The Syntax and Semantics of Differentiation

Richard Blute
University of Ottawa

June 14, 2021

Overview

- *Differential linear logic* (DLL), due to Ehrhard & Regnier, is an extension of linear logic (J.Y. Girard) via the addition of an inference rule modelling differentiation.
- It was inspired by models of linear logic discovered by Ehrhard, where morphisms have a natural smooth structure.
- The corresponding categorical structures are *differential categories*, due to RB, Cockett and Seely.
- Given this new syntactic/semantic way of thinking about differentiation, we should find models and apply ideas from (categorical) logic to fields where differentiation is fundamental, such as manifolds and tangent structures.

Monoidal categories

Definition

A *monoidal category* is a category $\mathcal{C}$ with a binary (operation) functor $\otimes\colon \mathcal{C} \times \mathcal{C} \to \mathcal{C}$ which is associative and has a unit object I. We'll also assume symmetry. Thus we have the following isomorphisms:

$$(A \otimes B) \otimes C \cong A \otimes (B \otimes C) \qquad A \otimes I \cong A \cong I \otimes A \qquad A \otimes B \cong B \otimes A$$

Two classes of examples we'll be interested in:

- Linear monoidal categories like $\mathbf{Vec}_k$, the category of k-vector spaces and linear maps. The monoidal structure is the usual tensor product and the unit is the base field.
- Cartesian monoidal categories like the category **Set** of sets and functions or **Top**, the category of topological spaces and continuous functions.. The monoidal structure is the cartesian product of sets, $\times$ and the tensor unit is the one point set, $\mathbf{1}$.

Monoidal categories II

- A crucial difference is that **Set** has canonical maps:
  $A \to A \times A \qquad A \times B \to A \qquad A \times B \to B \qquad A \to 1$
  while **Vec** and **Hilb** don't.
- A monoidal category is *closed* if the functor $A \otimes (-)\colon \mathcal{C} \to \mathcal{C}$ has a right adjoint $A \multimap (-)$:

$$Hom(A \otimes B, C) \cong Hom(B, A \multimap C)$$

.

- The categories **Set**, **Vec** and $\mathbf{Hilb}_{fd}$ are closed and in each case $A \multimap B$ is the appropriate function space. For **Set**, it's the set of functions and for **Vec** and $\mathbf{Hilb}_{fd}$, it's the set of linear functions.

Monoidal closed categories

- The category of finite-dimensional Hilbert spaces is closed, but the larger category of Hilbert spaces is not.
- The category of topological spaces is not closed, but the category of compactly generated Hausdorff spaces is.
- The category of smooth manifolds and smooth maps is not closed.

Closed categories are quite desirable. On the one hand, we can work with function spaces, and on the other hand, closed structure will allow us to model implication in various logics.

Categorical Proof Theory I

- **Categorical proof theory** (Lambek) begins with the idea of forming a category whose objects are formulas in a given logic and whose arrows are (equivalence classes of) proofs.
- Then we study the resulting category to determine its structure. Typically the category will be free in a certain sense.
- As a simple example, in *intuitionistic logic*, the logic of $\wedge, \vee$ and $\Rightarrow$, conjunction takes on the form of a categorical product and disjunction takes on the form of a coproduct.

  Thus for example, the projection $A \times B \to A$ is interpreted to mean that $A \wedge B$ logically entails A. The diagonal map $\Delta \colon A \to A \times A$ means that A logically entails $A \wedge A$, etc.

Categorical Proof Theory II

- Closed structure provides a model of logical implication.
- The counit of the adjunction

$$Hom(A \wedge B, C) \cong Hom(B, A \multimap C)$$

  is a map $\eta\colon A \wedge (A \multimap C) \to C$. This is the familiar *modus ponens* rule, that A and $A \multimap C$ logically entail C.
- In general, logical connectives become functors and inference rules will become natural transformations.
- Categories with the same structure can then be considered as models of that logical system.

- To go from intuitionistic logic to classical logic, you might try the following. For any two objects, A and B, in a cartesian closed category, there is a canonical map $A \rightarrow (A \Rightarrow B) \Rightarrow B$.
- Suppose we have an initial object, which we denote as F. Then define $\neg A = A \Rightarrow F$. The above map then becomes $A \rightarrow \neg\neg A$. To model classical logic, I could assume this map is an isomorphism.

Theorem (Joyal)

Any such category is a boolean algebra.

- The proof doesn't apply in linear logic, so we can have involutive negation and nontrivial categorical semantics. These are the $*$-*autonomous categories* of Barr.

- We use *sequent calculus* as our basic proof system. A *sequent* is something of the following form with the $\vdash$ representing logical entailment:

$$\Gamma \vdash A$$

  Here Γ is a finite list of formulas (the premises) in our logic and A is a single formula (the conclusion).

- Sequents are constructed and manipulated using *inference rules*. Here are two examples:

$$\frac{\Gamma \vdash A \qquad \Delta \vdash B}{\Gamma, \Delta \vdash A \wedge B}\ \wedge R$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \Rightarrow B} \Rightarrow R$$

Categorical Proof Theory V-Structural Rules

These rules are basically bookkeeping rules and allow us to manage premises:

- *Exchange* says we can rearrange the order of our premises as we like (σ is a permutation):

$$\frac{\Gamma \vdash A}{\sigma(\Gamma) \vdash A}\ \text{Ex}$$

- *Contraction* says that it is pointless to have duplicate premises:

$$\frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\ \text{Con}$$

- *Weakening* says that you can add additional premises.

$$\frac{\Gamma \vdash B}{\Gamma, A \vdash B}\ \text{Weak}$$

So the proof of a sequent $\Gamma \vdash A$ is a morphism in an appropriate category of the form

$$\bigwedge \Gamma \longrightarrow A$$

In traditional logics, the $\wedge$ is the cartesian product. The interpretation is built inductively. As one example, consider:

$$\frac{\Gamma, A, A \vdash B}{\Gamma, A \vdash B}\ \text{Con}$$

We suppose we have an arrow $\bigwedge \Gamma \wedge A \wedge A \to B$ interpreting the proof above the line. We build a map $\bigwedge \Gamma \wedge A \to B$ by precomposing with $\Delta\colon A \to A \wedge A$.

Similarly for weakening:

$$\frac{\Gamma \vdash B}{\Gamma, A \vdash B} \text{ Weak}$$

I assume I have an arrow:

$$\bigwedge \Gamma \longrightarrow B$$

and I precompose with the projection

$$\bigwedge \Gamma \wedge A \longrightarrow \bigwedge \Gamma \longrightarrow B$$

Categorical Proof Theory VIII-The Cut Rule

We also have a way of stringing proofs together:

$$\frac{\Delta \vdash B \quad \Gamma, B \vdash A}{\Gamma, \Delta \vdash A}\ \mathrm{Cut}$$

This is completely analogous to the composition of processes. We also have the crucial cut-elimination theorem:

Theorem (Gentzen)

Any sequent that can be proved can be proved without the CUT rule.

The proof is typically algorithmic and the algorithm is of great interest. When talking about categorical proof theory, we want that each proof is equivalent to a cut-free proof.

Linear Logic: A Resource-Sensitive Logic (Girard)

- The starting point for linear logic is the reinterpretation of sequents.
- The sequent $\Gamma \vdash A$ is traditionally interpreted as *From premises* Γ*, we can conclude* A.
- Think instead that formulas represent resources and $\Gamma \vdash A$ means that inputting the resources Γ into the system, it will output A.
- From this point of view, the rules contraction and weakening are wrong.
- Linear logic begins with the removal of these rules.

(Categorical) Linear Logic I

- The categorical product $\times$ becomes a tensor product $\otimes$, i.e we have no projection or diagonals. So we are talking about *monoidal (not cartesian) categories*.

- Instead of cartesian closed categories, we have *symmetric monoidal closed* categories as the basic structure. An excellent non-cartesian example is the category of vector spaces and linear maps. In this category, there are no canonical maps:

$$V \otimes W \longrightarrow V \qquad V \otimes W \longrightarrow W \qquad V \longrightarrow V \otimes V$$

- We can also remove the Exchange rule to get noncommutative logic. Or you can modify the exchange rule to only allow certain exchanges, as in *cyclic* linear logic (Yetter).

- Contraction and weakening are still available for formulas of the form $!X$, called *bang* X. So we have the following sequent rules:

$$\frac{\Gamma \vdash Y}{\Gamma, !X \vdash Y}$$

$$\frac{\Gamma, !X, !X \vdash Y}{\Gamma, !X \vdash Y}$$

- The operator $!\colon \mathcal{C} \to \mathcal{C}$ is a *comonad*, i.e. we have natural transformations:

$$! \xrightarrow{\delta} !! \qquad\qquad ! \xrightarrow{\epsilon} \mathrm{id}$$

satisfying associativity, and unit constraints.

(Categorical) Linear Logic III: Coalgebra Structure

- We also typically assume the category has products and that we have the isomorphisms:

$$!(A \times B) \cong !A \otimes !B \qquad\qquad !1 \cong I$$

  where I is the tensor unit and 1 is the terminal object. (But note this isn't necessary to model the fragment of the logic we have seen so far. But it does hold in most models.)

- Given this isomorphism, it is straightforward to endow the objects $!A$ with a *coalgebra* structure:

$$!A \xrightarrow{\Delta} !A \otimes !A \qquad\qquad !A \xrightarrow{\Delta} I$$

For example, to construct Δ, proceed as follows:

$$\cfrac{\cfrac{!(A \to A \times A)}{!A \to !(A \times A) \cong !A \otimes !A}}{!A \to !A \otimes !A}$$

- These maps satisfy the obvious (co)associative, (co)commutative and (co)unit axioms, making $!A$ a cocommutative counital coalgebra.
- They are used to model contraction and weakening with respect to formulas of the form $!A$,

CoKleisli Categories

Given a comonad $! \colon \mathcal{C} \to \mathcal{C}$, one can form the *coKleisli category* $CoK(!)$.

- Objects are the same as for $\mathcal{C}$.
- Arrows $X \to Y$ in $CoK(!)$ are arrows of the form $!X \to Y$ in $\mathcal{C}$. (How to compose?)
- Linear logic began with Girard's realization that his category of coherence spaces and stable maps (a model of simply-typed λ-calculus) was actually the coKleisli category of a more basic *linear* structure.
- So there was a decomposition:

$$Hom_{\mathrm{Stab}}(A, B) \cong Hom_{\mathrm{Lin}}(!A, B)$$

Differential Linear Logic I

- Again, this idea arose from semantic considerations. Ehrhard constructed two models of linear logic where there is just such a decomposition. These were the categories of *Köthe spaces* and *finiteness spaces*. Morphisms had a representation as power series, which could be differentiated.
- *Differential linear logic* (Ehrhard, Regnier) begins with the idea that there is a similar decomposition of a category of *smooth* maps into the coKleisli category of a category of linear maps.
- Categorically, we would like a category where the base category had linear maps and the coKleisli category had smooth maps.

- The important point is that differentiation is represented as an inference rule.
- To see what the inference rule would be, consider the following situation. I have two Euclidean spaces, X and Y, and a smooth map between them. In our model, it would be a map $f\colon !X \to Y$.
- At a point of X, its Jacobian matrix would be a linear map from X to Y. So the process of taking the Jacobian is a smooth map from X to linear maps from X to Y. This suggests an inference rule of the following form:

$$\frac{!X \vdash Y}{!X \vdash X \multimap Y}$$

- Or, equivalently:

$$\frac{!X \vdash Y}{X \otimes !X \vdash Y}$$

- The logic continues to satisfy cut-elimination.
- Proof nets are a graph-theoretic syntax for specifying proofs for various fragments of linear logic (Girard, Danos-Regnier). They satisfy remarkable properties. Cut-elimination becomes a graph-rewriting system satisfying good normalization properties. There is a good theory of proof-nets for DLL (Ehrhard & Regnier).
- The corresponding coKleisli category is cartesian closed, hence a model of λ-calculus, but with an additional operation of differentiation of terms. This leads to the *Differential λ-Calculus* (Ehrhard & Regnier).

Differential Categories I

- Categorically, I need an operator:

$$\frac{f\colon !X \to Y}{D(f)\colon X \otimes !X \to Y}$$

- It suffices to differentiate the identity map on $!X$. So we require a map

$$D(id_{!X}) = d\colon X \otimes !X \xrightarrow{d} !X$$

.

- Then an arbitrary smooth map $f\colon !X \to Y$ is differentiated by precomposition with d. So

$$D(f) = X \otimes !X \xrightarrow{d} !X \xrightarrow{f} Y$$

- To state axioms, we must have *additive structure on the Hom-sets*. We need to be able to add maps.

So a *differential category* (RB, Cockett, Seely) is a model of linear logic with a map of the above form satisfying basic differential identities, expressed coalgebraically. The necessary rules are:

- The derivative of a constant is 0.
- The derivative of a linear function is constant.
- Leibniz rule (Product rule).
- Chain rule.

Here's an example (product rule). The composite

$$X \otimes !X \xrightarrow{d} !X \xrightarrow{\Delta} !X \otimes !X$$

must equal:

$$X \otimes !X \xrightarrow{id \otimes \Delta} X \otimes !X \otimes !X \xrightarrow{d \otimes id} !X \otimes !X +$$
$$X \otimes !X \xrightarrow{id \otimes \Delta} X \otimes !X \otimes !X \cong !X \otimes X \otimes !X \xrightarrow{id \otimes d} !X \otimes !X$$

Differential Categories III

- There are many, many subtleties to this definition.
- There are some equations beyond the basic ones listed above which can be considered. We could assume the modality is *monoidal* or not, etc. In some cases, there is redundancy in the rules, etc.
- There are differing presentations which under some restrictions become equivalent.
- For example, one way to ensure that we have additive structure on the hom-sets is to assume our category has biproducts. In this case, each object $!X$ has the structure of a bialgebra. This leads to some new inference rules such as *cocontraction*:

Cocontraction:

$$\frac{\Gamma, !X \vdash Y}{\Gamma, !X, !X \vdash Y}$$

There is also a *coweakening*. In this case, to define a differential category, one can instead assume a map called *coderelection*, which is of the form

$$coder\colon X \to !X$$

satisfying similar equations to the above. Under certain hypotheses, this presentation becomes equivalent.

See J.S. Lemay's talk and *Differential categories revisited*, by RB, Cockett, Lemay and Seely, which cleans up these various issues.

Differential Categories V: Models

- There is a simple relational model, demonstrating that our axioms are consistent. In the category Rel whose objects are sets and whose arrows are binary relations, composition is the composition of relations, and the !-comonad is the finitary multiset functor and the tensor is the cartesian product of sets. Then the differentiation map $d_X : X \otimes !X \to !X$ is given by

$$(x, V) \quad \leftrightarrow \quad x \uplus V$$

  where $x \in X$, V is a finitary multiset of X and $\uplus$ is the multiset union.
- But we want models where the differential inference rule is actual differentiation.

A more interesting example is obtained by considering the category of (discrete) vector spaces. The *free symmetric algebra* on a vector space V is defined by

$$S(V) = k \oplus V \oplus V \otimes_s V \oplus \ldots$$

where $V \otimes_s V$ is the symmetrized tensor product. An explicit representation of $V \otimes_s V$ is the coequalizer of:

$$V \otimes V \xrightarrow{c} V \otimes V \quad \text{and} \quad V \otimes V \xrightarrow{id} V \otimes V$$

An explicit representation of $S(V)$ is given by choosing a basis for V, $\{x_i\}_{i \in I}$. Then $S(V)$ is the space of polynomials in $\{x_i\}_{i \in I}$.

It is standard that S is a monad on Vec, the category of vector spaces, and hence a comonad on Vec^{op}.

We need a differential on Vec^{op}, i.e. a map:

$$d \colon S(V) \longrightarrow V \otimes S(V)$$

It is easiest to describe using a basis:

$$f(x_1, x_2, x_3, \ldots, x_n) \mapsto \sum_{j=1}^{n} x_j \otimes \frac{\partial f}{\partial_{x_j}}$$

This is a finite sum, even if V is infinite-dimensional. It is straightforward to see all the equations are satisfied.

Differential Categories VIII

Daniel Murfet and James Clift show that Sweedler's construction of the cofree cocommutative coalgebra, which induces a comonad on **Vec** makes **Vec** a differential category.

Daniel Murfet's paper *Logic and linear algebra: an introduction* is highly recommended as an introduction to how linear logic proofs can be interpreted as linear maps between vector spaces. This paper is again using the cofree cocommutative coalgebra of Sweedler.

Looking for topological examples

- Ehrhard's two primary models are both differential categories. But we'd like models with a closer connection to analysis.
- While we don't require our categories to be closed, it is a desirable property. None of the standard examples of categories of spaces and smooth maps are closed, and there are no evident comonads.
- *Convenient vector spaces* provide an example which has all of the properties we are looking for, as shown by RB, Ehrhard and Tasson.
- For details, see the talk of Marie Kerjean. She and Yoann Dabrowski substantially expanded our initial work. In particular, they have a $*$-*autonomous example*.

Convenient vector spaces (Frölicher,Kriegl)

See their book *Linear Spaces and Differentiation Theory*.

- Convenient vector spaces are a special class of locally convex spaces.
- The category Con of convenient vector spaces and continuous linear maps forms a symmetric monoidal closed category. The tensor is a completion of the algebraic tensor. There is a convenient structure on the space of linear, continuous maps. There is a nice notion of smoothness in this category:

Definition

A function $f\colon E \to F$ with E, F being convenient vector spaces is *smooth* if it takes smooth curves in E to smooth curves in F.

This is inspired by *Boman's Theorem*.

Convenient vector spaces II (Boman's theorem)

This definition of smooth map is inspired by a theorem of Boman:

Theorem

A function $f\colon \mathbb{R}^m \longrightarrow \mathbb{R}^n$ is smooth if and only if its composite with every smooth curve $u\colon \mathbb{R} \longrightarrow \mathbb{R}^m$ is smooth.

In the framework of convenient vector spaces, Boman's theorem becomes our definition.

Convenient vector spaces III: More key points

Theorem (Frölicher,Kriegl)

- *The category of convenient vector spaces and smooth maps is cartesian closed.*
- *There is a comonad* $!$ *for which the smooth category is the coKleisli category.*
- $!(E \oplus F) \cong\ !E \otimes\ !F$.
- *Each object* $!E$ *has canonical bialgebra structure.*

Theorem (RB, Ehrhard, Tasson)

Con *is a differential category.*

A convenient differential category

The above results show that Con really is an optimal differential category.

- The differential inference rule is really modelled by a directional derivative.
- The coKleisli category really is a category of smooth maps.
- Both the base category and the coKleisli category are closed, so we can consider function spaces.

There is a well-established theory of convenient manifolds:

Kriegl, Michor-*The convenient setting for global analysis*

Cartesian Differential Categories (RB, Cockett, Seely)

The idea here is to work directly in the category of smooth maps, rather than consider it as a coKleisli category. This will provide more direct structure, and allow for more examples.

Definition

- A category is a *cartesian differential category* if it has finite products, has a differential operator:

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

- The operator must satisfy analogous axioms, i.e. it must be linear in the first variable (so in particular, I have to be able to define this), and then satisfy the same differentiation axioms, rewritten for this setting.

- The coKleisli category of a differential category is a cartesian differential category.
- The category whose objects are Euclidean spaces and arrows are smooth maps is a cartesian differential category, which does not arise as the coKleisli category of a differential category.

For details and many examples, see the talk of J.S. Lemay.

Cartesian differential categories are also simple examples of *tangent categories* (Rosicky, Cockett & Cruttwell). For more information on tangent categories:

ENJOY THE REST OF THE CONFERENCE

but especially the talk of Geoff Cruttwell.