Coherent Taylor expansion as a bimonad

Thomas Ehrhard, Aymeric Walch

Making differentiation functorial

In cartesian differential categories, differentiation is axiomatized as an operator

$$\frac{f : X \to Y}{Df : X \times X \to Y} \qquad Df(x,u) = f'(x)\cdot u$$

The chain rule of differentiation corresponds to an equation on D.

$$D(g \circ f) = Dg \circ \langle f \circ \mathrm{p}_1, Df \rangle$$
$$(g \circ f)'(x)\cdot u = g'(f(x)) \cdot f'(x) \cdot u$$

One of the main idea those last years: making this a functoriality equation. Let

$$\frac{f : X \to Y}{\mathrm{T}f : X \times X \to Y \times Y} \qquad \begin{array}{c}\mathrm{T}f = \langle f \circ \mathrm{p}_1, Df \rangle \\ \mathrm{T}f(x,u) = (f(x), f'(x)\cdot u)\end{array}$$

Chain rule becomes equivalent to the equation $\mathrm{T}(g \circ f) = \mathrm{T}g \circ \mathrm{T}f$.

Remarkable observation: other properties of differentiation become naturality equations. For example, let $\theta : (X \times X) \times (X \times X) \to X \times X$

$$\theta(x,u,v,w) = (x, u+v)$$

Then: $\theta \circ \mathrm{T}^2 f = \mathrm{T}f \circ \theta \iff f'(x)\cdot(u+v) = f'(x)\cdot u + f'(x) \cdot v$

Coherent differentiation

Issue with addition

In (cartesian) differential categories: unrestricted sum operation $\sigma : X \times X \to X$.

- Math pov: objects are vector spaces/modules
- Computer science pov: morphisms are non deterministic computation

Main idea: take $\mathrm{T}X \neq X \times X$ and restrain sum.

- If $\mathrm{T}X = X \times X$: cartesian differential categories
- If $\mathrm{T}X$ is the tangent bundle of X : tangent categories
- If $\mathrm{T}X$ is a **left summability structure** : coherent differentiation

Coherent differentiation extends differentiation to deterministic models of LL: coherence spaces, probabilistic coherence spaces, etc.

$$\begin{tikzcd} \begin{array}{c}\text{Differential}\\ \text{Category}\\ \text{(Linear Logic)}\end{array} \arrow[r, dotted, "\text{Generalizes}"] \arrow[d, "\text{Kleisli}"'] & \begin{array}{c}\text{Coherent Differential}\\ \text{Category}\\ \text{(Linear Logic)}\end{array} \arrow[d, dotted, "\text{Kleisli}"] \\ \begin{array}{c}\text{Cartesian Differential}\\ \text{Category}\\ \text{(first order } \lambda\text{-calculus)}\end{array} \arrow[r, dotted, "\text{Generalizes}"'] & \begin{array}{c}\text{Cartesian Coherent}\\ \text{Differential Category}\\ \text{(first order } \lambda\text{-calculus)}\end{array} \end{tikzcd}$$

A partial notion of summation

Let $\mathcal{L}$ be a model of linear logic with 0-morphisms.

Summability structure in models of LL

Let $\mathrm{S} : \mathcal{L} \to \mathcal{L}$ be a functor. Intuitively, $\mathrm{S}X = \{\langle\!\langle x_0, x_1 \rangle\!\rangle \mid x_0 + x_1 \text{ is defined}\}$
Assume that there exists the following natural transformations:

- $\pi_0, \pi_1 \in \mathcal{L}(\mathrm{S}X, X)$ jointly monic $\qquad \pi_i : \langle\!\langle x_0, x_1 \rangle\!\rangle \mapsto x_i$
- Sum $\sigma \in \mathcal{L}(\mathrm{S}X, X) \qquad \sigma : \langle\!\langle x_0, x_1 \rangle\!\rangle \mapsto x_0 + x_1$

$f_0, f_1 \in \mathcal{L}(X, Y)$ are **summable** if :

$$\exists \langle\!\langle f_0, f_1 \rangle\!\rangle \in \mathcal{L}(X, \mathrm{S}Y) \text{ s.t. } \pi_i \circ \langle\!\langle f_0, f_1 \rangle\!\rangle = f_i$$
$$\langle\!\langle f_0, f_1 \rangle\!\rangle : x \mapsto \langle\!\langle f_0(x), f_1(x) \rangle\!\rangle$$

+ some axioms on S: the hom-sets are finite **partially additive monoids**

Additivity and left additivity

- Left additivity: $(g_1 + g_2) \circ f = g_1 \circ f + g_2 \circ f$
- Additivity: $g \circ (f_1 + f_2) = g \circ f_1 + g \circ f_2$

This summability structure on $\mathcal{L}$ induces a **left summability structure** on $\mathcal{L}_!$ (similar, but there is no additivity).

Differentiation in $\mathcal{L}_!$

. Define ι_0, θ, c and l natural transformations in $\mathcal{L}$.

$$\begin{array}{ll} \iota_0 \in \mathcal{L}(X, \mathrm{S}X) & \iota_0 \circ x = \langle\!\langle x, 0 \rangle\!\rangle \\ \theta \in \mathcal{L}(\mathrm{S}^2 X, \mathrm{S}X) & \boxed{\theta \circ \langle\!\langle \langle\!\langle x, u \rangle\!\rangle, \langle\!\langle v, w \rangle\!\rangle \rangle\!\rangle = \langle\!\langle x, u+v \rangle\!\rangle} \\ \mathrm{c} \in \mathcal{L}(\mathrm{S}^2 X, \mathrm{S}^2 X) & \mathrm{c} \circ \langle\!\langle \langle\!\langle x, u \rangle\!\rangle, \langle\!\langle v, w \rangle\!\rangle \rangle\!\rangle = \langle\!\langle \langle\!\langle x, v \rangle\!\rangle, \langle\!\langle u, w \rangle\!\rangle \rangle\!\rangle \\ \mathrm{l} \in \mathcal{L}(\mathrm{S}X, \mathrm{S}^2 X) & \mathrm{l} \circ \langle\!\langle x, u \rangle\!\rangle = \langle\!\langle \langle\!\langle x, 0 \rangle\!\rangle, \langle\!\langle 0, u \rangle\!\rangle \rangle\!\rangle \end{array}$$

Then $(\mathrm{S}, \iota_0, \theta)$ is a monad, $(\mathrm{S}, \sigma, \mathrm{l})$ is a comonad, and $(\mathrm{S}, \iota_0, \theta, \sigma, \mathrm{l}, \mathrm{c})$ is a bimonad.
Differentiation: functor $\mathrm{T} : \mathcal{L}_! \to \mathcal{L}_!$ that **extends** S to $\mathcal{L}_!$ (chain rule).

$$\begin{array}{rcl} \mathrm{T}f : \quad !\mathrm{S}X & \to & \mathrm{S}Y \\ \langle\!\langle x, u \rangle\!\rangle & \mapsto & \langle\!\langle f(x), f'(x).u \rangle\!\rangle \end{array}$$

Other properties of differentiation: naturality of the families above in $\mathcal{L}_!$. Example:
$\mathrm{Der}\,(\theta) \in \mathcal{L}_!(\mathrm{T}^2 X, \mathrm{T}X)$ natural $\Leftrightarrow f'(x)\cdot(u+v) = f'(x)\cdot u + f'(x)\cdot v$

Differentiation as Distributive law

Everything boils down to a **distributive law**: a natural transformation $\partial : !\mathrm{S} \Rightarrow \mathrm{S}!$ with compatibility condition with regard to $\mathrm{der}, \mathrm{dig}, \iota_0, \theta, \mathrm{l}, \mathrm{c}, \mathrm{m}^2$.

Important observation

$(\mathrm{T}, \mathrm{Der}\,(\iota_0), \mathrm{Der}\,(\mathrm{l}))$ is a monad, but $(\mathrm{T}, \mathrm{Der}\,(\sigma), \mathrm{Der}\,(\mathrm{l}))$ is **not** a comonad because $\mathrm{Der}\,(\sigma)$ is not natural.

Coherent Taylor expansion

The functor T performs a first order Taylor expansion. What if it could perform the whole Taylor expansion? Remarkably, this can be done with a very similar theory.

ω-summability structures

We introduce an infinitary counterpart of summability structures.

$$\mathrm{S}X = \{\langle\!\langle x_i \rangle\!\rangle_{i=0}^{\infty} \mid \sum_{i=0}^{\infty} x_i \text{ is defined}\}$$

+some axioms on S; the hom-sets are **partially additive monoids**

It has a similar bimonad structure as before, and Taylor expansion is again a functor $\mathrm{T} : \mathcal{L}_! \to \mathcal{L}_!$ that extends S. Intuitively: if $f : \mathbb{R} \to \mathbb{R}$ is analytic we can write

$$f\Big(\sum_{n=0}^{\infty} x_n \epsilon^n\Big) = \sum_{n=0}^{\infty} f_n(x_0, \dots, x_n)\epsilon^n$$

where f_n can be computed by the Faà Di Bruno formula. Then $\mathrm{T}f$ can be seen as

$$\mathrm{T}f\, \langle\!\langle x_i \rangle\!\rangle_{i=0}^{\infty} = \langle\!\langle f_n(x_0, \dots, x_n) \rangle\!\rangle_{n=0}^{\infty}$$

The bimonad structure of T

We ask the same axioms as coherent differentiation, with an additional one that enforces the analyticity of maps: $\mathrm{Der}\,(\sigma)$ is natural in $\mathcal{L}_!$. T is now a **bimonad**!

Only **one** axiom is not about the bimonad T: the compatibility between ∂ and m^2.

The elementary theory

In models we know: $\mathrm{S} = \mathbb{D} \multimap \_$

- Coherent differentiation: $\mathbb{D} = (1 \times 1)$
- Coherent Taylor expansion: $\mathbb{D} = 1^{\mathbb{N}}$

This is similar to the notion of representability of the tangent functor found in tangent categories.

Theory of mates

This object has a bimonoid structure that relates to the bimonad structure of S.
Furthermore, ∂ boils down to a !-colagebra structure on $\mathbb{D}$.

A syntax for coherent Taylor expansion

- Using coherent differentiation, Ehrhard introduced a deterministic PCF with both fixpoints and differentiation, and a straightforward probabilistic extension.
- This axiomatization of Taylor expansion suggests that this calculus can feature Taylor expansion.
- This calculus should be related to the computation of higher order derivatives in automated differentiation.

LaTeX TikZposter