#### An overview on Dialectica and Differentiation

#### Davide Barbarossa

db2437@bath.ac.uk

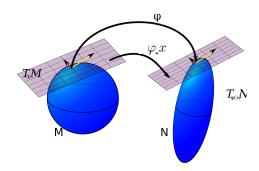
https://davidebarbarossa12.github.io/

Department of Computer Science



Seminar at University of Sussex, Brighton (UK)

13/11/2024

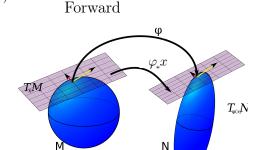


Cartesian differential categories ( $\sim$ '09)

$$\frac{f:A\to B}{Df:A\times A\to B}$$

Cartesian tangent categories ('14)

$$\frac{f:A\to B}{Tf:TA\to TB}$$

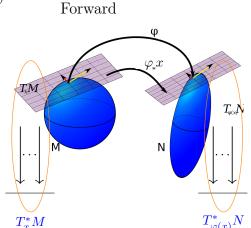


Cartesian differential categories ( $\sim$ '09)

$$\frac{f:A\to B}{Df:A\times A\to B}$$

Cartesian tangent categories ('14)

$$\frac{f:A\to B}{Tf:TA\to TB}$$

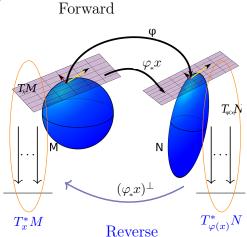


Cartesian differential categories ( $\sim$ '09)

$$\frac{f:A\to B}{Df:A\times A\to B}$$

Cartesian tangent categories ('14)

$$\frac{f:A\to B}{Tf:TA\to TB}$$



Cartesian differential categories ( $\sim$ '09)

$$\frac{f:A\to B}{Df:A\times A\to B}$$

Cartesian tangent categories ('14)

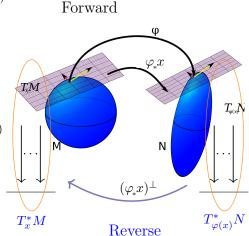
$$\frac{f:A\to B}{Tf:TA\to TB}$$

Cartesian reverse diff. categories ('20)

$$\frac{f:A\to B}{Rf:A\times B\to A}$$

Cart. reverse tangent categories ('24)

$$\frac{f:A\to B}{T^*f:f^*T^*B\to T^*A}$$



# Dialectica (overview)

	$Source \rightarrow Target$	
Gödel	$A \in \mathrm{HA}  \longmapsto  A_D\{w,c\} \in \mathrm{T}$	
('58)	$ \begin{array}{ccc} & such \ that \\ \vdash_{\mathrm{HA}} A & \Longrightarrow & \vdash_{\mathrm{T}} A_D \{\mathtt{M},c\} \ for \ some \ \mathtt{M} \in \mathrm{T} \end{array} $	

# Dialectica (overview)

		Sour	$rce \rightarrow Target$
Gödel ('58)	$A \in \mathrm{HA}$ $\vdash_{\mathrm{HA}} A$	$such\ that$	$A_D\{w,c\} \in \mathcal{T}$ $\vdash_{\mathcal{T}} A_D\{\mathtt{M},c\} \ \textit{for some} \ \mathtt{M} \in \mathcal{T}$
Pédrot (Diller- Nahm) ('15)	$A \in \Lambda \\ \mathtt{M} \in \Lambda$ $\mathtt{x}: A \vdash_{\Lambda} \mathtt{M}: B$	$\begin{array}{c} \longmapsto \\ \longmapsto \\ such \ that \\ \Longrightarrow \end{array}$	$\begin{split} W(A),  C(A) &\in \mathbf{P} \\ \mathbf{M}^{\bullet},  \mathbf{M}_{\mathbf{x}} &\in \mathbf{P} \ (\textit{for } \mathbf{x} \ \textit{variable}) \\ \\ \begin{cases} \mathbf{x} : W(A) \vdash_{\mathbf{P}} \mathbf{M}^{\bullet} : W(B) \\ \mathbf{x} : W(A) \vdash_{\mathbf{P}} \mathbf{M}_{\mathbf{x}} : C(B) \to \mathcal{M}[C(A)] \end{cases} \end{split}$

# Dialectica (Transformation)

	$\alpha$	$E \to F$
W	$lpha_W$	$W(E) \to W(F)$ $\times$ $W(E) \times C(F) \to \mathcal{M}[C(E)]$
$\mathbf{C}$	$\alpha_C$	$W(E) \times C(F)$

	x	$\lambda$ x.M	PQ
(_)•	x	$\left\langle\begin{array}{c} \lambda_{\mathtt{X}.\mathtt{M}^{\bullet}} \\ \lambda_{\pi.(\lambda_{\mathtt{X}.\mathtt{M}_{\mathtt{X}}})\pi^{1}\pi^{2} \end{array}\right\rangle$	P <sup>•1</sup> Q•
(_) <sub>y</sub>	$\begin{cases} \lambda \pi.[\pi], & \mathbf{x} = \mathbf{y} \\ \lambda \pi.0, & \mathbf{y} \neq \mathbf{y} \end{cases}$	$\lambda\pi.(\lambda {\tt x.M_y})\pi^1\pi^2$	$\lambda \pi. \begin{pmatrix} P_{y} \langle \mathbb{Q}^{\bullet}, \pi \rangle \\ + \\ P^{\bullet 2} \langle \mathbb{Q}^{\bullet}, \pi \rangle \gg \mathbb{Q}_{y} \end{pmatrix}$

Arrows in  $C: A \xrightarrow{f} B$  (linear) Arrows in  $C_!: A \xrightarrow{f} B := !A \xrightarrow{f} B$  (non-linear)

Arrows in  $C: A \xrightarrow{f} B$  (linear) Arrows in  $C_!: A \xrightarrow{f} B := !A \xrightarrow{f} B$  (non-linear)

$$Cartesian \\ +SMC \\ +Seely$$
  $A B A B A B$   $T$ 

$$\overline{\operatorname{ev}_{A,B} : [A \multimap B] \otimes A \multimap B}$$

$$\frac{f : A \multimap [E \multimap F]}{\lambda f : A \otimes E \multimap F}$$

 $\mathit{Arrows\ in}\ \mathcal{C}\colon A \xrightarrow{f} B\ (\mathit{linear}) \quad \mathit{Arrows\ in}\ \mathcal{C}_!\colon A \xrightarrow{f} B := !A \xrightarrow{f} B\ (\mathit{non-linear})$ 

$$\begin{array}{ccc} Cartesian \\ +SMC \\ +Seely \end{array} \quad \begin{array}{ccc} \underline{A} & \underline{B} & \underline{A} & \underline{B} \\ \overline{A \otimes B} & \overline{A \otimes B} & \overline{!\top} \\ \\ \star \text{-}autonomous & & \underline{f}: A \longrightarrow B \\ \underline{f^{\perp}: B^{\perp} \longrightarrow A^{\perp}} \end{array}$$

$$\overline{\operatorname{ev}_{A,B} : [A \multimap B] \otimes A \multimap B}$$

$$\frac{f : A \multimap [E \multimap F]}{\lambda f : A \otimes E \multimap F}$$

$$\overline{f^{\bot \bot} = f}$$

 $\mathit{Arrows\ in}\ \mathcal{C}\colon A \stackrel{f}{\multimap} B\ (\mathit{linear}) \quad \mathit{Arrows\ in}\ \mathcal{C}_!\colon A \stackrel{f}{\rightarrow} B := !A \stackrel{f}{\multimap} B\ (\mathit{non-linear})$ 

$$\begin{array}{ccc} +SMC \\ +Seely \end{array} & \begin{array}{ccc} A & B \\ \hline A \& B \end{array} & \begin{array}{c} A & B \\ \hline A \otimes B \end{array} & \begin{array}{c} F \\ \hline \end{array}$$

$$\star \text{-}autonomous \\ comm. \\ monoids \\ enriched \end{array} & \begin{array}{c} f:A \multimap B \\ \hline f^{\perp}:B^{\perp} \multimap A^{\perp} \end{array}$$

Cartesian

$$\overline{\operatorname{ev}_{A,B}: [A \multimap B] \otimes A \multimap B}$$

$$\frac{f: A \multimap [E \multimap F]}{\lambda f: A \otimes E \multimap F}$$

$$\overline{f^{\bot \bot} = f}$$

$$f: A \multimap B \quad g: A \multimap B$$

$$f+g: A \multimap B$$

Arrows in  $\mathcal{C}$ :  $A \stackrel{f}{\multimap} B$  (linear) Arrows in  $\mathcal{C}_!$ :  $A \stackrel{f}{\rightarrow} B := !A \stackrel{f}{\multimap} B$  (non-linear)

Arrows in  $\mathcal{C} \colon A \stackrel{f}{\multimap} B$  (linear) Arrows in  $\mathcal{C}_! \colon A \stackrel{f}{\rightarrow} B := !A \stackrel{f}{\multimap} B$  (non-linear)

 $\mathcal{C}_!$  is a model of differential  $\lambda$ -calculus where we can transpose linear arrows:

$$\frac{f:A\to B}{Df:A\times A\to B}\quad (in\ \mathcal{C}_!,\ E\times F:=E\&F)$$

# Dialectica and (Categorical) Differentiation

$$\sim_B \subseteq \{\vdash_{\mathbf{P}} \mathtt{M} : W(B)\} \times \mathcal{C}_!(\top, B)$$

$$\bowtie_B^A \ \subseteq \ \{\vdash_{\mathbf{P}} \mathtt{M} : C(B) \to \mathcal{M}[C(A)]\} \ \times \ \mathcal{C}_!(A,B) \times \mathcal{C}(B^\perp,A^\perp)$$

$\mathbf{M} \sim_{E \to F} f$	
$\mathbf{M}\bowtie_{E\to F}^{A}\binom{f}{g}$	for all $\mathtt{H} \sim_E e$ , we have $\lambda \pi.\mathtt{M} \langle \mathtt{H}, \pi \rangle \bowtie^A_F \left( \begin{array}{c} f _e : A \to F \\ g^\perp _e : F^\perp \multimap A^\perp \end{array} \right)$
	$\lambda^{\pi,\mathbb{M}(\mathbb{H},\pi)} \bowtie_F^{\perp} \left( g^{\perp}   _{a}^{\perp} : F^{\perp} \multimap A^{\perp} \right)$

#### The theorem

Let  $x : A \vdash_{\Lambda} M : B$ . Then:

$$1) \quad (\lambda \mathbf{x}.\mathbf{M})^{\bullet} \quad \sim_{A \to B} \quad [\![\lambda \mathbf{x}.\mathbf{M}]\!] : [A \to B]$$

2) 
$$(\lambda \mathbf{x} \cdot \mathbf{M}_{\mathbf{x}}) \mathbf{N} \bowtie_{B}^{A} \begin{pmatrix} [\![\mathbf{M}]\!] & : & A \to B \\ ([\![\mathbf{M}]\!] & a)^{\perp} & : & B^{\perp} \multimap A^{\perp} \end{pmatrix}$$
 for all  $\mathbf{N} \sim_{A} a$ .

Moral:

$$(\lambda x.M^{\bullet}, \lambda x.M_{x})$$

"represents" the pair ( $\llbracket M \rrbracket$ ,  $R \llbracket M \rrbracket$ ), where

$$R[M]: A \times B^{\perp} \to A^{\perp}$$

is the reverse differential of  $\llbracket M \rrbracket$ .

Yet, can we say that Dialectica really is (reverse) Differentiation?

• All Dialectica transformation/only the transformation of  $\lambda$ -calculus. What about quantifiers?

Yet, can we say that Dialectica really is (reverse) Differentiation?

All Dialectica transformation/only the transformation of λ-calculus.
 What about quantifiers?
 Counterexample (thanks to Thomas Powell), but that's another topic...

- All Dialectica transformation/only the transformation of λ-calculus.
   What about quantifiers?
   Counterexample (thanks to Thomas Powell), but that's another topic...
- The main feature of Dialectica is that in the target we have (e.g.) binary predicates (non-trivial subobjects in Dialectica categories). Here we don't: aren't we lose something about Dialectica?

- All Dialectica transformation/only the transformation of  $\lambda$ -calculus. What about quantifiers? Counterexample (thanks to Thomas Powell), but that's another topic...
- The main feature of Dialectica is that in the target we have (e.g.) binary predicates (non-trivial subobjects in Dialectica categories). Here we don't: aren't we lose something about Dialectica?

- All Dialectica transformation/only the transformation of  $\lambda$ -calculus. What about quantifiers? Counterexample (thanks to Thomas Powell), but that's another topic...
- The main feature of Dialectica is that in the target we have (e.g.) binary predicates (non-trivial subobjects in Dialectica categories). Here we don't: aren't we lose something about Dialectica?

  Not really, we are just formulating it differently.
- Is this correspondence astonishing/magic? Can we find some "reason" clarifying it?

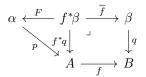
- All Dialectica transformation/only the transformation of  $\lambda$ -calculus. What about quantifiers? Counterexample (thanks to Thomas Powell), but that's another topic...
- The main feature of Dialectica is that in the target we have (e.g.) binary predicates (non-trivial subobjects in Dialectica categories). Here we don't: aren't we lose something about Dialectica?

  Not really, we are just formulating it differently.
- Is this correspondence astonishing/magic? Can we find some "reason" clarifying it?
   Definitely yes at first sight... but then we can clearly understand its reason by looking at the categorical framework behind it.

## Lens Categories

The category Lens( $\mathcal{L}$ ) of lenses over  $\mathcal{L}$  is defined as follows:

- ullet objects: arrows in  $\mathcal{L}$ , which we think as fibre bundles and we write  $p:\binom{\alpha}{A}$
- arrows from  $p:\binom{\alpha}{A}$  to  $q:\binom{\beta}{B}$  are the data of both a  $f:A\to B$  in  $\mathcal L$  and a span  $\alpha \xleftarrow{F} f^*\beta \xrightarrow{\overline{f}} \beta$  in  $\mathcal L$ , taken from the following pullback diagram:



## Lens Categories

The category Lens( $\mathcal{L}$ ) of lenses over  $\mathcal{L}$  is defined as follows:

- objects: arrows in  $\mathcal{L}$ , which we think as fibre bundles and we write  $p:\binom{\alpha}{A}$
- arrows from  $p:\binom{\alpha}{A}$  to  $q:\binom{\beta}{B}$  are the data of both a  $f:A\to B$  in  $\mathcal L$  and a span  $\alpha \xleftarrow{F} f^*\beta \xrightarrow{\overline{f}} \beta$  in  $\mathcal L$ , taken from the following pullback diagram:

$$\alpha \xleftarrow{F} f^*\beta \xrightarrow{\overline{f}} \beta$$

$$\downarrow^{p} f^*q \downarrow \qquad \downarrow^{q}$$

$$A \xrightarrow{f} B$$

Let  $\mathcal{E}Lens(\mathcal{L})$  be the full subcategory of Lens( $\mathcal{L}$ ) of trivial bundles, i.e. first projections. Concretely:

- Objects are first projections  $\pi_1:\binom{A\times X}{A}$
- An arrow from  $\pi_1: \binom{A\times X}{A}$  to  $\pi_1: \binom{B\times Y}{B}$  is given by an  $f: A\to B$  and a span  $A\times X\xleftarrow{F} A\times Y\xrightarrow{f\times 1} B\times Y$  such that  $F; \pi_1^{A,X}=\pi_1^{A,Y}$ .

Let  $\mathcal{L}$  be a Cartesian (closed, if you want  $\lambda$ -calculus) differential category where from the differential Df of a function f (a primitive data in  $\mathcal{L}$ ) we can define the reverse differential Rf of f. (Think of  $\mathcal{L} := \mathcal{C}_!$  of the first part).

Let  $\mathcal{L}$  be a Cartesian (closed, if you want  $\lambda$ -calculus) differential category where from the differential Df of a function f (a primitive data in  $\mathcal{L}$ ) we can define the reverse differential Rf of f. (Think of  $\mathcal{L} := \mathcal{C}_!$  of the first part).

We have a functor  $\mathcal{L} \to \mathcal{E} Lens(\mathcal{L})$  defined by:

$$A \qquad \mapsto \qquad \qquad \pi_1 : \binom{A \times A^{\perp}}{A}$$

$$A \xrightarrow{f} B \quad \mapsto \quad ( \quad f \quad , \quad A \times A^{\perp} \xleftarrow{\langle \pi_1, Rf \rangle} A \times B^{\perp} \xrightarrow{f \times 1} B \times B^{\perp} \quad ).$$

Let  ${\mathcal L}$  a Cartesian reverse differential category.

Let  $\mathcal{L}$  a Cartesian reverse differential category.

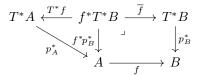
We have a functor  $\mathcal{L} \to \mathcal{E} Lens(\mathcal{L})$  defined by:

$$A \qquad \mapsto \qquad \qquad \pi_1 : {A \times A \choose A}$$

$$A \xrightarrow{f} B \qquad \mapsto \qquad ( f \quad , \quad A \times A \xrightarrow{\langle \pi_1, Rf \rangle} A \times B \xrightarrow{f \times 1} B \times B \quad )$$

where  $Rf: A \times B \to A$  is the reverse differential of f (a primitive data in  $\mathcal{L}$ ).

Let  $\mathcal{L}$  be a reverse tangent category. This means that  $\mathcal{L}$  has a tangent functor T giving tangent bundles  $p_A: \binom{TA}{A}$  of objects A and giving tangent arrows  $Tf: TA \to TA$  for arrows  $f: A \to B$ , and we can "reverse" T in order to get cotangent bundles  $p_A^*: \binom{T^*A}{A}$  and arrows in the pullback diagram below:



where  $T^*f$  is the diff. geometry formulation of the reverse differential of f.

Let  $\mathcal{L}$  be a reverse tangent category. This means that  $\mathcal{L}$  has a tangent functor T giving tangent bundles  $p_A: \binom{TA}{A}$  of objects A and giving tangent arrows  $Tf: TA \to TA$  for arrows  $f: A \to B$ , and we can "reverse" T in order to get cotangent bundles  $p_A^*: \binom{T^*A}{A}$  and arrows in the pullback diagram below:

$$T^*A \xleftarrow{T^*f} f^*T^*B \xrightarrow{\overline{f}} T^*B$$

$$\downarrow p_A^* \downarrow A \xrightarrow{f} B$$

where  $T^*f$  is the diff. geometry formulation of the reverse differential of f.

We have a functor  $\mathcal{L} \to \text{Lens}(\mathcal{L})$  defined by:

$$A \mapsto p_A^* : \binom{T^*A}{A}$$

$$A \xrightarrow{f} B \mapsto (f, T^*A \xleftarrow{T^*f} f^*T^*B \xrightarrow{\overline{f}} T^*B).$$

# Expressing Dialectica as a functor

 $\Lambda_{\rm cat} \to \mathcal{E} {
m Lens}({f P}_{\rm cat})$ 

# Expressing Dialectica as a functor

$$\Lambda_{\rm cat} \to \mathcal{E} \mathrm{Lens}(\mathbf{P}_{\rm cat})$$

- An object A is sent to the typed term  $\mathbf{z}: W(A) \times \mathcal{M}[C(A)] \vdash_{\mathbf{P}} \mathbf{z}^1 : W(A)$
- An arrow  $\mathbf{z}: A \vdash_{\Lambda} M : B$  in  $\Lambda_{\text{cat}}$  from A to B is sent to the arrow in  $\mathcal{E}\text{Lens}(\mathbf{P}_{\text{cat}})$  from  $\mathbf{z}: W(A) \times \mathcal{M}[C(A)] \vdash_{\mathbf{P}} \mathbf{z}^1 : W(A)$  to  $\mathbf{z}: W(B) \times \mathcal{M}[C(B)] \vdash_{\mathbf{P}} \mathbf{z}^1 : W(B)$  given by the following diagram:

$$W(A) \times \mathcal{M}[C(A)] \xleftarrow{\langle \mathbf{z}^1, (\mathbf{M}_{\mathbf{z}^1}) \mathbf{z}^2 \rangle} W(A) \times \mathcal{M}[C(B)] \xrightarrow{\langle \mathbf{M}^\bullet, \mathbf{z}^2 \rangle} W(B) \times \mathcal{M}[C(B)] \xrightarrow{\mathbf{z}^1} W(A) \xrightarrow{\mathbf{z}^1} W(B)$$

# Expressing Dialectica as a functor

 $\Lambda_{\rm cat} \to \mathcal{E} {\rm Lens}({\bf P}_{\rm cat})$ 

#### Moral:

The Dialectica transformation of  $\lambda$ -calculus encodes (reverse) Differentiation because it is a transformation into a category of Lenses, the latter being the abstract setting for Reverse Differentiation.

#### Final comments

- I didn't talk about Dialectica categories. I could have said something (ask me if you are interested)
- Explore categorical framework to reverse a Cartesian closed differential category in order to define Cartesian closed reverse differential/tangent categories
- Reverse differential  $\lambda$ -calculus? There is an interesting paper from Ong and Mak.

#### Final comments

- I didn't talk about Dialectica categories. I could have said something (ask me if you are interested)
- Explore categorical framework to reverse a Cartesian closed differential category in order to define Cartesian closed reverse differential/tangent categories
- Reverse differential  $\lambda$ -calculus? There is an interesting paper from Ong and Mak.

#### THANK YOU!