universität innsbruck

Partializations of Markov categories

A framework for partial stochastic maps

Areeb Shah Mohammed

Nondeterminism

Consider (finite) sets and relations: A point of the domain can have many "possible images" in the codomain — possibly even none!

Standard approach to compose relations $X \xrightarrow{R_1} Y \xrightarrow{R_2} Z$ is "go big": $x \sim z$ if for some y, $x \sim_1 y \sim_2 z$.

Consider a process that may produce one of many possible outputs, can be seen as a multivalued map defined on the points which have at least one output.

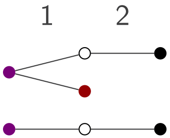
Question

What happens when it doesn't shutdown "gracefully"?

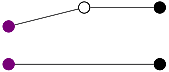
"Risk averse" composition: add condition that if $x \sim_1 y$, then $y \sim_2 z'$ for some z'.

An illustration

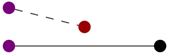
Two steps



Go Big



Play it safe



Expectations

Expectations on a compact interval $[a, b] \subseteq \mathbb{R}$ define an algebra for the distribution monad $\mathbf{E}[\_]\colon P[a, b] \to [a, b]$.

This can be extended to compact convex sets [*Świrszcz*].

However, expectations do not form an algebra $P\mathbb{R} \to \mathbb{R}$ on $\mathbb{R}$, as they are not defined over all distributions on $\mathbb{R}$.

They do define a *partial* map $P\mathbb{R} \rightharpoonup \mathbb{R}$ on the measurable subset $D \subset \mathbb{R}$ where defined, suggesting a "partial algebra" $P\mathbb{R} \rightharpoonup \mathbb{R}$.

This would even be "deterministic"!

Why not “failure” values?

Failed processes often represented by a “null” output $\perp$.

In a probabilistic setting, this has been successfully emulated using sub-stochastic distributions [*Di Lavore–Román(–Sobociński), Lorenz–Tull*].

However these follow the “go big” style of composition — the “possibilistic” analogue of sub-stochastic distributions is the category of relations `Rel`.

Involve a “probability of definition” for each point, not just a domain.

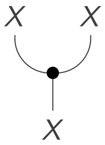
Also awkward to preserve useful properties like linearity of expectation.

Categories of stochastic maps

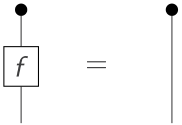
Definition (*Cho–Jacobs*)

CD categories: Symmetric monoidal categories with “copy and delete” commutative comonoid structures on each object that are compatible with tensoring.

copy_X =



del_X =



Definition (*Cho–Jacobs, Fritz*)

Total maps: commute with deletion.

Markov categories: all maps are total.

Recurring Examples

Examples (*Fritz*)

Various notions of “measurable spaces and stochastic maps”.

(i) `FinStoch`: finite sets and stochastic maps — here stochastic matrices;

(ii) `Dist`: sets and finitely supported distributions;

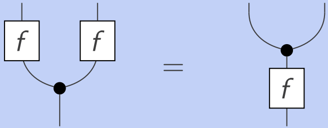
(iii) `SetMulti`: sets and multi-valued maps — possibilities rather than probabilities;

(iv) `BorelStoch`: standard Borel spaces and stochastic maps — Markov kernels.

Determinism

Definition (*Carboni–Walters*)

Copyable maps: commute with copying.



Wide subcategory $\mathtt{C}_{\mathrm{cop}}$ of copyable maps.

Deterministic maps: copyable and total.

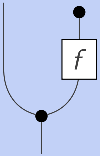
Warning

A CD category is Cartesian monoidal if and only if every map is deterministic [*Fox*].

Domains

Definition

Domain $\mathrm{dom}(f)$ of a map $f\colon X \to Y$: the endomorphism on X.



In the case of relations, the domain is the set of points that have at least one image.

$$\{(x,x) \ : \ \exists y \in Y, f(x,y)\}$$

Quasi-total maps: absorb domain, $f\,\mathrm{dom}(f) = f$ [*Di Lavore–Román*].

Quasi-Markov categories: all maps quasi-total.

Theorem (*Di Lavore–Román*)

Given positivity, quasi-totality is equivalent to $\mathrm{dom}(f)$ *being copyable.*

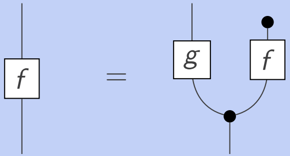
Poset enrichment

Theorem

The domain idempotents $\mathrm{dom}(f)$ *turn a positive quasi-Markov category into a restriction category.*

Corollary (*Cockett–Lack*)

Restriction category poset enrichment: $f \leq g \iff f = g\,\mathrm{dom}(f)$.



Partializable Markov categories

Definition

Partializable Markov categories:

(i) Positive;

(ii) Deterministic monomorphisms closed under:

    (a) pullback;

    (b) tensor.

Partialization

(finally!)

Definition

Partialization $\mathtt{Partial(C)}$

(i) Objects those of the original category $\mathtt{C}$;

(ii) Maps $X \to Y$ equivalence classes of spans

$$X \xleftarrow{i} D \xrightarrow{f} Y$$

with i a *deterministic* monomorphism;

Composition and tensor

Definition

(iii) Composition by pullback: For maps represented by spans $X \overset{i}{\hookleftarrow} D_f \xrightarrow{f} Y$ and $Y \overset{j}{\hookleftarrow} D_g \xrightarrow{g} Z$, the composite is represented by

$$X \overset{i}{\hookleftarrow} f^{-1}D_g \xrightarrow{gf} Z$$

(iv) Tensoring componentwise: for maps $X \overset{i}{\hookleftarrow} D_f \xrightarrow{f} Y$ and $X' \overset{j}{\hookleftarrow} D_g \xrightarrow{g} Y'$,

$$X \otimes X' \overset{i \otimes j}{\hookleftarrow} D_f \otimes D_g \xrightarrow{f \otimes g} Y \otimes Y'$$

(v) CD structure: inclusion of that of $\mathtt{C}$.

Composition in practice

Consider the composite of two maps $X \overset{i}{\hookleftarrow} D_f \overset{f}{\to} Y$ and $Y \overset{j}{\hookleftarrow} D_g \overset{g}{\to} Z$ in $\mathtt{Partial}(\mathtt{C})$. The domain of the composite is the pullback

$$\begin{array}{ccc} D & \xrightarrow{f|_T} & D_g \\ \big\downarrow & \lrcorner & \big\downarrow \\ D_f & \xrightarrow{f} & Y \end{array}$$

In $\mathtt{SetMulti}$

D is the $x \in D_f$ such that *all* images $f(x)$ belong to D_g.

$\mathtt{Dist}$

$$D = \{x \in D_f \ : \ \mathrm{Supp}(f(\_ \mid x)) \subseteq D_g\}$$

$\mathtt{BorelStoch}$

$$D = \{x \in D_f \ : \ f(D_g \mid x) = 1\}$$

Partializations are quasi-Markov

Theorem

(i) $\mathtt{C}$ *is the subcategory of total maps in* $\mathtt{Partial(C)}$ *([Cockett–Lack] and a little work);*

(ii) $\mathtt{Partial(C)}$ *is quasi-Markov;*

(iii) *The copyable maps of* $\mathtt{Partial(C)}$ *are* $X \xleftarrow{i} D \xrightarrow{f} Y$ *with* f *deterministic;*

(iv) $\mathtt{Partial(C)}$ *is positive;*

(v) *Given Kolmogorov products in* $\mathtt{C}$*, their inclusions into* $\mathtt{Partial(C)}$ *define Kolmogorov products.*

Warning

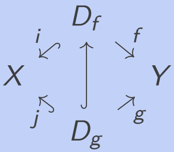
The usual notion of Kolmogorov product is no longer functorial in general.

Domains

Theorem (Consequence of [*Cockett–Lack*])

(i) $\texttt{Partial}(\mathtt{C})$ *is a split restriction category with the domain of a map* $X \overset{i}{\hookleftarrow} D \overset{f}{\to} Y$ *represented by the span* $X \overset{i}{\hookleftarrow} D \overset{i}{\hookrightarrow} X$.

(ii) *The restriction partial order* $(X \overset{i}{\hookleftarrow} D_f \overset{f}{\to} Y) \leq (X \overset{j}{\hookleftarrow} D_g \overset{g}{\to} Y)$ *on the hom-sets of* $\texttt{Partial}(\mathtt{C})$ *is equivalent to the existence of a factorization*



Representability

Definition (extends [*Fritz–Gonda–Perrone–Rischel*])

Representable quasi-Markov category: the inclusion $\mathtt{C}_{\mathsf{cop}} \hookrightarrow \mathtt{C}$ has a right adjoint $P\colon \mathtt{C} \to \mathtt{C}_{\mathsf{cop}}$, called the **distribution functor**.

$$\mathtt{C}(A, X) \cong \mathtt{C}_{\mathsf{cop}}(A, PX)$$

Denote

(i) The counit by $\mathtt{samp}_Y\colon PY \twoheadrightarrow Y$;

(ii) The copyable counterpart of a $f\colon X \to Y$ by $f^\sharp\colon X \to PY$.

Then,

$$f = \mathtt{samp}\, f^\sharp$$

Partialization and representability

Theorem

Consider a representable partializable Markov category $\mathtt{C}$. The sampling maps of $\mathtt{C}$ are also sampling maps for $\mathtt{Partial}(\mathtt{C})$

$$\mathtt{Partial}(\mathtt{C})_{\mathsf{cop}}(\_, PY) \xrightarrow{\mathtt{samp}_*} \mathtt{Partial}(\mathtt{C})(\_, Y)$$

Consequently, $\mathtt{Partial}(\mathtt{C})$ is representable.

Theorem

The copyable counterpart of a map $X \overset{i}{\hookleftarrow} D \xrightarrow{f} Y$ of $\mathtt{Partial}(\mathtt{C})$ is $X \overset{i}{\hookleftarrow} D \xrightarrow{f^\sharp} PY$.

The "pushforward" of a map $X \overset{i}{\hookleftarrow} D \xrightarrow{f} Y$ of $\mathtt{Partial}(\mathtt{C})$ is $PX \overset{Pi}{\hookleftarrow} PD \xrightarrow{Pf} PY$.

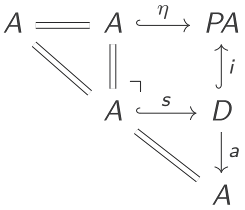
Partial algebras for the distribution monad

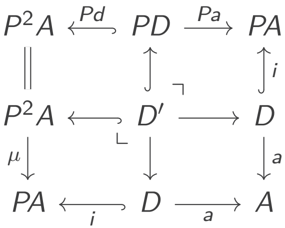
Consider a representable partializable Markov category $\mathtt{C}$.

Definition

Partial algebra: an algebra for the induced distribution monad P on $\mathtt{Partial(C)}_{\text{cop}}$.

A partial map $PA \to A$, represented in $\mathtt{C}_{\text{det}}$ by a span $PA \overset{i}{\hookleftarrow} D \overset{a}{\to} A$ such that





Expectation as a partial algebra

Theorem

In $\mathtt{Partial(BorelStoch)}$*, the expectation map gives* $\mathbb{R}_{\geq 0}$ *the structure of a partial algebra* $P\mathbb{R}_{\geq 0} \hookleftarrow D \xrightarrow{\mathbf{E}[\_]} \mathbb{R}_{\geq 0}$

$$\mathbf{E}[p] \coloneqq \int_{\mathbb{R}_{\geq 0}} x \, p(dx) \qquad\qquad D \coloneqq \{p \in P\mathbb{R}_{\geq 0} \ : \ \mathbf{E}[p] < \infty\}$$

Warning

The same expectation map does not make all of $\mathbb{R}$ a partial algebra.

Conditioning partial maps

Theorem

Consider a partializable Markov category $\mathtt{C}$ *with conditionals.*

Given a map $\varphi \colon A \to X \otimes Y$ *in* $\mathtt{Partial}(\mathtt{C})$ *represented by a span*

$$A \xleftarrow{i} D \xrightarrow{f} X \otimes Y$$

the conditional $\varphi_{|X} \colon X \otimes A \to Y$ *exists and is represented by the span*

$$X \otimes A \xleftarrow{X \otimes i} X \otimes D \xrightarrow{f_{|X}} Y$$

In particular, $\mathtt{Partial}(\mathtt{C})$ *has conditionals.*

Idempotent partial maps

Theorem

The idempotent partial maps are those that act as idempotents on their domain.

Explicitly, $\varepsilon = X \overset{i}{\hookleftarrow} D \xrightarrow{f} X$ *is idempotent if and only if it is* $X \overset{i}{\hookleftarrow} D \xrightarrow{ie} X$ *for an idempotent* e *of* D *in* $\mathtt{C}$.

Theorem

The idempotent ε *splits if and only if* e *does.*

For instance, every idempotent in $\mathtt{Partial}(\mathtt{BorelStoch})$ splits.

Theorem

The idempotent ε *is static/strong/balanced if and only if* e *is.*

Thank you for your attention!

Areeb Shah Mohammed

Subprobability measures

A notion of partiality — total probability intuitively the "probability of definition/existence".

Corresponding categories [*Di Lavore–Román, Lorenz–Tull*]

$\mathtt{Rel}$: sets and relations — hom-sets same as $\mathtt{Partial}(\mathtt{SetMulti})$;

$\mathtt{Kl}(D_{\leq 1})$: sets and subprobability measures — maps of $\mathtt{Partial}(\mathtt{Dist})$ are those that have probability 1 or 0;

$\mathtt{BorelStoch}_{\leq 1}$: standard Borel spaces and subprobability measures.

Problem

The sub-distribution composition law produces intermediate probabilities of definition. In particular quasi-totality is not preserved by sub-distribution composition.

Lax partial algebras

The analogous span $P\mathbb{R} \hookleftarrow D \xrightarrow{\mathbf{E}[\_]} \mathbb{R}$ defines a map of $\mathtt{Partial}(\mathsf{BorelMeas})$, and even satisfies the unit triangle condition.
But the multiplication square only commutes up to restriction of domain!

Defined on $\pi \in PD$ with

$$\int_p \int_x |x| \, p(dx) \pi(dp) < \infty$$

$$\begin{array}{ccccc} & & PPX & & \\ & \swarrow^{\mu} & & \searrow^{P\mathbf{E}[\_]} & \\ PX & & \leq & & PX \\ & \searrow_{\mathbf{E}[\_]} & & \swarrow_{\mathbf{E}[\_]} & \\ & & X & & \end{array}$$

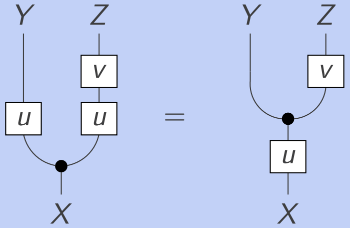
Defined on $\pi \in PD$ with

$$\int_p \left| \int_x x \, p(dx) \right| \pi(dp) < \infty$$

Positivity

Definition (extends [*Fritz*])

Positive CD categories: For every composable pair $X \xrightarrow{u} Y \xrightarrow{v} Z$ whose composite $v\,u \colon X \to Y$ is copyable,

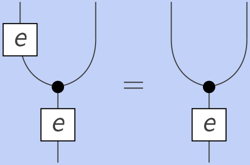


Idempotents

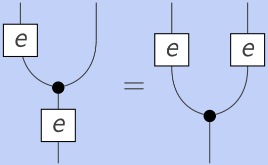
Definition (*Fritz–Gonda–Lorenzin–Perrone–Stein*)

An idempotent $e\colon X \to X$ in a quasi-Markov category C is:

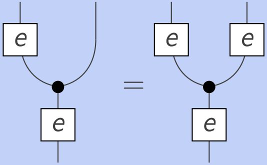
static if



strong if



balanced if

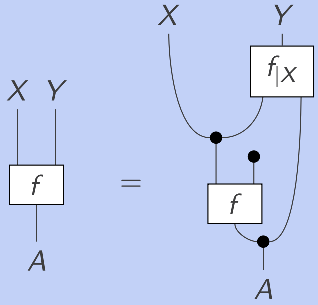


Copyable idempotents satisfy all three conditions.

Conditionals

Definition (*Cho–Jacobs, Fritz*)

Conditional of $f\colon A \to X \otimes Y$ with respect to an output X is an $f_{|X}\colon X \otimes A \to Y$ such that



As an equation,

$$f(x, y \mid a) = f(x \mid a) f_{|X}(y \mid x, a)$$

Special case of a state $p\colon I \to X \otimes Y$,

$$p(x, y) = p(x) p_{|X}(y \mid x)$$

(Strict) Kolmogorov products

Kolmogorov's extension theorem

A joint distribution on a family of random variables is uniquely characterized by a compatible family of "finite marginals".

Definition (*Fritz–Rischel*)

For a family of objects $(X_k)_{k \in K}$ let $\mathtt{FinSub}(K)$ be the poset of finite subsets of K and inclusions. This defines a diagram

$$X^{(-)} \colon \mathtt{FinSub}(K)^{\mathrm{op}} \to \mathtt{C} \qquad\qquad F \mapsto X^F := \bigotimes_{i \in F} X_i$$

Strict Kolmogorov product: A limit cone $(X^K \xrightarrow{\pi_F} X^F)_{F \subseteq K \text{ finite}}$ with deterministic legs preserved by tensoring with an arbitrary object Y.

Issues with strict Kolmogorov products

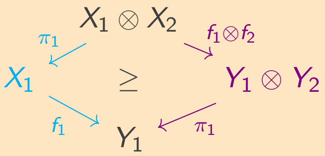
In the *Markov* case, a family $\left(X_k \xrightarrow{f_k} Y_k\right)_{k \in K}$ induces a universal $f \colon X^K \to Y^K$.

This is induced by the cone $\left(X^K \xrightarrow{\pi_F} X^F \xrightarrow{f^F} Y^F\right)_{F \subseteq K \text{ finite}}$.

Warning

In a general quasi-Markov category, the above cone maps do *not* form a strict cone, only a *lax* one. For instance, when $K = \{1, 2\}$,

Defined on $\mathrm{dom}(f_1) \otimes X_2$

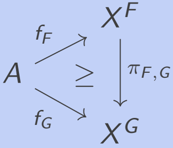


Defined on $\mathrm{dom}(f_1) \otimes \mathrm{dom}(f_2)$

Lax Kolmogorov products

Definition

(i) **Lax cone** over the diagram $X^{(-)}\colon \mathtt{FinSub}(K)^{\mathrm{op}} \to \mathtt{C}$: an object A and arrows $(f_F\colon A \to X^F)_{F \subseteq K \text{ finite}}$ such that for all $G \subseteq F \subseteq K$



(ii) A **lax Kolmogorov product** is a terminal lax cone $\left(X^K \xrightarrow{\pi_F} X^F\right)_{F \subseteq K \text{finite}}$: for any other lax cone $\left(A \xrightarrow{f_F} X^F\right)_{F \subseteq K \text{finite}}$ there is a greatest $A \xrightarrow{g} X^K$ such that each $\pi_F\, g \leq f_F$. We require it to have deterministic legs and be preserved by tensoring by arbitrary objects.

Infinite tensors of partial maps

Theorem

(i) Given Kolmogorov products in $\mathtt{C}$*, their inclusions into* $\mathtt{Partial}(\mathtt{C})$ *define both lax and strict Kolmogorov products.*

(ii) Given a family of maps $(X_k \overset{i_k}{\hookleftarrow} D_{f_k} \xrightarrow{f_k} Y_k)_{k \in K}$ *of* $\mathtt{Partial}(\mathtt{C})$*, the map* $X^K \to Y^K$ *induced by the universal product of the lax Kolmogorov product is* $(X^K \overset{i^K}{\hookleftarrow} \bigotimes_{k \in K} D_{f_k} \xrightarrow{f^K} Y^K)$*.*

Infinite copies in quasi-Markov categories

Consider a quasi-Markov category $\mathtt{C}$ with K-sized *strict* Kolmogorov products and a map $g \colon X \to Y$.

Construction

One would define the **infinite copy** $g^{(K)} \colon X \to Y^K$ by universal property as the unique map whose finite projections onto Y^F are given by F-many copies of g.

In a quasi-Markov category, these finite projections define a *strict* cone!

Theorem

Assume that a partializable $\mathtt{C}$ has Kolmogorov products.

For a $X \overset{i}{\hookleftarrow} D \xrightarrow{g} Y$ in $\mathtt{Partial}(\mathtt{C})$, the infinite copy is represented by $X \overset{i}{\hookleftarrow} D \xrightarrow{g^{(K)}} Y$.