# Contextual Semantics: From Quantum Mechanics to Logic, Databases, Constraints, and Complexity Lecture 2

Samson Abramsky

Department of Computer Science
The University of Oxford

# A Possibilistic Model Of An Experiment

# A Possibilistic Model Of An Experiment

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

# A Possibilistic Model Of An Experiment

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

The entry in row 1 column 1 says:

*If Alice looks at* $a_1$ *and Bob looks at* $b_1$*, then* **sometimes** *Alice sees a* 0 *and Bob sees a* 0*.*

# A Possibilistic Model Of An Experiment

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

The entry in row 1 column 1 says:

*If Alice looks at $a_1$ and Bob looks at $b_1$, then* **sometimes** *Alice sees a 0 and Bob sees a 0.*

The entry in row 2 column 1 says:

*If Alice looks at $a_1$ and Bob looks at $b_2$, then* **it never happens** *that Alice sees a 0 and Bob sees a 0.*

# A Possibilistic Model Of An Experiment

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

The entry in row 1 column 1 says:

> *If Alice looks at $a_1$ and Bob looks at $b_1$, then* **sometimes** *Alice sees a 0 and Bob sees a 0.*

The entry in row 2 column 1 says:

> *If Alice looks at $a_1$ and Bob looks at $b_2$, then* **it never happens** *that Alice sees a 0 and Bob sees a 0.*

Can we explain this behaviour using a classical source?

# What Do ‘Observables’ Observe?

# What Do 'Observables' Observe?

Surely **objective properties** of a physical system, which are independent of our choice of which measurements to perform — of our **measurement context**.

# What Do 'Observables' Observe?

Surely **objective properties** of a physical system, which are independent of our choice of which measurements to perform — of our **measurement context**.

More precisely, this would say that for each possible state of the system, there is a function $\lambda$ which for each measurement $m$ specifies an outcome $\lambda(m)$, **independently of which other measurements may be performed**.

# What Do ‘Observables’ Observe?

Surely **objective properties** of a physical system, which are independent of our choice of which measurements to perform — of our **measurement context**.

More precisely, this would say that for each possible state of the system, there is a function $\lambda$ which for each measurement $m$ specifies an outcome $\lambda(m)$, **independently of which other measurements may be performed**.

This point of view is called **non-contextuality**. It is equivalent to the assumption of a classical source.

# What Do 'Observables' Observe?

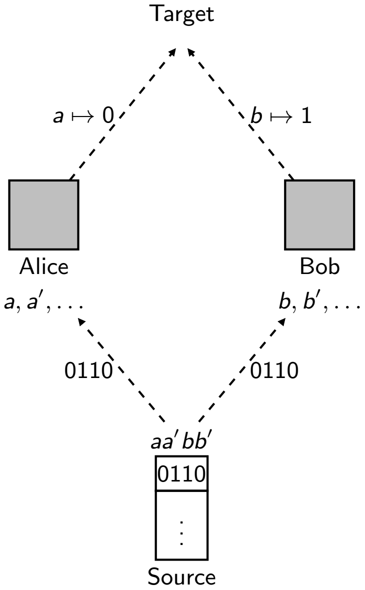
Surely **objective properties** of a physical system, which are independent of our choice of which measurements to perform — of our **measurement context**.

More precisely, this would say that for each possible state of the system, there is a function $\lambda$ which for each measurement $m$ specifies an outcome $\lambda(m)$, **independently of which other measurements may be performed**.

This point of view is called **non-contextuality**. It is equivalent to the assumption of a classical source.

However, this view is **impossible to sustain** in the light of our **actual observations of (micro)-physical reality**.

# Hidden Variables: The Mermin instruction set picture



# The ‘Hardy Paradox’

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

Which 'instruction set' $\lambda$ could the outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come? Clearly, we must have

$$\lambda : a_1 \mapsto 0, \quad b_1 \mapsto 0.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | ? | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

Which 'instruction set' $\lambda$ could the outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come? Clearly, we must have

$$\lambda : a_1 \mapsto 0, \quad b_1 \mapsto 0.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | | | 0 |

Which 'instruction set' $\lambda$ could the outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come? Clearly, we must have

$$\lambda : a_1 \mapsto 0, \quad b_1 \mapsto 0.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | ? | | 0 |

Which 'instruction set' $\lambda$ could the outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come? Clearly, we must have

$$\lambda : a_1 \mapsto 0, \quad b_1 \mapsto 0.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | 1 | | 0 |

Which 'instruction set' $\lambda$ could the outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come? Clearly, we must have

$$\lambda : a_1 \mapsto 0, \quad b_1 \mapsto 0.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | 1 | | 0 |

So there is a unique 'instruction set' $\lambda$ that outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come from:

$$\lambda : a_1 \mapsto 0, \quad a_2 \mapsto 0, \quad b_1 \mapsto 0, \quad b_2 \mapsto 1.$$

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | 1 | | 0 |

So there is a unique 'instruction set' $\lambda$ that outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come from:

$$\lambda : a_1 \mapsto 0, \quad a_2 \mapsto 0, \quad b_1 \mapsto 0, \quad b_2 \mapsto 1.$$

However, this would require the outcome $(0,0)$ for measurements $(a_2, b_1)$ to be possible, and this is **precluded**.

# The 'Hardy Paradox'

Hardy models: those whose support satisfies

| | $(0,0)$ | $(0,1)$ | $(1,0)$ | $(1,1)$ |
|---|---|---|---|---|
| $(a_1, b_1)$ | 1 | | | |
| $(a_1, b_2)$ | 0 | 1 | | |
| $(a_2, b_1)$ | 0 | | | |
| $(a_2, b_2)$ | | 1 | | 0 |

So there is a unique 'instruction set' $\lambda$ that outcomes $(0,0)$ for measurements $(a_1, b_1)$ could come from:

$$\lambda : a_1 \mapsto 0, \quad a_2 \mapsto 0, \quad b_1 \mapsto 0, \quad b_2 \mapsto 1.$$

However, this would require the outcome $(0,0)$ for measurements $(a_2, b_1)$ to be possible, and this is **precluded**.

Thus Hardy models are **contextual**. They cannot be explained by a classical source.

# 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** $s : C \rightarrow O$, 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:

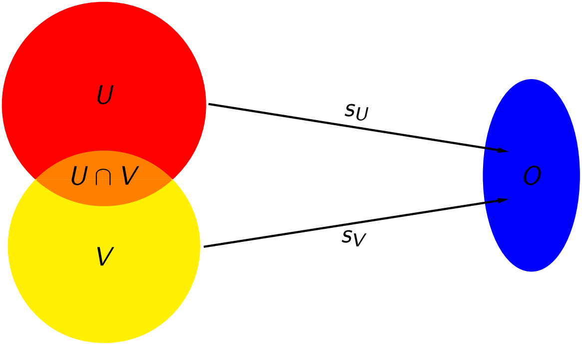
- A set of measurements $X$ (the 'space'). In our example, $X = \{a_1, a_2, b_1, b_2\}$.
- A family of subsets of $X$, the **measurement contexts** (a 'cover'); in our example, these are $\{\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}\}$.
- To each such set $C$ a probability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

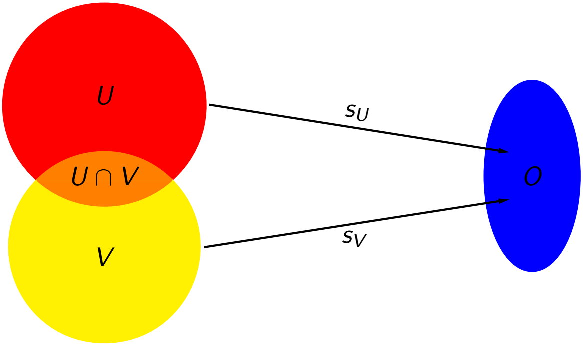
The different sets of compatible measurements correspond to the different contexts of measurement and observation of the physical system.

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**.

# 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 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.

# 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.

This geometric picture and the associated methods can also be applied to a wide range of situations in classical computer science.

# 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.

This geometric picture and the associated methods can also be applied to a wide range of situations in classical computer science.

We can vary the notion of distribution $d : X \to R$ by taking the weights in a commutative semiring $R$.

# 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.

This geometric picture and the associated methods can also be applied to a wide range of situations in classical computer science.

We can vary the notion of distribution $d : X \to R$ by taking the weights in a commutative semiring $R$.

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

# Restriction Of Distributions As Marginalization

# Restriction Of Distributions As Marginalization

We write the restriction maps for distributions explicitly:

# Restriction Of Distributions As Marginalization

We write the restriction maps for distributions explicitly:

Given $U \subseteq U'$ we have a map

$$\mathcal{D}_R(O^{U'}) \to \mathcal{D}_R(O^U) :: d \mapsto d|_U,$$

where for each $s : U \to O$:

$$d|_U(s) \; := \sum_{s' : U' \to O, s'|U = s} d(s').$$

# Restriction Of Distributions As Marginalization

We write the restriction maps for distributions explicitly:

Given $U \subseteq U'$ we have a map

$$\mathcal{D}_R(O^{U'}) \rightarrow \mathcal{D}_R(O^U) :: d \mapsto d|_U,$$

where for each $s : U \rightarrow O$:

$$d|_U(s) \; := \sum_{s' : U' \rightarrow O, s'|U = s} d(s').$$

Thus $d|_U$ is the **marginal** of the distribution $d$, which assigns to each section $s$ in the smaller context $U$ the sum of the weights of all sections $s'$ in the larger context which restrict to $s$.

# Restriction Of Distributions As Marginalization

We write the restriction maps for distributions explicitly:

Given $U \subseteq U'$ we have a map

$$\mathcal{D}_R(O^{U'}) \to \mathcal{D}_R(O^U) :: d \mapsto d|_U,$$

where for each $s : U \to O$:

$$d|_U(s) \ := \sum_{s' : U' \to O, s'|U = s} d(s').$$

Thus $d|_U$ is the **marginal** of the distribution $d$, which assigns to each section $s$ in the smaller context $U$ the sum of the weights of all sections $s'$ in the larger context which restrict to $s$.

(N.B.:This is functorial, hence defines a presheaf.)

# Restriction Of Distributions As Marginalization

We write the restriction maps for distributions explicitly:

Given $U \subseteq U'$ we have a map

$$\mathcal{D}_R(O^{U'}) \rightarrow \mathcal{D}_R(O^U) :: d \mapsto d|_U,$$

where for each $s : U \rightarrow O$:

$$d|_U(s) \ := \sum_{s' : U' \rightarrow O, s'|U = s} d(s').$$

Thus $d|_U$ is the **marginal** of the distribution $d$, which assigns to each section $s$ in the smaller context $U$ the sum of the weights of all sections $s'$ in the larger context which restrict to $s$.

(N.B.:This is functorial, hence defines a presheaf.)

In fact, it is the composition of $U \mapsto O^U$ and the covariant distribution functor $\mathcal{D}_R$.

# Measurement Structures

# Measurement Structures

A **measurement structure** $\mathcal{M}$ on $X$ is a family of measurement contexts which covers $X$, $\bigcup \mathcal{M} = X$.

# Measurement Structures

A **measurement structure** $\mathcal{M}$ on $X$ is a family of measurement contexts which covers $X$, $\bigcup \mathcal{M} = X$.

**Example: Bell scenarios**
If $X_A$, $X_B$ are disjoint sets of labels for measurements by Alice and Bob, the set of contexts will be

$$\{\{a, b\} \, : \, a \in X_A, b \in X_B\}.$$

This generalises immediately to multipartite scenarios.

# Measurement Structures

A **measurement structure** $\mathcal{M}$ on $X$ is a family of measurement contexts which covers $X$, $\bigcup \mathcal{M} = X$.

**Example: Bell scenarios**
If $X_A$, $X_B$ are disjoint sets of labels for measurements by Alice and Bob, the set of contexts will be

$$\{\{a, b\} \,:\, a \in X_A, b \in X_B\}.$$

This generalises immediately to multipartite scenarios.

**Example: Kochen-Specker configurations**
Given a set $X$ of vectors, the set of contexts will be

$$\{C \subseteq X \,:\, C \text{ forms an orthonormal basis}\}$$

# Measurement Structures

A **measurement structure** $\mathcal{M}$ on $X$ is a family of measurement contexts which covers $X$, $\bigcup \mathcal{M} = X$.

**Example: Bell scenarios**
If $X_A$, $X_B$ are disjoint sets of labels for measurements by Alice and Bob, the set of contexts will be

$$\{\{a, b\} \,:\, a \in X_A, b \in X_B\}.$$

This generalises immediately to multipartite scenarios.

**Example: Kochen-Specker configurations**
Given a set $X$ of vectors, the set of contexts will be

$$\{C \subseteq X \,:\, C \text{ forms an orthonormal basis}\}$$

N.B. Vorob'ev theorem.

# Example: the 18-vector configuration in $\mathbb{R}^4$

This uses the following measurement cover $\mathcal{U} = \{U_1, \ldots, U_9\}$:

| $U_1$ | $U_2$ | $U_3$ | $U_4$ | $U_5$ | $U_6$ | $U_7$ | $U_8$ | $U_9$ |
|---|---|---|---|---|---|---|---|---|
| $A$ | $A$ | $H$ | $H$ | $B$ | $I$ | $P$ | $P$ | $Q$ |
| $B$ | $E$ | $I$ | $K$ | $E$ | $K$ | $Q$ | $R$ | $R$ |
| $C$ | $F$ | $C$ | $G$ | $M$ | $N$ | $D$ | $F$ | $M$ |
| $D$ | $G$ | $J$ | $L$ | $N$ | $O$ | $J$ | $L$ | $O$ |

# Example: the 18-vector configuration in $\mathbb{R}^4$

This uses the following measurement cover $\mathcal{U} = \{U_1, \ldots, U_9\}$:

| $U_1$ | $U_2$ | $U_3$ | $U_4$ | $U_5$ | $U_6$ | $U_7$ | $U_8$ | $U_9$ |
|---|---|---|---|---|---|---|---|---|
| $A$ | $A$ | $H$ | $H$ | $B$ | $I$ | $P$ | $P$ | $Q$ |
| $B$ | $E$ | $I$ | $K$ | $E$ | $K$ | $Q$ | $R$ | $R$ |
| $C$ | $F$ | $C$ | $G$ | $M$ | $N$ | $D$ | $F$ | $M$ |
| $D$ | $G$ | $J$ | $L$ | $N$ | $O$ | $J$ | $L$ | $O$ |

We can label these letters with vectors in $\mathbb{R}^4$ such that the vectors in each column are orthogonal.

# Example: the 18-vector configuration in $\mathbb{R}^4$

This uses the following measurement cover $\mathcal{U} = \{U_1, \ldots, U_9\}$:

| $U_1$ | $U_2$ | $U_3$ | $U_4$ | $U_5$ | $U_6$ | $U_7$ | $U_8$ | $U_9$ |
|---|---|---|---|---|---|---|---|---|
| $A$ | $A$ | $H$ | $H$ | $B$ | $I$ | $P$ | $P$ | $Q$ |
| $B$ | $E$ | $I$ | $K$ | $E$ | $K$ | $Q$ | $R$ | $R$ |
| $C$ | $F$ | $C$ | $G$ | $M$ | $N$ | $D$ | $F$ | $M$ |
| $D$ | $G$ | $J$ | $L$ | $N$ | $O$ | $J$ | $L$ | $O$ |

We can label these letters with vectors in $\mathbb{R}^4$ such that the vectors in each column are orthogonal.

Yields a proof of the **Kochen-Specker theorem**.

# The Distribution Functor

# The Distribution Functor

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

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

# The Distribution Functor

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

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

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

# The Distribution Functor

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

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

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 Distribution Functor

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

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

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 Distribution Functor

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

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

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 $\mathcal{D}_R$ : **Set** $\longrightarrow$ **Set**.

# The Distribution Functor

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

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

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 $\mathcal{D}_R : \mathbf{Set} \longrightarrow \mathbf{Set}$.

We can compose this functor with $U \mapsto O^U$, to form a presheaf $\mathcal{F} : \mathcal{P}(X)^{\mathrm{op}} \longrightarrow \mathbf{Set}$.

# The Distribution Functor

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

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

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 $\mathcal{D}_R : \mathbf{Set} \longrightarrow \mathbf{Set}$.

We can compose this functor with $U \mapsto O^U$, to form a presheaf $\mathcal{F} : \mathcal{P}(X)^{\mathsf{op}} \longrightarrow \mathbf{Set}$.

**Contextual Probability Theory**!

# Basic Mathematical Setting

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"
- $\mathcal{M}$ is a family of subsets of $X$ with $\bigcup \mathcal{M} = X$; the "measurement contexts"

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"
- $\mathcal{M}$ is a family of subsets of $X$ with $\bigcup \mathcal{M} = X$; the "measurement contexts"
- $O$ is a set of "outcomes" or "values"

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"
- $\mathcal{M}$ is a family of subsets of $X$ with $\bigcup \mathcal{M} = X$; the "measurement contexts"
- $O$ is a set of "outcomes" or "values"

In addition, we have some commutative semiring $R$ of "weights".

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"
- $\mathcal{M}$ is a family of subsets of $X$ with $\bigcup \mathcal{M} = X$; the "measurement contexts"
- $O$ is a set of "outcomes" or "values"

In addition, we have some commutative semiring $R$ of "weights".

Then we can define the presheaf

$$\mathcal{F} : \mathcal{P}(X)^{\mathsf{op}} \longrightarrow \mathbf{Set} \;::\; U \mapsto \mathcal{D}_R(O^U)$$

# Basic Mathematical Setting

An **observational frame** is a structure $(X, \mathcal{M}, O)$ where:

- $X$ is a set of "measurement labels" or "variables"
- $\mathcal{M}$ is a family of subsets of $X$ with $\bigcup \mathcal{M} = X$; the "measurement contexts"
- $O$ is a set of "outcomes" or "values"

In addition, we have some commutative semiring $R$ of "weights".

Then we can define the presheaf

$$\mathcal{F} : \mathcal{P}(X)^{\mathsf{op}} \longrightarrow \mathbf{Set} \;::\; U \mapsto \mathcal{D}_R(O^U)$$

A setting for **contextual probability**.

# Empirical Models: Reconstructing Probability Tables

# Empirical Models: Reconstructing Probability Tables

We are given a **measurement structure** $\mathcal{M}$:

# Empirical Models: Reconstructing Probability Tables

We are given a **measurement structure** $\mathcal{M}$:

An **empirical model** for $\mathcal{M}$ 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 structure** $\mathcal{M}$:

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

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

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

# Empirical Models: Reconstructing Probability Tables

We are given a **measurement structure** $\mathcal{M}$:

An **empirical model** for $\mathcal{M}$ 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

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

# Empirical Models: Reconstructing Probability Tables

We are given a **measurement structure** $\mathcal{M}$:

An **empirical model** for $\mathcal{M}$ 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

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

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 structure** $\mathcal{M}$:

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

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

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

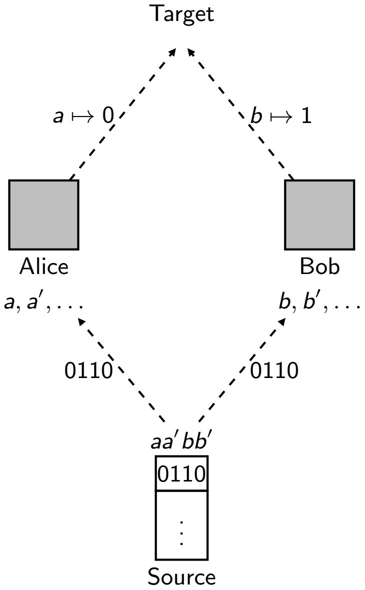
E.g. in the bipartite case, consider $C = \{m_a, m_b\}$, $C' = \{m_a, m'_b\}$.
Fix $s_0 \in \mathcal{E}(\{m_a\})$. Compatibility implies

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

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.

# Hidden Variables: The Mermin instruction set picture



# 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 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 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 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 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 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**:

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

# Global Sections Subsume Hidden-Variable Theories

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;=\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;\; = \;\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

The joint probabilities determined by $s$ factor as a product of the probabilities assigned to the individual measurements, independent of the context in which they appear. This subsumes **Bell locality** .

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;=\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

The joint probabilities determined by $s$ factor as a product of the probabilities assigned to the individual measurements, independent of the context in which they appear. This subsumes **Bell locality** .

So a global section **is** a deterministic local hidden-variable model.

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;=\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

The joint probabilities determined by $s$ factor as a product of the probabilities assigned to the individual measurements, independent of the context in which they appear. This subsumes **Bell locality** .

So a global section **is** a deterministic local hidden-variable model.

The general result is as follows:

**Theorem**

*Any factorizable (i.e. local) hidden-variable model defines a global section.*

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;=\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

The joint probabilities determined by $s$ factor as a product of the probabilities assigned to the individual measurements, independent of the context in which they appear. This subsumes **Bell locality** .

So a global section **is** a deterministic local hidden-variable model.

The general result is as follows:

**Theorem**

*Any factorizable (i.e. local) hidden-variable model defines a global section.*

So:

existence of a local hidden-variable model for a given empirical model
IFF
empirical model has a global section

# Global Sections Subsume Hidden-Variable Theories

Note also that this is a **local** model:

$$\delta_s|C(s') \;=\; \prod_{x \in C} \delta_{s|x}(s'|x).$$

The joint probabilities determined by $s$ factor as a product of the probabilities assigned to the individual measurements, independent of the context in which they appear. This subsumes **Bell locality** .

So a global section **is** a deterministic local hidden-variable model.

The general result is as follows:

**Theorem**

*Any factorizable (i.e. local) hidden-variable model defines a global section.*

Hence:

No such h.v. model exists (the empirical model is **non-local**/**contextual**)
IFF
there is an **obstruction to the existence of a global section**

# Methods for showing obstructions to global sections

# Methods for showing obstructions to global sections

1. S. Abramsky and L. Hardy, Logical Bell Inequalities, *Phys. Rev. A* 85, 062114 (2012).

**Theorem**

*Every Bell inequality is equivalent to a logical Bell inequality.*

# Methods for showing obstructions to global sections

1. S. Abramsky and L. Hardy, Logical Bell Inequalities, *Phys. Rev. A* 85, 062114 (2012).

**Theorem**

*Every Bell inequality is equivalent to a logical Bell inequality.*

2. Linear algebra/programming.
   S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(2011):113036, 2011.

# Methods for showing obstructions to global sections

1. S. Abramsky and L. Hardy, Logical Bell Inequalities, *Phys. Rev. A* 85, 062114 (2012).

**Theorem**

*Every Bell inequality is equivalent to a logical Bell inequality.*

2. Linear algebra/programming.
   S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(2011):113036, 2011.

**Theorem**

*Probabilistic models have local hidden-variable realizations with negative probabilities if and only if they satisfy no-signalling.*

# Methods for showing obstructions to global sections

1. S. Abramsky and L. Hardy, Logical Bell Inequalities, *Phys. Rev. A* 85, 062114 (2012).

**Theorem**

*Every Bell inequality is equivalent to a logical Bell inequality.*

2. Linear algebra/programming.
   S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(2011):113036, 2011.

**Theorem**

*Probabilistic models have local hidden-variable realizations with negative probabilities if and only if they satisfy no-signalling.*

3. Sheaf cohomology.
   S. Abramsky, S. Mansfield and R. Soares Barbosa, The Cohomology of Non-Locality and Contextuality, in *Proc. QPL 2011*, EPTCS v. 95:1–15, 2012.

# 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 non-locality among models:

# A Hierarchy

We can distinguish three degrees of non-locality among models:

- Strong non-locality implies logical non-locality, which implies (probabilistic) non-locality.

# A Hierarchy

We can distinguish three degrees of non-locality among models:

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

# A Hierarchy

We can distinguish three degrees of non-locality among models:

- Strong non-locality implies logical non-locality, which implies (probabilistic) non-locality.
- 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 non-locality among models:

- Strong non-locality implies logical non-locality, which implies (probabilistic) non-locality.
- 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

$$\text{non-locality} < \text{logical non-locality} < \text{strong non-locality}$$

# A Hierarchy

We can distinguish three degrees of non-locality among models:

- Strong non-locality implies logical non-locality, which implies (probabilistic) non-locality.
- 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

$$\text{non-locality} \; < \; \text{logical non-locality} \; < \; \text{strong non-locality}$$

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 non-locality among models:

- Strong non-locality implies logical non-locality, which implies (probabilistic) non-locality.
- 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

$$\text{non-locality} \; < \; \text{logical non-locality} \; < \; \text{strong non-locality}$$

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

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

# GHZ States

# GHZ States

In each finite dimension $n > 2$ we have the GHZ state, written in the $Z$ basis as

$$\frac{|\uparrow \cdots \uparrow\rangle \ + \ |\downarrow \cdots \downarrow\rangle}{\sqrt{2}}.$$

Physically, this corresponds to $n$ particles prepared in a certain entangled state.

# GHZ States

In each finite dimension $n > 2$ we have the GHZ state, written in the $Z$ basis as

$$\frac{|\uparrow \cdots \uparrow\rangle \; + \; |\downarrow \cdots \downarrow\rangle}{\sqrt{2}}.$$

Physically, this corresponds to $n$ particles prepared in a certain entangled state.

If we measure each particle with a choice of $X$ or $Y$ observable, the probability for each outcome is given by the inner product

$$|\langle \mathrm{GHZ} | b_1 \cdots b_n \rangle \rangle|^2.$$

# GHZ States

In each finite dimension $n > 2$ we have the GHZ state, written in the $Z$ basis as

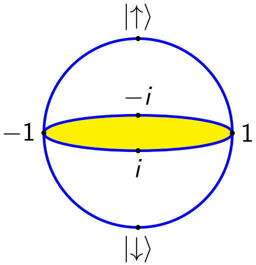
$$\frac{|\uparrow \cdots \uparrow\rangle \; + \; |\downarrow \cdots \downarrow\rangle}{\sqrt{2}}.$$

Physically, this corresponds to $n$ particles prepared in a certain entangled state.

If we measure each particle with a choice of $X$ or $Y$ observable, the probability for each outcome is given by the inner product

$$|\langle \mathrm{GHZ} | b_1 \cdots b_n \rangle\rangle|^2.$$

This computation is controlled by the product of the $|\downarrow\rangle$-coefficients of the basis vectors: cyclic group generated by $i \; \cong \; \mathbb{Z}_4$.



# Logical Specification Of GHZ Models

# Logical Specification Of GHZ Models

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1.

# Logical Specification Of GHZ Models

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1.

For each maximal context $C$, every $s$ in the support of the model satisfies the following conditions:

# Logical Specification Of GHZ Models

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1.

For each maximal context $C$, every $s$ in the support of the model satisfies the following conditions:

- If the number of $Y$ measurements in $C$ is a multiple of 4, the number of 1's in the outcomes specified by $s$ is even.

# Logical Specification Of GHZ Models

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1.

For each maximal context $C$, every $s$ in the support of the model satisfies the following conditions:

- If the number of $Y$ measurements in $C$ is a multiple of 4, the number of 1's in the outcomes specified by $s$ is even.
- If the number of $Y$ measurements is $4k + 2$, the number of 1's in the outcomes is odd.

# Logical Specification Of GHZ Models

The GHZ model of type $(n, 2, 2)$ can be specified as follows. We label the two measurements at each part as $X^{(i)}$ and $Y^{(i)}$, and the outcomes as 0 and 1.

For each maximal context $C$, every $s$ in the support of the model satisfies the following conditions:

- If the number of $Y$ measurements in $C$ is a multiple of 4, the number of 1's in the outcomes specified by $s$ is even.
- If the number of $Y$ measurements is $4k + 2$, the number of 1's in the outcomes is odd.

NB: a model with these properties can be realized in quantum mechanics.

# GHZ Models Are Strongly Contextual

We consider the case where $n = 4k$. Assume for a contradiction that we have a global section.

# GHZ Models Are Strongly Contextual

We consider the case where $n = 4k$. Assume for a contradiction that we have a global section.

If we take Y measurements at every part, the number of $R$ outcomes under the assignment has a parity $P$. Replacing any two $Y$'s by $X$'s changes the residue class mod 4 of the number of $Y$'s, and hence must result in the opposite parity for the number of $R$ outcomes under the assignment.

# GHZ Models Are Strongly Contextual

We consider the case where $n = 4k$. Assume for a contradiction that we have a global section.

If we take Y measurements at every part, the number of $R$ outcomes under the assignment has a parity $P$. Replacing any two $Y$'s by $X$'s changes the residue class mod 4 of the number of $Y$'s, and hence must result in the opposite parity for the number of $R$ outcomes under the assignment.

Thus for any $Y^{(i)}$, $Y^{(j)}$ assigned the **same** value, if we substitute X's in those positions they must receive **different** values. Similarly, for any $Y^{(i)}$, $Y^{(j)}$ assigned different values, the corresponding $X^{(i)}$, $X^{(j)}$ must receive the same value.

# GHZ Models Are Strongly Contextual

We consider the case where $n = 4k$. Assume for a contradiction that we have a global section.

If we take Y measurements at every part, the number of $R$ outcomes under the assignment has a parity $P$. Replacing any two $Y$'s by $X$'s changes the residue class mod 4 of the number of $Y$'s, and hence must result in the opposite parity for the number of $R$ outcomes under the assignment.

Thus for any $Y^{(i)}$, $Y^{(j)}$ assigned the **same** value, if we substitute X's in those positions they must receive **different** values. Similarly, for any $Y^{(i)}$, $Y^{(j)}$ assigned different values, the corresponding $X^{(i)}$, $X^{(j)}$ must receive the same value.

Suppose not all $Y^{(i)}$ are assigned the same value. Then for some i, j, k, $Y^{(i)}$ is assigned the same value as $Y^{(j)}$, and $Y^{(j)}$ is assigned a different value to $Y^{(k)}$. Thus $Y^{(i)}$ is also assigned a different value to $Y^{(k)}$. Then $X^{(i)}$ is assigned the same value as $X^{(k)}$, and $X^{(j)}$ is assigned the same value as $X^{(k)}$. By transitivity, $X^{(i)}$ is assigned the same value as $X^{(j)}$, yielding a contradiction.

# GHZ Models Are Strongly Contextual

We consider the case where $n = 4k$. Assume for a contradiction that we have a global section.

If we take Y measurements at every part, the number of $R$ outcomes under the assignment has a parity $P$. Replacing any two $Y$'s by $X$'s changes the residue class mod 4 of the number of $Y$'s, and hence must result in the opposite parity for the number of $R$ outcomes under the assignment.

Thus for any $Y^{(i)}$, $Y^{(j)}$ assigned the **same** value, if we substitute X's in those positions they must receive **different** values. Similarly, for any $Y^{(i)}$, $Y^{(j)}$ assigned different values, the corresponding $X^{(i)}$, $X^{(j)}$ must receive the same value.

Suppose not all $Y^{(i)}$ are assigned the same value. Then for some i, j, k, $Y^{(i)}$ is assigned the same value as $Y^{(j)}$, and $Y^{(j)}$ is assigned a different value to $Y^{(k)}$. Thus $Y^{(i)}$ is also assigned a different value to $Y^{(k)}$. Then $X^{(i)}$ is assigned the same value as $X^{(k)}$, and $X^{(j)}$ is assigned the same value as $X^{(k)}$. By transitivity, $X^{(i)}$ is assigned the same value as $X^{(j)}$, yielding a contradiction.

The remaining cases are where all Y's receive the same value. Then any pair of X's must receive different values. But taking any 3 X's, this yields a contradiction, since there are only two values, so some pair must receive the same value.

# Degrees of contextuality/non-locality for quantum states

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

In particular, we shall focus on $n$-qubit pure states. If we fix local observables for each party, such a state gives rise to a probability model as above.

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

In particular, we shall focus on $n$-qubit pure states. If we fix local observables for each party, such a state gives rise to a probability model as above.

We can lift the properties of models to states.

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

In particular, we shall focus on $n$-qubit pure states. If we fix local observables for each party, such a state gives rise to a probability model as above.

We can lift the properties of models to states.

- We say that a state is strongly non-local if for **some** choice of local observables for each party, the resulting empirical model is strongly non-local.

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

In particular, we shall focus on $n$-qubit pure states. If we fix local observables for each party, such a state gives rise to a probability model as above.

We can lift the properties of models to states.

- We say that a state is strongly non-local if for **some** choice of local observables for each party, the resulting empirical model is strongly non-local.
- We can similarly define logical non-locality for states; we say that a state is logically non-local if for some choice of local observables, the resulting empirical model is logically non-local; while the state is **not** strongly non-local.

# Degrees of contextuality/non-locality for quantum states

We can lift these concepts to define a novel way of classifying **quantum states** in terms of their **degree of non-locality**.

In particular, we shall focus on *n*-qubit pure states. If we fix local observables for each party, such a state gives rise to a probability model as above.

We can lift the properties of models to states.

- We say that a state is strongly non-local if for **some** choice of local observables for each party, the resulting empirical model is strongly non-local.
- We can similarly define logical non-locality for states; we say that a state is logically non-local if for some choice of local observables, the resulting empirical model is logically non-local; while the state is **not** strongly non-local.
- Finally, a state is weakly non-local if it is non-local, but neither of the previous two cases apply.

# The Characterization Problem

# The Characterization Problem

This gives rise to a natural and challenging problem:

## Problem

*Characterize the multipartite states in terms of their maximum degree of non-locality.*

# The Characterization Problem

This gives rise to a natural and challenging problem:

**Problem**

*Characterize the multipartite states in terms of their maximum degree of non-locality.*

We believe that an answer to this problem will shed considerable light on the structure of multipartite states, not least because it will necessitate solving the following task:

> *Given a multipartite state, find local observables which witness its highest degree of non-locality.*

# The bipartite case is a special case!

# The bipartite case is a special case!

Part of the thinking behind this conjecture is that the bipartite case may actually be **anomalous** within the landscape of multipartite entangled states.

# The bipartite case is a special case!

Part of the thinking behind this conjecture is that the bipartite case may actually be **anomalous** within the landscape of multipartite entangled states.

For example, the only strongly contextual bipartite models are the PR-boxes, which are of course not quantum realizable. By contrast, for all $n > 2$, the $n$-partite GHZ states are strongly contextual.

**Proposition**

*All bipartite entangled states* **except** *the maximally entangled ones are logically non-local.*

# The bipartite case is a special case!

Part of the thinking behind this conjecture is that the bipartite case may actually be **anomalous** within the landscape of multipartite entangled states.

For example, the only strongly contextual bipartite models are the PR-boxes, which are of course not quantum realizable. By contrast, for all $n > 2$, the $n$-partite GHZ states are strongly contextual.

**Proposition**

*All bipartite entangled states* **except** *the maximally entangled ones are logically non-local.*

The negative part of this result is 'folklore', but actually **proved** by Shane Mansfield and Carmen Constantin.

# The bipartite case is a special case!

Part of the thinking behind this conjecture is that the bipartite case may actually be **anomalous** within the landscape of multipartite entangled states.

For example, the only strongly contextual bipartite models are the PR-boxes, which are of course not quantum realizable. By contrast, for all $n > 2$, the $n$-partite GHZ states are strongly contextual.

**Proposition**

*All bipartite entangled states* **except** *the maximally entangled ones are logically non-local.*

The negative part of this result is 'folklore', but actually **proved** by Shane Mansfield and Carmen Constantin.

However, as we shall see, for $n > 2$ a different picture emerges.

# A Striking Result

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

**Theorem**

*For all $n \geq 1$, $P(n)$ and $L(n)$ partition the set of n-qubit pure states.*

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

**Theorem**

*For all $n \geq 1$, $P(n)$ and $L(n)$ partition the set of n-qubit pure states.*

Thus every pure state is either a state whose only form of entanglement is bipartite maximal entanglement in 2-qubit subsystems; or it is logically non-local.

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

**Theorem**

*For all $n \geq 1$, $P(n)$ and $L(n)$ partition the set of n-qubit pure states.*

Thus every pure state is either a state whose only form of entanglement is bipartite maximal entanglement in 2-qubit subsystems; or it is logically non-local.

Forthcoming papers by SA, Carmen Constantin and Shenggang Ying.

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

**Theorem**

*For all $n \geq 1$, $P(n)$ and $L(n)$ partition the set of n-qubit pure states.*

Thus every pure state is either a state whose only form of entanglement is bipartite maximal entanglement in 2-qubit subsystems; or it is logically non-local.

Forthcoming papers by SA, Carmen Constantin and Shenggang Ying.

Constructive proofs of some special cases (SA and Carmen Constantin, QPL 2013).

# A Striking Result

Let $P(n)$ be the class of $n$-qubit pure states which, up to permutation, can be written as tensor products of 1-qubit and 2-qubit maximally entangled states. Let $L(n)$ be the set of logically non-local $n$-qubit states.

**Theorem**

*For all $n \geq 1$, $P(n)$ and $L(n)$ partition the set of n-qubit pure states.*

Thus every pure state is either a state whose only form of entanglement is bipartite maximal entanglement in 2-qubit subsystems; or it is logically non-local.

Forthcoming papers by SA, Carmen Constantin and Shenggang Ying.

Constructive proofs of some special cases (SA and Carmen Constantin, QPL 2013).

The general result is proved non-constructively.

# Permutation Symmetric States

# Permutation Symmetric States

A permutation-symmetric $n$-qubit state is invariant under the action of $S_n$. A natural basis for the permutation-symmetric states is provided by the Dicke states.

# Permutation Symmetric States

A permutation-symmetric $n$-qubit state is invariant under the action of $S_n$. A natural basis for the permutation-symmetric states is provided by the Dicke states.

For each $n \geq 2$, $0 < k < n$ the Dicke state $S(n, k)$ is defined as:

$$S(n,k) \ := \ K \sum_{\text{perm}} |0^k 1^{n-k}\rangle.$$

where $K = \binom{n}{k}^{-1/2}$ is a normalization constant, and we sum over all products of $k$ 0-kets and $n - k$ 1-kets.

# Permutation Symmetric States

A permutation-symmetric $n$-qubit state is invariant under the action of $S_n$. A natural basis for the permutation-symmetric states is provided by the Dicke states.

For each $n \geq 2$, $0 < k < n$ the Dicke state $S(n,k)$ is defined as:

$$S(n,k) \; := \; K \sum_{\text{perm}} |0^k 1^{n-k}\rangle.$$

where $K = \binom{n}{k}^{-1/2}$ is a normalization constant, and we sum over all products of $k$ 0-kets and $n-k$ 1-kets.

The well-known $W$ state is the $S(3,2)$ Dicke state in the above notation.

# Permutation Symmetric States

A permutation-symmetric $n$-qubit state is invariant under the action of $S_n$. A natural basis for the permutation-symmetric states is provided by the Dicke states.

For each $n \geq 2$, $0 < k < n$ the Dicke state $S(n, k)$ is defined as:

$$S(n,k) \; := \; K \sum_{\text{perm}} |0^k 1^{n-k}\rangle.$$

where $K = \binom{n}{k}^{-1/2}$ is a normalization constant, and we sum over all products of $k$ 0-kets and $n - k$ 1-kets.

The well-known $W$ state is the $S(3, 2)$ Dicke state in the above notation.

## Proposition

*For each $n > 2$, and $0 < k < n$, the Dicke state $S(n, k)$ is logically non-local.*

# Permutation Symmetric States

A permutation-symmetric $n$-qubit state is invariant under the action of $S_n$. A natural basis for the permutation-symmetric states is provided by the Dicke states.

For each $n \geq 2$, $0 < k < n$ the Dicke state $S(n,k)$ is defined as:

$$S(n,k) \ := \ K \sum_{\text{perm}} |0^k 1^{n-k}\rangle.$$

where $K = \binom{n}{k}^{-1/2}$ is a normalization constant, and we sum over all products of $k$ 0-kets and $n-k$ 1-kets.

The well-known $W$ state is the $S(3,2)$ Dicke state in the above notation.

## Proposition

*For each $n > 2$, and $0 < k < n$, the Dicke state $S(n,k)$ is logically non-local.*

## Proposition

*All permutation symmetric states are logically non-local.*

# Functionally dependent balanced states

# Functionally dependent balanced states

A balanced $n+1$-qubit quantum state with a functional dependency given by an $n$-ary Boolean function $F : \{0,1\}^n \to \{0,1\}$has the form

$$\Psi_F = \frac{1}{\sqrt{2^n}} \sum_{q_1 q_2 \ldots q_n = 00\ldots 0}^{11\ldots 1} |q_1 q_2 \ldots q_n F(q1, q2, \ldots, q_n)\rangle$$

# Functionally dependent balanced states

A balanced $n+1$-qubit quantum state with a functional dependency given by an $n$-ary Boolean function $F : \{0,1\}^n \rightarrow \{0,1\}$has the form

$$\Psi_F = \frac{1}{\sqrt{2^n}} \sum_{q_1 q_2 \ldots q_n = 00\ldots0}^{11\ldots1} |q_1 q_2 \ldots q_n F(q1, q2, \ldots, q_n)\rangle$$

## Proposition

*All balanced functionally dependent states are in $P(n)$ or $L(n)$.*