Differential Categories and Machine Learning

JS PL (he/him)

Email: js.lemay@mq.edu.au
Website: https://sites.google.com/view/jspl-personal-webpage

Hello!

- Full Name: **JEAN-SIMON PACAUD LEMAY** but please just call me **JS**
- I'm a mathematician who sometimes plays with theoretical computer science.
- My Research: Category Theory, mostly with **Differential Categories**

What is the Theory of Differential Categories About?

- The theory of differential categories uses category theory to provide the foundations of differentiation and has been able to formalize numerous aspects of differential calculus.
- Differential categories are successful because they capture both the classical limit definition of differentiation and the more algebraic synthetic definition of differentiation.
- Differential categories have recently gained in popularity – especially in computer science!

TODAY'S STORY: An overview of how differential categories have been used for machine learning and automatic differentiation.

MAIN IDEA: Differential categories used for foundations for ML/AD, specifically for axiomatizing **REVERSE DIFFERENTIATION**

The Differential Category World: The Four Tomes

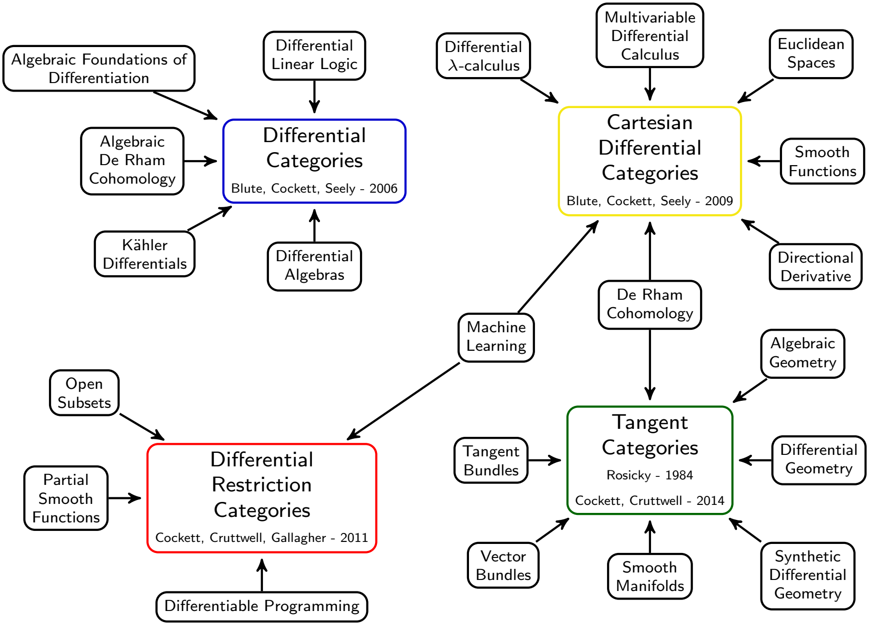
Differential Categories
Blute, Cockett, Seely - 2006

Cartesian Differential Categories
Blute, Cockett, Seely - 2009

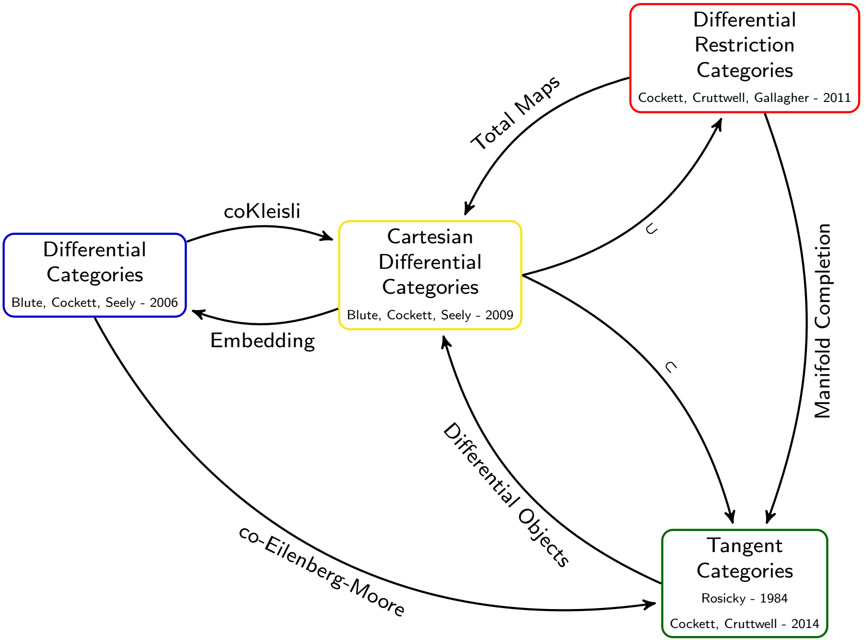
Differential Restriction Categories
Cockett, Cruttwell, Gallagher - 2011

Tangent Categories
Rosicky - 1984
Cockett, Cruttwell - 2014

The Differential Category World: A Taster



The Differential Category World: It's all connected!



Quick Overview

- Differential Categories :: **Deriving Transformation**

$$\mathsf{d} : !A \otimes A \to !A \qquad \frac{f : !A \to B}{\mathsf{D}[f] : !A \otimes A \to B}$$

- Cartesian Differential Categories :: **Differential Combinator**

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

Same thing for Differential Restriction Categories but with partiality (restriction categories!)

- Tangent Categories :: **Tangent Bundle Functor**

$$\mathsf{T}(A) \text{ abstract tangent bundle over } A \qquad \frac{f : A \to B}{\mathsf{T}(f) : \mathsf{T}(A) \to \mathsf{T}(B)}$$

Prologue: Differentiable Programming Language

M. Abadi & Plotkin, G. (2019). *A Simple Differentiable Programming Language.*

Martin Abadi

Gordon Plotkin

TL;DR: A programming language for automatic differentiation.

Gordon comes to FMCS

At FMCS 2019 (Kananaskis Valley, Alberta, Canada), the theme was **differential categories and differential programming**. Gordon Plotkin gave a talk on his differential programming language which he developed at Google and announced at the end:

"**You all need to develop a categorical theory of reverse differentiation!**"

Why Reverse Differential Categories

- There are two types of derivative operations used in automatic differentiation: the **forward** derivative (sum of partial derivatives) and the **reverse** derivative (tuple of partial derivatives).

$$\mathsf{D}[f] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \qquad \mathsf{D}[f](\vec{x}, \vec{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{x}) a_i$$

$$\mathsf{R}[f] : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n \qquad \mathsf{R}[f](\vec{x}, y) = \left\langle \frac{\partial f}{\partial x_1}(\vec{x}) a, \cdots, \frac{\partial f}{\partial x_n}(\vec{x}) a \right\rangle$$

- The theory of differential categories uses category theory to provide the foundations of differentiation and has been able to formalize numerous aspects of differential calculus from commutative algebra, differential geometry, and computer science.

Differential categories talked about the **FORWARD** derivative...

We wanted: Categorical version of **REVERSE** derivative.

Reverse Differential Categories

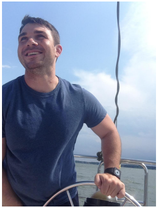
Cockett, R., Cruttwell, G., Gallagher, J., Lemay, J. S. P., MacAdam, B., Plotkin, G., & Pronk, D. (2020). *Reverse derivative categories.* In the proceedings of CSL2020.

Geoff Cruttwell

Gordon Plotkin

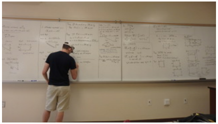
Robin Cockett

Dorette Pronk



Jonathan Gallager

Ben MacAdam



J.-S. P. L.

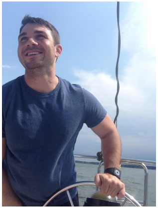
TL;DR: A category with a reverse differential operator on maps.

SDPL+

Cruttwell, G., Gallagher, J., & Pronk, D. (2020). *Categorical semantics of a simple differential programming language.*

Geoff Cruttwell

Dorette Pronk



Jonathan Gallager

TL;DR: Checks that Abadi and Plotkin's language is indeed a CRDC.

Reverse Derivative Ascent

P. Wilson & F. Zanasi (2020). *Reverse Derivative Ascent: A Categorical Approach to Learning Boolean Circuits.*

Paul Wilson

Fabio Zanasi

TL;DR: Algorithm on reverse differential categories, specifically for Boolean circuits – "learning the parameters of Boolean circuits" (STRING DIAGRAMS!)

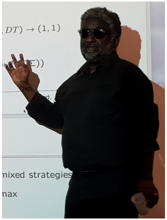
- Paul's website: `https://statusfailed.com/`

Categorical Foundations of Gradient-Based Learning

Cruttwell, G., Gavranović, B., Ghani, N., Wilson, P. & Zanasi, F. (2022). *Categorical Foundations of Gradient-Based Learning.*

Geoff Cruttwell

Bruno Gavranović



Neil Ghani

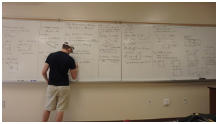
Paul Wilson

Fabio Zanasi

TL;DR: Foundations using lenses, parametric maps, and reverse differential categories.

Jacobians and Gradients

Lemay, J.-S. P. (2021). *Jacobians and Gradients for Cartesian Differential Categories.*



J.-S. P. L.

TL;DR: Studies Jacobians $\mathsf{J}(f) : A \to \mathsf{Linear}(A, B)$ and gradients $\nabla(f) : A \to \mathsf{Linear}(B, A)$ in a CRDC.

Geoff's Conference Talk

Geoff Cruttwell

- CogSci 2022 – Category Theory for Cognitive Science
- Mathematical Congress of the Americas

Main Idea about Reverse Differential Categories

We want to build a map of reverse differential categories...

So we need reverse versions of the four stages:

- Differential Categories;
- Cartesian Differential Categories;
- Differential Restriction Categories;
- Tangent Categories.

How we do we do this?

- Reverse differential operator = Forward differential operator + transpose/dagger
- Reverse differential category = Forward differential category + linear dagger

Cartesian Reverse Differential Categories

Cockett, R., Cruttwell, G., Gallagher, J., Lemay, J. S. P., MacAdam, B., Plotkin, G., & Pronk, D. (2020). *Reverse derivative categories.* In the proceedings of CSL2020.

- Cartesian reverse differential categories

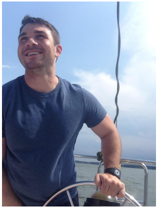
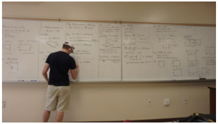
Cruttwell, G., Gallagher, J., & Pronk, D. (2020). *Categorical semantics of a simple differential programming language.*

- Reverse differential restriction categories

Monoidal Reverse Differential Categories

Cruttwell, G., Gallagher, J., Lemay, J.-S. P., & Pronk, D. (2020). *Monoidal reverse differential categories.*



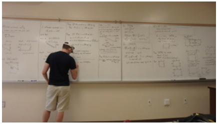



Geoff Cruttwell Dorette Pronk Jonathan Gallager J.-S. P. L.

- Reverse differential categories
- Cokleisli category is a Cartesian reverse differentiall category.

Reverse Tangent Categories *In the works*

Geoff Cruttwell



J.-S. P. L.

- Involves the cotangent bundle and duals of smooth vector bundles.
- Hopefully in some 2024 conference proceedings!

Quick Overview but in Reverse

- Reverse Differential Categories :: **Reverse Deriving Transformation**

$$\mathsf{r} : !A \otimes !A \to A \qquad \frac{f : !A \to B}{\mathsf{R}[f] : !A \otimes B \to A}$$

- Cartesian Reverse Differential Categories :: **Differential Combinator**

$$\frac{f : A \to B}{\mathsf{R}[f] : A \times B \to A} \qquad \mathsf{R}[f](a, b) = b \cdot \frac{\mathrm{d}f}{\mathrm{d}x}(a)$$

Same thing for Reverse Differential Restriction Categories

- Reverse Tangent Categories :: **Tangent Bundle Functor**

$\mathsf{T}^*(A)$ abstract cotangent bundle over A

$$\frac{f : A \to B}{\mathsf{T}^*(f) : A \times_B \mathsf{T}^*(B) \to \mathsf{T}^*(A)}$$

Mta + Strathclyde

Main idea: providing an axiomatization of differential structures purely in terms of lenses

Bruno's website: `https://github.com/bgavran/Category_Theory_Machine_Learning`

Some Ideas for future work

- "Infinite dimensional versions" of reverse differentiation
- Reverse differential algebra/geometry – what does this even mean?
- Reverse differential linear logic and reverse differential λ-calculus
- Axiomatization of (reverse) differential categories using Jacobians/gradients
- Couniversal construction of reverse differential categories
- Relating to other work (especially M. Vákár and his coauthors' work):

M. Huot, S. Staton, & M. Vákár (2020). *Correctness of Automatic Differentiation via Diffeologies and Categorical Gluing.*

M. Vákár & T. Smeding (2021). *CHAD: Combinatory Homomorphic Automatic Differentiation .*

M. Huot, S. Staton, & M. Vákár (2021). *Higher Order Automatic Differentiation of Higher Order Functions .*

- While using reverse differential categories for foundations is great, I would definitely like to see more applications! My hope is this continues to grow along with category theory and machine learning.

That's All Folks!

HOPE YOU ENJOYED
MY TALK!

THANK YOU FOR LISTENING!

MERCI!

Email: js.lemay@mq.edu.au

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

Bruno's website: https://github.com/bgavran/Category_Theory_Machine_Learning