The mathematical theory of contextuality
Lecture 2: sheaf-thereoretic formulation

Samson Abramsky

Department of Computer Science
University College London

Functorializing probability theory: the distribution functor

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

A **discrete distribution** on a set X is a function $d : X \to [0,1]$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

A **discrete distribution** on a set X is a function $d : X \to [0, 1]$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

Equivalently, discrete distributions define probability measures on subsets S of X:

$$d(S) \ = \ \sum_{x \in S} d(x).$$

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

A **discrete distribution** on a set X is a function $d : X \to [0, 1]$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

Equivalently, discrete distributions define probability measures on subsets S of X:

$$d(S) \;=\; \sum_{x \in S} d(x).$$

We write $\mathsf{D}(X)$ for the set of R-distributions on X.

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

A **discrete distribution** on a set X is a function $d : X \to [0,1]$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

Equivalently, discrete distributions define probability measures on subsets S of X:

$$d(S) \;=\; \sum_{x \in S} d(x).$$

We write $\mathsf{D}(X)$ for the set of R-distributions on X.

Functorial action: Given a function $f : X \to Y$, we define

$$\mathsf{D}(f) : \mathsf{D}(X) \to \mathsf{D}(Y) :: d \mapsto [y \mapsto \sum_{f(x)=y} d(x)].$$

Functorializing probability theory: the distribution functor

We shall consider discrete distributions only; in fact, everything can done in measure-theoretic generality. (Lawvere, Giry et al.)

A **discrete distribution** on a set X is a function $d : X \to [0,1]$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

Equivalently, discrete distributions define probability measures on subsets S of X:

$$d(S) \;=\; \sum_{x \in S} d(x).$$

We write $\mathsf{D}(X)$ for the set of R-distributions on X.

Functorial action: Given a function $f : X \to Y$, we define

$$\mathsf{D}(f) : \mathsf{D}(X) \to \mathsf{D}(Y) :: d \mapsto [y \mapsto \sum_{f(x)=y} d(x)].$$

In terms of measures:

$$\mathsf{D}(f)(d)(S) \;=\; d(f^{-1}(S)).$$

Functorial structure

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

To see this, take the projection function $\pi_1 : X \times Y \to X$.

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

To see this, take the projection function $\pi_1 : X \times Y \to X$.

Then given a joint distribution $d \in \mathsf{D}(X \times Y)$, $\mathsf{D}(\pi_1)(d)$ is the **marginal** of d:

$$\mathsf{D}(\pi_1)(d)(x) \;=\; \sum_{y \in Y} d(x, y).$$

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

To see this, take the projection function $\pi_1 : X \times Y \to X$.

Then given a joint distribution $d \in \mathsf{D}(X \times Y)$, $\mathsf{D}(\pi_1)(d)$ is the **marginal** of d:

$$\mathsf{D}(\pi_1)(d)(x) \;=\; \sum_{y \in Y} d(x, y).$$

More generally, the functorial action **pushes measures forward** along maps.

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

To see this, take the projection function $\pi_1 : X \times Y \to X$.

Then given a joint distribution $d \in \mathsf{D}(X \times Y)$, $\mathsf{D}(\pi_1)(d)$ is the **marginal** of d:

$$\mathsf{D}(\pi_1)(d)(x) \;=\; \sum_{y \in Y} d(x, y).$$

More generally, the functorial action **pushes measures forward** along maps.

Additional structure:
There are "canonical maps" (natural transformations)

$$\delta_X : X \to \mathsf{D}(X), \qquad \mu_X : \mathsf{D}(\mathsf{D}(X)) \to \mathsf{D}(X)$$

which make the distribution functor into a **monad**.

Functorial structure

This "functorial action" indeed yields a functor $\mathsf{D} : \mathbf{Set} \longrightarrow \mathbf{Set}$. (Exercise!)

The functorial action generalizes **marginalization**.

To see this, take the projection function $\pi_1 : X \times Y \to X$.

Then given a joint distribution $d \in \mathsf{D}(X \times Y)$, $\mathsf{D}(\pi_1)(d)$ is the **marginal** of d:

$$\mathsf{D}(\pi_1)(d)(x) \;=\; \sum_{y \in Y} d(x, y).$$

More generally, the functorial action **pushes measures forward** along maps.

Additional structure:
There are "canonical maps" (natural transformations)

$$\delta_X : X \to \mathsf{D}(X), \qquad \mu_X : \mathsf{D}(\mathsf{D}(X)) \to \mathsf{D}(X)$$

which make the distribution functor into a **monad**.

Normalization corresponds to this monad being **affine**

$$\mathsf{D}(\mathbf{1}) \;\cong\; \mathbf{1}.$$

Semirings

Semirings

A **semiring** is a structure $(R, +, 0, \times, 1)$ such that

- $(R, +, 0)$ is an abelian monoid
- $(R, \times, 1)$ is a monoid
- multiplication distributes over addition:

$$a \cdot (b + c) \;=\; a \cdot b + a \cdot c.$$

Semirings

A **semiring** is a structure $(R, +, 0, \times, 1)$ such that

- $(R, +, 0)$ is an abelian monoid
- $(R, \times, 1)$ is a monoid
- multiplication distributes over addition:

$$a \cdot (b + c) \;=\; a \cdot b + a \cdot c.$$

The semiring is commutative if $\times$ is.

Semirings

A **semiring** is a structure $(R, +, 0, \times, 1)$ such that

- $(R, +, 0)$ is an abelian monoid
- $(R, \times, 1)$ is a monoid
- multiplication distributes over addition:

$$a \cdot (b + c) \;=\; a \cdot b + a \cdot c.$$

The semiring is commutative if $\times$ is.

Examples

- Rings
- $\mathbb{N}$
- $\mathbb{R}_{\geq 0}$
- $(\{0, 1\}, \vee, 0, \wedge, 1)$, more generally any distributive lattice

Semirings

A **semiring** is a structure $(R, +, 0, \times, 1)$ such that

- $(R, +, 0)$ is an abelian monoid
- $(R, \times, 1)$ is a monoid
- multiplication distributes over addition:

$$a \cdot (b + c) \;=\; a \cdot b + a \cdot c.$$

The semiring is commutative if $\times$ is.

Examples

- Rings
- $\mathbb{N}$
- $\mathbb{R}_{\geq 0}$
- $(\{0, 1\}, \vee, 0, \wedge, 1)$, more generally any distributive lattice

Widely used in Computer Science, e.g. for path algorithms, weighted automata etc.

Semirings

A **semiring** is a structure $(R, +, 0, \times, 1)$ such that

- $(R, +, 0)$ is an abelian monoid
- $(R, \times, 1)$ is a monoid
- multiplication distributes over addition:

$$a \cdot (b + c) \ = \ a \cdot b + a \cdot c.$$

The semiring is commutative if $\times$ is.

Examples

- Rings
- $\mathbb{N}$
- $\mathbb{R}_{\geq 0}$
- $(\{0, 1\}, \vee, 0, \wedge, 1)$, more generally any distributive lattice

Widely used in Computer Science, e.g. for path algorithms, weighted automata etc.

Features in **tropical geometry** (the **max-plus** semiring).

The Parameterized Distribution Functor

The Parameterized Distribution Functor

Fix a commutative semiring R. An R**-distribution** on X is a function $d : X \to R$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

The Parameterized Distribution Functor

Fix a commutative semiring R. An R-**distribution** on X is a function $d : X \to R$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

We write $\mathcal{D}_R(X)$ for the set of R-distributions on X.

The Parameterized Distribution Functor

Fix a commutative semiring R. An R-**distribution** on X is a function $d : X \to R$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

We write $\mathcal{D}_R(X)$ for the set of R-distributions on X.

Examples: $\mathbb{R}_{\geq 0}$ (probability distributions), $\mathbb{B}$ (non-empty subsets), $\mathbb{R}$ (signed measures).

The Parameterized Distribution Functor

Fix a commutative semiring R. An R**-distribution** on X is a function $d : X \to R$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

We write $\mathcal{D}_R(X)$ for the set of R-distributions on X.

Examples: $\mathbb{R}_{\geq 0}$ (probability distributions), $\mathbb{B}$ (non-empty subsets), $\mathbb{R}$ (signed measures).

Functorial action: Given a function $f : X \to Y$, we define

$$\mathcal{D}_R(f) : \mathcal{D}_R(X) \to \mathcal{D}_R(Y) :: d \mapsto [y \mapsto \sum_{f(x)=y} d(x)].$$

The Parameterized Distribution Functor

Fix a commutative semiring R. An R-**distribution** on X is a function $d : X \to R$ which has finite support, and such that

$$\sum_{x \in X} d(x) = 1.$$

We write $\mathcal{D}_R(X)$ for the set of R-distributions on X.

Examples: $\mathbb{R}_{\geq 0}$ (probability distributions), $\mathbb{B}$ (non-empty subsets), $\mathbb{R}$ (signed measures).

Functorial action: Given a function $f : X \to Y$, we define

$$\mathcal{D}_R(f) : \mathcal{D}_R(X) \to \mathcal{D}_R(Y) :: d \mapsto [y \mapsto \sum_{f(x)=y} d(x)].$$

This yields a functor $\mathcal{D}_R : \mathbf{Set} \longrightarrow \mathbf{Set}$.

Presheaves

Presheaves

A **presheaf of sets** on a topological space $(X, \mathcal{T}_X)$ is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}.$$

Presheaves

A **presheaf of sets** on a topological space $(X, \mathcal{T}_X)$ is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}.$$

Spelling this out, for each open set $U \subseteq X$, we have a set $P(U)$, and whenever $U \subseteq V$, there is a function, the **restriction map**

$$\rho_U^V : P(V) \to P(U)$$

subject to the functoriality requirements: if $U \subseteq V \subseteq W$, then

$$\rho_U^V \circ \rho_V^W = \rho_U^W, \qquad \rho_U^U = \mathsf{id}_{P(U)}.$$

Presheaves

A **presheaf of sets** on a topological space $(X, \mathcal{T}_X)$ is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}.$$

Spelling this out, for each open set $U \subseteq X$, we have a set $P(U)$, and whenever $U \subseteq V$, there is a function, the **restriction map**

$$\rho_U^V : P(V) \to P(U)$$

subject to the functoriality requirements: if $U \subseteq V \subseteq W$, then

$$\rho_U^V \circ \rho_V^W = \rho_U^W, \qquad \rho_U^U = \mathsf{id}_{P(U)}.$$

Example: the presheaf of functions
For each open set U, $\mathcal{F}(U)$ is the set of **continuous functions** $f : U \to \mathbb{R}$.

Presheaves

A **presheaf of sets** on a topological space $(X, \mathcal{T}_X)$ is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}.$$

Spelling this out, for each open set $U \subseteq X$, we have a set $P(U)$, and whenever $U \subseteq V$, there is a function, the **restriction map**

$$\rho^V_U : P(V) \to P(U)$$

subject to the functoriality requirements: if $U \subseteq V \subseteq W$, then

$$\rho^V_U \circ \rho^W_V = \rho^W_U, \qquad \rho^U_U = \mathsf{id}_{P(U)}.$$

Example: the presheaf of functions
For each open set U, $\mathcal{F}(U)$ is the set of **continuous functions** $f : U \to \mathbb{R}$.

Restriction is function restriction!
If $U \subseteq V$ and $f : V \to \mathbb{R}$, $\rho^V_U(f) := f|_U$.

Presheaves

A **presheaf of sets** on a topological space $(X, \mathcal{T}_X)$ is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}.$$

Spelling this out, for each open set $U \subseteq X$, we have a set $P(U)$, and whenever $U \subseteq V$, there is a function, the **restriction map**

$$\rho_U^V : P(V) \to P(U)$$

subject to the functoriality requirements: if $U \subseteq V \subseteq W$, then

$$\rho_U^V \circ \rho_V^W = \rho_U^W, \qquad \rho_U^U = \mathsf{id}_{P(U)}.$$

Example: the presheaf of functions
For each open set U, $\mathcal{F}(U)$ is the set of **continuous functions** $f : U \to \mathbb{R}$.

Restriction is function restriction!
If $U \subseteq V$ and $f : V \to \mathbb{R}$, $\rho_U^V(f) := f|_U$.

Functoriality is easily verified: in this notation

$$(f|_V)_U = f|_U.$$

Some notes on presheaves

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.

- We can also define presheaves in categories other than **Set**. For example, a **presheaf of abelian groups** on a space X is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{AbGrp}$$

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.
- We can also define presheaves in categories other than **Set**. For example, a **presheaf of abelian groups** on a space X is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{AbGrp}$$

- If $P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf, and $F : \mathbf{Set} \to \mathbf{Set}$ is a functor, then $F \circ P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf.

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.
- We can also define presheaves in categories other than **Set**. For example, a **presheaf of abelian groups** on a space X is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{AbGrp}$$

- If $P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf, and $F : \mathbf{Set} \to \mathbf{Set}$ is a functor, then $F \circ P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf.
- **Morphisms** of presheaves are just natural transformations.

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.
- We can also define presheaves in categories other than **Set**. For example, a **presheaf of abelian groups** on a space X is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{AbGrp}$$

- If $P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf, and $F : \mathbf{Set} \to \mathbf{Set}$ is a functor, then $F \circ P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf.
- **Morphisms** of presheaves are just natural transformations.
- The category of all presheaves on a space X has a very rich structure — it is a **topos**. We shall not go into this aspect.

Some notes on presheaves

- Presheaves can be defined on any poset, and in fact on any (small) category. They are just contravariant functors to **Set**.
- We can also define presheaves in categories other than **Set**. For example, a **presheaf of abelian groups** on a space X is a functor

$$P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{AbGrp}$$

- If $P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf, and $F : \mathbf{Set} \to \mathbf{Set}$ is a functor, then $F \circ P : \mathcal{T}_X^{\mathsf{op}} \to \mathbf{Set}$ is a presheaf.
- **Morphisms** of presheaves are just natural transformations.
- The category of all presheaves on a space X has a very rich structure — it is a **topos**. We shall not go into this aspect.
- However, there is an important conceptual aspect which should be understood. Presheaves allow us to formalise the concept of **variable set**. The variation is essentially over **contexts**. So presheaves provide the natural setting for talking about contextuality!

Covers, gluing and the sheaf condition

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

A **family** on $\mathcal{U}$ for the presheaf P is a family $\{x_i\}_{i \in I}$ with $x_i \in P(U_i)$, $i \in I$.

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

A **family** on $\mathcal{U}$ for the presheaf P is a family $\{x_i\}_{i \in I}$ with $x_i \in P(U_i)$, $i \in I$.

The family is **compatible** if for all $i, j \in I$,

$$x_i|_{U_i \cap U_j} = x_j|_{U_i \cap U_j}$$

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

A **family** on $\mathcal{U}$ for the presheaf P is a family $\{x_i\}_{i \in I}$ with $x_i \in P(U_i)$, $i \in I$.

The family is **compatible** if for all $i, j \in I$,

$$x_i|_{U_i \cap U_j} = x_j|_{U_i \cap U_j}$$

The presheaf satisfies the **gluing condition** for the cover $\mathcal{U}$ if for every compatible family $\{x_i\}_{i \in I}$ on $\mathcal{U}$, there exists $x \in P(U)$ such that, for all $i \in I$,

$$x|_{U_i} = x_i$$

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

A **family** on $\mathcal{U}$ for the presheaf P is a family $\{x_i\}_{i \in I}$ with $x_i \in P(U_i)$, $i \in I$.

The family is **compatible** if for all $i, j \in I$,

$$x_i|_{U_i \cap U_j} = x_j|_{U_i \cap U_j}$$

The presheaf satisfies the **gluing condition** for the cover $\mathcal{U}$ if for every compatible family $\{x_i\}_{i \in I}$ on $\mathcal{U}$, there exists $x \in P(U)$ such that, for all $i \in I$,

$$x|_{U_i} = x_i$$

It satisfies the **unique gluing condition** for $\mathcal{U}$, or satisfies the **sheaf condition** with respect to $\mathcal{U}$, if the element satisfying the gluing condition for a given compatible family is unique.

Covers, gluing and the sheaf condition

Sheaf theory is about the passage from **local** to **global**; about piecing together consistent local information into global information.

Let $\mathcal{U} = \{U_i\}_{i \in I}$ be a family of open sets, covering $U = \bigcup_{i \in I} U_i$.

A **family** on $\mathcal{U}$ for the presheaf P is a family $\{x_i\}_{i \in I}$ with $x_i \in P(U_i)$, $i \in I$.

The family is **compatible** if for all $i, j \in I$,

$$x_i|_{U_i \cap U_j} = x_j|_{U_i \cap U_j}$$

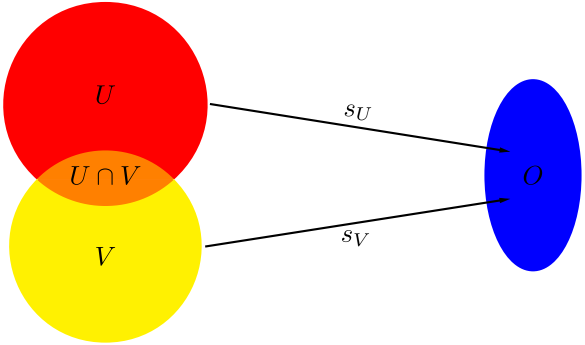
The presheaf satisfies the **gluing condition** for the cover $\mathcal{U}$ if for every compatible family $\{x_i\}_{i \in I}$ on $\mathcal{U}$, there exists $x \in P(U)$ such that, for all $i \in I$,

$$x|_{U_i} = x_i$$

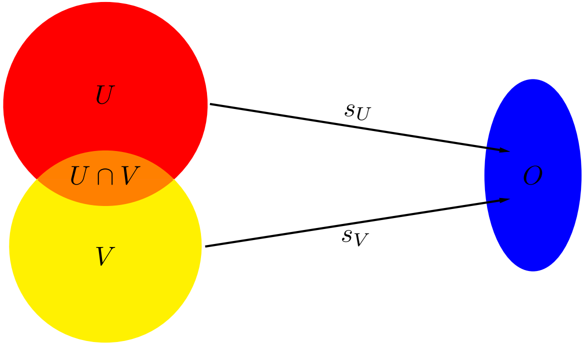
It satisfies the **unique gluing condition** for $\mathcal{U}$, or satisfies the **sheaf condition** with respect to $\mathcal{U}$, if the element satisfying the gluing condition for a given compatible family is unique.

The presheaf P is a **sheaf** if for every open cover $\mathcal{U}$, it satisfies the sheaf condition for $\mathcal{U}$.

Gluing functional sections



Gluing functional sections



If $s_U|_{U \cap V} = s_V|_{U \cap V}$, they can be glued to form

$$s : U \cup V \longrightarrow O$$

such that $s|_U = s_U$ and $s|_V = s_V$.

Obstructions

Obstructions

A major theme of modern mathematics is to identify and characterise situations where we **cannot** extend local information to global information.

Obstructions

A major theme of modern mathematics is to identify and characterise situations where we **cannot** extend local information to global information.

Mathematically, we seek to define invariants of the structure which provide **obstructions** to such extensions.

Obstructions

A major theme of modern mathematics is to identify and characterise situations where we **cannot** extend local information to global information.

Mathematically, we seek to define invariants of the structure which provide **obstructions** to such extensions.

In the case where it **is** possible to extend from local to global, these obstructions vanish.

Obstructions

A major theme of modern mathematics is to identify and characterise situations where we **cannot** extend local information to global information.

Mathematically, we seek to define invariants of the structure which provide **obstructions** to such extensions.

In the case where it **is** possible to extend from local to global, these obstructions vanish.

Where they do not vanish, they provide **witnesses** to this failure to extend from local to global — a constructive evidence for an impossibility.

Obstructions

A major theme of modern mathematics is to identify and characterise situations where we **cannot** extend local information to global information.

Mathematically, we seek to define invariants of the structure which provide **obstructions** to such extensions.

In the case where it **is** possible to extend from local to global, these obstructions vanish.

Where they do not vanish, they provide **witnesses** to this failure to extend from local to global — a constructive evidence for an impossibility.

In particular, this is one of the main intuitions behind **sheaf cohomology**.

The Sheaf of Events

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The presheaf of events:

$$\mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set} \quad :: \quad U \mapsto O^U$$

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The presheaf of events:

$$\mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set} \quad :: \quad U \mapsto O^U$$

Conceptually, a local section $s \in \mathcal{E}(U)$ represents the event of measuring or observing the variables $x \in U$, and observing the outcomes or values $s(x)$.

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The presheaf of events:

$$\mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set} \quad :: \quad U \mapsto O^U$$

Conceptually, a local section $s \in \mathcal{E}(U)$ represents the event of measuring or observing the variables $x \in U$, and observing the outcomes or values $s(x)$.

Restriction is by function restriction:

$$\rho^U_V(s) = s|_V, \qquad V \subseteq U$$

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The presheaf of events:

$$\mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set} \quad :: \quad U \mapsto O^U$$

Conceptually, a local section $s \in \mathcal{E}(U)$ represents the event of measuring or observing the variables $x \in U$, and observing the outcomes or values $s(x)$.

Restriction is by function restriction:

$$\rho^U_V(s) = s|_V, \qquad V \subseteq U$$

This presheaf is easily seen to satisfy the sheaf condition.

The Sheaf of Events

- Set of variables X (we will think of this as a discrete space)
- Set of outcomes O

The presheaf of events:

$$\mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set} \quad :: \quad U \mapsto O^U$$

Conceptually, a local section $s \in \mathcal{E}(U)$ represents the event of measuring or observing the variables $x \in U$, and observing the outcomes or values $s(x)$.

Restriction is by function restriction:

$$\rho^U_V(s) = s|_V, \qquad V \subseteq U$$

This presheaf is easily seen to satisfy the sheaf condition.

A useful generalization: we have a set O_x of outcomes for each measurement x. Then $\mathcal{E}(U) = \prod_{x \in U} O_x$. Restriction is by projection.

A Deterministic Universe is Classical

As we shall see, contextuality arises exactly where the sheaf property **fails**. Contextuality witnesses – Bell tests and other forms we will study – are **exactly** witnesses to this failure – **obstructions** to gluing.

A Deterministic Universe is Classical

As we shall see, contextuality arises exactly where the sheaf property **fails**. Contextuality witnesses – Bell tests and other forms we will study – are **exactly** witnesses to this failure – **obstructions** to gluing.

Thus a purely deterministic model, living on the event sheaf, is non-contextual. So if we want to find contextuality, we need to allow **distributions** over events.

A Deterministic Universe is Classical

As we shall see, contextuality arises exactly where the sheaf property **fails**. Contextuality witnesses – Bell tests and other forms we will study – are **exactly** witnesses to this failure – **obstructions** to gluing.

Thus a purely deterministic model, living on the event sheaf, is non-contextual. So if we want to find contextuality, we need to allow **distributions** over events.

This is in fact the content of the Conway-Kochen "Free Will Theorem".

Contextual Probability Theory

Contextual Probability Theory

We can form the presheaf $\mathcal{F} = \mathcal{D}_R \circ \mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set}$ by functor composition.

Contextual Probability Theory

We can form the presheaf $\mathcal{F} = \mathcal{D}_R \circ \mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set}$ by functor composition.

Explicitly: $\mathcal{F} : U \mapsto \mathcal{D}_R(O^U)$.

Contextual Probability Theory

We can form the presheaf $\mathcal{F} = \mathcal{D}_R \circ \mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set}$ by functor composition.

Explicitly: $\mathcal{F} : U \mapsto \mathcal{D}_R(O^U)$.

Restriction is by marginalization: if $U \subseteq V$ and $d \in \mathcal{F}(V)$,

$$d|_U : s \mapsto \sum_{t \in \mathcal{F}(V), t|_U = s} d(s)$$

Contextual Probability Theory

We can form the presheaf $\mathcal{F} = \mathcal{D}_R \circ \mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set}$ by functor composition.

Explicitly: $\mathcal{F} : U \mapsto \mathcal{D}_R(O^U)$.

Restriction is by marginalization: if $U \subseteq V$ and $d \in \mathcal{F}(V)$,

$$d|_U : s \mapsto \sum_{t \in \mathcal{F}(V), t|_U = s} d(s)$$

Rather than a fixed probability space (X, d), $d \in \mathcal{D}_R(X)$, we can now consider a **variable probability space**

$$(\mathcal{F}(U), d_U), \qquad U \subseteq X,\ d_U \in \mathcal{F}(U)$$

which varies functorially over the set of variables U.

Contextual Probability Theory

We can form the presheaf $\mathcal{F} = \mathcal{D}_R \circ \mathcal{E} : \mathcal{P}(X)^{\mathsf{op}} \to \mathbf{Set}$ by functor composition.

Explicitly: $\mathcal{F} : U \mapsto \mathcal{D}_R(O^U)$.

Restriction is by marginalization: if $U \subseteq V$ and $d \in \mathcal{F}(V)$,

$$d|_U : s \mapsto \sum_{t \in \mathcal{F}(V), t|_U = s} d(s)$$

Rather than a fixed probability space (X, d), $d \in \mathcal{D}_R(X)$, we can now consider a **variable probability space**

$$(\mathcal{F}(U), d_U), \qquad U \subseteq X,\ d_U \in \mathcal{F}(U)$$

which varies functorially over the set of variables U.

We shall now see how this arises naturally in some important situations.

A Probabilistic Model Of An Experiment

A Probabilistic Model Of An Experiment

Example: The Bell Model

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	$1/2$	0	0	$1/2$
a_1	b_2	$3/8$	$1/8$	$1/8$	$3/8$
a_2	b_1	$3/8$	$1/8$	$1/8$	$3/8$
a_2	b_2	$1/8$	$3/8$	$3/8$	$1/8$

A Probabilistic Model Of An Experiment

Example: The Bell Model

A	B	(0, 0)	(1, 0)	(0, 1)	(1, 1)
a_1	b_1	1/2	0	0	1/2
a_1	b_2	3/8	1/8	1/8	3/8
a_2	b_1	3/8	1/8	1/8	3/8
a_2	b_2	1/8	3/8	3/8	1/8

The entry in row 2 column 3 says:

If Alice looks at a_1 and Bob looks at b_2, then 1/8th of the time, Alice sees a 0 and Bob sees a 1.

Mathematical Structure of Probability Tables

Mathematical Structure of Probability Tables

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a	b	0	1/2	1/2	0
a'	b	3/8	1/8	1/8	3/8
a	b'	3/8	1/8	1/8	3/8
a'	b'	3/8	1/8	1/8	3/8

Mathematical Structure of Probability Tables

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a	b	0	$1/2$	$1/2$	0
a'	b	$3/8$	$1/8$	$1/8$	$3/8$
a	b'	$3/8$	$1/8$	$1/8$	$3/8$
a'	b'	$3/8$	$1/8$	$1/8$	$3/8$

The **measurement contexts** are

$$\{a, b\}, \quad \{a', b\}, \quad \{a, b'\}, \quad \{a', b'\}.$$

Mathematical Structure of Probability Tables

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a	b	0	$1/2$	$1/2$	0
a'	b	$3/8$	$1/8$	$1/8$	$3/8$
a	b'	$3/8$	$1/8$	$1/8$	$3/8$
a'	b'	$3/8$	$1/8$	$1/8$	$3/8$

The **measurement contexts** are

$$\{a, b\}, \quad \{a', b\}, \quad \{a, b'\}, \quad \{a', b'\}.$$

Each measurement has possible outcomes 0 or 1. The matrix entry at row (a', b) and column $(0, 1)$ indicates the **event**

$$\{a' \mapsto 0,\ b \mapsto 1\}.$$

Mathematical Structure of Probability Tables

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a	b	0	1/2	1/2	0
a'	b	3/8	1/8	1/8	3/8
a	b'	3/8	1/8	1/8	3/8
a'	b'	3/8	1/8	1/8	3/8

The **measurement contexts** are

$$\{a, b\}, \quad \{a', b\}, \quad \{a, b'\}, \quad \{a', b'\}.$$

Each measurement has possible outcomes 0 or 1. The matrix entry at row (a', b) and column $(0, 1)$ indicates the **event**

$$\{a' \mapsto 0,\ b \mapsto 1\}.$$

Each row of the table specifies a **probability distribution** on events O^C for a given choice of measurements C.

Presheaves, Sheaves and Gluing

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of X, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of X, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.
- To each such set C a probability distribution on **local sections** $s : C \to O$, where O is the set of **outcomes**. In our case, $O = \{0, 1\}$.

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of X, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.
- To each such set C a probability distribution on **local sections** $s : C \to O$, where O is the set of **outcomes**. In our case, $O = \{0, 1\}$.

These local sections correspond to the directly observable **joint outcomes** of **compatible measurements**, which can actually be performed jointly on the system.

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of X, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.
- To each such set C a probability distribution on **local sections** $s : C \to O$, where O is the set of **outcomes**. In our case, $O = \{0, 1\}$.

These local sections correspond to the directly observable **joint outcomes** of **compatible measurements**, which can actually be performed jointly on the system.

The different sets of compatible measurements correspond to the different contexts of measurement and observation of the physical system.

Presheaves, Sheaves and Gluing

Mathematically, this defines a **presheaf**. We have:

- A set of measurements X (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of X, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.
- To each such set C a probability distribution on **local sections** $s : C \to O$, where O is the set of **outcomes**. In our case, $O = \{0, 1\}$.

These local sections correspond to the directly observable **joint outcomes** of **compatible measurements**, which can actually be performed jointly on the system.

The different sets of compatible measurements correspond to the different contexts of measurement and observation of the physical system.

The fact that the behaviour of these observable outcomes cannot be accounted for by some context-independent global description of reality corresponds to the geometric fact that these local sections cannot be glued together into a **global section**.

Obstructions to gluing distributions

Obstructions to gluing distributions

In geometric language, Bell's theorem and related results corresponds to the fact that there is a **local section** which cannot be extended to a **global section** which is compatible with the family of distributions.

Obstructions to gluing distributions

In geometric language, Bell's theorem and related results corresponds to the fact that there is a **local section** which cannot be extended to a **global section** which is compatible with the family of distributions.

In other words, the space of **local probabilities/possibilities** is sufficiently logically 'twisted' to **obstruct** such an extension.

Obstructions to gluing distributions

In geometric language, Bell's theorem and related results corresponds to the fact that there is a **local section** which cannot be extended to a **global section** which is compatible with the family of distributions.

In other words, the space of **local probabilities/possibilities** is sufficiently logically 'twisted' to **obstruct** such an extension.

The quantum phenomena of **non-locality** and **contextuality** correspond exactly to the existence of obstructions to global sections in this sense.

Empirical Models: Reconstructing Probability Tables

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

An **empirical model** for μ is a family $\{e_C\}_{C \in \mathcal{M}}$, $e_C \in \mathcal{D}_R\mathcal{E}(C)$, which is **compatible**: for all $C, C' \in \mathcal{M}$,

$$e_C | C \cap C' = e_{C'} | C \cap C'.$$

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

An **empirical model** for μ is a family $\{e_C\}_{C \in \mathcal{M}}$, $e_C \in \mathcal{D}_R\mathcal{E}(C)$, which is **compatible**: for all $C, C' \in \mathcal{M}$,

$$e_C | C \cap C' = e_{C'} | C \cap C'.$$

Compatibility $\Longleftrightarrow$ No-Signalling

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

An **empirical model** for μ is a family $\{e_C\}_{C \in \mathcal{M}}$, $e_C \in \mathcal{D}_R\mathcal{E}(C)$, which is **compatible**: for all $C, C' \in \mathcal{M}$,

$$e_C | C \cap C' = e_{C'} | C \cap C'.$$

Compatibility $\iff$ No-Signalling

E.g. in the bipartite case, consider $C = \{m_a, m_b\}$, $C' = \{m_a, m'_b\}$.
Fix $s_0 \in \mathcal{E}(\{m_a\})$. Compatibility implies

$$\sum_{s \in \mathcal{E}(C), s|m_a = s_0} e_C(s) \quad = \quad \sum_{s' \in \mathcal{E}(C'), s'|m_a = s_0} e_{C'}(s').$$

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

An **empirical model** for μ is a family $\{e_C\}_{C \in \mathcal{M}}$, $e_C \in \mathcal{D}_R\mathcal{E}(C)$, which is **compatible**: for all $C, C' \in \mathcal{M}$,

$$e_C | C \cap C' = e_{C'} | C \cap C'.$$

$$\boxed{\text{Compatibility} \quad \Longleftrightarrow \quad \text{No-Signalling}}$$

E.g. in the bipartite case, consider $C = \{m_a, m_b\}$, $C' = \{m_a, m_b'\}$.
Fix $s_0 \in \mathcal{E}(\{m_a\})$. Compatibility implies

$$\sum_{s \in \mathcal{E}(C), s|m_a = s_0} e_C(s) \quad = \quad \sum_{s' \in \mathcal{E}(C'), s'|m_a = s_0} e_{C'}(s').$$

This says that the probability for Alice to get the outcome $s_0(m_a)$ is the same, whether we marginalize over the possible outcomes for Bob with measurement m_b, or with m_b'.

Empirical Models: Reconstructing Probability Tables

We are given a **measurement scenario** $\mu = \langle X, \mathcal{M}, O \rangle$:

An **empirical model** for μ is a family $\{e_C\}_{C \in \mathcal{M}}$, $e_C \in \mathcal{D}_R\mathcal{E}(C)$, which is **compatible**: for all $C, C' \in \mathcal{M}$,

$$e_C | C \cap C' = e_{C'} | C \cap C'.$$

$$\boxed{\text{Compatibility} \quad \Longleftrightarrow \quad \text{No-Signalling}}$$

E.g. in the bipartite case, consider $C = \{m_a, m_b\}$, $C' = \{m_a, m'_b\}$.
Fix $s_0 \in \mathcal{E}(\{m_a\})$. Compatibility implies

$$\sum_{s \in \mathcal{E}(C), s|m_a = s_0} e_C(s) \quad = \quad \sum_{s' \in \mathcal{E}(C'), s'|m_a = s_0} e_{C'}(s').$$

This says that the probability for Alice to get the outcome $s_0(m_a)$ is the same, whether we marginalize over the possible outcomes for Bob with measurement m_b, or with m'_b.

In other words, Bob's choice of measurement cannot influence Alice's outcome.

Global Sections

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R\mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R\mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

A joint distribution, defined on **all** measurements, which marginalizes to yield the empirically observed probabilities?

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R \mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

A joint distribution, defined on **all** measurements, which marginalizes to yield the empirically observed probabilities?

Note that $s \in \mathcal{E}(X) = O^X$ specifies an outcome for every measurement simultaneously, independent of the measurement context.

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R\mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

A joint distribution, defined on **all** measurements, which marginalizes to yield the empirically observed probabilities?

Note that $s \in \mathcal{E}(X) = O^X$ specifies an outcome for every measurement simultaneously, independent of the measurement context.
For every context C, it restricts to yield $s|C$.

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R\mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

A joint distribution, defined on **all** measurements, which marginalizes to yield the empirically observed probabilities?

Note that $s \in \mathcal{E}(X) = O^X$ specifies an outcome for every measurement simultaneously, independent of the measurement context.
For every context C, it restricts to yield $s|C$.

Thus it can be seen as a **deterministic hidden variable** — an **instruction set**!

Global Sections

We are given an empirical model $\{e_C\}_{C \in \mathcal{M}}$.

Question: does there exist a **global section** for this family?

I.e. $d \in \mathcal{D}_R\mathcal{E}(X)$ such that, for all $C \in \mathcal{M}$

$$d|C = e_C.$$

A joint distribution, defined on **all** measurements, which marginalizes to yield the empirically observed probabilities?

Note that $s \in \mathcal{E}(X) = O^X$ specifies an outcome for every measurement simultaneously, independent of the measurement context.
For every context C, it restricts to yield $s|C$.

Thus it can be seen as a **deterministic hidden variable** — an **instruction set!**

If d is a global section for the model $\{e_C\}$, we recover the predictions of the model by **averaging over the values of these hidden variables**:

$$e_C(s) \;=\; d|C(s) \;=\; \sum_{s' \in \mathcal{E}(X), s'|C = s} d(s') \;=\; \sum_{s' \in \mathcal{E}(X)} \delta_{s'|C}(s) \cdot d(s').$$

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

- X is a set of variables or measurement labels. Sufficient to consider finite discrete space — the base space of the bundle.

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

- X is a set of variables or measurement labels. Sufficient to consider finite discrete space — the base space of the bundle.

- $\mathcal{M} = \{C_i\}_{i \in I}$ set of **contexts** *i.e.* co-measurable variables. In quantum terms, compatible observables.

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

- X is a set of variables or measurement labels. Sufficient to consider finite discrete space — the base space of the bundle.

- $\mathcal{M} = \{C_i\}_{i \in I}$ set of **contexts** *i.e.* co-measurable variables. In quantum terms, compatible observables.

- O is set of outcomes or values for the variables, which we take to be the same in each fibre.

We have a sheaf of sets over $\mathcal{P}(X)$, namely $\mathcal{E} :: U \longmapsto O^U$ with restriction

$$\mathcal{E}(U \subseteq U') \colon \mathcal{E}(U') \longrightarrow \mathcal{E}(U) :: s \longmapsto s|U \ .$$

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

- X is a set of variables or measurement labels. Sufficient to consider finite discrete space — the base space of the bundle.

- $\mathcal{M} = \{C_i\}_{i \in I}$ set of **contexts** *i.e.* co-measurable variables. In quantum terms, compatible observables.

- O is set of outcomes or values for the variables, which we take to be the same in each fibre.

We have a sheaf of sets over $\mathcal{P}(X)$, namely $\mathcal{E} :: U \longmapsto O^U$ with restriction

$$\mathcal{E}(U \subseteq U') : \mathcal{E}(U') \longrightarrow \mathcal{E}(U) :: s \longmapsto s|U \; .$$

Each $s \in \mathcal{E}(U)$ is a **section**, and, in particular, $g \in \mathcal{E}(X)$ is a **global section**.

Sheaf formulation of contextuality

Measurement scenarios $\langle X, \mathcal{M}, O \rangle$:

- X is a set of variables or measurement labels. Sufficient to consider finite discrete space — the base space of the bundle.
- $\mathcal{M} = \{C_i\}_{i \in I}$ set of **contexts** *i.e.* co-measurable variables. In quantum terms, compatible observables.
- O is set of outcomes or values for the variables, which we take to be the same in each fibre.

We have a sheaf of sets over $\mathcal{P}(X)$, namely $\mathcal{E} :: U \longmapsto O^U$ with restriction

$$\mathcal{E}(U \subseteq U') \colon \mathcal{E}(U') \longrightarrow \mathcal{E}(U) :: s \longmapsto s|U \ .$$

Each $s \in \mathcal{E}(U)$ is a **section**, and, in particular, $g \in \mathcal{E}(X)$ is a **global section**.

A probability table can be represented by a family $\{p_C\}_{C \in \mathcal{M}}$ with p_C a probability distribution on $\mathcal{E}(C) = O^C$, where contexts C corresponds to the rows of the table.

Empirical Models

Empirical Models

The logical and strong forms of contextuality are concerned with **possibilities**, which can be represented by a subpresheaf $\mathcal{S}$ of $\mathcal{E}$, where for each context $U \subseteq X$, $\mathcal{S}(U) \subseteq O^U$ is the set of all possible outcomes.

Empirical Models

The logical and strong forms of contextuality are concerned with **possibilities**, which can be represented by a subpresheaf $\mathcal{S}$ of $\mathcal{E}$, where for each context $U \subseteq X$, $\mathcal{S}(U) \subseteq O^U$ is the set of all possible outcomes.

Explicitly, $\mathcal{S}$ is defined as follows, where $\mathsf{supp}(p_C|U \cap C)$ is the support of the marginal of p_C at $U \cap C$.

$$\mathcal{S}(U) := \left\{ s \in O^U \;\middle|\; \forall C \in \mathcal{M}.\, s|_{U \cap C} \in \mathsf{supp}(p_C|_{U \cap C}) \right\}$$

Empirical Models

The logical and strong forms of contextuality are concerned with **possibilities**, which can be represented by a subpresheaf $\mathcal{S}$ of $\mathcal{E}$, where for each context $U \subseteq X$, $\mathcal{S}(U) \subseteq O^U$ is the set of all possible outcomes.

Explicitly, $\mathcal{S}$ is defined as follows, where $\mathsf{supp}(p_C|U \cap C)$ is the support of the marginal of p_C at $U \cap C$.

$$\mathcal{S}(U) := \left\{ s \in O^U \;\middle|\; \forall C \in \mathcal{M}.\, s|_{U \cap C} \in \mathsf{supp}(p_C|_{U \cap C}) \right\}$$

We can use this formalisation to characterize contextuality as follows.

Definition

For any empirical model $\mathcal{S}$:

- For all $C \in \mathcal{M}$ and $s \in \mathcal{S}(C)$, $\mathcal{S}$ is **logically contextual** at s, written $\mathsf{LC}(\mathcal{S}, s)$, if s is not a member of any compatible family.
- $\mathcal{S}$ is **strongly contextual**, written $\mathsf{SC}(\mathcal{S})$, if $\mathsf{LC}(\mathcal{S}, s)$ for all s. Equivalently, if it has no global section, *i.e.* if $\mathcal{S}(X) = \varnothing$.

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \;:=\; \coprod_{i \in I} X_i \;:=\; \{(i, x) \,:\, i \in I, x \in X_i\}$$

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \ := \ \coprod_{i \in I} X_i \ := \ \{(i, x) \ : \ i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \ = \ \{(i, x) \ : \ x \in X_i\}$.

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \;:=\; \coprod_{i \in I} X_i \;:=\; \{(i, x) \,:\, i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \;=\; \{(i, x) \,:\, x \in X_i\}$.

There is also a natural projection function

$$p : X \to I \qquad p : (i, x) \mapsto i$$

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \;:=\; \coprod_{i \in I} X_i \;:=\; \{(i, x) \;:\; i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \;=\; \{(i, x) \;:\; x \in X_i\}$.

There is also a natural projection function

$$p : X \to I \qquad p : (i, x) \mapsto i$$

Conversely, given $p : X \to I$, we can form the indexed family $\{X_i\}_{i \in I}$, where $X_i := p^{-1}(\{i\})$.

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \;:=\; \coprod_{i \in I} X_i \;:=\; \{(i, x) \,:\, i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \;=\; \{(i, x) \,:\, x \in X_i\}$.

There is also a natural projection function

$$p : X \to I \qquad p : (i, x) \mapsto i$$

Conversely, given $p : X \to I$, we can form the indexed family $\{X_i\}_{i \in I}$, where $X_i := p^{-1}(\{i\})$.

These are equivalent ways of looking at the same idea.

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \; := \; \coprod_{i \in I} X_i \; := \; \{(i, x) \, : \, i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \; = \; \{(i, x) \, : \, x \in X_i\}$.

There is also a natural projection function

$$p : X \to I \qquad p : (i, x) \mapsto i$$

Conversely, given $p : X \to I$, we can form the indexed family $\{X_i\}_{i \in I}$, where $X_i := p^{-1}(\{i\})$.

These are equivalent ways of looking at the same idea.

With additional structure we get:
- topological bundles, fibre bundles, principal bundles, ...
- fibrations vs. indexed categories, ...

Two views of variation: indexed and fibred

Indexed family of sets $\{X_i\}_{i \in I}$.

Formally, take

$$X \;:=\; \coprod_{i \in I} X_i \;:=\; \{(i, x) \;:\; i \in I, x \in X_i\}$$

The family is $\phi : I \to \mathcal{P}(X)$, $\phi(i) \;=\; \{(i, x) \;:\; x \in X_i\}$.

There is also a natural projection function

$$p : X \to I \qquad p : (i, x) \mapsto i$$

Conversely, given $p : X \to I$, we can form the indexed family $\{X_i\}_{i \in I}$, where $X_i := p^{-1}(\{i\})$.

These are equivalent ways of looking at the same idea.

With additional structure we get:

- topological bundles, fibre bundles, principal bundles, ...
- fibrations vs. indexed categories, ...

Sheaves on X are equivalently formulated as continuous maps $p : Y \to X$ which are **local homeomorphisms** (*espaces étalé*).

Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

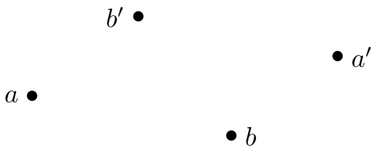
	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

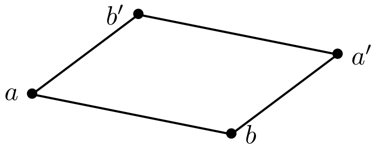


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

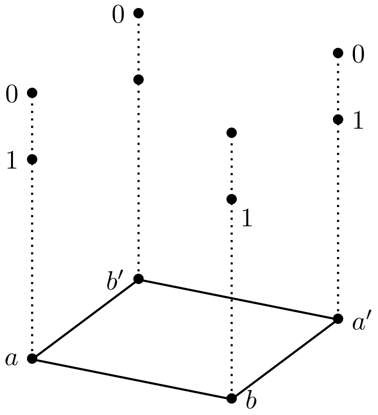


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

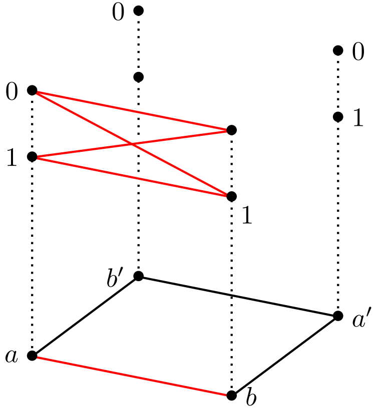


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

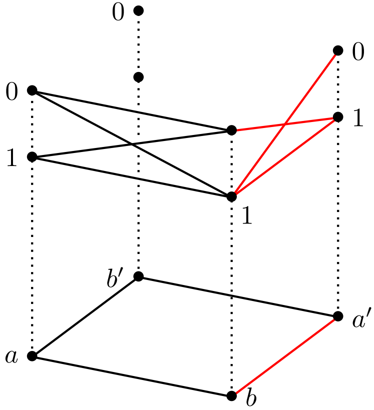


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

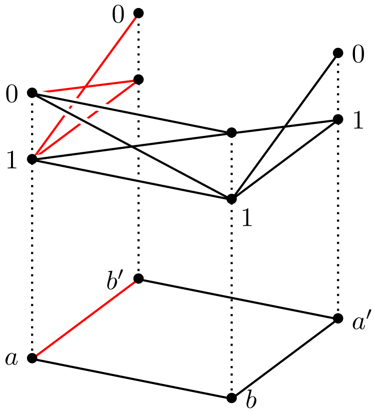


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

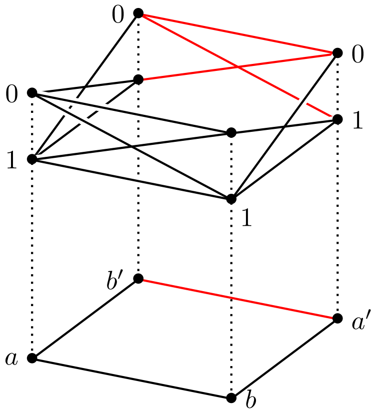


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

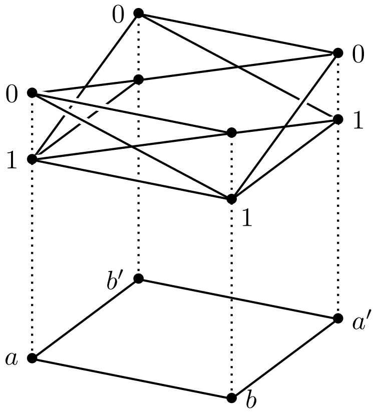


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

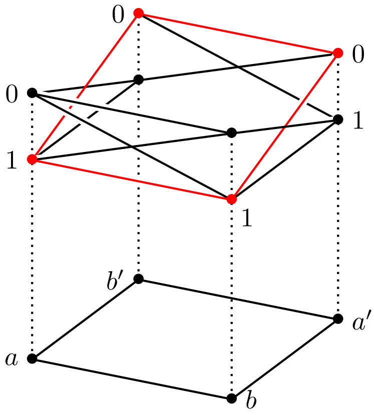


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

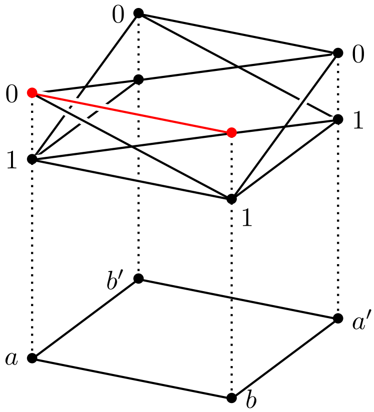


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

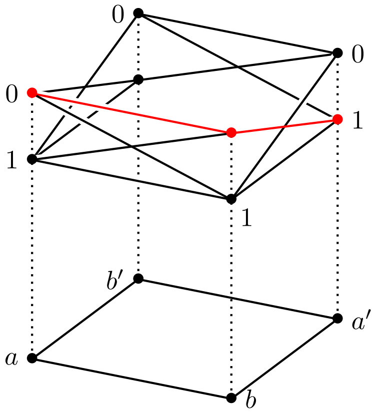


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

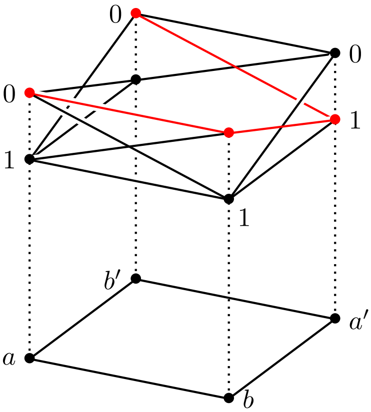


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

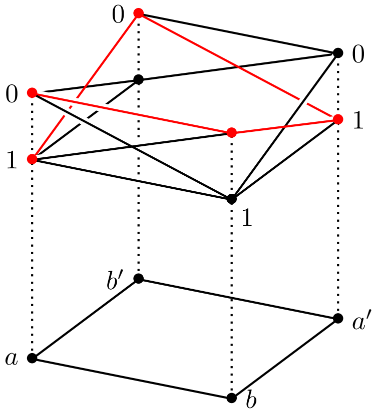


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

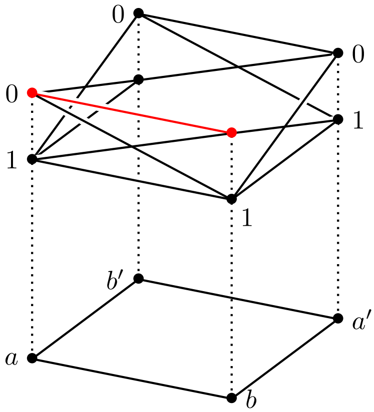


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

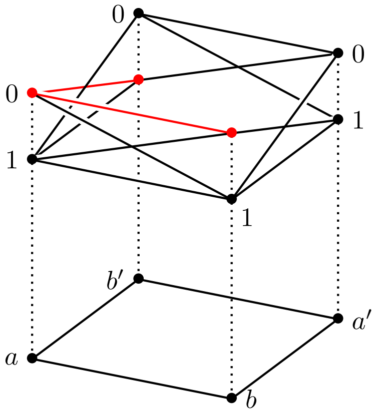


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

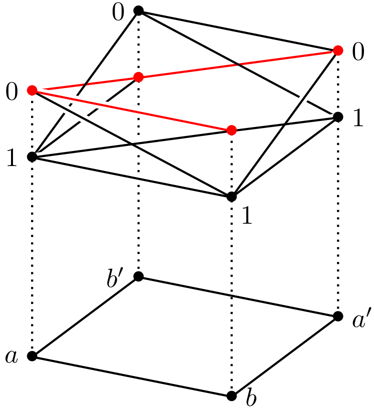


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×

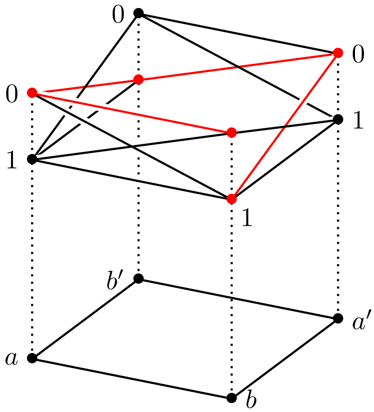


Bundle Pictures

Logical Contextuality

- Ignore precise probabilities
- Events are possible or not
- E.g. the Hardy model:

	00	01	10	11
ab	✓	✓	✓	✓
ab'	×	✓	✓	✓
$a'b$	×	✓	✓	✓
$a'b'$	✓	✓	✓	×



Strong Contextuality

A	B	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
a_1	b_1	1	0	0	1
a_1	b_2	1	0	0	1
a_2	b_1	1	0	0	1
a_2	b_2	0	1	1	0

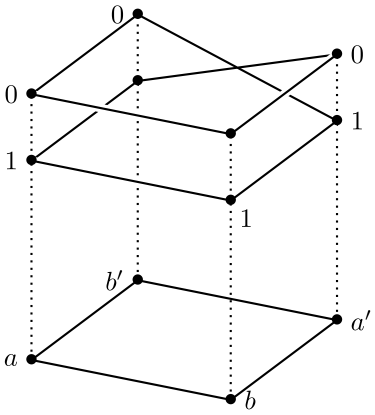
The PR Box

Bundle Pictures

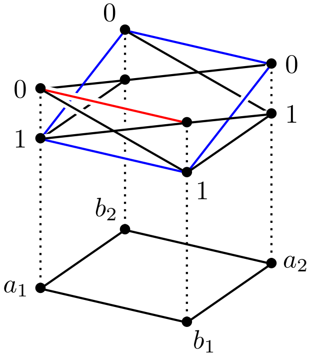
Strong Contextuality

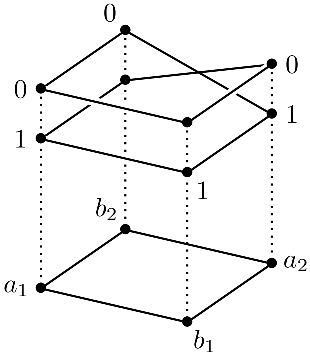
- E.g. the PR box:

	00	01	10	11
ab	✓	×	×	✓
ab'	✓	×	×	✓
$a'b$	✓	×	×	✓
$a'b'$	×	✓	✓	×



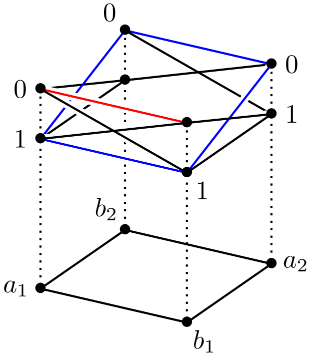
Visualizing Contextuality

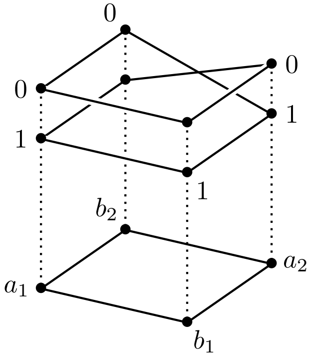




The Hardy table and the PR box as bundles

Visualizing Contextuality



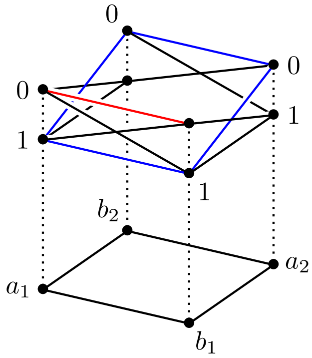


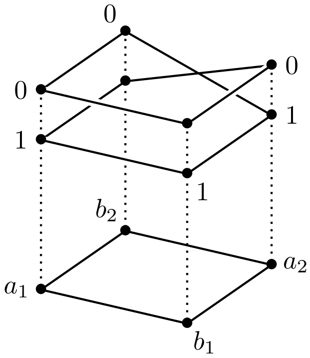
The Hardy table and the PR box as bundles

A hierarchy of degrees of contextuality:

$$\text{Bell} \;<\; \text{Hardy} \;<\; \text{GHZ}$$

Visualizing Contextuality





The Hardy table and the PR box as bundles

Degrees of contextuality

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

- **logically contextual** if some possible joint outcome $s \in O^C$ in the support is not accounted for by any global assignment $t \in O^X$ which is consistent with the support of the model. That is, for no such t do we have $t|C = s$.

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

- **logically contextual** if some possible joint outcome $s \in O^C$ in the support is not accounted for by any global assignment $t \in O^X$ which is consistent with the support of the model. That is, for no such t do we have $t|C = s$.

  Geometrically, this is saying that some local section cannot be extended to a global one. Equivalently, that the support of the model cannot be covered by the consistent global assignments.

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

- **logically contextual** if some possible joint outcome $s \in O^C$ in the support is not accounted for by any global assignment $t \in O^X$ which is consistent with the support of the model. That is, for no such t do we have $t|C = s$.

  Geometrically, this is saying that some local section cannot be extended to a global one. Equivalently, that the support of the model cannot be covered by the consistent global assignments.

- It is **strongly contextual** if its support has **no global section**; that is, there is no consistent global assignment.

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

- **logically contextual** if some possible joint outcome $s \in O^C$ in the support is not accounted for by any global assignment $t \in O^X$ which is consistent with the support of the model. That is, for no such t do we have $t|C = s$.

  Geometrically, this is saying that some local section cannot be extended to a global one. Equivalently, that the support of the model cannot be covered by the consistent global assignments.

- It is **strongly contextual** if its support has **no global section**; that is, there is no consistent global assignment.

  This says that **no** possible joint outcome is accounted for by **any** global section!

Degrees of contextuality

Firstly, we say that a global assignment $t \in O^X$ is **consistent with the support** of a model if for all $C' \in \mathcal{M}$, $t|_{C'}$ is in the support at C'.

An empirical model is

- **logically contextual** if some possible joint outcome $s \in O^C$ in the support is not accounted for by any global assignment $t \in O^X$ which is consistent with the support of the model. That is, for no such t do we have $t|C = s$.

  Geometrically, this is saying that some local section cannot be extended to a global one. Equivalently, that the support of the model cannot be covered by the consistent global assignments.

- It is **strongly contextual** if its support has **no global section**; that is, there is no consistent global assignment.

  This says that **no** possible joint outcome is accounted for by **any** global section!

  Obviously, strong non-locality implies logical non-locality.

A Hierarchy

A Hierarchy

We can distinguish three degrees of contextuality among models:

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.
- The Bell model is non-local, but not logically non-local.

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.
- The Bell model is non-local, but not logically non-local.
- The Hardy model is logically non-local, but not strongly non-local.

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.
- The Bell model is non-local, but not logically non-local.
- The Hardy model is logically non-local, but not strongly non-local.

Thus we have a strict hierarchy

$$\text{contextuality} \;<\; \text{logical contextuality} \;<\; \text{strong contextuality}$$

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.
- The Bell model is non-local, but not logically non-local.
- The Hardy model is logically non-local, but not strongly non-local.

Thus we have a strict hierarchy

$$\text{contextuality} \; < \; \text{logical contextuality} \; < \; \text{strong contextuality}$$

The model arising from the GHZ state (with 3 or more parties) with X, Y measurements at each site is strongly non-local.

A Hierarchy

We can distinguish three degrees of contextuality among models:

- Strong contextuality implies logical contextuality, which implies (probabilistic) contextuality.
- The Bell model is non-local, but not logically non-local.
- The Hardy model is logically non-local, but not strongly non-local.

Thus we have a strict hierarchy

$$\text{contextuality} \;<\; \text{logical contextuality} \;<\; \text{strong contextuality}$$

The model arising from the GHZ state (with 3 or more parties) with X, Y measurements at each site is strongly non-local.

Thus in terms of well-known examples, we have

$$\text{Bell} < \text{Hardy} < \text{GHZ}$$