ACCEPTING NORMALIZATION

MARIO ROMÁN
j.w.w. Elena Di Lavore, Paweł Sobociński, Márk Széles

Topos Oxford Seminar

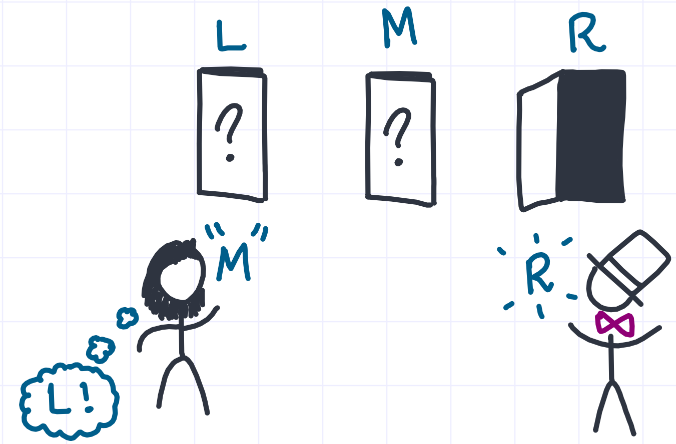
Monty Hall Problem

$$\overset{(i)}{\rightsquigarrow} \quad \tfrac{1}{3}|L\rangle + \tfrac{1}{3}|M\rangle + \tfrac{1}{3}|R\rangle$$

$$\overset{(ii)}{\rightsquigarrow} \quad \tfrac{1}{3}|LR\rangle + \tfrac{1}{6}|ML\rangle + \tfrac{1}{6}|MR\rangle + \tfrac{1}{3}|RL\rangle$$

$$\overset{(iii)}{\rightsquigarrow} \quad \tfrac{1}{3}\cancel{|LR\rangle} + \tfrac{1}{6}|ML\rangle + \tfrac{1}{6}\cancel{|MR\rangle} + \tfrac{1}{3}|RL\rangle$$

$$\overset{(iv)}{\rightsquigarrow} \quad \tfrac{1}{3}|ML\rangle + \tfrac{2}{3}|RL\rangle$$



- Subdistributions appear when updating.
- Normalization is implicit at the end.
- Can we do without subdistributions?

The Quest for Normalized Kernel Composition

(1) Work up-to-scalar.

$$\lambda \cdot d_1(x) = d_2(x);$$
$$\lambda \cdot f_1(x;y) = f_2(x;y);$$

(2) Work up-to-parameterized-scalar.

$$\lambda(x) \cdot f_1(x;y) = f_2(x;y);$$

(3) Work up-to-validity.

$$dist? : DMX \to MDX$$

$$dist?(d) = \begin{cases} \perp, & \text{if } d(\perp) > 0, \\ d, & \text{otherwise.} \end{cases}$$

(1) Quotienting preserved by composition. Normalizes states, but not kernels.

e.g. $f(x_1) = \frac{1}{3}|y_1\rangle$

$f(x_2) = \frac{1}{4}|y_1\rangle + \frac{1}{4}|y_2\rangle$

(2) Normalizes kernels, but it is not preserved by composition.

(3) Distributive law. Black-hole semantics instead of updating semantics.

Accepting Normalization

Def. Distribution monad, with Kleisli category Stoch.

$$DX = \left\{ d : X \xrightarrow{f.s.} [0,1] \;\middle|\; \textstyle\sum_x d(x) = 1 \right\}.$$

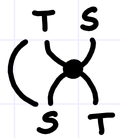
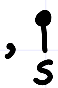
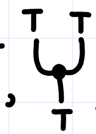
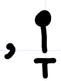
Def. Maybe monad, with Kleisli category Par.

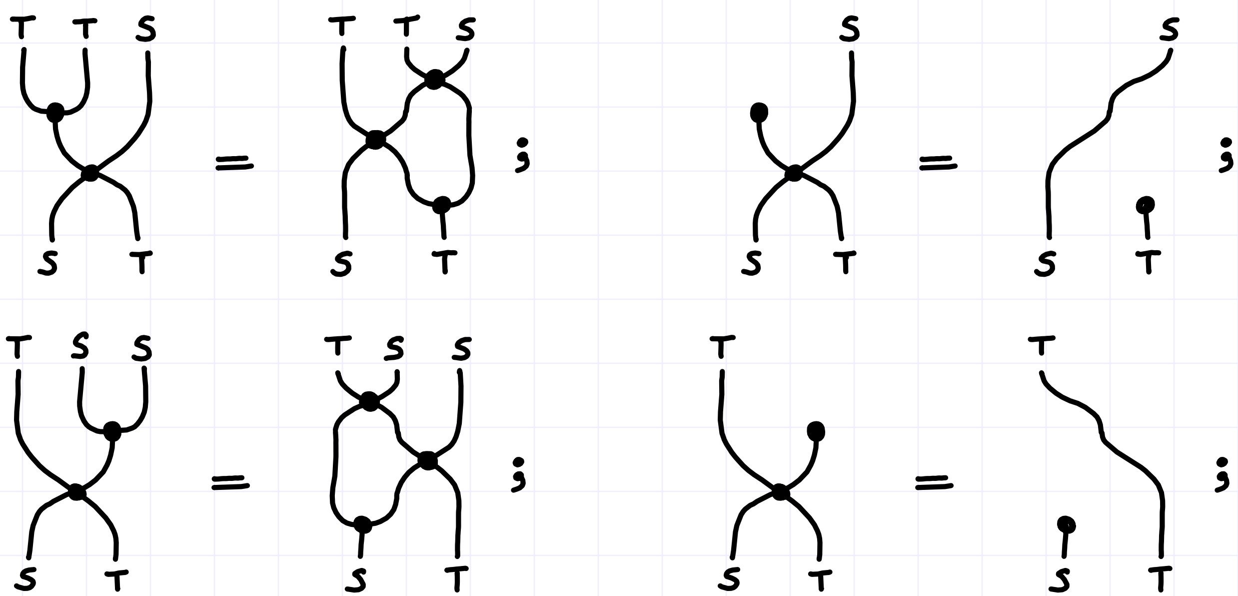
$$MX = X + \{\perp\}.$$

Def. Normalization, distributive law candidate. $n_X : DMX \to MDX$.

$$n(f)(x) = \frac{f(x)}{\sum_{x'} f(x')} ; \quad \text{with } n(f) = \perp \text{ when } \textstyle\sum_{x'} f(x') = 0.$$

Distributive Law

( : $TS \to ST$) between (S, , ) and (T, , ), satisfying



Normalization is not Associative

Prop. Normalized channels do not have an associative composition.

Proof. Indeed, by their formula, left associated composition, $(f ⨾ g) ⨾ h$, gives the 'right' result.

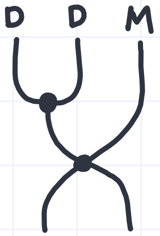
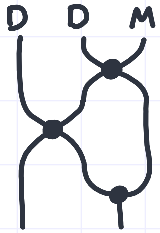
$$\frac{\Sigma z. (f ⨾ g)(x;z) \cdot h(z;w)}{\Sigma zw. (f ⨾ g)(x;z) \cdot h(z;w)} = \frac{\Sigma z. \frac{\Sigma y. f(x;y) \cdot g(y;z)}{\Sigma yz. f(x;y) \cdot g(y;z)} \cdot h(z;w)}{\Sigma zw. \frac{\Sigma y. f(x;y) \cdot g(y;z)}{\Sigma yz. f(x;y) \cdot g(y;z)} \cdot h(z;w)} = \frac{\Sigma yz. f(x;y) \cdot g(y;z) \cdot h(z;w)}{\Sigma yzw. f(x;y) \cdot g(y;z) \cdot h(z;w)}$$

While right associated composition, $f ⨾ (g ⨾ h)$, does not.

$$(f ⨾ (g ⨾ h))(x;w) = \frac{\Sigma y. f(x;y) \cdot (g ⨾ h)(y;w)}{\Sigma y,w. f(x;y) \cdot (g ⨾ h)(y;w)} = \frac{\Sigma y. f(x;y) \cdot \frac{\Sigma z. g(y,z) \cdot h(z,w)}{\Sigma zw. g(y,z) \cdot h(z,w)}}{\Sigma yw. f(x;y) \cdot \frac{\Sigma z. g(y,z) \cdot h(z,w)}{\Sigma zw. (y,z) \cdot h(z,w)}}$$

Normalization is not a Distributive Law

PROP. Normalization fails D-multiplicativity.


$\neq$

;

$$\begin{tikzcd} DDMX \arrow[r, "\mu"] \arrow[d, "\eta"'] & DMX \arrow[dr, "\eta"] & \\ DMDX \arrow[r, "D\eta"] & MDDX \arrow[r, "M\mu"] & MDX \end{tikzcd} \;;$$

PROOF. With $X = \{x, y\}$, insert $\tfrac{1}{2}\big|\tfrac{1}{3}|x\rangle + \tfrac{2}{3}|\bot\rangle\big\rangle + \tfrac{1}{2}\big|\tfrac{1}{3}|y\rangle + \tfrac{2}{3}|\bot\rangle\big\rangle \in DDMA$.

$$\tfrac{1}{2}\big|\tfrac{1}{3}|x\rangle + \tfrac{2}{3}|\bot\rangle\big\rangle + \tfrac{1}{2}\big|\tfrac{2}{3}|y\rangle + \tfrac{1}{3}|\bot\rangle\big\rangle \in DDMX$$

$\xrightarrow{D\eta}$

$$\tfrac{1}{2}\big||x\rangle\big\rangle + \tfrac{1}{2}\big||y\rangle\big\rangle + |\bot\rangle \in DMDX$$

$\eta \downarrow$

$$\tfrac{1}{2}\big||x\rangle\big\rangle + \tfrac{1}{2}\big||y\rangle\big\rangle \in MDDX$$

$\mu \downarrow$

$$\tfrac{1}{2}|x\rangle + \tfrac{1}{2}|y\rangle \in MDX$$

$\xrightarrow{\mu}$

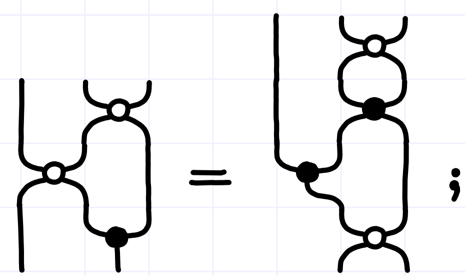
$$\tfrac{1}{6}|x\rangle + \tfrac{1}{3}|y\rangle + \tfrac{1}{2}|\bot\rangle \in DMX$$

$\downarrow \eta$

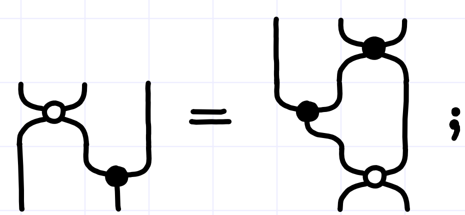
$$\tfrac{1}{3}|x\rangle + \tfrac{2}{3}|y\rangle \in MDX$$

Distributive Sesquilaw

DEF. A *distributive sesquilaw* between two monads, $(\times_\bullet, \times_\circ, S, T)$, consists of a distributive law $(\times_\bullet) : ST \to TS$ and an almost distributive law, $(\times_\circ) : TS \to ST$, such that it is its right inverse, $(\times_\bullet) ; (\times_\circ) = |\,|$, and satisfies multiplicativity up to its idempotent.



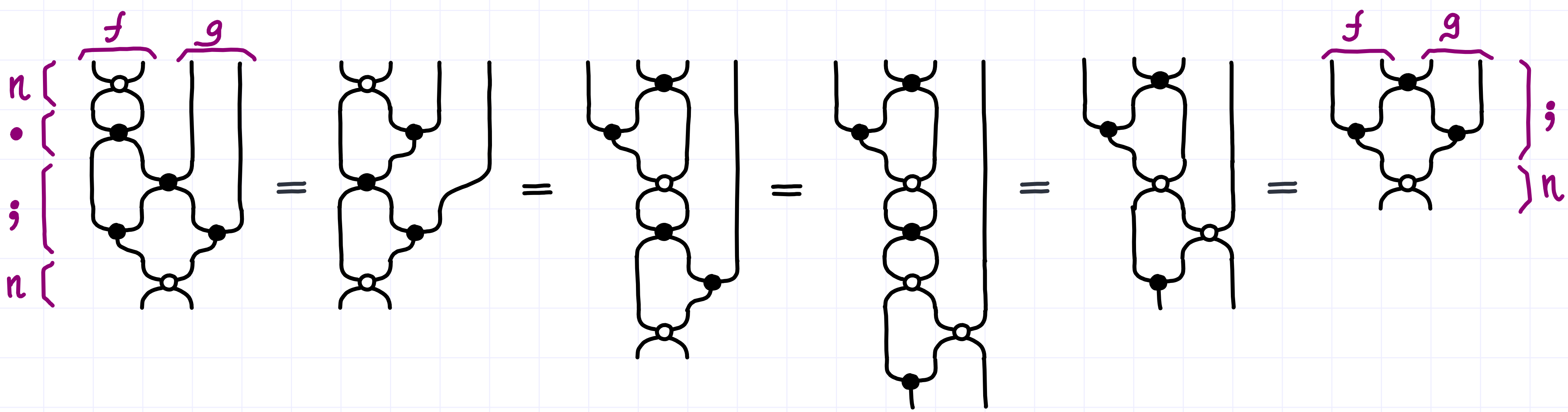
The equation may be equivalently written as follows.



Renormalization

Prop. Normalization is left absorptive, $n(n(f)\,;g) = n(f\,;g)$.

Proof. In general, for any distributive sesquilaw; we reason by string diagrams.

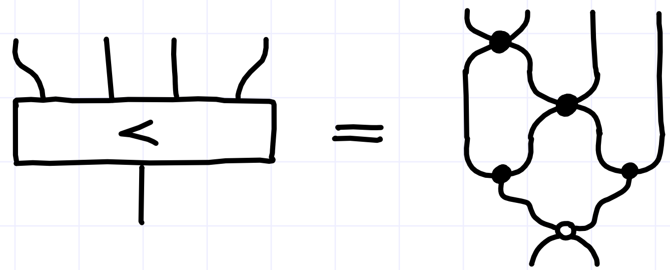


Updates Act on Priors

Prop. Substochastic kernels act on normalized kernels:

$$(<) : \mathrm{Norm}(X ; Y) \times \mathrm{Subd}(Y ; Z) \longrightarrow \mathrm{Norm}(X ; Z),$$

defined by $p < f = n(p^{\bullet} ; f)$, satisfies $p < (f ; g) = (p < f) < g$ and $p < id = p$.



;

THE QUEST FOR STOCHASTIC RELATIONS.

NON-EMPTY RELATIONS. Kleisli of the non-empty powerset monad, N.

MAY-MUST RELATIONS. Morphisms are relations, $R : X \to NMY$.

$$(R;S)(x;z) = \exists y \in Y.\ R(x;y) \wedge S(y;z)$$
$$(R;S)(x;\perp) = R(x;\perp) \vee (\exists y \in Y.\ R(x;y) \wedge S(y;\perp)).$$

DIJKSTRA RELATIONS. Morphisms are relations, $R : X \to MNY$.

$$(R * S)(x;z) = \exists y \in Y.\ R(x;y) \wedge S(y;z),$$

but failing whenever $\exists y \in Y.\ f(x;y) \wedge g(y;\perp)$.