CAT L4: Quantum Non-Locality and Contextuality

Samson Abramsky

Department of Computer Science, University of Oxford
Non-Locality and Contextuality

The concepts of non-locality and contextuality play a central rôle in quantum foundations: Bell’s theorem, the Kochen-Specker theorem etc.

They also play an important rôle in quantum information: entanglement as a resource, now contextuality as a resource . . .

- These notions are **not** inherently quantum-mechanical in nature. Indeed, the importance of Bell’s theorem is that it is about the entire space of physical theories. We shall study non-locality and contextuality in a general setting.
- The structures we shall expose arise in many different contexts: from quantum mechanics to relational databases, (in)dependence logics, and social choice.

We use the mathematical language of **sheaf theory**. We show that non-locality and contextuality can be characterized precisely in terms of the existence of **obstructions to global sections**.

We give linear algebraic methods for **computing** these obstructions.

Direct path from sheaf theory to computing global sections using Mathematica™!
The Basic Scenario

Alice

Bob
The Stern-Gerlach Experiment
A Probabilistic Model Of An Experiment

Example: The Bell Model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>(0, 0)</th>
<th>(1, 0)</th>
<th>(0, 1)</th>
<th>(1, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$b_1$</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>$a_1$</td>
<td>$b_2$</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$b_1$</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$b_2$</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
</tbody>
</table>

Important note: this is **physically realizable**!

Generated by a quantum state subjected to spin measurements with certain choices of angle.

Extensively tested experimentally. (Shimony: ‘experimental metaphysics’).
Structural properties of probability tables

Constraints between rows: forms of independence.

- **No-signalling**: the probability distribution Alice sees on outcomes of her chosen measurement cannot depend on Bob’s choice of measurement.

  Necessary for consistency with SR (?! : ‘eat my shorts’).

  Satisfied by QM.

- **Locality/non-contextuality**: Probability of joint outcomes of (Alice, Bob) measurement factors as a product of the probabilities observed by Alice and Bob individually (i.e. ‘locally’).

  This is, famously, **not** satisfied by QM (Bell’s theorem).
Compatibility

It may *not* be possible, in general, to perform all measurements together. This is implicit in the idea that each agent makes a choice of measurement from several alternatives; only the measurements which are chosen are actually performed.

If measurements reveal objective properties of the systems being measured, it seems that it should be the case that for any combination of measurements, it makes sense to ask at least for a probability distribution on their possible outcomes, which is consistent with the actually observed outcomes.

Quantum mechanics *denies this*.

Moreover, as we shall see, there are probability tables for which, as a mathematical fact, there is *no* consistent extension to a joint distribution on outcomes; so we must consider certain combinations of measurements as not jointly performable *in principle*, under any physical theory whatever.
Anatomy of a Probability Table

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>a'</td>
<td>b</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>b'</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
<td></td>
</tr>
<tr>
<td>a'</td>
<td>b'</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
<td></td>
</tr>
</tbody>
</table>

The measurement contexts are

\[ \{a, b\}, \{a', b\}, \{a, b'\}, \{a', b'\}. \]

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 \(C\).
The Presheaf of Distributions

We fix a set of measurements $X$, and a set of outcomes $O$.

For each set of measurements $U \subseteq X$, we define $D_R \mathcal{E}(U)$ to be the set of probability distributions on events $s : U \rightarrow O$. Such an event specifies that outcome $s(m)$ occurs for each measurement $m \in U$.

Given $U \subseteq U'$, we have an operation of restriction:

$$D_R \mathcal{E}(U') \rightarrow D_R \mathcal{E}(U) : d \mapsto d|U,$$

where for each $s \in \mathcal{E}(U)$:

$$d|U(s) := \sum_{s' \in \mathcal{E}(U'), s'|U=s} d(s').$$

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 $s$.

Mathematical notes: (i) This is functorial, hence defines a presheaf. (ii) We could vary $R$. 
Empirical Models: Reconstructing Probability Tables

Corresponding to the choices of measurements by agents, or more generally to the idea that it may not be possible to perform all measurements together, we consider a **measurement structure** \( \mathcal{M} \): a family of subsets of \( X \) which covers \( X \), \( \bigcup \mathcal{M} = X \).

The sets \( C \in \mathcal{M} \) are the **measurement contexts**; the sets of measurements which can be performed together.

These are the sets which index the rows of a generalized probability table.

An **empirical model** for \( \mathcal{M} \) is a family \( \{ e_C \}_{C \in \mathcal{M}}, e_C \in \mathcal{D}_R^E(C) \).

Thus each \( e_C \) is a probability distribution on the row indexed by \( C \); it specifies a probability for the events corresponding to the observation of an outcome for each measurement in \( C \).
Compatibility And No-Signalling

We shall consider models \( \{ e_C \mid C \in \mathcal{M} \} \) which are compatible in the sense of agreeing on overlaps: for all \( C, C' \in \mathcal{M} \),

\[
e_C | C \cap C' = e_{C'} | C \cap C'.
\]

This ‘geometric’ compatibility condition corresponds to the physical condition of 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.
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**.

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

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

Hence:
Existence of Global Sections

Linear algebraic method.

Define system of linear equations $\mathbf{M}\mathbf{X} = \mathbf{V}$.

Solutions $\leftrightarrow$ Global sections

Incidence matrix $\mathbf{M}$ (0/1 entries). Depends only on $\mathcal{M}$ and $\mathcal{E}$.

Enumerate $\bigsqcup_{C \in \mathcal{M}} \mathcal{E}(C)$ as $s_1, \ldots, s_p$.

Enumerate $O^X$ as $t_1, \ldots, t_q$.

$$\mathbf{M}[i,j] = 1 \iff t_j|C = s_i \quad (s_i \in \mathcal{E}(C)).$$

Conceptually, boolean matrix representation of the map

$$\mathcal{E}(X) \longrightarrow \prod_{C \in \mathcal{M}} \mathcal{E}(C) :: s \mapsto (s|C)_{C \in \mathcal{M}}.$$

Bell scenarios $(n, k, l)$: matrix is $(kl)^n \times l^{kn}$.

Incidence matrix for $(2, 2, 2)$ is $16 \times 16$. 
The (2, 2, 2) Incidence Matrix

\[
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

This matrix has rank 9.
The Linear System

A model $e$ determines a vector $V = [e(s_1), \ldots, e(s_p)]$.

Solve

$$MX = V$$

for $X$ over the semiring $R$.

The solution yields weights in $R$ for the global assignments in $O^X$; i.e. a distribution in $D_R\mathcal{E}(X)$.

The equations enforce the constraints that this distribution marginalizes to yield the probabilities of the empirical model.

Hence solutions correspond exactly to global sections — which as we have seen, correspond exactly to local hidden-variable realizations!
The Bell Model

<table>
<thead>
<tr>
<th></th>
<th>(0, 0)</th>
<th>(1, 0)</th>
<th>(0, 1)</th>
<th>(1, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a, b)</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>(a', b)</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
<tr>
<td>(a, b')</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
<tr>
<td>(a', b')</td>
<td>3/8</td>
<td>1/8</td>
<td>1/8</td>
<td>3/8</td>
</tr>
</tbody>
</table>

Solutions in the non-negative reals: this corresponds to solving the linear system over $\mathbb{R}$, subject to the constraint that $X \geq 0$ (linear programming problem).
Bell’s Theorem

**Proposition**  
The Bell model has no global section.

**Proof**  
We focus on 4 out of the 16 equations, corresponding to rows 3, 7, 11 and 14 of the incidence matrix. We write $X_i$ rather than $X[i]$.

\[
\begin{align*}
X_9 + X_{10} + X_{11} + X_{12} &= 1/2 \\
X_9 + X_{11} + X_{13} + X_{15} &= 1/8 \\
X_3 + X_4 + X_{11} + X_{12} &= 1/8 \\
X_2 + X_6 + X_{10} + X_{14} &= 1/8
\end{align*}
\]

Adding the last three equations yields

\[
X_2 + X_3 + X_4 + X_6 + X_9 + X_{10} + 2X_{11} + X_{12} + X_{13} + X_{14} + X_{15} = 3/8.
\]

Since all these numbers must be non-negative, the left-hand side of this equation must be greater than or equal to the left-hand side of the first equation, yielding the required contradiction.

□
The Hardy Model

We consider the possibilistic version of the Hardy model, specified by the following table.

<table>
<thead>
<tr>
<th></th>
<th>(0, 0)</th>
<th>(1, 0)</th>
<th>(0, 1)</th>
<th>(1, 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((a, b))</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>((a', b))</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>((a, b'))</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>((a', b'))</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

This is obtained from a standard probabilistic Hardy model by replacing all positive entries by 1; thus it can be interpreted as the support of the probabilistic model.

Now we are interested in solutions over the boolean semiring, i.e. a boolean satisfiability problem. E.g. the equation specified by the first row of the incidence matrix gives the clause

\[ X_1 \lor X_2 \lor X_3 \lor X_4 \]

while the fifth yields the formula

\[ \neg X_1 \land \neg X_3 \land \neg X_5 \land \neg X_7. \]
The ‘Hardy paradox’

A solution is an assignment of boolean values to the variables which simultaneously satisfies all these formulas. Again, it is easy to see by a direct argument that no such assignment exists.

Proposition

The possibilistic Hardy model has no global section over the booleans.

Proof

We focus on the four formulas corresponding to rows 1, 5, 9 and 16 of the incidence matrix:

\[
\begin{align*}
X_1 \lor X_2 \lor X_3 \lor X_4 \\
\neg X_1 \land \neg X_3 \land \neg X_5 \land \neg X_7 \\
\neg X_1 \land \neg X_2 \land \neg X_9 \land \neg X_{10} \\
\neg X_4 \land \neg X_8 \land \neg X_{12} \land \neg X_{16}
\end{align*}
\]

Since every disjunct in the first formula appears as a negated conjunct in one of the other three formulas, there is no satisfying assignment.
Boolean obstructions are stronger than probabilistic ones

**Proposition**

Let $\mathbf{V}$ be the vector over $\mathbb{R}_{\geq 0}$ for a probabilistic model, $\mathbf{V}_b$ the boolean vector obtained by replacing non-zero elements of $\mathbf{V}$ by 1. If $\mathbf{M} \mathbf{X} = \mathbf{V}$ has a solution over $\mathbb{R}_{\geq 0}$, then $\mathbf{M} \mathbf{X} = \mathbf{V}_b$ has a solution over the booleans.

**Proof**

Simply because $0 \mapsto 0, \quad r > 0 \mapsto 1$

is a semiring homomorphism.

So:

non-existence of solution over booleans

$\Rightarrow$

non-existence of solution over $\mathbb{R}_{\geq 0}$

Bell: no solution over $\mathbb{R}_{\geq 0}$; solution over the booleans.

Hardy: no solution over the booleans.

Conclusion: Bell $<\text{ Hardy}$. 

---

Samson Abramsky (Department of Computer Science, University of Oxford)
Solutions over $\mathbb{R}$

Distributions over $\mathbb{R}$: signed measures (‘negative probabilities’).
Wigner, Dirac, Feynman, Sudarshan, . . .

This simply involves solving linear system over $\mathbb{R}$ with no additional constraints.

**Theorem**

*Probability models of type $(n, 2, 2)$, for all $n \geq 1$, have local hidden-variable realizations with negative probabilities if and only if they satisfy no-signalling.*

Thus we have a striking equivalence between no-signalling models, and those admitting local hidden-variable realizations with negative probabilities, for all Bell-type $(n, 2, 2)$-scenarios.

Can be extended to $(n, k, l)$ scenarios.

Should extend to arbitrary measurement covers.
Reasons

This result follows from elegant self-similarity properties of the inductively defined incidence matrices $M(n)$:

\[
M(1) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad M(n + 1) = \begin{bmatrix} M(n) & M(n) & 0 & 0 \\ 0 & 0 & M(n) & M(n) \\ M(n) & 0 & M(n) & 0 \\ 0 & M(n) & 0 & M(n) \end{bmatrix}
\]

and of the probability vectors $V$ corresponding to no-signalling models, from which it follows that

\[
\text{rank}(M(n)) = \text{rank}([M(n)|V]) = 3^n.
\]
Example: The ‘Popescu-Rohrlich box’

\[
\begin{array}{c|cccc}
(a, b) & (0, 0) & (1, 0) & (0, 1) & (1, 1) \\
\hline
1/2 & 0 & 0 & 1/2 \\
(a', b) & 1/2 & 0 & 0 & 1/2 \\
(a, b') & 1/2 & 0 & 0 & 1/2 \\
(a', b') & 0 & 1/2 & 1/2 & 0 \\
\end{array}
\]

The PR boxes exhibit super-quantum correlations, and cannot be realized in quantum mechanics.

Example solution:

\[\begin{bmatrix} 1/2, 0, 0, 0, -1/2, 0, 1/2, 0, -1/2, 1/2, 0, 0, 1/2, 0, 0, 0 \end{bmatrix}.\]

This vector can be taken as giving a local hidden-variable realization of the PR box using negative probabilities. Similar explicit realizations can be given for the other PR boxes.
Strong Contextuality

If we wish to maintain a realistic view of the nature of physical reality, then when we measure a system with respect to some quantity, there should be a definite value possessed by the system for this quantity, **independent of the measurement which we perform**.

This value may be influenced by some unseen factors, and hence our measurements yield only frequencies, not certain outcomes. Nevertheless, these definite, objective values should exist.

From this perspective, the following fact is shocking:

It is **not possible to assign definite values to all measurements**, independently of the selected measurement context (*i.e.* the set of measurements which we perform), consistently with the predictions of quantum mechanics.

Equivalently, the model has **no global section compatible with its support**.

Note that this is a very weak requirement: just that **some** assignment is possible. The negative result is correspondingly very strong.
Strong Contextuality

Given an empirical model $e$, we define the set

$$S_e := \{ s \in \mathcal{E}(X) : \forall C \in \mathcal{M}. s|C \in \text{supp}(e_C) \}.$$ 

A consequence of the extendability of $e$ is that $S_e$ is non-empty.

We say that the model $e$ is strongly contextual if this set $S_e$ is empty. Thus strong non-contextuality implies non-extendability.

In fact, it is strictly stronger. The Hardy model, which as we saw in the previous section is possibilistically non-extendable, is not strongly contextual. The Bell model similarly fails to be strongly contextual.

The question now arises: are there models arising from quantum mechanics which are strongly contextual in this sense?

We shall now show that the well-known GHZ models, of type $(n, 2, 2)$ for all $n > 2$, are strongly contextual. This will establish a strict hierarchy

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

of increasing strengths of obstructions to non-contextual behaviour for these salient models.
Spin Measurements

Spin Right or Left along the $x$-axis.
Spin Forward or Back along the $y$-axis.

These directions determine observables $X$ and $Y$.

Note that $X$ and $Y$ do not commute; hence according to quantum mechanics, they are incompatible; they cannot be measured together.
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

\[ |\langle \text{GHZ} | b_1 \cdots b_n \rangle |^2. \]

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

Samson Abramsky (Department of Computer Science, University of Oxford)

CAT L4: Quantum Non-Local and Contextuality

28 / 31
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.

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.
Final Remarks

- Our approach is independent of quantum mechanics, since we aim to study the general structure of physical theories. No Hilbert spaces in this talk!
- Still, all the ideas we have discussed can be represented faithfully in quantum mechanics. Leads to some interesting developments, e.g. a Generalized No-Signalling Theorem.
- A unified approach to non-locality and contextuality. Kochen-Specker theorem also falls within the scope of our theory; it is exactly about the non-existence of global sections.
- The mathematical aspects can be pursued much more deeply. Opens the prospect of applying the powerful tools developed in sheaf theory to the study of quantum (and computational) foundations.
- The same methods and structures can be applied to the study of notions of locality and contextuality in other areas, e.g. relational databases, logics of independence, social choice theory.
- Interplay between abstract mathematics, foundations of physics, and computational exploration.