$. This is a smooth function from R^3 to R^2. Its Jacobian (matrix) is $\begin{pmatrix} 2x + yz & xz & xy \\ -y & -x & 3z^2 \end{pmatrix}$, *i.e.* a *linear* map from R^3 to R^2.

This gives a smooth assignment of a linear map from a point; *i.e.* given a map $f\colon A \subseteq R^n \longrightarrow R^m$, one gets a smooth map $D[f]\colon A \longrightarrow L(R^n, R^m)$: for a point $x \in A$, $D[f](x)$ is given by the Jacobian of f at x.

The formalism

So the type of the differential should be $D[f]\colon !A \longrightarrow A \multimap B$. To avoid assuming our categories are closed, we transpose this to obtain a differential combinator of the form:

$$\frac{f\colon !A \longrightarrow B}{D[f]\colon A \otimes !A \longrightarrow B}$$

Finally, it turns out that we can replace the differential combinator D with a deriving transformation $d\colon A \otimes !A \longrightarrow !A$. This will be the formalism we'll present today. And we'll use S instead of the ! of linear logic ...

So we arrive at the following definition.

Monoidal differential categories

A monoidal differential category is a monoidal category with a coalgebra modality equipped with a natural deriving transformation

$$d_\otimes \colon A \otimes S(A) \longrightarrow S(A)$$

[d.1] Constant maps: $d_\otimes e = 0$

[d.2] Linearity: $d_\otimes \epsilon_A = (1 \otimes e)u_\otimes$

[d.3] Product rule:
$d_\otimes \Delta = (1 \otimes \Delta_\otimes) a_\otimes^{-1} (d_\otimes \otimes 1) + (1 \otimes \Delta_\otimes) a_\otimes^{-1} (c_\otimes \otimes 1) a_\otimes (1 \otimes d_\otimes)$

[d.4] Chain rule: $d_\otimes \delta = (1 \otimes \Delta_\otimes) a_\otimes^{-1} (d_\otimes \otimes \delta) d_\otimes$

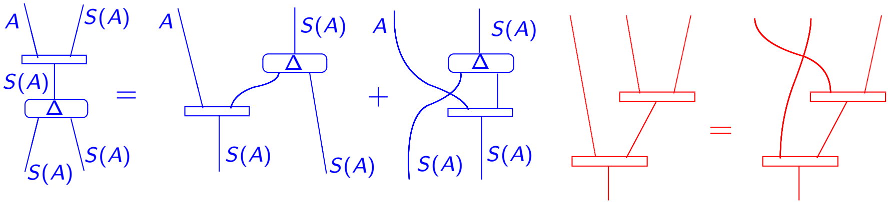
[d.5] Interchange rule: $(1 \otimes d_\otimes) d_\otimes = a_\otimes (c_\otimes \otimes 1) a_\otimes^{-1} (1 \otimes d_\otimes) d_\otimes$

Interchange rule?

The interchange rule was not one of the original axioms in [2006], and is not necessary for the basic theory of monoidal differential categories (for example, first year differential calculus can be modelled as a codifferential category without **[d.5]**.) However, we introduced it in [2009] in order to ensure that the coKleisli category was a Cartesian differential category. It also turns out that interchange law is necessary to characterize monoidal differential storage categories as the linear maps of a Cartesian differential category [2015].

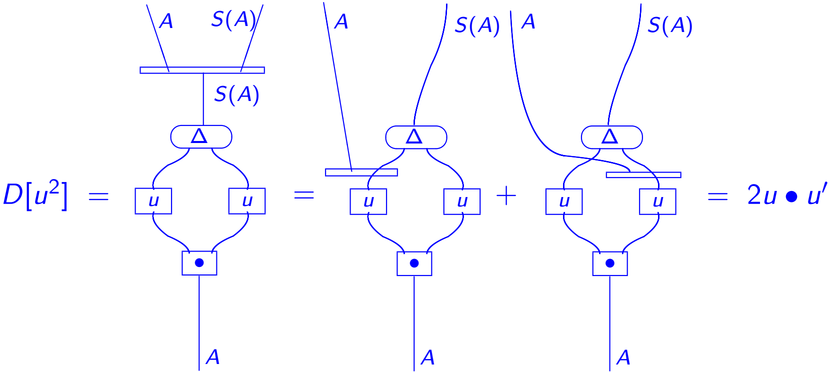
Circuits

As an example of the graphical calculus we use with monoidal differential categories, here are the product rule and the interchange rule, graphically. The thin horizontal "box" represents the deriving transformation $d_\otimes$, the "wires" are objects & other boxes are morphisms (some labelled to help the reader!).



An example calculation

To illustrate the circuit calculus[1], here is the calculation of the derivative of u^2 for a (smooth) function u: $!A \longrightarrow A$; we assume the existence of a (linear, commutative) "multiplication" operator $\bullet$: $A \otimes A \longrightarrow A$ so that u^2 means $u \bullet u$.



[1] Pun intended

Storage

A storage modality S on a symmetric monoidal category $\mathbb{X}$ is a symmetric monoidal comonad $\langle S, \delta, \epsilon \rangle$ so that each cofree object is naturally a commutative comonoid $\langle S(A), \Delta, e \rangle$, and so that the comonoid structure is a coalgebra morphism for the comonad.

Equivalently, a symmetric monoidal category has a storage modality iff the induced symmetric tensor on the category of coalgebras for the comonad is a product.

This implies the existance of the "exponential" or "storage" isomorphisms:

$$S(A) \otimes S(B) \longrightarrow S(A \times B) \quad \text{and} \quad \top \longrightarrow S(1)$$

Cartesian differential categories

Now we focus on the category of smooth maps, regarding linear maps as constituting a full subcategory. The idea is that we are describing what was the coKleisli category in the previous monoidal differential setting.

We suppose the ambient category $\mathbb{X}$ is a Cartesian left additive category, meaning that each hom-set is a commutative monoid, $f(g+h) = fg + fh, f0 = 0$ (left-additivity), and $\mathbb{X}$ has finite products (including 1) whose structure maps are (left and right) additive (*i.e.* commutative monoid homomorphisms). Note that in such a setting, products are in fact biproducts.

Given such an $\mathbb{X}$, the differential operator will be typed thus:

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

Cartesian differential categories

In this setting, we require $D_\times$ to satisfy the following axioms:

[CD.1] $D_\times[f+g] = D_\times[f] + D_\times[g]$ and $D_\times[0] = 0$ (D is linear)

[CD.2] $\langle h+k, v\rangle D_\times[f] = \langle h, v\rangle D_\times[f] + \langle k, v\rangle D_\times[f]$ and $\langle 0, v\rangle D_\times[f] = 0$ (D is additive in 1st argument)

[CD.3] $D_\times[1] = \pi_0$, $D_\times[\pi_0] = \pi_0\pi_0$ and $D_\times[\pi_1] = \pi_0\pi_1$

[CD.4] $D_\times[\langle f, g\rangle] = \langle D_\times[f], D_\times[g]\rangle$ (D is “coherent” wrt $\times$)

[CD.5] $D_\times[fg] = \langle D_\times[f], \pi_1 f\rangle D_\times[g]$ (Chain rule)

[CD.6] $\langle\langle g, 0\rangle, \langle h, k\rangle\rangle D_\times[D_\times[f]] = \langle g, k\rangle D_\times[f]$ ($D[f]$ is “linear” in 1st argument)

[CD.7] $\langle\langle 0, h\rangle, \langle g, k\rangle\rangle D_\times[D_\times[f]] = \langle\langle 0, g\rangle, \langle h, k\rangle\rangle D_\times[D_\times[f]]$ (Equality of mixed partials)

Properties of CDCs

The main "first fact" about Cartesian differential categories is

Theorem: The coKleisli category of a monoidal differential category is a Cartesian differential category. [2009]

And secondly, they arise naturally from the structure of differentiation. Specifically, one can construct a comonad Faà on the category of Cartesian left additive categories and sum-product-preserving (on the nose) functors, essentially given by the higher order chain rules, for which Cartesian differential categories are precisely the coalgebras. [2011]

Example of a CDC

The canonical example of a Cartesian differential category is, of course, from ordinary college calculus:

Finite dimensional vector spaces and smooth maps form a Cartesian differential category, with D given by the Jacobian.

$$D_\times[\langle f_1, \ldots, f_n \rangle](\vec{x}, \vec{y}) = [(\partial_{x_i} f_j)(\vec{y})]_{i,j=1,1}^{m,n} \ \vec{x}$$

To parse this, consider $f(x_1, x_2, x_3) = x_1 x_2 x_3 \colon R^3 \longrightarrow R$.

Its Jacobian is the linear function given by the matrix $[x_2 x_3, x_1 x_3, x_1 x_2] \colon R^3 \longrightarrow R$.

Given $\langle y_1, y_2, y_3 \rangle \in R^3$, $D_\times[f](\vec{x}, \vec{y}) = y_2 y_3 x_1 + y_1 y_3 x_2 + y_1 y_2 x_3$.

Linear maps

We think of the maps of a CDC as being "smooth"; we must also ask "what are the linear maps?".

Definition: In a Cartesian differential category, a map f is linear if $D_\times[f] = \pi_0 f$.

Proposition: In a CDC:

- Every linear map is additive
- 0 is linear; if f, g are linear, so is $f + g$
- Linear maps compose; identity maps are linear (so linear maps form a subcategory)
- Projections are linear; pairs of linear maps are linear
- $\langle 1, 0 \rangle D_\times[f]$ is linear
- a, b linear and $af' = fb$ (some f, f'), then $(a \times a) D_\times[f'] = D_\times[f]\, b$
- If g is linear and a retraction and gh is linear, then h is linear
- If f is linear and an isomorphism, then f^{-1} is linear.

Linear maps

Hence:

Corollary: The linear maps of a CDC $\mathbb{X}$ form an additive subcategory $\mathbb{X}_{\mathrm{lin}}$ which has biproducts;

$\mathbb{X}_{\mathrm{lin}} \hookrightarrow \mathbb{X}$ reflects isos and creates products.

Term calculus

One tool we have found useful is a term calculus for Cartesian differential categories [2009]. Here's a hint at its flavour. Consider $f\colon R^3 \longrightarrow R^2\colon \langle x, y, z\rangle \mapsto \langle x^2 + xyz, z^3 - xy\rangle$
We may think of the Jacobian evaluated at (r, s, t) thus:

$$\frac{\partial\langle x^2 + xyz, z^3 - xy\rangle}{\partial(x, y, z)}(r, s, t) = \left(\begin{array}{ccc} 2r + st & rt & rs \\ -s & -r & 3t^2 \end{array}\right)$$

We apply this Jacobian to a vector to obtain a point in R^2:

$$\left(\begin{array}{ccc} 2r + st & rt & rs \\ -s & -r & 3t^2 \end{array}\right)\cdot(u, v, w) = ((2r+st)u+rtv+rsw, -su-rv+3t^2w)$$

This is what we write in the term logic as

$$\frac{\partial(x^2 + xyz, z^3 - xy)}{\partial(x, y, z)}(r, s, t) \cdot (u, v, w)$$

The Chain Rule

Then, in this term calculus, the chain rule (for example) looks like this:

$$\frac{\partial t[t'/p']}{\partial p}(s) \cdot u = \frac{\partial t}{\partial p'}(t'[s/p]) \cdot \left(\frac{\partial t'}{\partial p}(s) \cdot u\right)$$

(where no variable of p may occur in t)

Compare with the combinator version:

$$D_\times[fg] = \langle D_\times[f], \pi_1 f\rangle D_\times[g]$$

In [2009] we show this term calculus is sound and complete for CDCs.

Generalized CDCs

More recently, Geoff Cruttwell [MSCS 2015] has generalized the definition of CDCs, equipping each object X with a commutative monoid $\langle L_0(X), +_X, 0_X \rangle$ so that every $f \colon X \to Y$ induces a map $D[f] \colon L_0(X) \times X \to L_0(Y)$ (instead of $D[f] \colon X \times X \to Y$).

One should think of $L_0(X)$ as "vectors" (and X as "points"). In R^n these are usually identified, an identification carried over in the abstraction of CDCs.

Cruttwell's approach makes the distinction, which enables him to get somewhat "better" results for generalized CDCs (especially when one passes to restriction structure, which we shall not mention further here). In particular, the generalized CDCs are comonadic over Cartesian categories (ordinary CDCs are only monadic over Cartesian left additive categories).

Is this what Ehrhard & Regnier had in mind?

The original work (by Ehrhard and Regnier) on abstract differential structures that inspired our research was partly aimed at the differential λ-calculus, and related structures. Our original work (on monoidal differential categories [2006]) was rather more "ascetic", using less structure to capture as much of simple differential calculus as possible—though we did consider the situation of adding "just about everything short of closedness" to compare our work with theirs.

But can we now say more? Can we add closed structure, for example, and if we do, is the result different in any essentials from the (recent) work by Ehrhard and others on, say, the differential λ calculus, or the resource λ calculus?

Well ...

Closed structure

It is in fact straightforward to extend these structures to be (monoidally or Cartesian, as appropriate) closed. With the addition of storage, we arrive then at the categorical semantics for what is essentially Ehrhard and Regnier's original notion of differential structure. In [2015] we gave a detailed analysis of how storage works in various contexts, culminating in the following facts:

Theorem: The linear maps of a Cartesian closed differential storage category (in which linear idempotents split) form a monoidal closed differential storage category.

Theorem: (A strengthening of a previous result) The coKleisli category of a monoidal closed differential storage category is a Cartesian closed differential storage category.