THE LOGIC IN COMPUTER SCIENCE COLUMN

BY

YURI GUREVICH

Microsoft Research

One Microsoft Way, Redmond WA 98052, USA

gurevich@microsoft.com

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

Samson Abramsky

Department of Computer Science

The University of Oxford

`samson.abramsky@cs.ox.ac.uk`

Abstract

We discuss quantum non-locality and contextuality, emphasising logical and structural aspects. We also show how the same mathematical structures arise in various areas of classical computation.

1 Introduction

In this paper we shall discuss some fundamental concepts in quantum mechanics: **non-locality**, **contextuality** and **entanglement**. These concepts play a central rôle in the rapidly developing field of quantum information, in delineating how quantum resources can transcend the bounds of classical information processing. They also have profound consequences for our understanding of the very nature of physical reality.

Our aim is to present these ideas in a manner which should be accessible to any computer scientist, and which emphasises the logical and structural aspects. We shall also show how the same mathematical structures which arise in our analysis of these ideas appear in a range of contexts in classical computation.

2 Alice and Bob look at bits

We consider the following scenario, depicted in Figure 1. Alice and Bob are agents positioned at nodes of a network. Alice can access local bit registers a_1
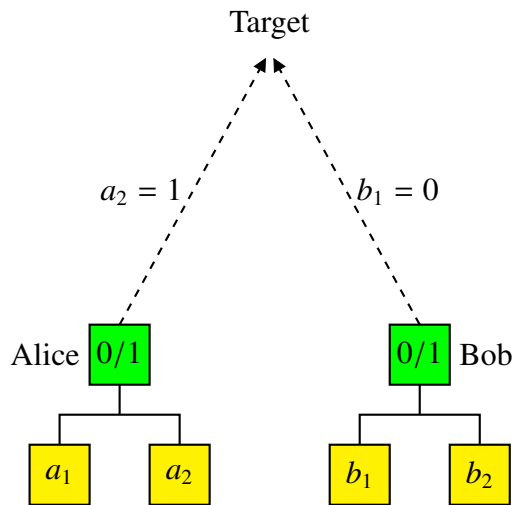
Figure 1: Alice and Bob look at bits

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	$1/8$	$3/8$	$3/8$	$1/8$

Figure 2: The Bell table

and a_2, while Bob can access local bit registers b_1, b_2. Alice can load one of her bit registers into a processing unit, and test whether it is 0 or 1. Bob can perform the same operations with respect to his bit registers. They send the outcomes of these operations to a common target, which keeps a record of the joint outcomes.

We now suppose that Alice and Bob perform repeated rounds of these operations. On different rounds, they may make different choices of which bit registers to access, and they may observe different outcomes for a given choice of register. The target can compile statistics for this series of data, and infer probability distributions on the outcomes. The probability table in Figure 2 records the outcome of such a process.

Consider for example the cell at row 2, column 3 of the table. This corresponds to the following event:
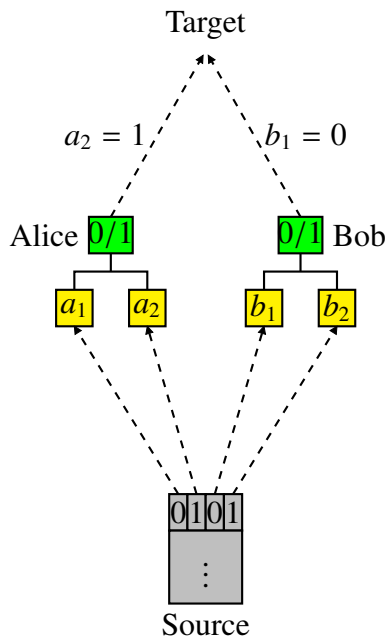
Figure 3: A Source

- Alice loads register a_1 and observes the value 0.

- Bob loads register b_2 and observes the value 1.

This event has the probability 1/8, conditioned on Alice's choice of a_1 and Bob's choice of b_2.

Each row of the table specified a probability distribution on the possible joint outcomes, conditioned on the indicated choice of bit registers by Alice and Bob.

We can now ask:

> How can such an observational scenario be realised?

The obvious classical mechanism we can propose to explain these observations is depicted in Figure 3.

We postulate a **source** which on each round chooses values for each of the registers a_1, a_2, b_1, b_2, and loads each register with the chosen value. Alice and Bob will then observe the values which have been loaded by the source. We can suppose that this source is itself randomised, and chooses the values for the registers according to some probability distribution P on the set of 2^4 possible assignments.

We can now ask the question: is there any distribution P which would give rise to the table specified in Figure 2?

Important Note   A key observation is that, in order for this question to be non-trivial, we must assume that the choices of bit registers made by Alice and Bob are **independent** of the source.[1] If the source could determine which registers are to be loaded on each round, as well as their values, then it becomes a trivial matter to achieve **any** given probability distribution on the joint outcomes.

Under this assumption of independence, it becomes natural to think of this scenario as a kind of **correlation game**. The aim of the source is to achieve as high a degree of correlation between the outcomes of Alice and Bob as possible, whatever the choices made by Alice and Bob on each round.

3   Logic rings a Bell

We shall now make a very elementary and apparently innocuous deduction in elementary logic and probability theory, which could easily be carried out by students in the first few weeks of a Probability 101 course.

Suppose we have propositional formulas $\varphi_1, \ldots, \varphi_N$. We suppose further that we can assign a probability p_i to each φ_i.

In particular, we have in the mind the situation where the boolean variables appearing in φ_i correspond to empirically testable quantities, such as the values of bit registers in our scenario; φ_i then expresses a condition on the outcomes of an experiment involving these quantities. The probabilities p_i are obtained from the statistics of these experiments.

Now suppose that these formulas are **not simultaneously satisfiable**. Then (e.g.)

$$\bigwedge_{i=1}^{N-1} \phi_i \; \rightarrow \; \neg \phi_N, \quad \text{or equivalently} \quad \phi_N \; \rightarrow \; \bigvee_{i=1}^{N-1} \neg \phi_i.$$

Using elementary probability theory, we can calculate:

$$p_N \; \leq \; \mathsf{Prob}(\bigvee_{i=1}^{N-1} \neg \phi_i) \; \leq \; \sum_{i=1}^{N-1} \mathsf{Prob}(\neg \phi_i) \; = \; \sum_{i=1}^{N-1}(1 - p_i) \; = \; (N-1) - \sum_{i=1}^{N-1} p_i.$$

The first inequality is the monotonicity of probability, and the second is sub-additivity.

Hence we obtain the inequality

$$\sum_{i=1}^{N} p_i \; \leq \; N - 1.$$

---

[1]This translates formally into a conditional independence assumption, which we shall not spell out here; see e.g. [14, 17].

We shall refer to this as a **logical Bell inequality**, for reasons to be discussed later. Note that it hinges on a purely logical consistency condition.

3.1 Logical analysis of the Bell table

We return to the probability table from Figure 2.

	$(0,0)$	$(1,0)$	$(0,1)$	$(1,1)$
(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

If we read 0 as true and 1 as false, the highlighted entries in each row of the table are represented by the following propositions:

$$
\begin{aligned}
\varphi_1 &= (a_1 \wedge b_1) \quad \vee \quad (\neg a_1 \wedge \neg b_1) \quad = \quad a_1 \leftrightarrow b_1 \\
\varphi_2 &= (a_1 \wedge b_2) \quad \vee \quad (\neg a_1 \wedge \neg b_2) \quad = \quad a_1 \leftrightarrow b_2 \\
\varphi_3 &= (a_2 \wedge b_1) \quad \vee \quad (\neg a_2 \wedge \neg b_1) \quad = \quad a_2 \leftrightarrow b_1 \\
\varphi_4 &= (\neg a_2 \wedge b_2) \quad \vee \quad (a_2 \wedge \neg b_2) \quad = \quad a_2 \oplus b_2.
\end{aligned}
$$

The events on first three rows are the correlated outcomes; the fourth is anticorrelated. These propositions are easily seen to be jointly unsatisfiable. Indeed, starting with φ_4, we can replace a_2 with b_1 using φ_3, b_1 with a_1 using φ_1, and a_1 with b_2 using φ_2, to obtain $b_2 \oplus b_2$, which is obviously unsatisfiable.

It follows that our logical Bell inequality should apply, yielding the inequality

$$
\sum_{i=1}^{4} p_i \leq 3.
$$

However, we see from the table that $p_1 = 1$, $p_i = 6/8$ for $i = 2, 3, 4$. Hence the table yields a violation of the Bell inequality by $1/4$.

This rules out the possibility of giving an explanation for the observational behaviour described by the table in terms of a classical source. We might then conclude that such behaviour simply cannot be realised. However, as we shall now see, **in the presence of quantum resources, this is no longer the case**.

3.2 A crash course in qubits

We shall now very briefly give enough information about some of the primitives of quantum computing to show how these can be used to realise behaviour such as that in the Bell table in Figure 2. There are a number of excellent introductions to quantum computing aimed at or accessible to computer scientists (see e.g. [31, 36, 32]), and we refer the reader seeking more detailed information to these.

A classical bit register of the kind we began our discussion with can hold the values 0 or 1; we can say that it has two possible states. The operations we can perform on such a register, or an array of such registers, include:

- Reading the value currently held in the register without changing the state of the register.

- Using the values currently held in one or more such registers to compute a new value according to any boolean function.

In quantum computing, we introduce a new object, the qubit, with very different properties. The key features of the qubit are best explained using the beautiful geometric representation in terms of the "Bloch sphere" (the unit 2-sphere), as illustrated in Figure 4.

Note the following key features:

- States of the qubit[2] are represented as points on the surface of the sphere. In Figure 4, a state $|\psi\rangle$ is depicted. Note that there are a continuum of possible states.

- Each pair (Up, Down) of antipodal points on the sphere define a possible measurement that we can perform on the qubit. Each such measurement has two possible outcomes, corresponding to Up and Down in the given direction. We can think of this physically e.g. as measuring Spin Up or Spin Down in a given direction in space.

- When we subject a qubit to a measurement (Up, Down), the state of the qubit determines a probability distribution on the two possible outcomes. For a geometric view on this see Figure 5. The probabilities are determined by the **angles** between the qubit state $|\psi\rangle$ and the points ($|\text{Up}\rangle$, $|\text{Down}\rangle$) which specify the measurement. In algebraic terms, $|\psi\rangle$, $|\text{Up}\rangle$ and $|\text{Down}\rangle$ are unit vectors in the complex vector space $\mathbb{C}^2$, and the probability of observing Up

_____

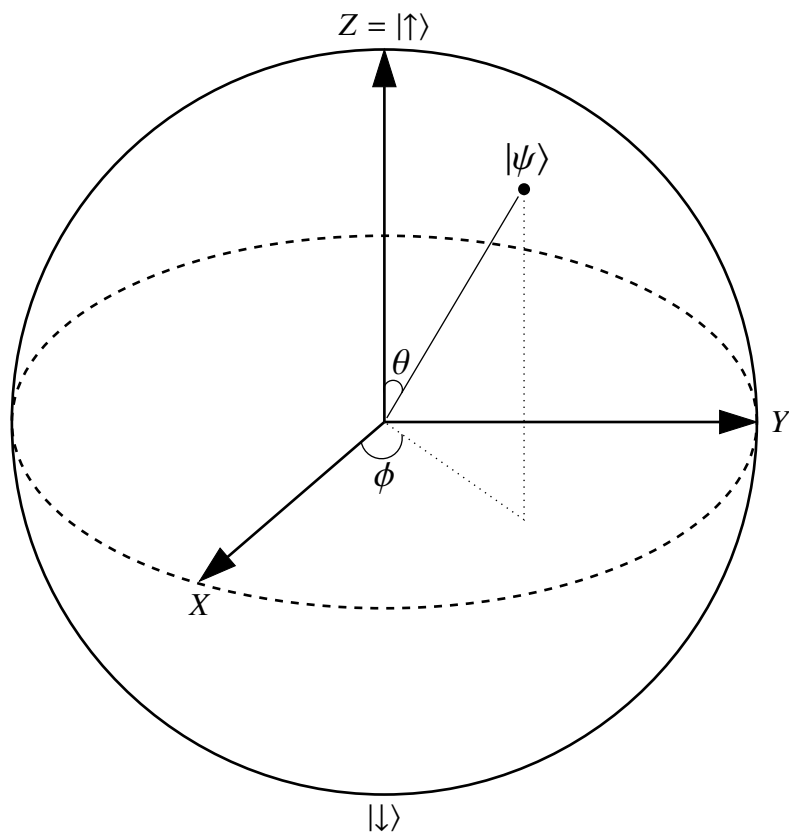[2]More precisely, the pure states; mixed states are represented as points in the interior of the sphere.

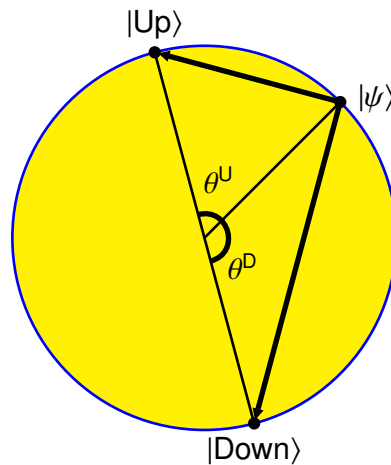Figure 4: The Bloch sphere representation of qubits

Figure 5: Truth makes an angle with reality

when in state $|\psi\rangle$ is given by the square modulus[3] of the inner product:

$$|\langle\psi|\mathsf{Up}\rangle|^2.$$

This is known as the **Born rule**. It gives the basic predictive content of quantum mechanics.

- Note in addition that a measurement has an **effect** on the state, which will no longer be the original state $|\psi\rangle$, but rather one of the states Up or Down, in accordance with the measured value.

The sense in which the qubit generalises the classical bit is that, for each question we can ask — *i.e.* for each measurement — there are just two possible answers. We can view the states of the qubit as superpositions of the classical states 0 and 1, so that we have a probability of getting each of the answers for any given state.

But in addition, we have the important feature that there are a continuum of possible questions we can ask. However, note that on each run of the system, we can only ask **one** of these questions. We cannot simultaneously observe Up or Down in two different directions. Note that this corresponds to the feature of the scenario we discussed in Section 1, that Alice and Bob could only look at one their local registers on each round.

---

[3]Recall that the square modulus of a complex number $z = a + ib$ is given by $|z|^2 = zz^* = (a + ib)(a - ib)$.

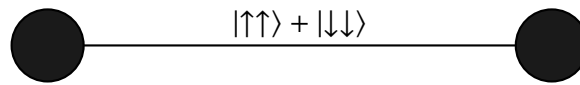$$|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle$$

Figure 6: The Bell state

3.3 Compound systems and entanglement

The deeper features of quantum behaviour are revealed when we look at **compound systems** of multiple qubits.[4] It is here that we find the phenomena of quantum entanglement and non-locality.

Consider for example the 2-qubit system shown in Figure 6. We can think of Alice holding one qubit, and Bob the other. The combined state of the system is described by the vector $|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle$.[5] According to the standard postulates of quantum mechanics, when Alice measures her qubit, she may, with equal probability, get either answer (Spin Up or Down). If she gets the answer Spin Up, then the state of the entangled qubit becomes $|\uparrow\uparrow\rangle$, so that if Bob now measures his qubit, he can **only** get the answer Spin Up; while if she gets the answer Spin Down, the state becomes $|\downarrow\downarrow\rangle$, and Bob can only get the answer Spin Down. This is regardless of the fact that Bob may be far away from Alice (spacelike separated). This is the phenomenon that Einstein famously referred to as "spooky action at a distance", and which Schrödinger named **entanglement**.

How can the world be this way? This remains a challenge to our understanding of the nature of physical reality. Meanwhile, though, the field of quantum information seeks to understand how entanglement can be **used** as a new kind of resource, opening up new possibilities which transcend those of the classical models of information and computation.

3.4 From the Bell state to the Bell table

We refer again to the table in Figure 2. This table can be physically realised, using the Bell state $(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)/\sqrt{2}$, with Alice and Bob performing 1-qubit local measurements corresponding to directions in the XY-plane of the Bloch sphere, at relative angle $\pi/3$. Thus this behaviour is **physically realisable** using quantum entanglement, although, as we have seen, it has no realisation by means of a classical source.

---

[4]More generally, we can consider d-dimensional quantum systems for any positive integer d. A system of n qubits has dimension 2^n. Contextuality emerges already at dimension $d = 3$.

[5]We are ignoring normalisation constants.

Computing the Bell table

Some readers may find it helpful to see in detail how a table such as that in Figure 2 is computed. We shall now explain this. Nothing following this subsection will depend on this material, so it can safely be skipped.

We shall consider spin measurements lying in the equatorial plane of the Bloch sphere, *i.e.* the XY-plane as shown in Figure 4. For such a measurement at an angle ϕ to the X-axis, the Spin Up outcome is specified by the vector $(|{\uparrow}\rangle + e^{i\phi}|{\downarrow}\rangle)/\sqrt{2}$, while the Spin Down outcome is specified by $(|{\uparrow}\rangle + e^{i(\phi+\pi)}|{\downarrow}\rangle)/\sqrt{2}$. For the X direction itself, we have $\phi = 0$, and these are the vectors $(|{\uparrow}\rangle + |{\downarrow}\rangle)/\sqrt{2}$ and $(|{\uparrow}\rangle - |{\downarrow}\rangle)/\sqrt{2}$ respectively.

We shall use the measurement in the X direction for Alice's measurement a_1 and Bob's measurement b_1; while a_2 and b_2 will be interpreted by the measurements at angle $\phi = \pi/3$ to the X axis. Note that Alice's measurements are applied to the first qubit of the Bell state, while Bob's measurements are applied to the second qubit.

For example, consider the following situation: Alice performs the measurement a_1 on the first qubit and observes the outcome 0 (Spin Up), while Bob performs the measurement b_2 on the second qubit and observes outcome 1 (Spin Down). This corresponds to the cell in row 2, column 3 of the table in Figure 2. This event is represented by taking the tensor product of the vectors representing the outcomes for the local measurements by Alice and Bob on their qubits:

$$\frac{|{\uparrow}\rangle + |{\downarrow}\rangle}{\sqrt{2}} \otimes \frac{|{\uparrow}\rangle + e^{i4\pi/3}|{\downarrow}\rangle}{\sqrt{2}} \quad = \quad \frac{|{\uparrow}{\uparrow}\rangle + e^{i4\pi/3}|{\uparrow}{\downarrow}\rangle + |{\downarrow}{\uparrow}\rangle + e^{i4\pi/3}|{\downarrow}{\downarrow}\rangle}{2}.$$

Call this vector M. The probability of observing this event when performing the joint measurement (a_1, b_2) on the Bell state $B = (|{\uparrow}{\uparrow}\rangle + |{\downarrow}{\downarrow}\rangle)/\sqrt{2}$ is given, using the Born rule, by $|\langle B|M\rangle|^2$. Since the vectors $|{\uparrow}{\uparrow}\rangle, |{\uparrow}{\downarrow}\rangle, |{\downarrow}{\uparrow}\rangle, |{\downarrow}{\downarrow}\rangle$ are pairwise orthogonal, this simplifies to

$$\left|\frac{1 + e^{i4\pi/3}}{2\sqrt{2}}\right|^2 \quad = \quad \frac{|1 + e^{i4\pi/3}|^2}{8}.$$

Using the Euler identity $e^{i\theta} = \cos\theta + i\sin\theta$, we have

$$|1 + e^{i\theta}|^2 \quad = \quad 2 + 2\cos\theta.$$

Hence

$$\frac{|1 + e^{i4\pi/3}|^2}{8} \quad = \quad \frac{2 + 2\cos(4\pi/3)}{8} \quad = \quad \frac{1}{8},$$

the value given in the table in Figure 2. The other entries can be computed similarly.

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

Figure 7: The Hardy Paradox

Summary

More broadly, we can say that this shows that quantum mechanics predicts correlations which exceed those which can be achieved by any classical mechanism. This is the content of **Bell's theorem** [15], a famous result in the foundations of quantum mechanics, and in many ways the starting point for the whole field of quantum information. Moreover, these predictions have been confirmed by many experiments which have been performed [11, 10].

4 The "Hardy Paradox"

We shall now see how the same phenomena manifest themselves in a stronger form, which highlights a direct connection with logic. Consider the table in Figure 7.

This table depicts the same kind of scenario we considered previously. However, the entries are now either 0 or 1. The idea is that a 1 entry represents a positive probability. Thus we are distinguishing only between **possible** (positive probability) and **impossible** (zero probability). In other words, the rows correspond to the **supports** of some (otherwise unspecified) probability distributions. Moreover, only four entries of the table are filled in. Our claim is that just from these four entries, referring only to the supports, we can deduce that there is no classical explanation for the behaviour recorded in the table. Moreover, this behaviour can again be realised in quantum mechanics, yielding a stronger form of Bell's theorem, due to Lucien Hardy [20].[6]

---

[6]For a detailed discussion of realisations of the Bell and Hardy models in quantum mechanics, see Section 7 of [2]. Further details on the Hardy construction can be found in [20, 30].

4.1 What Do "Observables" Observe?

Classically, we would take the view that physical observables directly reflect properties of the physical system we are observing. These are objective properties of the 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 λ 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**, and may seem self-evident. However, this view is **impossible to sustain** in the light of our **actual observations of (micro)-physical reality**.

Consider once again the Hardy table depicted in Figure 7. Suppose there is a function λ which accounts for the possibility of Alice observing value 0 for a_1 and Bob observing 0 for b_1, as asserted by the entry in the top left position in the table. Then this function λ must satisfy

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

Now consider the value of λ at b_2. If $\lambda(b_2) = 0$, then this would imply that the event that a_1 has value 0 and b_2 has value 0 is possible. However, **this is precluded** by the 0 entry in the table for this event. The only other possibility is that $\lambda(b_2) = 1$. Reasoning similarly with respect to the joint values of a_2 and b_2, we conclude, using the bottom right entry in the table, that we must have $\lambda(a_2) = 0$. Thus the only possibility for λ consistent with these entries is

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

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

We are thus forced to conclude that the Hardy models are contextual. Moreover, we can say that they are contextual in a logical sense, stronger than the probabilistic form we saw with the Bell tables, since we only needed information about possibilities to infer the contextuality of this behaviour.

5   Mathematical Structure of Possibility Tables

Consider again a table such as

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

Let us anatomise the structure of this table. There are **measurement contexts**

$$\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}.$$

These are the possible combinations of measurements which can be made together, yielding the directly accessible empirical observations.[7] Each measurement has possible outcomes 0 or 1. More generally, we write O for the set of possible outcomes. Thus for example the matrix entry at row (a_2, b_1) and column $(0, 1)$ indicates the **event**

$$\{a_2 \mapsto 0, \ b_1 \mapsto 1\}.$$

The set of events relative to a context C is the set of functions O^C. Each row of the table specifies a **Boolean distribution** on events O^C for a given choice of measurement context C. Such a Boolean distribution is just a non-empty set of events.

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

$$\{a_1, b_1\}, \quad \{a_2, b_1\}, \quad \{a_1, b_2\}, \quad \{a_2, b_2\}$$

as already discussed.

- To each such set C a boolean distribution (finite non-empy subset) on **local sections** $s : C \to O$, where O is the set of **outcomes**. Each row of the above table specifies such a distribution. Note that this notion of distribution generalises naturally to distributions valued in a **commutative semiring**. We assume that the distributions have finite support, and are normalised (have total weight 1). In our case, we are using the idempotent semiring of the booleans. We use the notation $\mathcal{D}_B(X)$ for the set of boolean distributions on a set X.

---

[7]In quantum mechanics, these correspond to compatible families of observables.
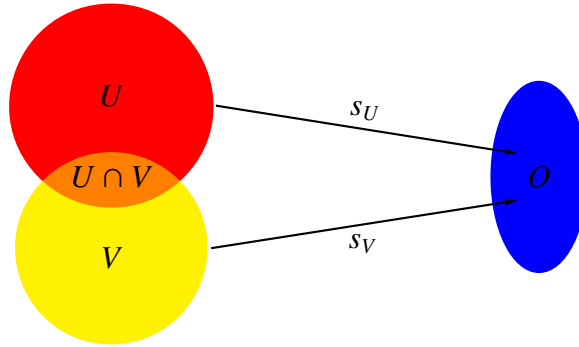
Figure 8: Gluing functions

Note that, if we use the semiring of non-negative reals instead, we obtain probability distributions with finite support.

- A distribution on C restricts to $C' \subseteq C$ by pointwise restriction of the local sections. More precisely, given such a distribution d on O^C, we restrict it to C' by defining, for $s \in O^{C'}$:

$$d|_{C'}(s) \;=\; \sum_{s' \in O^C, s'|_{C'} = s} d(s').$$

This definition makes sense for any semiring. In the boolean case, where addition is disjunction, it can be expressed equivalently as **projection**, where we think of the distribution as a finite set:

$$d|_{C'} \;=\; \{s|_{C'} \;:\; s \in d\}.$$

In the probability case, it gives the usual notion of **marginalisation**.

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**.

For a picture of the familiar and simple situation of gluing functions together, consider the diagram in Figure 8. If $s_U|_{U \cap V} = s_V|_{U \cap V}$, they can be glued to form

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

branch-name	account-no	customer-name	balance
Cambridge	10991-06284	Newton	£2,567.53
Hanover	10992-35671	Leibniz	€11,245.75
. . .	. . .	. . .	. . .

Figure 9: A relation table

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

In geometric language, the Hardy paradox 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 boolean distributions. In other words, the space of **local 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 language is substantiated by the results in [7], which show that **sheaf cohomology** can be used to characterise these obstructions, and to witness contextuality in a wide range of cases.

This geometric picture and the associated methods can be applied to a wide range of situations in classical computer science, which do not seem to have anything in common with the quantum realm. In particular, as we shall now see, there is an isomorphism between the formal description we have given for the quantum notions of non-locality and contextuality, and basic definitions and concepts in relational database theory.

6   Relational Databases and Bell's Theorem

Consider an example of a table in a relational database, as in Figure 9.

Let us anatomise such tables:

- The columns are determined by a set A of **attributes**. Assume $A \subset \mathcal{A}$ for some global set $\mathcal{A}$ specified by the database schema.

- For each attribute a, there is a possible set of **data values** D_a. For simplicity, we collect these into a global set $D = \bigsqcup_{a \in \mathcal{A}} D_a$.

- An **A-tuple** is specified by a function $t : A \to D$.

- A **relation instance** or **table** of schema A is a set of A-tuples.

- A **database schema** is given by a family $\Sigma = \{A_1, \ldots, A_k\}$ of finite subsets of $\mathcal{A}$.

- A database **instance** of schema Σ is given by a family of relation instances $\{R_i\}$ where R_i is of schema A_i.

Does this look familiar? In fact, it is straightforward to express this structure in the language of presheaves:

- An A-tuple t is just a local section over A: $t \in D^A$.

- A relation table R of schema A is a boolean distribution on A-tuples:

$$R \in \mathcal{D}_B(D^A).$$

- Note that if $A \subseteq B$, then restriction is just **projection**. For $R \in \mathcal{D}_B(D^B)$

$$R|_A := \{t|_A \ : \ t \in R\}.$$

- We can regard a schema Σ as a cover of $\mathcal{A}$.

- A database instance of schema Σ is a family of elements $\{R_A\}_{A \in \Sigma}$.

- The compatibility condition for an instance is **projection consistency**:

$$R_A|_{A \cap B} = R_B|_{A \cap B}$$

means that the two relations have the same projections onto their common set of attributes.

6.1 Universal Relations

A **universal relation** for an instance $\{R_A \ : \ A \in \Sigma\}$ of a schema Σ is a relation $R \in \mathcal{D}_B(D^{\mathcal{A}})$ such that, for all $A \in \Sigma$:

$$R|_A = R_A.$$

Thus it is a relation defined on the whole set of attributes $\mathcal{A}$ from which each of the relations in the instance can be recovered by projection.

This notion, and various related ideas, played an important rôle in early developments in relational database theory; see e.g. [27, 19, 24, 26, 34]. Note that a universal relation instance corresponds exactly to the notion of **global section** for the database instance viewed as a compatible family. (Compatibility is obviously a necessary condition for such an instance to exist).

It is also standard that a universal relation need not exist in general, and even if it exists, it need not be unique. There is a substantial literature devoted to the issue of finding conditions under which these properties do hold.

There is a simple connection between universal relations and lossless joins.

Proposition. Let $(R_1, \ldots, R_k)$ be an instance for the schema $\Sigma = \{A_1, \ldots, A_k\}$. Define $R := \bowtie_{i=1}^{k} R_i$. Then a universal relation for the instance exists if and only if $R|_{A_i} = R_i$, $i = 1, \ldots, k$, and in this case R is the largest relation in $\mathcal{R}(\bigcup_i A_i)$ satisfying the condition for a global section. $\square$

We can summarise the striking correspondence we have found between the realms of quantum contextuality and database theory in the following dictionary:

Relational databases	measurement scenarios
attribute	measurement
set of attributes defining a relation table	compatible set of measurements
database schema	measurement cover
tuple	local section (joint outcome)
relation/set of tuples	boolean distribution on joint outcomes
universal relation instance	global section/hidden variable model
acyclicity	Vorob'ev condition [35]

This dictionary goes beyond what we have discussed so far. The last entry concerns Vorob'ev's Theorem [35], a remarkable result motivated by game theory which provides a necessary and sufficient combinatorial condition on a set cover or hypergraph (formulated equivalently in terms of abstract simplicial complexes) such that any compatible family of probability distributions over this cover can be glued together into a global section — a joint distribution on the whole set of vertices which marginalises to yield the given distribution over each simplex. This condition is equivalent to the well-studied notion of **acyclicity** of database schemes [13, 25].

It seems that there is considerable scope for taking these connections and common structures further. For example, we can consider probabilistic databases, and more generally distributions valued in semirings. See [1] for a more detailed discussion.
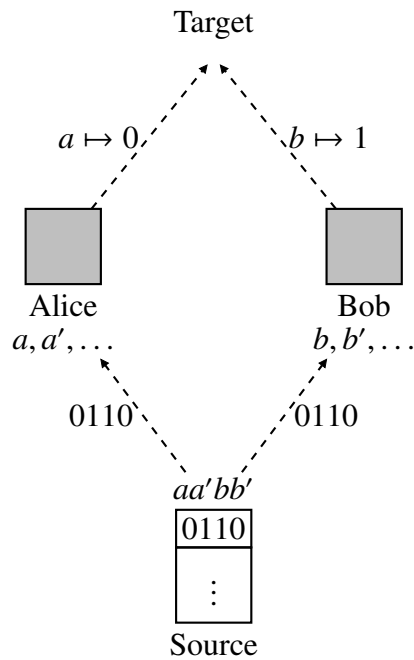
Figure 10: The Mermin instruction set picture

6.2 Hidden variables and all that

We mentioned hidden variable models in the above table, but have not otherwise done so in this article. Traditionally, such models have played a leading rôle in discussions of quantum non-locality and contextuality. Essentially, a local hidden-variable model is what we called a "classical source" in Section 3. Indeed, a standard way of picturing such a model, due to David Mermin [29], is shown in Figure 10. This is essentially the same picture as Figure 3. Mermin calls the hidden variables "instruction sets"; these correspond exactly to the global assignments we have been discussing, which can be considered as canonical forms of hidden variables. It is shown in [3, Theorem 8.1] that these are equivalent to the more general forms of hidden variable models which have been considered in the literature.

6.3 Contextual semantics

Why do such similar structures arise in such apparently different settings? The phenomenon of contextuality is pervasive. Once we start looking for it, we can find it everywhere! Examples already considered include: physics [3], computation [5], and natural language [8].

This leads to what we may call the **Contextual semantics hypothesis**: we can find common mathematical structure in all these diverse manifestations, and develop a widely applicable theory.

7   Kochen-Specker Models

We now return to quantum mechanics, and discuss another fundamental result, the Kochen-Specker theorem [23].[8] This result shows the contextuality of quantum mechanics in an even stronger form than Bell's theorem, in the sense that the argument is independent of any particular quantum state. Whereas our arguments for the Bell and Hardy theorems hinged on realising contextual behaviours using certain entangled quantum states, the Kochen-Specker argument rests on properties of certain families of measurements which hold for **any** quantum state.

There is, however, a trade-off. Whereas the conclusion of the Kochen-Specker theorem is stronger than that of Bell's theorem, its assumptions are also stronger, in that it assumes (for a contradiction) non-contextuality for measurements **in general**. By contrast, Bell's theorem applies to a particular class of measurement scenarios where Alice and Bob are spacelike separated; in these situations, the assumption of non-contextuality is supported by relativistic considerations, which imply that there can be no direct causal influence by the measurements on each other.

The stronger form of state-independent contextuality given by the Kochen-Specker theorem is nevertheless of great interest, and has been the subject of a number of recent experimental verifications [12, 22]. It is also a topic of current interest to develop methods for exploiting contextuality as a resource in quantum information, extending what has been done for non-locality. A feature of our sheaf-theoretic framework, as described in Section 5, is that it provides a unified setting for Bell's theorem, the Kochen-Specker theorem, and other results relating to non-locality and contextuality.

We recall the general setting discussed in Section 5. We have a set X of measurement labels, and a family $\mathcal{U}$ of subsets of X — a "measurement cover". The sets $C \in \mathcal{U}$ are the **measurement contexts**; those combinations of measurements which can be performed together. Formally speaking, $(X, \mathcal{U})$ is just a hypergraph.

For convenience we fix our set of outcomes as $O = \{0, 1\}$. Given $C \in \mathcal{U}$, we say that $s \in O^C$ satisfies the **KS property** if $s(x) = 1$ for exactly one $x \in C$. The **Kochen-Specker model** over $(X, \mathcal{U})$ is defined by setting d_C, for each $C \in \mathcal{U}$, to be the set of all $s \in O^C$ which satisfy the KS property. Note that the model is uniquely determined once we have given $(X, \mathcal{U})$.

---

[8]Since Bell independently proved a version of this result [16], it is often called the Bell-Kochen-Specker theorem.

Note that, if we regard the elements of X as propositional variables, we can think of $s \in O^C$ as a truth-value assignment.[9] Then the KS property for an assignment s is equivalent to s satisfying the following formula:

$$\mathsf{ONE}(C) := \bigvee_{x \in C} (x \ \wedge \bigwedge_{x' \in C \setminus \{x\}} \neg x')$$

We say that the Kochen-Specker model over $(X, \mathcal{U})$ is contextual if there is no global assignment $s : X \to O$ on the whole set of variables X such that $s|_C \in d_C$ for all $C \in \mathcal{U}$. Equivalently, we can say that the model is contextual if the formula

$$\bigwedge_{C \in \mathcal{U}} \mathsf{ONE}(C)$$

is unsatisfiable.[10]

It is interesting to compare this with the property of the Hardy models discussed in Section 5. As we saw there, the contextuality property exhibited by these models was that there was a local section in the support at some $C \in \mathcal{U}$ which was not extendable to a global assignment on X which was compatible with the support. By contrast, the form of contextuality we are considering here is much stronger; that **there is no global assignment at all** which is consistent with the support. In fact, the Hardy models do not satisfy this stronger property.

The simplest example of a contextual Kochen-Specker model is the triangle, *i.e.* the cover

$$\{a, b\}, \{b, c\}, \{a, c\}$$

on $X = \{a, b, c\}$. For a more elaborate example, consider the set $X = \{m_1, \ldots, m_{18}\}$, and the measurement cover $\mathcal{M}$ whose elements are the columns of the following table:

m_1	m_1	m_8	m_8	m_2	m_9	m_{16}	m_{16}	m_{17}
m_2	m_5	m_9	m_{11}	m_5	m_{11}	m_{17}	m_{18}	m_{18}
m_3	m_6	m_3	m_7	m_{13}	m_{14}	m_4	m_6	m_{13}
m_4	m_7	m_{10}	m_{12}	m_{14}	m_{15}	m_{10}	m_{12}	m_{15}

How we do we show that a model such as this is contextual? We shall give a combinatorial criterion on $(X, \mathcal{U})$ which can be used for most of the examples which have appeared in the literature.

---

[9] Interpreting 1 as true and 0 as false.

[10] Note that in the general case where O is some finite set, this becomes a constraint satisfaction problem. Contextuality means that the problem has no solution.

For each $x \in X$, we define

$$\mathcal{U}(x) := \{C \in \mathcal{U} \ : \ x \in C\}.$$

Proposition [3, Proposition 7.1]. If the Kochen-Specker model on $(X, \mathcal{U})$ is non-contextual, then every common divisor of $\{|\mathcal{U}(x)| \ : \ x \in X\}$ must divide $|\mathcal{U}|$. □

Applying this to the above example, we note that the cover $\mathcal{M}$ has 9 elements, while each element of X appears in two members of $\mathcal{M}$. Thus the Kochen-Specker model on $(X, \mathcal{M})$ is contextual.

Quantum representations

What do these combinatorial questions have to do with quantum mechanics? A contextual Kochen-Specker model $(X, \mathcal{U})$ gives rise to a quantum mechanical witness of contextuality whenever we can label X with unit vectors in $\mathbb{R}^n$, for some fixed n, such that $\mathcal{U}$ consists exactly of those subsets C of X which form orthonormal bases of $\mathbb{R}^n$. The point of our example $\mathcal{M}$ above is that it is possible to label the 18 elements of X with vectors in $\mathbb{R}^4$ such that the four-element subsets in $\mathcal{M}$ are orthogonal [18]. This yields one of the most economical known quantum witnesses for contextuality.[11]

To connect this directly to quantum measurement, note that such a family of vectors can be used to define corresponding measurements, such that the measurements corresponding to orthogonal sets are compatible, and moreover for any quantum state $|\psi\rangle$, the support of the distribution on outcomes induced by performing this joint measurement on $|\psi\rangle$ will satisfy the KS property. Thus contextuality of the model yields a state-independent witness of quantum contextuality. For a detailed discussion of this point, see Section 9.2 of [3].

The smallest dimension for which contextuality witnesses appear in this form is $n = 3$. Currently, the smallest known Kochen-Specker model providing a contextuality witness in dimension 3 has 31 vectors [33]. Computational methods have established a lower bound of 18 [9].

8 Discussion and Further Reading

One aim of this paper has been to present some central concepts of quantum information and foundations in a form which will be accessible to computer scientists, in particular those with an interest in logical and structural methods. At the same time, we have also aimed to provide an introduction to recent research by

---

[11]By contrast, the triangle does **not** yield a quantum witness, since orthogonality is a pairwise notion; if all the pairs are orthogonal, the whole set must be also.

the author and a number of colleagues, which aims to use tools which have been developed within computer science logic and semantics to study these quantum notions. This "high-level" approach has led to a number of developments, both within quantum information, and in identifying the same formal structures in a number of classical computational situations; we have seen an example of this in the case of relational database theory.

We shall conclude by discussing some references where the interested reader can find further information, and see these ideas developed in greater depth.[12]

8.1 The sheaf-theoretic approach

As discussed briefly in Section 5, our analysis of non-locality and contextuality uses the mathematical framework of sheaves and presheaves. The issue of finding "local realistic" explanations of correlated behaviour is interpreted geometrically in terms of finding global sections in the sense of sheaf theory. These ideas, and many basic results, are developed in the paper [3] with Adam Brandenburger which laid the basis for this approach.

This leads to a number of developments in quantum information and foundations:

- The sheaf-theoretic language allows a unified treatment of non-locality and contextuality, in which results such as Bell's theorem [15] and the Kochen-Specker theorem [23] fit as instances of more general results concerning obstructions to global sections. In recent work [28], it has been shown how this framework can be used to **transform** contextuality scenarios into non-locality scenarios.

- A hierarchy of degrees of non-locality or contextuality is identified in [3]. This explains and generalises the notion of "inequality-free" or "probability-free" non-locality proofs, and makes a strong connection to logic, as developed in [2]. This hierarchy is lifted to a novel classification of multipartite entangled states, leading to some striking new results concerning multipartite entanglement, which is currently poorly understood. These results will appear in forthcoming joint publications with Carmen Constantin and Shenggang Ying.

- The obstructions to global sections witnessing contextuality are characterised in terms of sheaf cohomology in [7] with Shane Mansfield and Rui Barbosa, and a range of examples are treated in this fashion.

---

[12]The papers by the author which are referenced can be found at `arXiv.org`.

- A striking connection between no-signalling models and global sections with signed measures ("negative probabilities") is established in [3]. An operational interpretation of such negative probabilities, involving a signed version of the strong law of large numbers, is developed in [4].

8.2 Logical Bell inequalities

The discussion in Section 3 is based on [6]. Bell inequalities are a central technique in quantum information. In [6] with Lucien Hardy, a general notion of "logical Bell inequality", based on purely logical consistency conditions, is introduced, and it is shown that every Bell inequality (*i.e.* every inequality satisfied by the "local polytope") is equivalent to a logical Bell inequality. The notion is developed at the level of generality of [3], and hence applies to arbitrary contextuality scenarios, including multipartite Bell scenarios and Kochen-Specker configurations.

8.3 Contextual semantics in classical computation

We discussed the isomorphism between the basic concepts of quantum contextuality and those of relational database theory in Section 6. A number of other connections have been studied:

- In [2] connections between non-locality and logic are emphasised. A number of natural complexity and decidability questions are raised in relation to non-locality.

- Our discussion of the Hardy paradox in Section 5 showed that the key issue was that a local section (assignment of values) could not be extended to a global one consistently with some constraints (the "support table"). This directly motivated some joint work with Georg Gottlob and Phokion Kolaitis [5], in which we studied a refined version of **constraint satisfaction**, dubbed "robust constraint satisfaction", in which one asks if a partial assignment of a given length can always be extended to a solution. The tractability boundary for this problem is delineated in [5], and this is used to settle one of the complexity questions posed in [2].

- Application of the contextual semantics framework to natural language semantics is initiated in [8] with Mehrnoosh Sadrzadeh. In this paper, a basic part of the Discourse Representation Structure framework [21] is formulated as a presheaf, and the gluing of local sections into global ones is used to represent the resolution of anaphoric references.

Further connections and applications of contextual semantics are currently being studied, and it seems likely that more will be forthcoming.

References

[1] S. Abramsky. Relational databases and Bell's theorem. In V. Tannen, L. Wong, L. Libkin, W. Fan, W.C. Tan, and M. Fourman, editors, *In Search of Elegance in the Theory and Practice of Computation: Essays Dedicated to Peter Buneman*, pages 13–35. Springer, 2013.

[2] S. Abramsky. Relational hidden variables and non-locality. *Studia Logica*, 101(2):411–452, 2013.

[3] S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(2011):113036, 2011.

[4] S. Abramsky and A. Brandenburger. An operational interpretation of negative probabilities and no-signalling models. In F. van Breugel, E. Kashefi, C. Palamidessi, and J. Rutten, editors, *Horizons of the Mind: A Tribute to Prakash Panagaden*, pages 59–75. Springer, 2014.

[5] S. Abramsky, G. Gottlob, and P.G. Kolaitis. Robust constraint satisfaction and local hidden variables in quantum mechanics. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, pages 440–446. AAAI Press, 2013.

[6] S. Abramsky and L. Hardy. Logical Bell inequalities. *Physical Review A*, 85(6):062114, 2012.

[7] S. Abramsky, S. Mansfield, and R.S. Barbosa. The cohomology of non-locality and contextuality. *Electronic Proceedings in Theoretical Computer Science*, 95:1–15, 2012.

[8] S. Abramsky and M. Sadrzadeh. Semantic Unification: A sheaf theoretic approach to natural language. In C. Casadio, B. Coecke, M. Moortgat, and P.J. Scott, editors, *Categories and Types in Logic, Language, and Physics, A Festshrift for Jim Lambek*, volume 8222 of *Lecture Notes in Computer Science*, pages 1–13. Springer, 2014.

[9] F. Arends, J. Ouaknine, and C.W. Wampler. On searching for small Kochen-Specker vector systems. In *Graph-Theoretic Concepts in Computer Science*, pages 23–34. Springer, 2011.

[10] A. Aspect. Bell's inequality test: more ideal than ever. *Nature*, 398(6724):189–190, 1999.

[11] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Physical review letters*, 49(25):1804, 1982.

[12] H. Bartosik, J. Klepp, C. Schmitzer, S. Sponar, A. Cabello, H. Rauch, and Y. Hasegawa. Experimental test of quantum contextuality in neutron interferometry. *Physical Review Letters*, 103(4):40403, 2009.

[13] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis. On the desirability of acyclic database schemes. *Journal of the ACM (JACM)*, 30(3):479–513, 1983.

[14] J. Bell. An Exchange on Local Beables. *Dialectica*, 39:85–96, 1985.

[15] J.S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964.

[16] J.S. Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447–452, 1966.

[17] A. Brandenburger and N. Yanofsky. A classification of hidden-variable properties. *Journal of Physics A: Mathematical and Theoretical*, 41:425302, 2008.

[18] A. Cabello, J.M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996.

[19] R. Fagin, A.O. Mendelzon, and J.D. Ullman. A simplified universal relation assumption and its properties. *ACM Transactions on Database Systems (TODS)*, 7(3):343–360, 1982.

[20] L. Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Physical Review Letters*, 71(11):1665–1668, 1993.

[21] H. Kamp and U. Reyle. *From Discourse to Logic: Introduction to model-theoretic semantics of natural language, formal logic and discourse representation theory.* Springer, 1993.

[22] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and CF Roos. State-independent experimental test of quantum contextuality. *Nature*, 460(7254):494–497, 2009.

[23] S. Kochen and E.P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.

[24] H.F. Korth, G.M. Kuper, J. Feigenbaum, A. Van Gelder, and J.D. Ullman. SYSTEM/U: a database system based on the universal relation assumption. *ACM Transactions on Database Systems (TODS)*, 9(3):331–347, 1984.

[25] D. Maier. *The Theory of Relational Databases.* Computer Science Press Rockville, 1983.

[26] D. Maier and J.D. Ullman. Maximal objects and the semantics of universal relation databases. *ACM Transactions on Database Systems (TODS)*, 8(1):1–14, 1983.

[27] D. Maier, J.D. Ullman, and M.Y. Vardi. On the foundations of the universal relation model. *ACM Transactions on Database Systems (TODS)*, 9(2):283–308, 1984.

[28] S. Mansfield and R.S. Barbosa. Extendability in the Sheaf-theoretic Approach: Construction of Bell Models from Kochen-Specker Models. *arXiv preprint arXiv:1402.4827*, 2014.

[29] N.D. Mermin. Quantum mysteries revisited. *Am. J. Phys*, 58(8):731–734, 1990.

[30] N.D. Mermin. Quantum mysteries refined. *American Journal of Physics*, 62:880, 1994.

[31] N.D. Mermin. *Quantum Computer Science*. Cambridge: Cambridge University Press, 2007.

[32] M.Q.C. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[33] A. Peres. *Quantum theory: concepts and methods*, volume 57. Springer, 1995.

[34] J.D. Ullman. *Principles of Database Systems*. Prentice Hall, 1983.

[35] N.N. Vorob'ev. Consistent families of measures and their extensions. *Theory of Probability & Its Applications*, 7(2):147–163, 1962.

[36] N.S. Yanofsky and M.A. Mannucci. *Quantum Computing for Computer Scientists*, volume 20. Cambridge University Press Cambridge, 2008.