Research

Author for correspondence:
Samson Abramsky
e-mail: samson.abramsky@cs.ox.ac.uk

Combining contextuality and causality: a game semantics approach

Samson Abramsky[1], Rui Soares Barbosa[2] and Amy Searle[3]

[1]Department of Computer Science, University College London, 66–72 Gower Street, London WC1E 6EA, UK
[2]INL—International Iberian Nanotechnology Laboratory, Av. Mestre José Veiga, 4715-330 Braga, Portugal
[3]Department of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK

SA, 0000-0003-3921-6637

We develop an approach to combining contextuality with causality, which is general enough to cover causal background structure, adaptive measurement-based quantum computation and causal networks. The key idea is to view contextuality as arising from a game played between Experimenter and Nature, allowing for causal dependencies in the actions of both the Experimenter (choice of measurements) and Nature (choice of outcomes).

This article is part of the theme issue 'Quantum contextuality, causality and freedom of choice'.

1. Introduction

Contextuality is a key non-classical feature of quantum theory. Besides its importance in quantum foundations, it has been linked to quantum advantage in information-processing tasks. It also arises beyond quantum mechanics, cf. [1].

We wish to generalize contextuality to accommodate *causality* and *adaptivity*. These features may arise from:

— fundamental aspects of the physical setting, in particular the causal structure of spacetime;

— the causal structure of an experiment, where measurements are performed in some causal order, and moreover, which measurements are performed may depend on the outcomes of previous measurements;
— feed-forward in measurement-based quantum computation (MBQC) [2], and more generally, adaptive computation.

Our objectives include:

— A more fine-grained analysis of contextuality. Signalling should be allowed from the causal past, i.e. the backward light cone, and thus no-signalling/no-disturbance should be imposed only from *outside* it. This in turn modifies the scope of classicality (non-contextuality), which now becomes relative to this weaker form of no-signalling constraints.
— A better connection with computational models such as circuits and MBQC. Explicitly representing causal flows of information, e.g. outputs of gates feeding into inputs of other gates, enables a deeper analysis of the relationships between contextuality and quantum advantage.

It turns out that capturing these different manifestations of causality and their interactions with contextuality is rather subtle. The perspective we adopt here is to view contextuality as a two-person game played between Experimenter and Nature. The Experimenter's moves are the measurements; i.e. the actions of the Experimenter are to choose the next measurement to be performed. Nature's moves are the outcomes. We can capture the various forms of causal dependency which may arise in terms of *strategies* for Experimenter or for Nature.

The game format is already familiar in the form of non-local games. There, the Verifier plays the role of the Experimenter, and Nature responds with outcomes according to the probability distributions corresponding to Alice–Bob strategies. Non-local games are one-shot games, with a single round of interaction. By considering more general games, causal structure can be incorporated.

Our treatment builds upon the sheaf-theoretic approach to contextuality. A pleasing feature is that once one modifies the basic sheaf of events to take causal structure into account, the further definitions and treatment of contextuality follow *automatically*. This illustrates the advantages of a compositional and functorial approach.

2. Previous work

Pearl had already noted the connection with Bell inequalities in his seminal paper on testability of causal models with latent and instrumental variables [3]. The extension of causal networks to allow for quantum resources, or more generally the operations offered by Generalized Probabilistic Theories, has been studied e.g. in [4,5].

Our starting point is the sheaf-theoretic treatment of contextuality introduced in [6], and extensively developed subsequently. This is a general, mathematically robust approach, which provides a basis for:

— the contextual fraction as a measure of contextuality [7];
— a general characterization of non-contextuality inequalities in terms of consistency conditions ('logical Bell inequalities', Boole's 'conditions of possible experience') [8,9];
— resource theory of contextuality, and simulations between contextual systems [7,10–13];
— cohomological criteria for contextuality, the topology of contextuality [14–16];
— connections with logic and computation, database theory, constraint satisfaction [17,18];
— generalizations [19,20] and applications [21] of Vorob'ev's theorem [22].

The aim is to develop a refined version incorporating causality for which all these features will carry over.

There have been some prior works in this direction:

— Mansfield [23] introduced a refinement of the sheaf-theoretic approach with an order on the measurements, and used it to study the two-slit experiment and the Leggett–Garg scenario.
— Gogioso & Pinzani [24,25] developed a causal refinement of the sheaf-theoretic approach to non-locality, i.e. for the case of *Bell-type scenarios*. They introduce an order on the sites or agents in the Bell scenario.

In both cases, the order is used to refine the no-signalling or no-disturbance condition which guarantees that joint distributions have consistent marginals. In the presence of causality, signalling is allowed from within the backwards light cone or causal past of an event, and thus no-signalling is only required outside it.

One may contrast this with the Contextuality-by-Default (CbD) approach introduced by Dzhafarov and co-workers [26,27]. In CbD, *every* variable is regarded as contextual, differently labelled in each context. Classicality is characterized by the existence of a joint distribution under which different occurrences of variables with the same 'content' have the same value with the maximum probability consistent with their individual marginals. This allows for the analysis of arbitrary signalling systems, which has applications, e.g. in the behavioural sciences, where signalling is the norm. Moreover, this signalling may in general be impossible to characterize or control.

By contrast, both in the above work by Mansfield and Gogioso–Pinzani (GP) and in the present paper, the aim is to *explicitly describe* a given causal background—which might arise from the structure of an experiment, circuit or physical system—and to characterize contextuality relative to such a background.

In this paper, we extend the scope of previous work in several directions. First, we allow more general dependencies of events on their prior causal histories. In particular, the choice of which measurement to perform can depend on previous outcomes as well as on which measurements have been performed. This is an important feature of MBQC (feed-forward), and more generally of adaptive computation. Second, we extend general contextuality scenarios with causality, not just the non-locality Bell scenarios as in the GP approach. Finally, and most subtly, we recognize the different roles played by Nature and Experimenter in their causal interactions, highlighting an important difference between causal background and adaptivity.

An interesting feature of our approach, in common with that of GP, is that it proceeds essentially by modifying the sheaf of events from Abramsky & Brandenburger [6] to reflect the refined signalling constraints in the presence of causality. Once this has been done, the remainder of the analysis of contextuality follows exactly the same script as in [6]. In particular, the appropriate definition of empirical model, the relaxed no-signalling constraints, and the notion of classicality/non-contextuality follow automatically.

3. Examples

As we have already suggested, causality in relation to contextuality has dual aspects. It may be imposed by Nature, in the form of a causal background against which the contextual behaviour plays out; or it may be imposed by the Experimenter, e.g. to achieve computational effects (adaptive computation). We illustrate these two sources of causality in two basic examples.

(a) Example I: causal background à la GP

Consider a standard bipartite non-locality scenario, e.g. the Bell–CHSH scenario: two experimenters, Alice and Bob, with sets of local measurements I_A and I_B, and outcome sets O_A and O_B. We may think of these as 'inputs' and 'outputs'.

We now introduce a variation, in which we assume that Alice's events *causally precede* those of Bob. Thus Bob's backward light cone includes the events where Alice chooses a measurement and observes an outcome.

Whereas in a standard, causally 'flat' scenario, we would have deterministic outcomes given by functions

$$s_A : I_A \to O_A \quad \text{and} \quad s_B : I_B \to O_B,$$

with these causal constraints, we have functions

$$s_A : I_A \to O_A \quad \text{and} \quad s_B : I_A \times I_B \to O_B.$$

That is, the responses by Nature to Bob's measurement may depend on the previous measurement made by Alice.[1]

If we have measurements $x_1, x_2 \in I_A$, $y \in I_B$, then $\{(x_1, 0), (y, 0)\}$ and $\{(x_2, 0), (y, 1)\}$ are valid histories in a single deterministic model. If we now go to distributions over such histories, say $d_{\{x,y\}}$ as a distribution over outcomes for the Alice measurement x and the Bob measurement y, then of the usual no-signalling/compatibility equations

$$d_{\{x,y\}}|_{\{x\}} = d_{\{x\}} \tag{3.1}$$

and

$$d_{\{x,y\}}|_{\{y\}} = d_{\{y\}}, \tag{3.2}$$

only (3.1) remains. In fact, $d_{\{y\}}$ is not even defined, since $\{y\}$ is not a 'causally secured' context: the measurement y can never occur on its own without a preceding Alice measurement.

Thus no-signalling is relaxed in a controlled fashion.

(b) Example II: Anders–Browne

The Anders–Browne construction [28] shows how we can use a form of Experimenter-imposed causality to promote two sub-universal computational models (Pauli measurements and mod-2 linear classical processing) to universal MBQC.

It uses the GHZ state as a resource state:

$$\text{GHZ} = \frac{|\uparrow\uparrow\uparrow\rangle + |\downarrow\downarrow\downarrow\rangle}{\sqrt{2}}.$$

Performing local Pauli X and Y measurements, we obtain the following table of possible joint outcomes[2]

	$+++$	$++-$	$+-+$	$+--$	$-++$	$-+-$	$--+$	$---$
XXX	1	0	0	1	0	1	1	0
XYY	0	1	1	0	1	0	0	1
YXY	0	1	1	0	1	0	0	1
YYX	0	1	1	0	1	0	0	1

In terms of parities, i.e. products of ± 1 outputs under the correspondence given by the group isomorphism $\langle \{0, 1\}, \oplus \rangle \cong \langle \{+1, -1\}, \cdot \rangle$, the support satisfies the following equations:

$$
\begin{aligned}
X_1 \quad X_2 \quad X_3 &= +1 \\
X_1 \quad Y_2 \quad Y_3 &= -1 \\
Y_1 \quad X_2 \quad Y_3 &= -1 \\
Y_1 \quad Y_2 \quad X_3 &= -1.
\end{aligned}
$$

---

[1]Note that, in a deterministic model, Nature 'knows' what response it would have given for Alice's measurement, so there is no real dependency on this outcome.

[2]The table shows only the possibilistic information, i.e. the supports of the probability distributions on joint outcomes, which are uniform on each row.

The idea is to use an Experimenter causal flow to implement OR.[3] Taking X as 0, Y as 1, we consider the measurements for Alice and Bob as inputs to an OR gate. We then use the following simple mod-2 linear mapping (XOR on the bit representations) from the Alice–Bob measurements to determine Charlie's measurement:

$$
\begin{array}{ccc}
0,0 & \mapsto & 0 \\
0,1 & \mapsto & 1 \\
1,0 & \mapsto & 1 \\
1,1 & \mapsto & 0
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
X,X & \mapsto & X \\
X,Y & \mapsto & Y \\
Y,X & \mapsto & Y \\
Y,Y & \mapsto & X.
\end{array}
$$

The output of the OR function is read off from the XOR of the three outcome bits.

We draw attention to the following two remarks.

— This example illustrates causality that is purely employed by the Experimenter. From Nature's point of view, it is just the standard (causally flat) GHZ construction.
— The above describes a simplified 'one-shot' implementation of a single OR gate. To represent general logical circuits with embedded OR gates, using this construction as a building block, really requires (classically computed) feed-forward of measurement settings. This means that there is full adaptivity at work, i.e. dependence of measurement choices on prior measurement outcomes.

4. Game semantics of causality

We conceptualize the dual nature of causality as a two-person game, played between Experimenter and Nature:

— Experimenter's moves are measurements to be performed;
— Nature's moves are the outcomes.

By formalizing this, we develop a theory of causal contextuality that recovers:

— the usual theory of contextuality in the 'flat' case,
— the GP theory of non-locality in a causal background,
— MBQC with adaptive computation,
— classical causal networks,

as special cases, and more.

(a) Measurement scenarios

We begin by briefly reviewing some basic ingredients of the sheaf-theoretic formulation of contextuality. For further details, see e.g. [6].

A (flat) measurement scenario is a pair (X, O), where:

— X is a set of *measurements*.
— $O = \{O_x\}_{x \in X}$ is the set of possible *outcomes* for each measurement.

An *event* has the form (x, o), where $x \in X$ and $o \in O_x$. It corresponds to the measurement x being performed, with outcome o being observed.

---

[3] Anders–Browne in fact implemented NAND, using a different version of the GHZ state. Implementing any nonlinear Boolean function in this setting suffices to achieve universality.

Given a set of events s, its domain is the set of measurements performed:

$$\mathrm{dom}(s) := \pi_1 s = \{x \mid \exists o. (x, o) \in s\}.$$

We say that s is *consistent* if $(x, y), (x, y') \in s$ implies $y = y'$. In this case, s defines a function from the measurements in its domain to outcomes.

A consistent set of events is a *section*. We define the *event sheaf* $\mathscr{E}$ over sets of measurements: for each set $U \subseteq X$ of measurements, $\mathscr{E}(U)$ is the set of sections whose domain is U; when $U \subseteq V$, there is a restriction map $\mathscr{E}(V) \to \mathscr{E}(U)$. The functoriality of these restriction maps formalizes the no-disturbance condition, or 'generalized no-signalling', at the level of deterministic models. Generalized no-signalling of probabilistic (or possibilistic) models will then follow automatically when we compose with the appropriate distribution monad, cf. [6].

The sheaf property of the event sheaf—that compatible families of local sections glue together to yield unique global sections—corresponds to the fact that *deterministic models are non-contextual*.[4] When we pass to distributions over the event sheaf, the sheaf property no longer holds, and this is exactly how contextuality arises. More precisely, we extend the measurement scenario to a *contextuality scenario* by specifying a cover of X; a failure of the sheaf property with respect to this cover constitutes a witness to contextuality.

Our general strategy to accommodate causality is to modify the definition of the event sheaf. After this, we essentially follow the same script as above to give an account of contextuality in the causal setting. A similar procedure is followed in [24,25].

(b) Causal measurement scenarios

A causal measurement scenario is a tuple $M = (X, O, \vdash)$, where the additional ingredient is an *enabling relation* that expresses causal constraints. The intended interpretation of $s \vdash x$, where $s \in \bigcup_{U \subseteq X} \mathscr{E}(U)$ is a consistent set of events and $x \in X$ a measurement, is that it is possible to perform x after the events in s have occurred. Note that this constraint refers to the measurement outcomes as well as the measurements that have been performed. This allows adaptive behaviours to be described.

Given such a causal measurement scenario M, we use it to generate a set of *histories*. A history is a set of events that can happen in a causally consistent fashion. We associate each measurement x with a unique event occurrence, so histories are required to be consistent.

To formalize this, we first define the *accessibility relation* $\rhd$ between consistent sets of events s and measurements x: $s \rhd x$ if and only if $x \notin \mathrm{dom}(s)$ and for some $t \subseteq s$, $t \vdash x$. The intuition is that x may be performed if the events in s have occurred. Now, $\mathscr{H}(M)$, the set of histories over M, is defined inductively as the least family H of consistent sets of events which contains the empty set and is closed under accessibility, meaning that if $s \in H$ and $s \rhd x$, then for all $o \in O_x$, $s \cup \{(x, o)\} \in H$. Note that if a measurement can be performed, then any of its outcomes may occur, forming a valid history.

We can give a more explicit description of $\mathscr{H}(M)$ as a least fixed point. We define an increasing family of sets of histories $\{H_k\}$ inductively:

$$H_0 := \{\varnothing\}$$

and

$$H_{k+1} := H_k \cup \{s \cup \{(x, o)\} \mid s \in H_k, s \rhd x, o \in O_x\}.$$

If X is finite, then for some k we have $H_k = H_{k+1}$, and $\mathscr{H}(M) = H_k$ for the least such k.

---

[4]Note that if we drop no-signalling, as in the CbD approach, this no longer holds.

(c) Strategies

We regard a causal measurement scenario as specifying a game between Experimenter and Nature. Events (x, o) correspond to the Experimenter choosing a measurement x, and Nature responding with outcome o. The histories correspond to the *plays* or runs of the game.

Given this interpretation, we consider the notion of strategy. We focus first on the player Nature, whose strategies one may think of as hidden variables. This is in line with the usual discussion of contextuality, where the experimenter may freely choose which measurements to perform and such choices are beyond the scope of analysis. Later, in §8, we also consider the parallel notion of strategy for the Experimenter, which can express adaptivity.

We define a *strategy for Nature* over the game M as a set of histories $\sigma \subseteq \mathscr{H}(M)$ satisfying the following conditions:

— σ is downwards closed: if $s, t \in \mathscr{H}(M)$ and $s \subseteq t \in \sigma$, then $s \in \sigma$.
— σ is deterministic and total: if $s \in \sigma$ and $s \vartriangleright x$, then there is a unique $o \in O_x$ such that $s \cup \{(x, o)\} \in \sigma$.

Thus at any position s reachable under the strategy σ, the strategy determines a unique response to any measurement that can be chosen by the Experimenter.

We note an important property of strategies.

Proposition 4.1 (Monotonicity). *If $s, t \in \sigma$, $s \subseteq t$, and $s \vartriangleright x$, then*

$$s \cup \{(x, o)\} \in \sigma \Rightarrow t \cup \{(x, o)\} \in \sigma.$$

Proof. Under the given assumptions, since $t \vartriangleright x$, we must have $t \cup \{(x, o')\} \in \sigma$ for some $o' \in O_x$. Since $s \vartriangleright x$, we have that $s \cup \{(x, o')\}$ is a history (i.e. in $\mathscr{H}(M)$), and by down-closure, $s \cup \{(x, o')\} \in \sigma$. Since σ is deterministic, we must have $o = o'$. ∎

Monotonicity says that the outcomes for a measurement x under strategy σ are determined at the minimal histories at which x can occur. This still leaves open the possibility of σ assigning different outcomes to x relative to incomparable causal pasts.

We note another useful property, which follows immediately from totality and determinism.

Proposition 4.2 (Maximality). *If σ, τ are strategies with $\sigma \subseteq \tau$, then $\sigma = \tau$.*

(d) The presheaf of strategies

Given a causal measurement scenario $M = (X, O, \vdash)$ and a set of measurements $U \subseteq X$, we define M_U, the restriction of M to U, as the causal measurement scenario $(U, \{O_x\}_{x \in U}, \vdash_U)$, where $s \vdash_U x$ iff $s \vdash x$ and $\mathrm{dom}(s) \cup \{x\} \subseteq U$. Note that $M_X = M$.

Proposition 4.3. *If $U \subseteq V$, then $\mathscr{H}(M_U)$ is a down-closed subset of $\mathscr{H}(M_V)$ under set inclusion.*

Given a strategy σ over M_V, and $U \subseteq V$, we define $\sigma|_U$, the restriction of σ to U, as the intersection $\sigma|_U := \sigma \cap \mathscr{H}(M_U)$.

Proposition 4.4. *If σ is a strategy over M_V and $U \subseteq V$, then $\sigma|_U$ is a strategy over M_U.*

Proof. The restriction $\sigma|_U$ inherits down-closure from σ. For the second condition, if $s \in \sigma|_U$ and $s \vartriangleright_U x$, then $s \in \sigma$ and $s \vartriangleright_V x$. So, there is a unique $o \in O_x$ such that $s \cup \{(x, o)\} \in \sigma$. But since $x \in U$, we have $s \cup \{(x, o)\} \in \mathscr{H}(M_U)$, and so $s \cup \{(x, o)\} \in \sigma|_U$. ∎

Given a causal measurement scenario $M = (X, O, \vdash)$, we can now define a presheaf

$$\Gamma : \mathcal{P}(X)^{\mathrm{op}} \to \mathbf{Set},$$

of strategies over M. For each $U \subseteq X$, $\Gamma(U)$ is the set of strategies for M_U. Given $U \subseteq V$, the restriction map $\Gamma(U \subseteq V): \Gamma(V) \to \Gamma(U)$ is given by $\sigma \mapsto \sigma|_U$.

The following is immediate:

Proposition 4.5. Γ *is a presheaf.*

(e) Historical note

Causal measurement scenarios are a renaming and repurposing of Kahn–Plotkin information matrices [29], which were introduced *ca* 1975 to represent concrete domains.[5]

We have changed the terminology to reflect the intuitions and applications motivating the present paper:

Kahn–Plotkin	Here
information matrix	causal measurement scenario
cell	measurement
value	outcome
decision	event
configuration	history

The interpretation of causal measurement scenarios as Experimenter–Nature games, the notion of strategy and the presheaf of strategies, are all new to the present paper.

5. Causal contextuality

Our plan now is to follow the script from Abramsky & Brandenburger [6], replacing the event sheaf $\mathscr{E}$ by the presheaf of strategies Γ. Thus local sections are replaced by strategies, whose assignments of outcomes to measurements are sensitive to the previous history of the game.

A *causal contextuality scenario* is a structure $(M, \mathcal{C})$, where $M = (X, O, \vdash)$ is a causal measurement scenario and $\mathcal{C}$ is a cover of X, i.e. a family $\mathcal{C} = \{C_i\}_{i \in I}$ of subsets of measurements $C_i \subseteq X$ satisfying $\bigcup \mathcal{C} = \bigcup_{i \in I} C_i = X$. We work with the presheaf Γ of strategies over M, as described in the previous section.

Recall the distribution monad $\mathscr{D}_R$ from Abramsky & Brandenburger [6], where R is a semiring. When R is the non-negative reals, it yields the usual discrete probability distributions. We construct the presheaf $\mathscr{D}_R \Gamma$, obtained by composing the endofunctor part of the monad with the sheaf of strategies Γ.

An *empirical model* on the scenario $(M, \mathcal{C})$ is a compatible family for the presheaf $\mathscr{D}_R \Gamma$ over the cover $\mathcal{C} = \{C_i\}_{i \in I}$. That is, it is a family $\{e_i\}_{i \in I}$, where $e_i \in \mathscr{D}_R \Gamma(C_i)$, subject to the compatibility conditions: for all $i, j \in I$, $e_i|_{C_i \cap C_j} = e_j|_{C_i \cap C_j}$. Each distribution e_i assigns probabilities to the strategies over M_{C_i}, i.e. to those strategies over M that only perform measurements drawn from the context C_i. As usual, the compatibility conditions require that the marginal distributions agree. This follows the definition of empirical model in [6], replacing the event sheaf by the presheaf of strategies.

The empirical model is *causally non-contextual* if this compatible family extends to a global section of the presheaf $\mathscr{D}_R \Gamma$, i.e. if there is a distribution $d \in \mathscr{D}_R \Gamma(X)$ such that, for all $i \in I$, $d|_{C_i} = e_i$.

If a causal contextuality scenario is finite, then so is the set of histories and therefore that of strategies. The causally non-contextual models thus form a convex polytope, the convex hull of the empirical models on $(M, \mathcal{C})$ corresponding to deterministic strategies $\sigma \in \Gamma(X)$. This is in keeping with the usual setup of 'flat' non-locality and contextuality (i.e. without causality), where such classical polytopes are studied. The classicality of a given model, i.e. membership in this polytope, can be checked by linear programming; and this also suggests a generalization of the contextual fraction [7] to the causal setting.

Similarly, causal contextuality is witnessed by violations of the linear inequalities defining the facets of the polytope. An open question is to find a logical characterization of such inequalities in the spirit of 'logical Bell inequalities' [8].

---

[5]For a historical perspective, see [30].

6. Special cases

To check that these notions make sense, we look at two special cases: flat scenarios and GP scenarios.

(a) Flat scenarios

A contextuality scenario from Abramsky & Brandenburger [6] is $(X, O, \mathcal{C})$. We define the trivial enabling relation where all measurements are initially enabled: $\varnothing \vdash x$ for all $x \in X$. This yields a causal measurement scenario $(M, \mathcal{C})$, where $M = (X, O, \vdash)$.

For any set of measurements $U \subseteq X$, the histories over M_U have support contained in U. Using the monotonicity property and the fact that all measurements are enabled by $\varnothing$, any strategy σ in $\Gamma(U)$ assigns the same outcome to each measurement across all its histories. Hence, it will correspond to a section in $\mathscr{E}(U) = \prod_{x \in U} O_x$. In fact, these will be in bijective correspondence.

Because of this bijective correspondence between Γ and $\mathscr{E}$, we see that the notions of empirical model, global section and contextuality defined for the game-based scenario coincide with the usual notions in this case.

As this example illustrates, the restrictions on which measurements can be performed together are imposed by the cover, not by the causal structure.

(b) GP scenarios

In recent work, Gogioso & Pinzani [24] studied a causal refinement of the sheaf-theoretic approach to non-locality over Bell scenarios.

A GP scenario is given by $((\Omega, \leq), \{I_\omega\}_{\omega \in \Omega}, \{O_\omega\}_{\omega \in \Omega})$, where:

— Ω is a set of sites or agents (Alice, Bob, etc.), with a causal ordering.
— I_ω is the set of inputs (or measurement settings) at ω.
— O_ω is the set of outputs (or measurement outcomes) at ω.

Given such a scenario, we define a causal measurement scenario $M = (X, O, \vdash)$. This mirrors the usual encoding of Bell non-locality scenarios as contextuality scenarios. First, we set:

— $X := \sum_{\omega \in \Omega} \prod_{\omega' \leq \omega} I_{\omega'} = \{(\omega, \mathbf{i}) \mid \omega \in \Omega, \mathbf{i} \in \{\mathbf{i}(\omega') \in I_{\omega'}\}_{\omega' \leq \omega}\}$;
— $O_{(\omega, i)} := O_\omega$.

Given a set of events

$$s = \{((\omega_1, \mathbf{i}_1), o_1), \ldots, ((\omega_n, \mathbf{i}_n), o_n)\}$$

and a measurement $(\omega, \mathbf{i}) \in X$, we define $s \vdash (\omega, \mathbf{i})$ if and only if the support of s has a measurement for each site strictly preceding ω, i.e. $\{\omega_1, \ldots, \omega_n\} = \{\omega' \in \Omega \mid \omega' < \omega\}$, and moreover $\mathbf{i}(v) = \mathbf{i}_j(v)$ for all $v \leq \omega_j$. The vector $\mathbf{i}$ thus encodes all prior choices, as Nature's strategies are allowed to depend on them. So, a measurement $(\omega, \mathbf{i})$ can only be played after a measurement from each site in the causal past of ω has been played. Consequently, the support of any history is a set of measurements per site for some lower subset $\lambda \subseteq \Omega$. This corresponds to the usual notion of context for Bell scenarios, refined to ensure that such contexts are 'causally secured'.

We consider a simple example to illustrate the comparison between Γ defined over $(X, O, \vdash)$, and the 'sheaf of sections' from Gogioso & Pinzani [24].

We take Ω to be the two-chain $\omega_1 < \omega_2$. This is a variation on a standard bipartite Bell–CHSH type scenario, with Alice causally preceding Bob, and hence allowed to signal to Bob. We take the standard Bell scenario cover, where the maximal contexts correspond to choosing one measurement per site, and focus our analysis on the contexts below the cover.[6]

---

[6]The equivalence between sections of Γ and those of the presheaf from Gogioso & Pinzani [24] actually extends more generally to all subsets of measurements, but this is sufficient to illustrate our main point.

Now consider a strategy $\sigma \in \Gamma(X)$. The non-empty histories in M which are compatible in the standard Bell cover have the form

$$\{(\omega_1, \{\omega_1 \mapsto z_1\}), o_1\} \quad \text{or} \quad \{((\omega_1, \{\omega_1 \mapsto z_1\}), o_1), ((\omega_2, \{\omega_1 \mapsto z_1, \omega_2 \mapsto z_2\}), o_2)\},$$

where $z_i \in \{x, y\}$, $o_i \in \{0, 1\}$, $i = 1, 2$. Using monotonicity, the strategy σ assigns a unique o_1 for each (ω_1, z_1) and a unique o_2 for each (ω_1, z_1) and (ω_2, z_2). Thus σ determines a pair of functions of type

$$(I_{\omega_1} \to O_{\omega_1}) \times (I_{\omega_1} \times I_{\omega_2} \to O_{\omega_2}).$$

This accords with the description given in [24]; see in particular the discussion in §5. It extends to an equivalence between Γ and the sheaf of sections of Gogioso & Pinzani [24].

Thus, if we take the standard Bell cover we obtain the same empirical models and notion of contextuality as in [24]. In an extended version of the present paper, we show that this analysis carries over to general GP scenarios. Hence, we recover the GP theory as a special case of our framework.

7. The sheaf property for the strategy presheaf

The strategy presheaf Γ plays the role in our causal theory of the event sheaf $\mathscr{E}$ in [6]. The sheaf property of $\mathscr{E}$ has some conceptual significance since it shows that for deterministic models local consistency implies global consistency. It is only when we introduce distributions, whether probabilistic or possibilistic, that the sheaf property fails and contextuality arises. This raises the question of whether Γ is also a sheaf.

Let $\{U_i\}_{i \in I}$ be a family of subsets of X covering $U = \bigcup_{i \in I} U_i$. Suppose we are given a compatible family $\{\sigma_i\}_{i \in I}$, with $\sigma_i \in \Gamma(U_i)$ and $\sigma_i|_{U_i \cap U_j} = \sigma_j|_{U_i \cap U_j}$ for all $i, j \in I$. The sheaf property requires that there exists a unique strategy $\sigma \in \Gamma(U)$ such that $\sigma|_{U_i} = \sigma_i$ for all $i \in I$.

From the definition of restriction, if such a gluing σ exists, it must contain the union $\sigma' := \bigcup_{i \in I} \sigma_i$. So, if this σ' happens to be a strategy, by maximality it must be the required unique gluing of the family $\{\sigma_i\}_{i \in I}$. The union of down-closed sets is down-closed. Thus σ' can only fail to be a strategy if determinacy or totality fails. We show that the first of these can never arise.

Proposition 7.1. *If $\{\sigma_i\}_{i \in I}$ is a compatible family for the presheaf Γ, then $\sigma' := \bigcup_{i \in I} \sigma_i$ is deterministic.*

Proof. Suppose that $s \cup \{(x, o_k)\} \in \sigma'$ for $k = 1, 2$. For some $i, j \in I$ we have $s \cup \{(x, o_1)\} \in \sigma_i$ and $s \cup \{(x, o_2)\} \in \sigma_j$. This implies that $\mathrm{dom}(s) \cup \{x\} \subseteq U_i \cap U_j$, and hence $s \cup \{(x, o_1)\} \in \sigma_i|_{U_i \cap U_j}$ and $s \cup \{(x, o_2)\} \in \sigma_j|_{U_i \cap U_j}$. By compatibility and determinacy of σ_i and σ_j, this implies $o_1 = o_2$. ∎

In general, it may not be possible to complete σ' to a total strategy, and if such an extension does exist, it may not be unique. We give simple examples to show how these can happen.

Example 7.2. Fix $X = \{x, y, z\}$, $O_w = \{0, 1\}$ for all $w \in \{x, y, z\}$, and the following enabling relation:

$$\varnothing \vdash x, \quad \varnothing \vdash y, \quad \{(x, 0)\} \vdash z, \quad \{(y, 0)\} \vdash z.$$

Consider the cover consisting of $U_1 := \{x, z\}$ and $U_2 := \{y, z\}$, and take strategies

$$\sigma_1 := \{\varnothing, \{(x, 0)\}, \{(x, 0), (z, 0)\}\} \quad \text{and} \quad \sigma_2 := \{\varnothing, \{(y, 0)\}, \{(y, 0), (z, 1)\}\}.$$

Note that σ_1 and σ_2 are compatible since they both restrict to the empty strategy over $U_1 \cap U_2 = \{z\}$, as the measurement z is not enabled. Similarly, σ_1 and σ_2 are both total. However, $\sigma_1 \cup \sigma_2$ is not total, nor can it be completed to a total strategy. Note that y is accessible from $s = \{(x, 0), (z, 0)\}$ so an extension of s must assign an outcome to y, which must be equal to zero due to downward-closedness. Following the same reasoning for x, we are forced to include both $\{(x, 0), (y, 0), (z, 0)\}$ and $\{(x, 0), (y, 0), (z, 1)\}$ in the strategy, contradicting determinism.

Example 7.3. Fix $X = \{x, y, z\}$, $O_w = \{0, 1\}$ for all $w \in \{x, y, z\}$, and the following enabling relation:

$$\varnothing \vdash x, \qquad \varnothing \vdash y, \qquad \{(x, 0), (y, 0)\} \vdash z.$$

Consider the cover consisting of $U_1 := \{x, z\}$ and $U_2 := \{y, z\}$, and take strategies

$$\sigma_1 := \{\, \varnothing,\ \{(x, 0)\}\,\} \quad \text{and} \quad \sigma_2 := \{\, \varnothing,\ \{(y, 0)\}\,\}.$$

Note that σ_1 and σ_2 are compatible since they both restrict to the empty strategy over $U_1 \cap U_2 = \{z\}$, as the measurement z is not enabled. Similarly, σ_1 and σ_2 are both total, since z is not accessible from any history over U_1 or U_2. However, $\sigma_1 \cup \sigma_2$ is not total, since y is accessible from $\{(x, 0)\}$ and x is accessible from $\{(y, 0)\}$. There is a unique choice of outcomes that can be assigned to these variables leading to restrictions to σ_1 and σ_2 as required for a gluing. Both lead to the history $\{(x, 0), (y, 0)\}$. However, z is now accessible from this history, and there are no constraints on the value assigned to it, so we lose uniqueness.

This example is rather pathological, as it hinges on the inaccessibility of z in the cover, leading to the following question.

Question 7.4. Is there a notion of 'good cover' which implies that gluings exist and are unique?

Note added in proof. We have found a positive answer to this question. If we require that the cover comprises sets of measurements which are causally secured in an appropriate sense, then the sheaf property holds. This will be described in detail in a sequel to the present paper.

8. Experimenter strategies and adaptive computation

The strategies considered so far have been strategies for Nature. These prescribe a response—an outcome—for each measurement that can be chosen by the Experimenter. Using the duality inherent in game theory, there is also a notion of strategy for Experimenter. To formulate this, we use the following observation.

Proposition 8.1. *For a history* $s \in \mathscr{H}(M)$*, the following are equivalent:*

(1) s *is maximal in* $(\mathscr{H}(M), \subseteq)$*;*
(2) *no measurement is accessible from* s*, i.e. for all* $x \in X$*,* $\neg(s \rhd x)$*.*

We now define a *strategy for Experimenter* over the game M to be a set of histories $\tau \subseteq \mathscr{H}(M)$ satisfying the following conditions:

— τ is downwards closed: if $s, t \in \mathscr{H}(M)$ and $s \subseteq t \in \tau$, then $s \in \tau$.
— τ is co-total: if $s \in \tau$ and s is not maximal, then there is a measurement x with $s \rhd x$ such that $s \cup \{(x, o)\} \in \tau$ for some $o \in O_x$. Moreover, for all such x, $s \cup \{(x, o')\} \in \tau$ for all $o' \in O_x$.

Thus at each stage, the strategy determines which measurements may be performed. Note that it may allow more than one measurement, so some nondeterminism remains.

For each such measurement, it must then accept any possible response from Nature. The future choices of the Experimenter can then depend on Nature's responses, allowing for adaptive protocols.

If we are given a strategy for Nature σ and a strategy for the Experimenter τ, we can play them off against each other, resulting in $\langle \sigma \mid \tau \rangle := \sigma \cap \tau$. This is the down-set of a set of maximal histories. This operation can be extended to distributions on strategies, i.e. to *mixed strategies*, in a bilinear fashion.[7]

We refer to strategies for Nature as N-strategies, and to strategies for Experimenter as E-strategies.

---

[7] The extension to mixed strategies hinges on the fact that the distribution monad is commutative.

(a) Anders–Browne revisited

We now show how the Anders–Browne construction of an OR gate discussed in §3b can be formalized using an Experimenter strategy.

First, we have the description of the standard GHZ construction. This is given by a flat measurement scenario with $X = \{A_i, B_j, C_k \mid i, j, k \in \{0, 1\}\}$, and $O_x = \{0, 1\}$ for all $x \in X$. The maximal compatible sets of measurements are all sets of the form $\{A_i, B_j, C_k\}$ with $i, j, k \in \{0, 1\}$, i.e. a choice of one measurement per each site or agent. We regard each measurement as initially enabled. The N-strategies for this scenario form the usual sections assigning an outcome to each choice of measurement for each site, and the GHZ model assigns distributions on these strategies as in the table shown in §3b.

To get the Anders–Browne construction, we consider the E-strategy which initially allows any A or B measurement to be performed, and after a history $\{(A_i, o_1), (B_j, o_2)\}$ chooses the C-measurement $C_{i \oplus j}$. Playing this against the GHZ model results in a strategy that computes the OR function with probability 1.

The full power of adaptivity is required when using this as a building block to implement a more involved logical circuit. Suppose that the output of the OR gate above is to be fed as the first input of a second OR gate, built over a GHZ scenario with measurements labelled $\{A'_i, B'_j, C'_k \mid i, j, k \in \{0, 1\}\}$. The E-strategy implements the first OR gate as above, with any B' measurement also enabled, being a free input. After that, the A'-measurement can be determined: after a history containing $\{(A_i, o_1), (B_j, o_2), (C_{i \oplus j}, o_3)\}$, the E-strategy chooses the A'-measurement $A'_{o_1 \oplus o_2 \oplus o_3}$. The second OR gate is then implemented like the first. Note that the choice of A'-measurement depends not only on previous measurement choices, but on outcomes provided by Nature.

9. Outlook

In a forthcoming extended version of this paper, we show how a number of additional examples, including Leggett–Garg, can be handled in our approach. We also show that our formalism faithfully represents a number of others, including GP scenarios, adaptive MBQC and causal networks. In forthcoming related work, we incorporate a form of memory (or look-back) restriction in some simple scenarios whereby Nature may only remember the k most recent events, and obtain a Vorob'ev-type theorem [19,22] in that setup.

In future work, we aim to employ our formalism to describe unconditional quantum advantage in shallow circuits, building on [31,32]. We will also investigate other applications to quantum advantage.

We also aim to clarify how our approach can be related to the currently very active study of indefinite causal orders [33,34].

The game formulation opens up the possibility of tapping into the sophisticated literature in game theory, e.g. on properties of game trees [35,36]. This is likely to offer important pointers for further development of our framework. For example, the concepts of randomized strategies (and the relationship between global and local randomizations) or memory properties are likely to be fruitful. A similar source of potential inspiration is the literature on game semantics, where game concepts are used to model a wide array of programming language features; see e.g. [37] for a recent overview.

References

1. Dzhafarov EN, Jordan JS, Zhang R, Cervantes VH. 2015 *Contextuality from quantum physics to psychology*, vol. 6. Singapore: World Scientific.
2. Briegel HJ, Browne DE, Dür W, Raussendorf R, Van den Nest M. 2009 Measurement-based quantum computation. *Nat. Phys.* **5**, 19–26. (doi:10.1038/nphys1157)
3. Pearl J. 1995 On the testability of causal models with latent and instrumental variables. In *UAI '95: Proc. of the Eleventh Annual Conf. on Uncertainty in Artificial Intelligence, Montreal, Quebec, Canada, August 18–20, 1995* (eds P Besnard, S Hanks), pp. 435–443. Cambridge, MA: Morgan Kaufmann.
4. Henson J, Lal R, Pusey MF. 2014 Theory-independent limits on correlations from generalized Bayesian networks. *New J. Phys.* **16**, 113043. (doi:10.1088/1367-2630/16/11/113043)
5. Chaves R, Carvacho G, Agresti I, Di Giulio V, Aolita L, Giacomini S, Sciarrino F. 2018 Quantum violation of an instrumental test. *Nat. Phys.* **14**, 291–296. (doi:10.1038/s41567-017-0008-5)
6. Abramsky S, Brandenburger A. 2011 The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.* **13**, 113036. (doi:10.1088/1367-2630/13/11/113036)
7. Abramsky S, Barbosa RS, Mansfield S. 2017 Contextual fraction as a measure of contextuality. *Phys. Rev. Lett.* **119**, 050504. (doi:10.1103/PhysRevLett.119.050504)
8. Abramsky S, Hardy L. 2012 Logical Bell inequalities. *Phys. Rev. A* **85**, 062114. (doi:10.1103/PhysRevA.85.062114)
9. Abramsky S. 2020 Classical logic, classical probability, and quantum mechanics. In *Quantum, probability, logic: The work and influence of Itamar Pitowsky* (eds H Meir, O Shenker), pp. 1–17. Cham, Switzerland: Springer International Publishing.
10. Karvonen M. 2019 Categories of empirical models. In *15th Int. Conf. on Quantum Physics and Logic* (*QPL 2018*), Halifax, Nova Scotia, 3-7 June 2018 (eds P Selinger, G Chiribella), vol. 287, *Electronic Proceedings in Theoretical Computer Science*, pp. 239–252. (http://arXiv.org/abs/1901.09476v1)
11. Abramsky S, Barbosa RS, Karvonen M, Mansfield S. 2019 A comonadic view of simulation and quantum resources. In *34th Annual ACM/IEEE Symp. on Logic in Computer Science* (*LICS*), Vancouver, BC, 24-27 June 2019, pp. 1–12. New York, NY: IEEE.
12. Karvonen M. 2021 Neither contextuality nor nonlocality admits catalysts. *Phys. Rev. Lett.* **127**, 160402. (doi:10.1103/PhysRevLett.127.160402)
13. Barbosa RS, Karvonen M, Mansfield S. 2021 Closing Bell: boxing black box simulations in the resource theory of contextuality. (http://arxiv.org/abs/2104.11241)
14. Abramsky S, Mansfield S, Barbosa RS. 2012 The cohomology of non-locality and contextuality. In *8th Int. Workshop on Quantum Physics and Logic* (*QPL 2011*), Nijmegen, 27-29 October 2011. *Electronic Proceedings in Theoretical Computer Science* (eds B Jacobs, P Selinger, B Spitters), pp. 1–14. (http://arXiv.org/abs/1210.0298v1)
15. Abramsky S, Barbosa RS, Kishida K, Lal R, Mansfield S. 2015 Contextuality, cohomology and paradox. In *24th EACSL Annual Conf. on Computer Science Logic* (*CSL 2015*) (ed S Kreutzer), Berlin, 7-10 September 2015 *Leibniz International Proceedings in Informatics* (*LIPIcs*), vol. 41, pp. 211–228. Wadern, Germany: Schloss Dagstuhl–Leibniz-Zentrum für Informatik.
16. Carù G. 2017 On the cohomology of contextuality. In *13th Int. Conf. on Quantum Physics and Logic* (*QPL 2016*) (eds R Duncan, C Heunen), Glasgow, UK, 6-10 June 2016, *Electronic Proceedings in Theoretical Computer Science*, vol. 236, pp. 21–39. Open Publishing Association.

17. Abramsky S. 2015 Contextual semantics: from quantum mechanics to logic, databases, constraints, and complexity. *Context. Quantum Phys. Psychol.* **6**, 23–50.
18. Conghaile AÓ. 2022 Cohomology in constraint satisfaction and structure isomorphism. (http://arxiv.org/abs/2206.15253)
19. Barbosa RS. 2015 *Contextuality in quantum mechanics and beyond*. DPhil thesis, University of Oxford.
20. Atserias A, Kolaitis PG. 2020 Consistency, acyclicity, and positive semirings. (http://arxiv.org/abs/2009.09488)
21. Barbosa RS. 2014 On monogamy of non-locality and macroscopic averages: examples and preliminary results. In *11th Int. Workshop on Quantum Physics and Logic* (*QPL 2014*) (eds B Coecke, I Hasuo, P Panangaden), Kyoto, Japan, 4-6 June 2014, *Electronic Proceedings in Theoretical Computer Science*, vol. 172, pp. 36–55. Open Publishing Association.
22. Vorob'ev NN. 1962 Consistent families of measures and their extensions. *Theory Probab. Appl.* **7**, 147–163 (English)/153–159 (Russian).
23. Mansfield S. 2017 A unified approach to contextuality and violations of macrorealism. Lecture given at QCQMB workshop, Prague.
24. Gogioso S, Pinzani N. 2021 The sheaf-theoretic structure of definite causality. In *18th Int. Conf. on Quantum Physics and Logic* (*QPL 2021*) (eds C Heunen, M Backens), Gdansk, Poland, 7-11 June 2021, *Electronic Proceedings in Theoretical Computer Science*, vol. 343, pp. 301–324. Open Publishing Association.
25. Gogioso S, Pinzani N. 2023 The topology of causality. (http://arxiv.org/abs/2303.07148)
26. Dzhafarov EN, Kujala JV. 2016 Context–content systems of random variables: the Contextuality-by-Default theory. *J. Math. Psychol.* **74**, 11–33. (doi:10.1016/j.jmp.2016.04.010)
27. Dzhafarov EN, Kujala JV, Cervantes VH. 2016 Contextuality-by-default: a brief overview of ideas, concepts, and terminology. In *Quantum Interaction: 9th Int. Conf. (QI 2015),* Filzbach, Switzerland, July 15–17 2015, *Revised Selected Papers 9*, pp. 12–23. Cham, Switzerland: Springer.
28. Anders J, Browne DE. 2009 Computational power of correlations. *Phys. Rev. Lett.* **102**, 050502. (doi:10.1103/PhysRevLett.102.050502)
29. Kahn G, Plotkin GD. 1993 Concrete domains. *Theor. Comput. Sci.* **121**, 187–277. (doi:10.1016/0304-3975(93)90090-G)
30. Brookes S. 1993 Historical introduction to 'Concrete domains' by G. Kahn and G.D. Plotkin. *Theor. Comput. Sci.* **121**, 179–186. (doi:10.1016/0304-3975(93)90089-C)
31. Bravyi S, Gosset D, König R. 2018 Quantum advantage with shallow circuits. *Science* **362**, 308–311. (doi:10.1126/science.aar3106)
32. Aasnæss S. 2022 Comparing two cohomological obstructions for contextuality, and a generalised construction of quantum advantage with shallow circuits. (http://arxiv.org/abs/2212.09382)
33. Oreshkov O, Costa F, Brukner Č. 2012 Quantum correlations with no causal order. *Nat. Commun.* **3**, 1092. (doi:10.1038/ncomms2076)
34. Chiribella G, D'Ariano GM, Perinotti P, Valiron B. 2013 Quantum computations without definite causal structure. *Phys. Rev. A* **88**, 022318. (doi:10.1103/PhysRevA.88.022318)
35. Kuhn HW. 1953 Extensive games and the problem of information. In *Contributions to the Theory of Games, Volume II* (eds HW Kuhn, AW Tucker), Annals of Mathematics Studies, vol. 28, pp. 193–216. Princeton, NJ: Princeton University Press.
36. Isbell JR. 1958 Finitary games. In *Contributions to the Theory of Games, Volume III* (eds M Drescher, AW Tucker, P Wolfe), Annals of Mathematics Studies, vol. 39, pp. 79–96. Princeton, NJ: Princeton University Press.
37. Murawski AS, Tzevelekos N. 2016 An invitation to game semantics. *ACM SIGLOG News* **3**, 56–67. (doi:10.1145/2948896.2948902)