Hypernormalisation, linear exponential monads and the Giry tricocycloid (extended abstract)

Richard Garner
Centre of Australian Category Theory, Macquarie University, Australia

Background  A basic construction in probability theory is that of normalising a sub-probability distribution of weight $\leq 1$ to a probability distribution of weight 1. The simplest case is that of finitely supported, discrete probability sub-distributions on a set $A$, i.e., finitely supported functions $\omega: A \to [0, 1]$ with $\omega(A) := \sum_{a \in A} \omega(a) \leq 1$. If $\omega(A) \neq 0$, then the normalisation $\bar{\omega}$ of $\omega$ is defined by $\bar{\omega}(a) = \omega(a)/\omega(A)$. This is, of course, a probability distribution, i.e., $\bar{\omega}(A) = 1$. But if $\omega(A) = 0$, then we cannot normalise $\omega$; so normalisation is only a partial operation. In [2], Jacobs introduces hypernormalisation which, among other things, addresses this defect.

Hypernormalisation is a total function $N: D(A_1 + \cdots + A_n) \to D(DA_1 + \cdots + DA_n)$ where $D(X)$ will denote the set of finitely supported probability distributions on $X$. To define $N$ at $\omega \in D(A_1 + \cdots + A_n)$, we first restrict $\omega$ along the $n$ coproduct injections to get sub-distributions $\omega_i$ on $A_i$; we then select the non-zero sub-distributions among these, say $\omega_{i_1}, \ldots, \omega_{i_m}$; finally, we define $N(\omega)$ to take the value $\omega_{i_k}(A_{i_k})$ at the element $\bar{\omega}_{i_k}$ in the $DA_{i_k}$-summand of $DA_1 + \cdots + DA_n$, and to be zero elsewhere. So $N(\omega)$ “normalises the non-zero distributions among $\omega_1, \ldots, \omega_n$ and records the weights”.

In [1], I establish links between hypernormalisation, and structures arising in monoidal category theory, linear logic and quantum algebra—as I will now explain.

Convex coproducts  The assignation $X \mapsto DX$ underlies the finite Giry monad $D$ on the category of sets, whose algebras are convex spaces. A (abstract) convex space is a set $A$ with with a “convex combination” operation $(0, 1) \times A \times A \to A$, which we write as $r, a, b \mapsto r(a, b)$ or $r, a, b \mapsto r \cdot a + r^* \cdot b$, where $r^* := 1 - r$. The axioms are that $r(a,a) = a$, $r(a, b) = r^*(b, a)$ and $r(s(a, b), c) = (rs)(a, (r/s)^*\cdot (b, c))$ for $a, b, c \in A$ and $r, s \in (0, 1)$.

The first recasting of hypernormalisation is in terms of coproducts in the category $\textbf{Conv}$ of convex spaces. These are unusually simple; the binary coproduct is:

$$A \star B = A + (0, 1) \times A \times B + B$$  \hspace{1cm} (1)

with a suitable convex structure. The outer summands give the coproduct inclusions $\iota_1: A \to A \star B \leftarrow B: \iota_2$, and the middle summand gives elements of the form $r \cdot a + r^* \cdot b$.
Now the free functor $\textbf{Set} \to \textbf{Conv}$ sends a set $A$ to $DA$ with the convex structure induced pointwise from $[0, 1]$. Being a left adjoint, $F$ preserves coproducts, and so we have an isomorphism

$$\varphi: D(A + B) \cong DA \star DB$$

of convex spaces. Working through the definitions, we see that $\varphi$ is very close to being (binary) hypernormalisation:

$$\varphi(\omega) = \begin{cases} 
\iota_1(\omega|_A) & \text{if } \omega(A) = 1; \\
\iota_2(\omega|_B) & \text{if } \omega(B) = 1; \\
\omega(A) \cdot \omega|_A + \omega(B) \cdot \omega|_B & \text{otherwise.}
\end{cases}$$

Recapturing $N$ Nice as it is, this map $\varphi$ is not quite hypernormalisation. How do we close the gap? Since hypernormalisation $D(A + B) \to D(DA + DB)$ fails to be a map of convex spaces, we must for this go outside the category $\textbf{Conv}$ of convex spaces, and we do so in a seemingly simple-minded manner, by passing to the category $\textbf{Conv}_{arb}$ of convex spaces and arbitrary maps.

The key point is that the coproduct monoidal structure $(\star, 0)$ on $\textbf{Conv}$ extends to a monoidal structure on $\textbf{Conv}_{arb}$. On objects this is (necessarily) defined as before; while the tensor of maps in $\textbf{Conv}_{arb}$ is given by $f \star g = f + ((0, 1) \times f \times g) + g$, i.e., exactly the same formula as in $\textbf{Conv}$.

Using this tensor, we obtain for any convex spaces $A$ and $B$ a map in $\textbf{Conv}_{arb}$:

$$A \star B \xrightarrow{\eta_A \star \eta_B} DA \star DB \xrightarrow{\varphi^{-1}} D(A + B)$$

where $\eta_X: X \to D(X)$, the unit of the finite Giry monad, sends $x \in X$ to the Dirac distribution at $x$. Working through the definitions, the displayed composite sends elements $\iota_1(a)$ and $\iota_2(b)$ of $A \star B$ to the Dirac distributions on $A + B$ concentrated at $a$, respectively $b$; while an element $r \cdot a + r^* \cdot b$ of $A \star B$ is sent to the two-point distribution with weight $r$ at $a$ and weight $r^*$ at $b$. Combined with our description of $\varphi$, this shows that $N$ is the composite:

$$D(A + B) \xrightarrow{N} D(DA + DB) \xrightarrow{\varphi} DA \star DB \xrightarrow{\eta_A \star \eta_B} DDA \star DDB.$$

Linear exponential monads This re-derivation of hypernormalisation leaves one question unanswered: why should there be an extension of the coproduct monoidal structure on $\textbf{Conv}$ to $\textbf{Conv}_{arb}$? A moment’s thought shows the fundamental reason to be that the underlying set of $A \star B$ depends only on the underlying sets of $A$ and $B$, and not on their convex space structure.

This suggests that the symmetric monoidal structure on $\textbf{Conv}$ could be a lifting of one on $\textbf{Set}$; i.e., that $\textbf{Set}$ could have a symmetric monoidal structure $(\star, 0)$ making $U: (\textbf{Conv}, \star) \to (\textbf{Set}, \star)$ strict symmetric monoidal. Were this so, then we could re-find the monoidal structure on $\textbf{Conv}_{arb}$ by factorising $U$ as (bijective on objects, fully faithful) in the category of symmetric monoidal categories.

In fact, this is what happens; we describe the relevant monoidal structure on $\textbf{Set}$—the Giry monoidal structure—below. However, first we note that this monoidal structure’s lifting to $\textbf{Conv}$ is really struc-
ture on the monad $D$: it says that it is a linear exponential monad.

A linear exponential monad $T$ on a symmetric monoidal category $(C, \otimes, I)$ is a monad for which $(\otimes, I)$ lifts to $T\text{-}\text{Alg}$, and there becomes finite coproduct. Such monads interpret the connective $\otimes$ (“why not?”) of linear logic. In fact, they also interpret abstract hypernormalisation. Indeed, if $C$ has finite sums, then we get invertible maps (“Seely isomorphisms”) $\varphi: TA + TB \to TA \otimes TB$ from the fact that $TA \otimes TB$ is a coproduct of free $T$-algebras $TA$ and $TB$. Mimicking (2), we get hypernormalisation maps $N: T(A + B) \to T(TA + TB)$ by taking $N = \varphi^{-1} \circ (\eta_{TA} \otimes \eta_{TB}) \circ \varphi$.

These generalise precisely the maps $N$ of the motivating case, and I show in [1] that many pleasant algebraic properties of that case carry over to the general one.

The Giry tricocycloid  We now construct the Giry monoidal structure on $\text{Set}$. Remarkably, a construction from quantum algebra provides just what is needed.

An abelian tricocycloid [4] in a symmetric monoidal category $C$ comprises an object $H$; an isomorphism $v: H \otimes H \to H \otimes H$ satisfying $(v \otimes 1)(1 \otimes v)(v \otimes 1) = (1 \otimes v)(v \otimes 1)(1 \otimes v)$; and an involution $\gamma: H \to H$ satisfying $(1 \otimes \gamma)v(1 \otimes \gamma) = v(1 \otimes 1)v$. If $C$ has finite coproducts distributing over $\otimes$, then $(H, v, \gamma)$ induces a symmetric monoidal structure on $C$, with unit $0$ and binary tensor

$$A \star B = A + H \otimes A \otimes B + B.$$  (3)

The maps $v$ and $\gamma$ appear in the associativity and symmetry constraints respectively.

Comparing (1) with (3) suggests instantiating this in $\text{Set}$ with $H = (0, 1)$. Indeed, defining $v$ by $v(r, s) = (rs, r^*s^*)$—the terms appearing the third convex space axiom—and $\gamma$ by $\gamma(r) = r^*$ yields an abelian tricocycloid, whose induced monoidal structure is the Giry one.

Other examples  In [1] I examine the force of hypernormalisation for a range of linear exponential monads. In particular, I consider the expectation monad [3] on $\text{Set}$, involving involves finitely additive rather than finitely supported measures. This is linear exponential for the Giry monoidal structure; in fact, I conjecture that the expectation monad is terminal among such linear exponential monads.

References


