

Random Variables, Independence Structures and Dagger Categories of Relations

Dario Stein, Radboud University Nijmegen

joint work with Chris Heunen¹, Matthew Di Meglio¹, Paolo Perrone²

¹ University of Edinburgh, ² University of Oxford

Topos Institute Colloquium – 22 May 2025

Outline

Section 1 – Model

Random Variables and Probability Sheaves

Section 2 – Generalize

Markov categories \mathbb{C} , probability spaces $\mathbb{P}(\mathbb{C})$, sample spaces $\mathbb{S}(\mathbb{C})$
Examples: isometries, nondeterminism, fresh name generation

Section 3 – Understand

What's the relationship between the categories \mathbb{P} and \mathbb{S} ?
Generalizing the regular category \leftrightarrow tabular allegory equivalence

Section 1

Random Variables

Random variables

Let's desugar the following statement

"Let X, Y be independent standard normal variables, then $P(X \geq Y) = \frac{1}{2}$ "

- 1 There exists a sample space $(\Omega, \Sigma, \mathbf{P})$ and two measurable functions

$$X, Y : (\Omega, \Sigma) \rightarrow (\mathbb{R}, \mathcal{B})$$

- 2 The laws $P_X(A) = P(X^{-1}(A)), P_Y(A) = P(Y^{-1}(A))$ satisfy

$$P_X(A) = P_Y(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{1}{2}x^2} dx$$

- 3 Independence

$$P(\{\omega : X(\omega) \in A, Y(\omega) \in B\}) = P_X(A) \cdot P_Y(B)$$

- 4 Then conclude that $P(\{\omega : X(\omega) \geq Y(\omega)\}) = \frac{1}{2}$

Random variables

What's nice about random variables?

- 1 Can be manipulated like values
- 2 measures are constructed implicitly by pushforward
- 3 close connection with functional analysis [$L^p(\Omega, \Sigma, P)$]
- 4 $X = Y$ equality almost surely
- 5 conditional expectation
- 6 **dependence on Ω implicit**

Random variables

What's awkward about random variables?

- 1 **dependence on Ω implicit**
- 2 type-safety: $\mathbb{E}[(X - \mathbb{E}[X])^2]$, $\mathbb{E}[X|Y = y]$, $\mathbb{E}[X|Y]$, \dots
- 3 constructing explicit distributions
- 4 $X \stackrel{d}{=} Y$ equality in distribution
- 5 conditional distributions

Questions [Tao]

- 1 What is the formal status of the sample space (Ω, Σ, P) ?
- 2 How can we silently enlarge it (allocate new random variables)?
- 3 How to make sure everything stays consistent?

A Convenient Setting for Random Variables

Convenient Setting

Can we find a typed setting which

- 1 needs no explicit tracking of measurability or sample spaces
- 2 faithfully includes standard Borel spaces X
- 3 has an object $\text{RV}(X)$ of random variables valued in X
- 4 allows reasoning by higher-order logic?

A Convenient Setting for Random Variables

Some desiderata

- 1 $(X \rightarrow Y) \rightarrow (RV(X) \rightarrow RV(Y))$
- 2 $RV(X \times Y) \cong RV(X) \times RV(Y)$
- 3 $\text{Law} : RV(X) \rightarrow \mathcal{G}(X)$ ← Giry monad
- 4 $(\perp) \subseteq RV(X) \times RV(Y)$
- 5 $(\sim) \subseteq RV(X) \times \mathcal{G}(X)$
- 6 $\mathbb{E} : RV([0, 1]) \rightarrow [0, 1]$
- 7 $\mathbb{E}[-|-] : RV([0, 1]) \times RV(X) \rightarrow RV([0, 1])$
- 8 $\forall F : RV(X) \forall \mu : \mathcal{G}(Y) \exists G : RV(Y) : G \sim \mu \wedge G \perp F$

A Convenient Setting for Random Variables

Convenient Setting

Can we find a typed setting which

- 1 needs no explicit tracking of measurability or sample spaces
- 2 faithfully includes standard Borel spaces X
- 3 has an object $\text{RV}(X)$ of random variables valued in X
- 4 allows reasoning by higher-order logic?

[Simpson'17] There is a boolean topos satisfying these desiderata, namely **Probability Sheaves**

Probability Sheaves

Probability Sheaves

Definition

Let **SBP** be the category of **standard Borel sample spaces**

- 1 objects are standard Borel probability spaces $\Omega = (X, \Sigma, P)$
- 2 morphisms $(X, \Sigma_X, P) \rightarrow (Y, \Sigma_Y, Q)$ are equivalence classes $[f]_P$ of measure-preserving measurable functions

$$(X, \Sigma_X) \rightarrow (Y, \Sigma_Y), f_*P = Q$$

up to p -almost sure equality.

Interpretation: Morphisms $\pi : \Omega' \rightarrow \Omega$ are projections or coarse-grainings, e.g.

- 1 $(X \times Y, \Sigma_{X \times Y}, P) \rightarrow (X, \Sigma_X, P_X)$
- 2 $(X, \mathcal{F}, P) \rightarrow (X, \mathcal{E}, P|_{\mathcal{E}})$ where $\mathcal{E} \subseteq \mathcal{F}$

Probability Sheaves

Definition

A *probability presheaf* is a functor $F : \mathbf{SBP}^{\text{op}} \rightarrow \mathbf{Set}$.

Idea: Every element $X \in F(\Omega)$ can be **extended** along $\Omega' \xrightarrow{\pi} \Omega$ to $X \cdot \pi \in F(\Omega')$.

Definition

For every standard Borel space V , we have a presheaf $\text{RV}(V) : \mathbf{SBP}^{\text{op}} \rightarrow \mathbf{Set}$ with

$$\text{RV}(V)(\Omega, p) = \{X : \Omega \rightarrow V \text{ measurable}\} / \text{p-a.s.}$$

The extension action is $(\Omega \xrightarrow{X} V) \cdot (\Omega' \xrightarrow{\pi} \Omega) = X \circ \pi$.

Proposition

RV defines a cartesian functor $\mathbf{Sbs} \rightarrow [\mathbf{SBP}^{\text{op}}, \mathbf{Set}]$.

- $\text{RV}(V \xrightarrow{f} W)_{\Omega} : (\Omega \xrightarrow{X} V) \mapsto f \circ X$
- the desired operations are equivariant (natural transformations), e.g.

$$\text{Law} : \text{RV}(V) \rightarrow \Delta\mathcal{G}(V), (\Omega \xrightarrow{X} V) \mapsto X_* p_{\Omega}$$

Probability Sheaves

Theorem

Each presheaf $\text{RV}(V)$ is a sheaf for the atomic topology on \mathbf{SBP} .

Given $\pi : \Omega' \rightarrow \Omega$,

- 1 $X \cdot \pi = Y \cdot \pi \Rightarrow X = Y$
- 2 $X' \in \text{RV}(V)(\Omega')$ descends to $X \in \text{RV}(V)(\Omega)$ iff X' is π -invariant, i.e.

$$\forall \rho_1, \rho_2, \pi \circ \rho_1 = \pi \circ \rho_2 \Rightarrow X' \cdot \rho_1 = X' \cdot \rho_2$$

What's the deeper meaning of the atomic sheaf property?

Sheaves and Conditional Independence

Definition (Independent Square)

We call a commutative square in \mathbf{SBP} *independent*

$$\begin{array}{ccc} \Omega' & \xrightarrow{f_1} & \Omega_1 \\ f_2 \downarrow & \perp\!\!\!\perp & \downarrow g_1 \\ \Omega_2 & \xrightarrow{g_2} & \Omega \end{array}$$

if f_1 and f_2 are conditionally independent given $g_1 f_1 (= g_2 f_2)$. ← regular conditional probabilities

Theorem (Simpson)

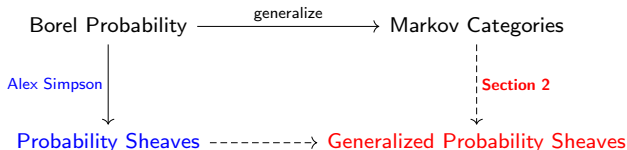
TFAE for a presheaf $F : \mathbf{SBP}^{\text{op}} \rightarrow \mathbf{Set}$

- 1 F is a sheaf for the **atomic topology** on \mathbf{SBP}
- 2 F sends *independent squares* in \mathbf{SBP} to *pullback squares* in \mathbf{Set}

Probability Sheaves

Summary Section I: The random variable formalism lives in **atomic sheaves** over **sample spaces**.

We want to: Generalize the situation, and understand what makes it work.



Section 2

Markov Categories and Sample Spaces

Markov categories

Markov categories [Fritz'20]

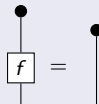
Markov categories $(\mathbb{C}, \otimes, I, \text{copy}, \text{del})$ are an axiomatization of stochastic maps (Markov kernels)



copy



delete



normalization

Markov categories: Examples

FinStoch (discrete probability)

Finite sets X , and stochastic matrices $p(y|x)$ (Kleisli maps $X \rightarrow D(Y)$)

BorelStoch (Borel probability)

Standard Borel spaces (X, Σ_X) and Markov kernels (Kleisli maps $(X, \Sigma_X) \rightarrow \mathcal{G}(Y, \Sigma_Y)$)

Gauss (Gaussian probability)

Euclidean spaces \mathbb{R}^n , and affine-linear maps with Gaussian noise $f(x) = Ax + \mathcal{N}(\mu, \Sigma)$

SetMulti (nondeterminism)

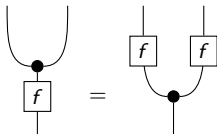
Sets X , and left-total relations $R \subseteq X \times Y$ (Kleisli maps $X \rightarrow \mathcal{P}_{\supseteq \emptyset}(Y)$)

StrongName (fresh name generation)

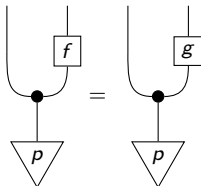
Strong nominal sets X , and name-generating equivariant functions (Kleisli maps $X \rightarrow N(Y)$); e.g. $f(a) = \langle a \rangle$, $g(a) = \langle b \rangle b$.

Markov categories

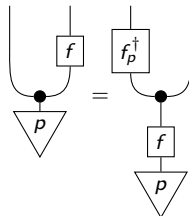
Many probabilistic concepts can be captured precisely in the language of Markov categories:



f deterministic



$f = g$ p -almost surely



Bayesian inversion

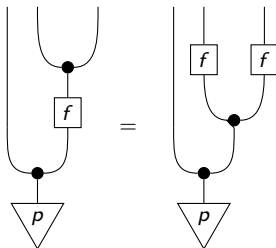
Probability- and Sample Spaces

Let \mathbb{C} be a Markov category with conditionals.

Definition

A probability space is a pair (X, p) with $p : I \rightarrow X$. A morphism $(X, p) \rightarrow (Y, q)$

- in $\mathbb{P}(\mathbb{C})$ is $[f]_p : X \rightarrow Y$ with $fp = q$
- in $\mathbb{S}(\mathbb{C})$ is $[f]_p : X \rightarrow Y$ with $fp = q$ and f is *p-a.s. deterministic*



Classifying Sample Spaces

We wish to understand $\mathbb{S}(\mathbb{C})$ in our examples:

Definition (Simplifying assumption)

Call a sample space (X, p) *faithful* if $f =_p g \Leftrightarrow f = g$. [← forget about a.s. equality](#)

Theorem

The following are equivalent

- 1 (X, p) is isomorphic in $\mathbb{P}(\mathbb{C})$ to a faithful sample space (S, σ)
- 2 (X, p) is isomorphic in $\mathbb{S}(\mathbb{C})$ to a faithful sample space (S, σ)
- 3 (S, σ) is a *split support* for p .

In all examples except **BorelStoch**, every sample space is isomorphic to a faithful one.

Probability

$\mathbb{P}(\mathbf{FinStoch})$ is equivalent to the category of couplings

- objects are (X, p) with $p(x) > 0$
- morphisms $(X, p_X) \rightarrow (Y, p_Y)$ are joint distribution $p(x, y)$ with $p_1(x) = p_X(x)$, $p_2(y) = p_Y(y)$.

$\mathbb{S}(\mathbf{FinStoch})$ is equivalent to **FinRV**

- objects are (X, p) with $p(x) > 0$
- morphisms $(X, p) \rightarrow (Y, q)$ are surjective functions f with

$$q(y) = \sum_{x \in f^{-1}(y)} p(x)$$

Relations

$\mathbb{P}(\mathbf{SetMulti})$ is equivalent to

- objects are sets X
- morphisms are bi-total relations $R \subseteq X \times Y$

$\mathbb{S}(\mathbf{SetMulti})$ is equivalent to **Surj**

- objects are sets X
- morphisms are surjective functions

Gaussian Probability

$\mathbb{P}(\mathbf{Gauss})$ is equivalent to **Con**

- objects are euclidean spaces \mathbb{R}^n
- morphisms are matrices $A \in \mathbb{R}^{n \times m}$ with $\|Ax\| \leq \|x\|$ (contractions)!

$\mathbb{S}(\mathbf{Gauss})$ is equivalent to **Colso**

- objects are euclidean spaces \mathbb{R}^n
- morphisms are $A \in \mathbb{R}^{n \times m}$ with $AA^T = I_n$.

Name Generation

$\mathbb{S}(\mathbf{StrongName})$ is equivalent to $\mathbf{FinInj}^{\text{op}}$

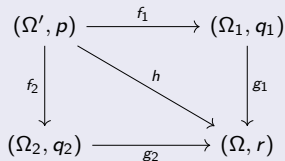
- objects are finite sets n
- morphisms $m \rightarrow n$ are injections $n \rightarrow m$

$$\mathbf{Nom}(\mathbb{A}^{\#m}, \mathbb{A}^{\#n}) \cong \mathbf{Inj}(n, m)$$

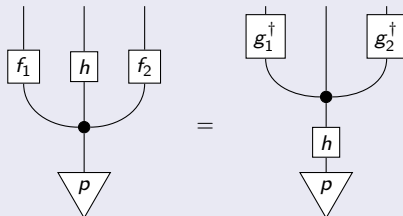
Conditional independence

Definition

We call a commutative square in $\mathbb{S}(\mathbb{C})$



independent if $f_1 \perp f_2 \mid h$, i.e.



Probability- and Sample Spaces

Can we develop a theory of probability sheaves for $\mathbb{S}(\mathbb{C})$?

- 1 Yes, Simpson postulated axioms (IP1)-(IP5) which can be verified by hand ([Stein, LICS'25])
- 2 more elegant route: let's understand the relationship between $\mathbb{S}(\mathbb{C})$ and $\mathbb{P}(\mathbb{C})$ using categorical logic

Some Properties [Perrone & al]

- 1 $\mathbb{P}(\mathbb{C})$ and $\mathbb{S}(\mathbb{C})$ are both semicartesian monoidal (not Markov!)
- 2 Bayesian inversion is a dagger functor on $\mathbb{P}(\mathbb{C})$

$$(X, p) \begin{array}{c} \xrightarrow{f} \\ \xleftarrow{f^\dagger} \end{array} (Y, q)$$

- 3 f p-a.s. deterministic $\Leftrightarrow f \circ f^\dagger = \text{id}$ (\dagger -coisometry!)
- 4 That is $\mathbb{S}(\mathbb{C}) = \text{CoIsom}(\mathbb{P}(\mathbb{C}))$.

Idea: If \mathbb{P} has a nice enough \dagger -structure, then $\text{CoIsom}(\mathbb{P})$ has a nice independence structure.

Section 3

Independence Structures

Quick Recap

Classic story

locally regular categories \leftrightarrow tabular allegories

New story

epiregular independence categories \leftrightarrow dilar \dagger -categories

Regular Categories and Allegories

Rel-construction

Every locally regular category \mathbb{C} has a category of relations $\mathbf{Rel}(\mathbb{C})$.

- 1 same objects as \mathbb{C}
- 2 morphisms are equivalence classes of jointly monic spans $[f \xleftarrow{f} \Omega \xrightarrow{g} Y]$
- 3 composition by pullback, followed by image factorization

Question: How to characterize the categories $\mathbf{Rel}(\mathbb{C})$, and recover \mathbb{C} from them?

Regular Categories and Allegories

Allegory

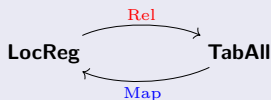
An **allegory** is a poset-enriched \dagger -category \mathbb{R} where each hom-set carries a lattice structure $R \cap S$, satisfying the modular law.

- 1 a **map** $f : X \rightarrow Y$ satisfies $1 \subseteq f^\dagger f$ (left-total) and $ff^\dagger \subseteq 1$ (subfunctional).
- 2 a **tabulation** of R is a span of maps $[f, g]$ with $gf^\dagger = R$
- 3 a **tabulator** is a tabulation satisfying $f^\dagger f \cap g^\dagger g = 1$ (think joint monicity)

A **tabular allegory** is one where every morphism has a tabulator.

A correspondence

- 1 For every locally regular category \mathbb{C} , $\mathbf{Rel}(\mathbb{C})$ is a tabular allegory
- 2 For every tabular allegory \mathbb{R} , $\mathbf{Map}(\mathbb{R})$ is a locally regular category
- 3 we have a 2-equivalence



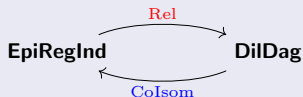
Our Solution

For intuitions, keep in mind the following dictionary

Allegories	Our story
map	coisometry
tabulation	dilation
tabulator	dilator
pullback	independent pullback

Theorem

- 1 For every **dilar \dagger -category** \mathbb{P} , $\text{CoIsom}(\mathbb{P})$ is an **epiregular independence category**
- 2 For every **epiregular independence category** \mathbb{S} , $\text{Rel}(\mathbb{S})$ is a **dilar \dagger -category**
- 3 We have a 2-equivalence



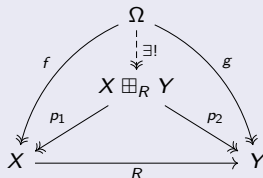
Dilar \dagger -categories

Definition

In a \dagger -category \mathbb{P} , we say $f : X \rightarrow Y$ is

- 1 a **coisometry** if $ff^\dagger = 1$ (\dagger -epi)
- 2 an **isometry** if $f^\dagger f = 1$ (\dagger -mono)

A **dilation** of $R : X \rightarrow Y$ is a span of coisometries with $R = gf^\dagger$. A **dilator** is a terminal dilation



We say \mathbb{P} is **dilar** if every morphism has a dilator.

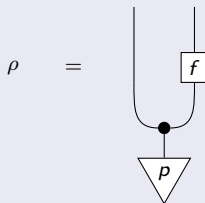
Dilar \dagger -categories

Proposition

If \mathbb{C} is a Markov category with conditionals, then $\mathbb{P}(\mathbb{C})$ is a dilar \dagger -category. The dilator of $f : (X, p) \rightarrow (Y, q)$ is given by the span

$$(X, p) \xleftarrow{\pi_1} (X \otimes Y, \rho) \xrightarrow{\pi_2} (Y, q)$$

where



Independence Structures

Theorem

For any \dagger -category \mathbb{P} , $\text{CoIsom}(\mathbb{P})$ carries an independence structure

$$\begin{array}{ccc} \bullet & \xrightarrow{f} & \bullet \\ g \downarrow & \perp & \downarrow u \\ \bullet & \xrightarrow{v} & \bullet \end{array} : \Leftrightarrow fg^\dagger = u^\dagger v$$

Examples:

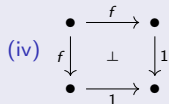
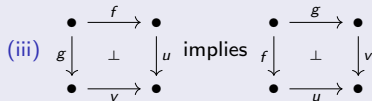
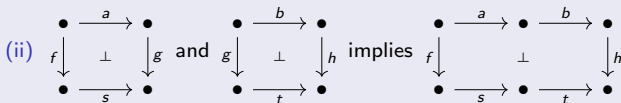
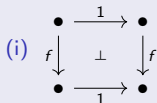
- 1 in Hilbert spaces: relative orthogonality
- 2 in $\mathbb{S}(\mathbb{C})$: conditional independence!

If \mathbb{P} has dilators, this should make the independence structure extra nice ...

Independence Structures

Definition [Alex Simpson'18, ~]

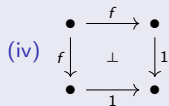
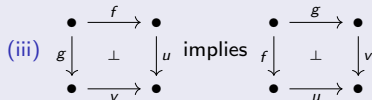
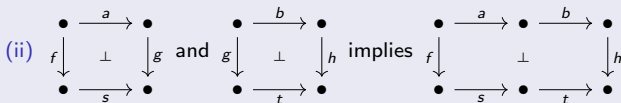
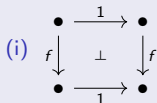
An *independence category* is equipped with a predicate \perp on commuting squares s.t



Independence Structures

Definition [Alex Simpson'18, ~]

An *independence category* is equipped with a predicate \perp on commuting squares s.t

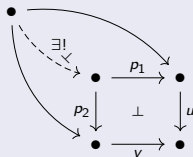


← (iv) not self-dual: independence \perp vs co-independence \top

Independent Pullback

Definition [Simpson]

In an independence category (\mathbb{S}, \perp) , an **independent pullback** is terminal among independent squares.



Epiregular independence category

Definition

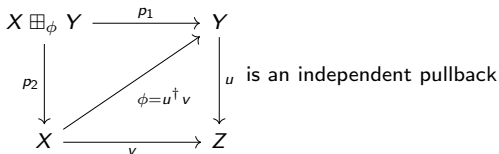
An independence category (\mathbb{S}, \perp) is **epiregular** if

- 1 every morphism is strong epi (orthogonal to jointly monic spans)
- 2 every span has a (strong epi, jointly mono)-factorization
- 3 every span completes to an independent pullback

Theorem

For every dilar \dagger -category \mathbb{P} , $\mathbf{CoIsom}(\mathbb{P})$ is an epiregular independence category.

Idea: a span of coisometries $[f, g]$ is jointly monic iff it is a dilator of gf^\dagger .



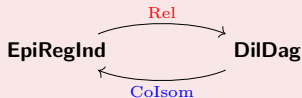
Back and Forth

Definition

For an epiregular independence category (\mathbb{S}, \perp) , define $\text{Rel}(\mathbb{S}, \perp)$ as usual, but compose by **independent pullback**.

Main Theorem

For every epiregular independence category (\mathbb{S}, \perp) , $\text{Rel}(\mathbb{S}, \perp)$ is a dilar \dagger -category, and we have a 2-equivalence



Examples: Surjections

Conditional variation independence

A commuting square of surjections

$$\begin{array}{ccc} \Omega & \xrightarrow{f} & X \\ g \downarrow & \perp & \downarrow u \\ Y & \xrightarrow{v} & Z \end{array}$$

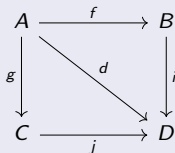
is independent if for all $x \in X$, $y \in Y$ with $u(x) = v(y)$, there exists $\omega \in \Omega$ with $f(\omega) = x$, $g(\omega) = y$.

Here, independence \Leftrightarrow weak pullback in **Set** (but not in **Surj**).
Independent pullback is the pullback in **Set**.

Examples: Injections

Relative disjointness

A commuting square of injections is co-independent



if $\text{im}(i) \cap \text{im}(j) = \text{im}(d)$.

Inj^{op} is an epiregular independence category.

In separation logic, think of Inj as heap layouts. Independence pushouts are **Set**-pushouts, and have been used in the semantic of local state [Kammar&al,LICS'17]

Examples: Isometries of Hilbert spaces

Relative Orthogonality

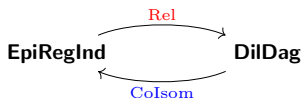
A bounded linear map f between Hilbert spaces is an isometry if $\|fx\| = \|x\|$. A commuting square of isometries is co-independent if

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ g \downarrow & \top & \downarrow i \\ C & \xrightarrow{j} & D \end{array} \quad :\Leftrightarrow \quad i^*j = fg^*$$

$\mathbf{Isom}^{\text{op}} \cong \mathbf{CoIsom}$ is an epiregular independence category (essentially by Sz.-Nagy's dilation theorem).

To summarize

We can extend the regular category/allegory 2-equivalence to a 2-equivalence



under which $\mathbb{S}(\mathbb{C})$ and $\mathbb{P}(\mathbb{C})$ recover each other. A theory of probability sheaves can be developed for any epiregular independence category \mathbb{S} .

$\mathbb{S}(\mathbb{C})$	$\mathbb{P}(\mathbb{C})$	\mathbb{C}
FinRV	Coupl	FinStoch
SBP	BorelCoupl	BorelStoch
Surj	TotRel	SetMulti
Isom^{op}	Con	Gauss
Inj^{op}	plnj	StrongName

Outlook

Outlook

- 1 let \mathcal{E} be the topos of atomic sheaves on \mathbb{S}
- 2 the inclusion $J : \mathbb{S} \rightarrow \mathbb{P} = \mathbf{Rel}(\mathbb{S})$ induces an adjunction

$$\begin{array}{ccc} & \xrightarrow{\Sigma_J} & \\ & \perp & \\ [\mathbb{S}^{\text{op}}, \mathbf{Set}] & \xleftarrow{\Delta_J} & [\mathbb{P}^{\text{op}}, \mathbf{Set}] \end{array}$$

whose monad restricts to $\mathcal{M} : \mathcal{E} \rightarrow \mathcal{E}$ and is commutative and affine

- 3 the Kleisli category $\mathcal{Kl}(\mathcal{M})$ is a Markov category
- 4 if $\mathbb{S} = \mathbb{S}(\mathbb{C})$, then $\mathbb{C} \rightarrow \mathcal{Kl}(\mathcal{M})$ a Markov embedding.

Application in Computer Science: “A Nominal Approach to Probabilistic Separation Logic” [Li&al, LICS’24] (Lilac), “A Monad for Full Ground Reference Cells” [Kammar&al, LLICS’17]

Appendix

The 2-category **DilDag** consists of

0. dilar \dagger -categories
1. functors preserving daggers and dilators
2. natural coisometries

The 2-category **EpiRegInd** consists of

0. epiregular independence categories
1. functors preserving independent squares
2. natural transformations with independent naturality squares