Cartesian Differential Kleisli Categories

JS PL (he/him)

Email: js.lemay@mq.edu.au

Website: https://sites.google.com/view/jspl-personal-webpage

Cartesian Differential Categories

- Cartesian differential categories (CDC) provide the categorical foundations of multivariable differential calculus over Euclidean spaces.

  R. Blute, R. Cockett, R.A.G. Seely, **Cartesian Differential Categories**

- CDC provide the categorical semantics of the differential λ-calculus.

  T. Ehrhard, L. Regnier **The differential λ-calculus.**

  Manzonetto, G., **What is a categorical model of the differential and the resource λ-calculi?**

  R. Cockett, J. Gallagher **Categorical models of the differential λ-calculus.**

Applications of CDC

- Causal computation:
  - Sprunger, D. and Katsumata, S., **Differentiable causal computations via delayed trace**
- Incremental Computation:
  - Alvarez-Picallo, M. and Ong, C.-H. L., **Change actions: models of generalised differentiation**
- Game Theory:
  - Laird, J. and Manzonetto, G. and McCusker, G., **Constructing differential categories and deconstructing categories of games**
- Differentiable Programming:
  - Abadi, M. and Plotkin, G., **A simple differentiable programming language**
  - Cruttwell, G. and Gallagher, J. and Pronk, D. **Categorical semantics of a simple differential programming language.**
- Machine Learning and Automatic Differentiation:
  - Cockett, R., Cruttwell, G., Gallagher, J., Lemay, J. S. P., MacAdam, B., Plotkin, G., & Pronk, D. **Reverse derivative categories.**
  - Cruttwell, G. and Gavranović, B. and Ghani, N. and Wilson, P. and Zanasi, F. **Categorical Foundations of Gradient-Based Learning.**
  - Wilson, P. and Zanasi, F. **Reverse Derivative Ascent: A Categorical Approach to Learning Boolean Circuits.**

TODAY'S STORY: Understanding when the **Kleisli category** of a monad is a CDC.

Motivation:

- There is an Abelian functor calculus CDC

  Bauer, K., Johnson, B., Osborne, C., Riehl, E. and Tebbe, A., 2018. **Directional derivatives and higher order chain rules for abelian functor calculus.**

  and it is built from the Kleisli category of the chain complex monad.

- This is an important example but not built in the usual way...

- Important kind examples of CDC are the **coKleisli category** of differential categories (categorical semantics of Differential Linear Logic)

  R. Blute, R. Cockett, R.A.G. Seely, **Differential Categories** (2006)

- So it is natural to ask how to build CDC from Kleisli categories.

ANSWER: LIFTING

TODAY'S STORY: Monads on CDC that **lift** the differential structure to their Kleisli categories – meaning that the Kleisli category is a CDC with the same differentiation as the base CDC.

- Lifting structure to Kleisli categories is always an interesting question in category theory;
- Also an important concept in computer science since solutions to the lifting problem allow us to extend desirable structures or properties of a base programming language to the effectful programming language.
- By extending the differential combinator to effectul programs, we would be able to apply differential calculus based techniques and algorithms on effectul programs.

Cartesian Differential Categories - Definition

Definition

For a commutative semiring k, a **Cartesian k-differential category** is a category $\mathbb{X}$ with finite products such that:

- Hom-sets $\mathbb{X}(A, B)$ are k-modules such that pre-composition preserves the k-linear structure:

$$(r \cdot f + s \cdot g) \circ x = r \cdot (f \circ x) + s \cdot (g \circ x)$$

- A **differential combinator** D, which is a family of operators $\mathbb{X}(A, B) \xrightarrow{\mathsf{D}} \mathbb{X}(A \times A, B)$,

$$\frac{f : A \to B}{\mathsf{D}[f] : A \times A \to B}$$

  where $\mathsf{D}[f]$ is called the derivative of f, and which satisfies seven axioms which capture the basics of the derivative from differential calculus (such as the chain rule, etc.)

To help us, we will use the following term logic [1]:

$$\mathsf{D}[f](a, b) := \frac{\mathsf{d}f(x)}{\mathsf{d}x}(a) \cdot b$$

So for example the chain rule is:

$$\frac{\mathsf{d}g\,(f(x))}{\mathsf{d}x}(a) \cdot b = \frac{\mathsf{d}g(y)}{\mathsf{d}y}(f(x)) \cdot \left(\frac{\mathsf{d}f(x)}{\mathsf{d}x}(a) \cdot b \right)$$

[1] There is a sound & complete term logic for CDC. Anything we can prove using the term logic, holds in any CDC. Super useful!

Cartesian Differential Categories - Main Example

Example

Define SMOOTH as the category whose objects are the Euclidean real vector spaces $\mathbb{R}^n$ and whose maps are the real smooth functions $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^m$ between them. SMOOTH is a CDC where the differential combinator is defined as the directional derivative of a smooth function.

A smooth function $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^m$ is in fact a tuple $F = \langle f_1, \cdots, f_m \rangle$ of smooth functions $\mathbb{R}^n \xrightarrow{f_i} \mathbb{R}$. Using the convention that $\vec{x} \in \mathbb{R}^n$ are column vectors, the derivative $\mathbb{R}^n \times \mathbb{R}^n \xrightarrow{\mathsf{D}[F]} \mathbb{R}^m$ is defined as:

$$\mathsf{D}[F](\vec{x}, \vec{y}) := \left\langle \sum_{i=1}^{n} \frac{\partial f_1}{\partial x_i}(\vec{x}) y_i, \cdots, \sum_{i=1}^{n} \frac{\partial f_n}{\partial x_i}(\vec{x}) y_i \right\rangle$$

Cartesian Differential Categories - Linear Maps

Definition

In a CDC a map $A \xrightarrow{f} B$ is **linear** if:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x}(a) \cdot b = f(b)$$

Example

In SMOOTH, being linear in the CDC sense is the same as being $\mathbb{R}$-linear in the classical sense.

For example, $\mathbb{R}^n \xrightarrow{F} \mathbb{R}^m$ is linear in the CDC sense if and only if it is $\mathbb{R}$-linear in the classical sense:

$$F(s\vec{x} + t\vec{y}) = sF(\vec{x}) + tF(\vec{y})$$

Cartesian Differential Monads

Objective: Define a monad on a CDC whose Kleisli category is a CDC.

Definition

A **Cartesian differential monad** on a CDC is a monad $\mathbb{S} := (\mathsf{S}, \mu, \eta)$:

$$\mathsf{S} : \mathbb{X} \to \mathbb{X} \qquad \mu_A : \mathsf{SS}(A) \to \mathsf{S}(A) \qquad \eta_A : A \to \mathsf{S}(A)$$

such that:

- S preserves products, that is, $\omega_{A,B} := \mathsf{S}(A \times B) \xrightarrow{\langle \mathsf{S}(\pi_1), \mathsf{S}(\pi_2) \rangle} \mathsf{S}(A) \times \mathsf{S}(B)$ is an isomorphism.
  (which we need for the Kleisli category to have products)
- S preserves the k-linear structure:

  $$\mathsf{S}(r \cdot f + s \cdot g) = r \cdot \mathsf{S}(f) + s \cdot \mathsf{S}(g)$$

  (which we need for the Kleisli category to have k-linear structure)
- S is compatible with the differential combinator:

  $$\mathsf{D}[\mathsf{S}(f)] = \mathsf{S}(A) \times \mathsf{S}(A) \xrightarrow{\omega^{-1}_{A,A}} \mathsf{S}(A \times A) \xrightarrow{\mathsf{S}(\mathsf{D}[f])} \mathsf{S}(B)$$

- μ and η are differential linear.
  (which we need for the axioms of a differential combinator in the Kleisli category)

Kleisli Categories of Cartesian Differential Monads

To help with Kleisli categories we use interpretation brackets $[\![-]\!]$ to help distinguish between composition in the base category and Kleisli composition.

For a monad $\mathbb{S}$, its Kleisli category $\mathsf{KL}(\mathbb{S})$ is the category whose objects are the same as $\mathbb{X}$ and where a map $f : A \to B$ in the Kleisli category is a map of type $[\![f]\!] : A \to \mathsf{S}(B)$ in the base category.

Identity and composition is given by:

$$[\![1_A]\!] := \eta_A \qquad\qquad [\![g \circ f]\!] := \mu_C \circ \mathsf{S}\left([\![g]\!]\right) \circ [\![f]\!]$$

Theorem (JS PL)

For a Cartesian differential monad $\mathbb{S}$, the Kleisli category $\mathsf{KL}(\mathbb{S})$ is a CDC where in particular the differential combinator is defined as follows:

$$\frac{[\![f]\!] : A \to \mathsf{S}(B)}{[\![\mathsf{D}_{\mathbb{S}}[f]]\!] := \mathsf{D}\left[[\![f]\!]\right] : A \times A \to \mathsf{S}(B)}$$

and the canonical adjunction between $\mathbb{X}$ and $\mathsf{KL}(\mathbb{S})$ preserves the CDC structure (up to isomorphism).

It's remarkable that the differential combinator still satisfies the chain rule even with the Kleisli composition.

Example: Tangent Monad

Example

Every CDC always has a Cartesian differential monad provided by its tangent bundle monad.

The monad $\mathbb{T} := (\mathsf{T}, \mu, \eta)$ is defined as follows:

$$\mathsf{T}(A) = A \times A \qquad \mathsf{T}(f) = \langle f \circ \pi_1, \mathsf{D}[f] \rangle \qquad \mu_A = \langle \pi_1, \pi_2 + \pi_3 \rangle \qquad \eta_A = \langle 1_A, 0 \rangle$$

In term calculus notation, the tangent bundle on maps is given by:

$$\mathsf{T}(f)(a, b) = \left(f(a), \frac{\mathsf{d}f(x)}{\mathsf{d}x}(a) \cdot b \right)$$

$\mathbb{T}$ is always a Cartesian differential monad.

Maps in $\mathsf{KL}(\mathbb{T})$ is actually a pair of maps $[\![f]\!] = \langle f_1, f_2 \rangle : A \to B \times B$ and should be thought of as generalized vector fields. Composition is given by:

$$[\![g \circ f]\!] = \langle g_1 \circ f_1, g_2 \circ f_1 + \mathsf{D}[g_1] \circ \langle f_1, f_2 \rangle + \mathsf{D}[g_2] \circ \langle f_1, f_2 \rangle \rangle$$

and the differential combinator is given by:

$$[\![\mathsf{D}_{\mathbb{T}}[f]]\!] := \langle \mathsf{D}[f_1], \mathsf{D}[f_2] \rangle$$

Example: Tangent Monad

Example

Let's consider the tangent bundle on SMOOTH, which is given by:

$$\mathsf{T}(\mathbb{R}^n) = \mathbb{R}^n \times \mathbb{R}^n \qquad \mathsf{T}(F)(\vec{x}, \vec{y}) = (F(\vec{x}), \mathsf{D}[F](\vec{x}, \vec{y}))$$

$$\mu_{\mathbb{R}^n}(\vec{x}, \vec{y}, \vec{z}, \vec{w}) = (\vec{x}, \vec{y} + \vec{z}) \qquad \eta_{\mathbb{R}^n}(\vec{x}) = (\vec{x}, \vec{0})$$

Maps in the Kleisi category are smooth functions $[\![F]\!] : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ where:

$$[\![F]\!](\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}))$$

A vector field is of the form:

$$[\![V]\!](\vec{x}) = (\vec{x}, v_2(\vec{x}))$$

and these correspond precisely to the vector fields of $\mathbb{R}^n$ in the classical differential geometry.

Example: Reader Monad

Example

In a Cartesian *closed* differential category, reader monads are Cartesian differential monads.

A Cartesian closed k-differential category CDC which is also Cartesian closed such that differential combinator is compatible with the closed structure:

$$\frac{\mathsf{d}\lambda y.f(y,x)}{\mathsf{d}x}(a)\cdot b = \lambda y.\frac{\mathsf{d}f(y,x)}{\mathsf{d}x}(a)\cdot b$$

Every model of the differential λ-calculus induces a Cartesian closed differential category, and conversely, every Cartesian closed differential category gives rise to a model of the differential λ-calculus.

Every object C induces a monad $\mathbb{R}(C) := \left([C,-],\mu^C,\eta^C\right)$ defined as follows:

$$[C,-](A) := [C,A] \qquad [C,-](f)\,(g(-)) = [C,f]\,(g(-)) = \lambda x.f(g(x))$$

$$\mu^C_A\,(F(-)(-)) = \lambda x.F(x)(x) \qquad \eta^C_A(a) = \lambda x.a$$

$\mathbb{R}(C)$ is a Cartesian differential monad.

Example: Kleisli Category of the Reader Monad

Example

For reader monads, their Kleisli categories capture partial differentiation.

Maps in $\mathsf{KL}(\mathbb{R}^C)$, which are of the form $A \to [C, A]$, correspond precisely to maps of the form $C \times A \to B$ via Currying and unCurrying.

The Kleisli category of the reader monad of C is isomorphic to the coKleisli category of the comonad $C \times -$, which is better known as the simple slice category over C, denoted $\mathbb{X}[C]$.

$\mathbb{X}[C]$ is a CDC but whose differential combinator D^C is given by partial differentiation:

$$\mathsf{D}^C[f](c, a, b) = \frac{\mathsf{d}f(c, x)}{\mathsf{d}x}(a) \cdot b$$

However, the partial derivative of a map $f : C \times A \to B$ is the same as taking the total derivative of its Curry $\lambda(f) : A \to [C, B]$ and then unCurrying, that is:

$$\mathsf{D}^C[f] = \lambda^{-1}(\mathsf{D}[\lambda(f)])$$

As such, the lifting of the differential combinator D to $\mathsf{KL}(\mathbb{R}^C)$ does indeed recapture the fact that we can define partial differentiation in context C.

Cartesian Differential Abstract Kleisli Categories

- We can precisely characterize which CDC are the Kleisli category of a Cartesian differential monad using **abstract Kleisli categories**, which give a direct description of Kleisli categories without a starting monad.

Führmann **Direct models of the computational lambda-calculus**

Proposition

A CDC is the Kleisli category of a Cartesian differential monad if and only if it is a Cartesian differential abstract Kleisli category.

Interested? See my paper for more details.

Kleisli Differential Combinator

It is possible to start from a slightly more general setting, one where the base category does not necessarily have a differential combinator.

Definition

For a monad $\mathbb{S}$ which preserves products and the k-linear structure, a **Kleisli differential combinator** which is a family of operators $\mathbb{X}(A, B) \xrightarrow{\mathsf{B}} \mathbb{X}(A \times A, \mathsf{S}(B))$,

$$\frac{f : A \to B}{\mathsf{B}[f] : A \times A \to \mathsf{S}(B)}$$

satisfying certain axioms, which are essentially the axioms of a differential combinator in the Kleisli category.

Kleisli Differential Combinator

Proposition

If $\mathbb{S}$ has a Kleisli differential combinator, then $\mathsf{KL}(\mathbb{S})$ *is a CDC where in particular the differential combinator is defined as follows:*

$$\frac{A \xrightarrow{[\![f]\!]} B}{[\![\mathsf{D}_{\mathbb{S}}[f]]\!] := A \times A \xrightarrow{\mathsf{B}[[\![f]\!]]} \mathsf{SS}(B) \xrightarrow{\mu_B} \mathsf{S}(B)}$$

Proposition

Every Cartesian differential monad has a Kleisli differential combinator:

$$\frac{A \xrightarrow{f} B}{[\![\mathsf{D}_{\mathbb{S}}[f]]\!] := A \times A \xrightarrow{\mathsf{D}[f]} B \xrightarrow{\eta_B} \mathsf{S}(B)}$$

and the resulting differential combinator for the Kleisli category is the same as before.

- For every monad $\mathbb{S}$ with a Kleisli differential combinator, it is always possible to build a CDC with a Cartesian differential monad $\mathbb{S}'$ such that the Kleisli categories are the same $\mathsf{KL}(\mathbb{S}) = \mathsf{KL}(\mathbb{S}')$. So an argument can be made we could have just started with a CDC...
- Also I don't have any good examples that aren't just Cartesian differential monads.
- Kleisli differential combinators may play a role in better understanding the Abelian functor calculus model and possibly higher categorical CDCs.

Eilenberg-Moore Category of a Cartesian Differential Monad

- Whenever we have a monad $\mathbb{S}$, an important question is always what are its algebras $(A, \mathsf{S}(A) \xrightarrow{\alpha} A)$ and Eilenberg-Moore category $\mathsf{EM}(\mathbb{S})$?

- Unfortunately, $\mathsf{EM}(\mathbb{S})$ is not a CDC since the derivative of an S-algebra morphism may not be an S-algebra morphism.

- **ANSWER:** Instead $\mathsf{EM}(\mathbb{S})$ is a **TANGENT CATEGORY**.
  So $\mathbb{S}$-algebras are abstract smooth manifolds.

- Main idea:
  - CDC = Differential Calculus over Euclidean spaces $\mathbb{R}^n$
  - Tangent Category = Differential calculus over tangent bundles of smooth manifolds

  So very briefly a tangent category is a category equipped with an endofunctor T which encodes the essential properties of the tangent bundle. So for an object A, $\mathsf{T}(A)$ is an abstract tangent bundle of A.

J. Rosický (1984) **Abstract tangent functors**

R. Cockett, G. Cruttwell (2014) **Differential structure, tangent structure, and SDG**

Theorem

For a Cartesian differential monad $\mathbb{S}$, its Eilenberg-Moore $\mathsf{EM}(\mathbb{S})$ *is a tangent category where the tangent bundle of an S-algebra is given by:*

$$\mathsf{T}(A, \mathsf{S}(A) \xrightarrow{\alpha} A) = (A \times A, \mathsf{S}(A \times A) \xrightarrow{\omega_{A,A}} \mathsf{S}(A) \times \mathsf{S}(A) \xrightarrow{\mathsf{T}(\alpha)} A \times A)$$

Examples are a bit mysterious

Example

- For both Cartesian differential categories and tangent categories, algebras of the tangent bundle monad are not necessarily well studied and somewhat mysterious.
- Even in specific examples from differential calculus and differential geometry, there does not in general seem to be a precise characterization beyond the categorical definition.
- Yet, the above theorem tells us that the Eilenberg-Moore category of the tangent bundle monad is always a Cartesian tangent category – which may have interesting consequences in future work.

Example

- In general, it is understood that algebras of reader monads are not easily characterized.
- However, we now know that for Cartesian closed differential categories, and thus for any model of the differential λ-calculus, the algebras of the reader monads form a tangent category.

Concluding Thoughts

- Using Cartesian differential monad we were able to lift the differential combinator to the Kleisli category. YAY!
- Tangent bundle monads and reader monads are very nice examples of Cartesian differential monads.

However there are lots of other examples of monads of interest such as:

- State monads: $[S, S \times -]$
- Selection monads: $[[-, S], -]$
- Continuation monads: $[-, [-, S]]$
- Writer monads: $M \times -$

But these monads don't preserve the CDC structure! So their Kleisli categories can't be CDCs...

The axioms of a Cartesian differential monad are quite strong/restrictive...

CONJECTURE: Instead, their Kleisli categories might still be tangent categories thanks to possible distributive laws between the monad and the tangent bundle monad!

- Thus effectful programs may still form a tangent category, further motivating the need for a term calculus for tangent categories and developing "tangent programming languages".

HOPE YOU ENJOYED MY TALK!

THANKS FOR LISTENING!

MERCI!

Email: jslemay@mq.edu.au

Website: https://sites.google.com/view/jspl-personal-webpage