Sweedler Semantics and the propogation of uncertainty through Turing machines

Will Troiani

November 2023

A provoking question...

- Vector space semantics of linear logic:

$$!A \rightsquigarrow \mathrm{Sym}\, A$$

- Vector space semantics of *differential* linear logic:

$$!A \rightsquigarrow \bigoplus_{P \in A} \mathrm{Sym}\, A$$

Thesis of James Clift "Turing Machines and Differential Linear Logic", adapted Girard's translation of Turing machines into linear logic to *not* require quantifiers.

Question

What do these derivatives compute?

Coalgebraic geometry

Recall: given a finite dimensional $\mathbb{R}$-algebra A, the dual A^* is a *coalgebra*

$$\begin{aligned} A^* &\longrightarrow \mathbb{R} \\ \varphi &\longmapsto \varphi(1) \end{aligned} \qquad \begin{aligned} A^* &\longrightarrow (A \otimes A)^* \longrightarrow A^* \otimes A^* \\ \varphi &\longmapsto (a \otimes a' \mapsto \varphi(aa')) \end{aligned}$$

Given a vector space V, there is a *universal* coalgebra $!V$ and linear map $d : !V \longrightarrow V$, which are universal amongst $\mathbb{R}$-linear maps $\phi : C \longrightarrow V$ where C is a coalgebra:

$$\begin{tikzcd} & !V \arrow[d] \\ C \arrow[ur, dashed] \arrow[r] & V \end{tikzcd}$$

If V is finite dimensional, then

$$!V = \bigoplus_{P \in V} \operatorname{Sym} V, \quad |v_1, \ldots, v_n\rangle_P = v_1 \otimes \ldots \otimes v_n \in \operatorname{Sym}_P V$$

Syntactic derivative

For us, "linear logic" means "intuitionistic, multiplicative, exponential linear logic with &", and similarly for "differential linear logic" (though we omit & for the latter).

Definition

Given a proof π *of* $!A \vdash B$ *in linear logic, the derivative* $\frac{\partial}{\partial A}\pi$ *is the proof*

$$\dfrac{\dfrac{\begin{array}{c}\pi\\ \vdots\\ \Gamma, !A \vdash B\end{array}}{\Gamma, !A, !A \vdash B}\ \text{coctr}}{\Gamma, !A, A \vdash B}\ \text{coder}$$

whose denotation is the composite

$$[\![\Gamma]\!] \otimes ![\![A]\!] \otimes [\![A]\!] \xrightarrow{\mathrm{id} \otimes D} [\![\Gamma]\!] \otimes ![\![A]\!] \xrightarrow{[\![\pi]\!]} [\![B]\!]$$

Interpretation of the derivative

The map D_V, for a vector space V, is defined by

$$D_V(|v_1, \dots, v_s\rangle_P \otimes v) = |v, v_1, \dots, v_s\rangle_P$$

so in particular

$$D_V(|\varnothing\rangle_P \otimes v) = |v\rangle_P$$

The vector $|v\rangle_P$ is an example of a **primitive element** of the cofree coalgebra. It is well understood how these relate to coalgebraic derivatives.

$$\mathrm{Prim}!V \cong V \oplus V\epsilon$$

where we think of an element $P + v\epsilon$ as tangent vector

$$(-1, 1) \longrightarrow p + vt \in V$$

So the primitive elements of $!V$ are the tangent vectors of V.

Primitive elements

It can be shown that $[\![\psi]\!]$ maps primitive elements to primitive elements (ie, maps tangent vectors to tangent vectors). This gives us a way of understanding the derivative in the Sweedler semantics, at least in a special case, say $\psi : !A \vdash B$:

$$\left[\!\!\left[\frac{\partial}{\partial A}\psi\right]\!\!\right](|\varnothing\rangle_P \otimes v) = [\![\mathrm{prom}\,\psi]\!]\,|v\rangle_P \in \mathrm{Prim}!B$$

where $\mathrm{prom}\,\psi$ is given by appending a promotion rule to ψ.

$$\begin{tikzcd}
\bigotimes_{i=1}^{r} ![\![A_i]\!] \arrow[rr, "{[\![\mathrm{prom}]\!]}"] & & ![\![B]\!] \\
\mathrm{Prim}\left(\bigotimes_{i=1}^{r} ![\![A_i]\!]\right) \arrow[u, "\text{inclusion}"] \arrow[rr, "\text{Coalgebraic}", "\text{derivative}"'] & & \mathrm{Prim}![\![B]\!] \arrow[u, "\text{inclusion}"']
\end{tikzcd}$$

The coalgebraic derivative is completely understood.

Computational interpretation

So we understand these derivatives syntactically, and mathematically, but

Question

What do these derivatives mean computationally?

We let $\mathcal{P}_i$ a finite set of proofs of A_i, $\mathcal{Q}$ of B, assume $\{[\![\nu]\!]\}_{\nu\in\mathcal{Q}} \subseteq [\![B]\!]$ is linearly independent and $\left\{\pi(X_1,\ldots,X_r) | X_i \in \mathcal{P}_i^{n_i}, 1 \leq i \leq r\right\} \subseteq \mathcal{Q}$ and consider

$$\begin{array}{ccc} ![\![A_1]\!] \otimes \ldots \otimes ![\![A_r]\!] & \xrightarrow{[\![\psi]\!]} & [\![B]\!] \\ \uparrow{\scriptstyle \iota} & & \\ \mathbb{R}P_1 \times \ldots \times \mathbb{R}P_r & & \end{array}$$

where

$$\iota(\omega_1,\ldots,\omega_r) = \bigotimes_{i=1}^{r} |\varnothing\rangle_{[\![\omega_i]\!]}$$

A collection of explicit polynomials

Let $\mathcal{P}_i$ be a finite set of proofs of A_i, $\mathcal{Q}$ of B, assume $\{[\![\nu]\!]\}_{\nu\in\mathcal{Q}} \subseteq [\![B]\!]$ is linearly independent and

$$\left\{\pi(X_1,\ldots,X_r)|X_i \in \mathcal{P}_i^{n_i}, 1 \leq i \leq r\right\} \subseteq \mathcal{Q}$$

There exists a unique $\mathbb{R}$-algbera homomorphism F_ψ such that the following diagram commutes

$$\begin{tikzcd}
![\![A_1]\!] \otimes \ldots \otimes ![\![A_r]\!] \arrow[r, "[\![\psi]\!]"] & [\![B]\!] \\
\mathbb{R}\mathcal{P}_1 \times \ldots \times \mathbb{R}\mathcal{P}_r \arrow[u, "\iota"] \arrow[r, "F_\psi"] & \mathbb{R}\mathcal{Q} \arrow[u, "[\![-]\!]"']
\end{tikzcd}$$

where F_ψ is determined by a $\mathbb{R}$-algebra homomorphism $f_\psi : \mathbb{R}[\mathcal{Q}] \longrightarrow \mathbb{R}[\mathcal{P}_1 \sqcup \ldots \sqcup \mathcal{P}_r]$ in the sense that

$$F_\psi(\mathbf{a})_\tau = \mathrm{Eval}_{x^i_\rho = a^i_\rho}\, f_\psi(\tau) = f_\psi(\tau)(\mathbf{a}), \quad \forall \mathbf{a} \in \prod_i \mathbb{R}\mathcal{P}_i$$

An idea emerges...

Perhaps these linear combinations $\sum_i a_i \alpha_i$ of proofs $\alpha_i \vdash A$ are being encoded as a vector

$$|\varnothing\rangle_{\sum_i a_i \alpha_i} \in [\![!A]\!] \tag{1}$$

So in the special case where $a_i \geq 0$ and $\sum_i a_i = 1$, these vectors (1) can be interpreted as *distributions* over the proofs α_i of A.

Question

Are probability distributions sent to probability distributions?

Answer: not sure, but they definitely are sometimes...

A particular type of derivative

Recall that categorical semantics attempt to identify proofs up to natural permutation.

$$\dfrac{\dfrac{A, B \vdash C}{!A, B \vdash C}\,\text{der}}{!A \vdash B \multimap C}\,\text{R}\multimap \qquad = \qquad \dfrac{\dfrac{A, B \vdash C}{A \vdash B \multimap C}\,\text{R}\multimap}{!A \vdash B \multimap C}\,\text{der}$$

Assume $!A$ is a hypothesis of the proof introduced by a dereliction rule. Bring the dereliction down as far to the bottom of the proof as possible.

$$\dfrac{\dfrac{!A, A \vdash B}{!A, !A \vdash B}\,\text{der}}{!A \vdash B}\,\text{ctr} \qquad \dfrac{\dfrac{!\Gamma, A \vdash B}{!\Gamma, !A \vdash B}\,\text{der}}{!\Gamma, !A \vdash !B}\,\text{prom}$$

Plain proofs

Definition

A proof equivalent under cut-elimination to one of the form

$$\dfrac{\dfrac{\begin{array}{c}\pi\\ \vdots\\ n_1 A_1, \dots, n_r A_r \vdash B\end{array}}{n_1 !A_1, \dots, n_r !A_r \vdash B}\ \text{der}}{!A_1, \dots, !A_r \vdash B}\ \text{ctr / weak}$$

is plain.

Let ψ be plain, the papers of Murfet and Clift give a computational interpretation to $[\![\frac{\partial}{\partial A_i} \psi]\!]$.

Probabilistic semantics

Given a set Z, write

$$\Delta Z = \left\{ \sum_{z \in Z} \lambda_z z \in \mathbb{R}Z \middle| \sum_{z \in Z} \lambda_z = 1 \text{ and } \lambda_z \geq 0 \text{ for all } z \in Z \right\}$$

Proposition

$\exists!$ *function* $\Delta\psi$ *rendering the following diagram commutative*

$$\begin{tikzcd}
![\![A_1]\!] \otimes \ldots \otimes ![\![A_r]\!] \arrow[r, "[\![\psi]\!]"] & [\![B]\!] \\
\Delta\mathcal{P}_1 \times \ldots \times \Delta\mathcal{P}_r \arrow[u, "\iota"] \arrow[r, "\Delta\psi"] & \Delta\mathcal{Q} \arrow[u, "[\![-]\!]"']
\end{tikzcd}$$

Derivatives, primitive elements, and probability

The main Theorem of Murfet and Clift's papers is that the derivative of $\Delta\psi$ is calculated by the coalgebraic derivative, which we have already related to the syntactic derivative. For any point $\mathbf{w} \in \prod_{i=1}^{r} \mathcal{P}_i$ and proofs $\zeta, \rho \in \mathcal{P}_i$

$$\left[\!\left[\frac{\partial}{\partial X_i} \psi(\zeta, w_1, \dots, w_r) \right]\!\right] - \left[\!\left[\frac{\partial}{\partial X_i} \psi(\rho, w_1, \dots, w_r) \right]\!\right]$$
$$= [\![\psi]\!] \Big(|\varnothing\rangle_{[\![w_1]\!]} \otimes \dots \otimes |[\![\zeta]\!] - [\![\rho]\!]_{[\![w_i]\!]} \otimes \dots \otimes |\varnothing\rangle_{[\![w_r]\!]}\rangle \Big)$$

Theorem

$$[\![\psi]\!] \Big(|\varnothing\rangle_{[\![w_1]\!]} \otimes \dots \otimes |[\![\zeta]\!] - [\![\rho]\!]_{[\![w_i]\!]} \otimes \dots \otimes |\varnothing\rangle_{[\![w_r]\!]}\rangle \Big) = T_{\mathbf{w}}(\Delta\psi)(B_\rho^\zeta)$$

where $\{B_\rho^\zeta\}_{\rho \neq \zeta \in \mathcal{P}}$ is a particular choice of basis for $T_{\mathbf{w}}(\Delta\mathcal{P})$.

So, do we believe in plain proofs...?

In Clift's (impressive) masters thesis, he constructs an embedding of Turing machines into Linear Logic which is similar to Girard's embedding, but does *not* require quantifiers. This means that we can use the material constructed today to calculate the derivatives of these Turing machines *as long as they are interpreted as plain proofs*......... which they ARE!