Generalised Winograd Schema and its Contextuality

Kin Ian Lo[1] Mehrnoosh Sadrzadeh[1] Shane Mansfield[2]

[1]Department of Computer Science, University College London

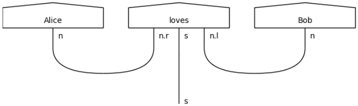
[2]Quandela

Quantum Physics and Logic,
19 July 2023

Introduction

This project is about **quantum** marrying **natural language**.

- Two decades ago, categorial quantum mechanics met Lambek's grammar, DisCoCat was born.



Question

Are there more ways to connect quantum mechanics and language?

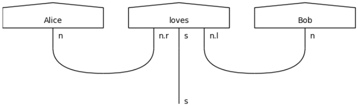
Context is a common theme in both quantum and language:

- **contextuality** is a key feature of quantum that enables some of its computational power,
- disambiguating language requires **context**

Introduction

This project is about **quantum** marrying **natural language**.

- Two decades ago, categorial quantum mechanics met Lambek's grammar, DisCoCat was born.



Question

Are there more ways to connect quantum mechanics and language?

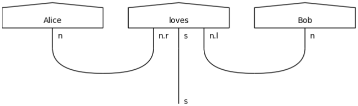
Context is a common theme in both quantum and language:

- **contextuality** is a key feature of quantum that enables some of its computational power,
- disambiguating language requires **context**

Introduction

This project is about **quantum** marrying **natural language**.

- Two decades ago, categorial quantum mechanics met Lambek's grammar, DisCoCat was born.



Question

Are there more ways to connect quantum mechanics and language?

Context is a common theme in both quantum and language:

- **contextuality** is a key feature of quantum that enables some of its computational power,
- disambiguating language requires **context**

What is contextuality?

Quantum is (was) super weird:

- measurements give probabilistic outcomes,
- some measurements are incompatible (uncertainty principle).

People[1] tried to fix it with **non-contextual hidden variable theories**:

- measurements outcomes are pre-existing and deterministic,
- the pre-existing outcomes are independent of the **context**, i.e. what other measurements are performed (non-contextuality),
- a quantum state is a probabilistic mixture of hidden states, explaining the probabilistic nature of quantum mechanics.

[1] Einstein, Podolsky, Rosen, et al.

What is contextuality?

Quantum is (was) super weird:

- measurements give probabilistic outcomes,
- some measurements are incompatible (uncertainty principle).

People[1] tried to fix it with **non-contextual hidden variable theories**:

- measurements outcomes are pre-existing and deterministic,
- the pre-existing outcomes are independent of the **context**, i.e. what other measurements are performed (non-contextuality),
- a quantum state is a probabilistic mixture of hidden states, explaining the probabilistic nature of quantum mechanics.

[1] Einstein, Podolsky, Rosen, et al.

What is contextuality?

Theorem (Bell-Kochen-Specker)

There is no non-contextual hidden variable theory that reproduces the predictions of quantum mechanics.

Proof.

Proof by counterexample. The idea is to construct a quantum system that does not admit a non-contextual hidden variable theory. (details omitted) □

The computational power of contextuality

- Contextuality supplies the magic in magic state distillation
- Magic states are important for making fault-tolerant quantum computing universal

ARTICLE

doi:10.1038/nature13460

Contextuality supplies the 'magic' for quantum computation

Mark Howard[1,2], Joel Wallman[2], Victor Veitch[2,3] & Joseph Emerson[2]

Quantum computers promise dramatic advantages over their classical counterparts, but the source of the power in quantum computing has remained elusive. Here we prove a remarkable equivalence between the onset of contextuality and the possibility of universal quantum computation via 'magic state' distillation, which is the leading model for experimentally realizing a fault-tolerant quantum computer. This is a conceptually satisfying link, because contextuality, which precludes a simple 'hidden variable' model of quantum mechanics, provides one of the fundamental characterizations of uniquely quantum phenomena. Furthermore, this connection suggests a unifying paradigm for the resources of quantum information: the non-locality of quantum theory is a particular kind of contextuality, and non-locality is already known to be a critical resource for achieving advantages with quantum communication. In addition to clarifying these funda-

The computational power of contextuality

- Contextuality lifts (classic) linear computers to non-linear ones in MBQC

Computational Power of Correlations

Janet Anders* and Dan E. Browne†
Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom
(Received 7 May 2008; published 4 February 2009)

We study the intrinsic computational power of correlations exploited in measurement-based quantum computation. By defining a general framework, the meaning of the computational power of correlations is made precise. This leads to a notion of resource states for measurement-based *classical* computation. Surprisingly, the Greenberger-Horne-Zeilinger and Clauser-Horne-Shimony-Holt problems emerge as optimal examples. Our work exposes an intriguing relationship between the violation of local realistic models and the computational power of entangled resource states.

DOI: 10.1103/PhysRevLett.102.050502 PACS numbers: 03.67.Lx, 03.65.Ud, 89.70.Eg

Contextuality in measurement-based quantum computation

Robert Raussendorf*
Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada
(Received 1 May 2013; revised manuscript received 11 July 2013; published 19 August 2013)

We show, under natural assumptions for qubit systems, that measurement-based quantum computations (MBQCs) which compute a nonlinear Boolean function with a high probability are contextual. The class of contextual MBQCs includes an example which is of practical interest and has a superpolynomial speedup over the best-known classical algorithm, namely, the quantum algorithm that solves the “discrete log” problem.

Contextuality in language

The word **bank** has 9 different meanings in the Oxford English Dictionary.

- He sat near the **bank**, reviewing his account details.

- He sat near the **bank**, observing the flowing water.

Images generated on Stable Diffusion Online

Contextuality in language

The word **bank** has 9 different meanings in the Oxford English Dictionary.

- He sat near the **bank**, reviewing his account details.
- He sat near the **bank**, observing the flowing water.

Images generated on Stable Diffusion Online

Contextuality in language

The word **bank** has 9 different meanings in the Oxford English Dictionary.

- He sat near the **bank**, reviewing his account details.

- He sat near the **bank**, observing the flowing water.

Images generated on Stable Diffusion Online

Contextuality in language

Definition (contextuality in language (roughly))

The intended meaning of an ambiguous word depends on the words around it, i.e. the context.

Quantum contextuality vs. contextuality in language

In summary, we have 2 notions of contextuality:

- Quantum contextuality
  - The failure of having a non-contextual hidden variable theory for quantum mechanics
  - Have been shown to possess some computational power
- Contextuality in language
  - In the common usage of language, the meaning of a word depends on the context

Goal

Can we exploit this connection to develop new quantum algorithms for natural language processing?

We constructed two families of linguistic schemas that are "quantum contextual".

Quantum contextuality vs. contextuality in language

In summary, we have 2 notions of contextuality:

- Quantum contextuality
  - The failure of having a non-contextual hidden variable theory for quantum mechanics
  - Have been shown to possess some computational power
- Contextuality in language
  - In the common usage of language, the meaning of a word depends on the context

Goal

Can we exploit this connection to develop new quantum algorithms for natural language processing?

We constructed two families of linguistic schemas that are "quantum contextual".

Quantum contextuality vs. contextuality in language

In summary, we have 2 notions of contextuality:

- Quantum contextuality
  - The failure of having a non-contextual hidden variable theory for quantum mechanics
  - Have been shown to possess some computational power
- Contextuality in language
  - In the common usage of language, the meaning of a word depends on the context

Goal

Can we exploit this connection to develop new quantum algorithms for natural language processing?

We constructed two families of linguistic schemas that are "quantum contextual".

Sheaf-theoretic framework for contextuality

Definition

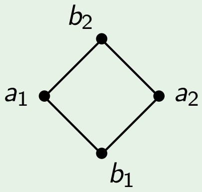
A **measurement scenario** is a tuple $\langle X, \mathcal{M}, O \rangle$ where

- X is a set of observables,
- $\mathcal{M}$ is a set of contexts (compatible observables),
- O is a set of outcomes.

Example

The Bell-CHSH experiments can be described as a measurement scenario $\langle X, \mathcal{M}, O \rangle$ where

- $X = \{a_1, a_2, b_1, b_2\}$;
- $\mathcal{M} = \{\{a_1, b_1\}, \{a_1, b_2\}, \{a_2, b_1\}, \{a_2, b_2\}\}$;
- $O = \{+1, -1\}$.



Empirical model

A measurement scenario $\langle X, \mathcal{M}, O \rangle$ tells us what measurements can be made and what outcomes can be observed.

Definition (Empirical model)

An **empirical model** e for a measurement scenario $\langle X, \mathcal{M}, O \rangle$ is the collection of probability distributions $\{e_C\}_{C \in \mathcal{M}}$.

Empirical model

Example

An empirical model of a Bell-CHSH experiment is

	e_C			
C	$(+,+)$	$(+,-)$	$(-,+)$	$(-,-)$
(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)	$1/8$	$3/8$	$3/8$	$1/8$

Violation to Bell-CHSH inequality

$$\langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle = 2\sqrt{2} \approx 2.828 > 2$$

Empirical model

Example

A beyond quantum empirical model of a Bell-CHSH measurement scenario is the PR box:

	e_C			
C	$(+,+)$	$(+,-)$	$(-,+)$	$(-,-)$
(a_1, b_1)	$1/2$	0	0	$1/2$
(a_1, b_2)	$1/2$	0	0	$1/2$
(a_2, b_1)	$1/2$	0	0	$1/2$
(a_2, b_2)	0	$1/2$	$1/2$	0

Violation to Bell-CHSH inequality

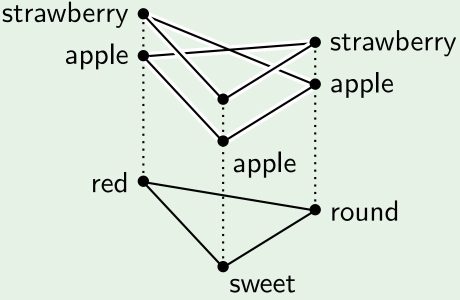
$$\langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle = 4$$

Previous work: PR-like models

- PR-boxes are the most contextual models
- First attempt at constructing contextual examples of natural language

Example

- There is an apple and a strawberry.
  - **One of them** is red and **the same one** is sweet.
  - **One of them** is sweet and **the same one** is round.
  - **One of them** is round and **the other one** is red.

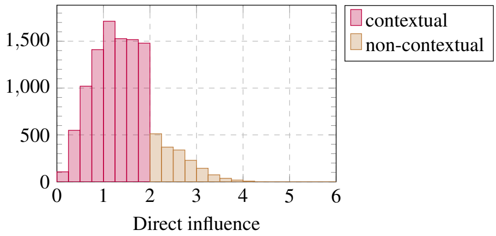


Previous work: PR-like models

- We ask BERT for probabilities to construct probabilistic models

	e_C			
C	(apl, apl)	(apl, sby)	(str, sby)	(sby, sby)
(red, round)	0.4	0	0	0.6
(round, sweet)	0.3	0	0	0.7
(sweet, red)	0	0.6	0.4	0

11,000 models created for different choices of words.



Winograd Schema Challenge

- Winograd Schema Challenge (WSC) is a test of machine intelligence, proposed by Hector Levesque in 2011.
- Intended to be an alternative to the Turing Test.

Example (councilmen-demonstrators)

- The city councilmen refused the demonstrators a permit because **they** *feared* violence.
- The city councilmen refused the demonstrators a permit because **they** *advocated* violence.


Generated on Stable Diffusion Online with prompts: (left) *The city councilmen considering whether to grant the demonstrators a permit.* (right) *Demonstrators without a permit advocating violence*

Winograd Schema Challenge

Example (trophy-suitcase)

- The trophy doesn't fit into the brown suitcase because **it** is too *small*.
- The trophy doesn't fit into the brown suitcase because **it** is too *big*.

Generated on Stable Diffusion Online with prompts: (left) *A golden World Cup trophy placed inside an opened small brown suitcase.* (right) *A huge golden World Cup trophy placed upright inside an opened brown suitcase.*

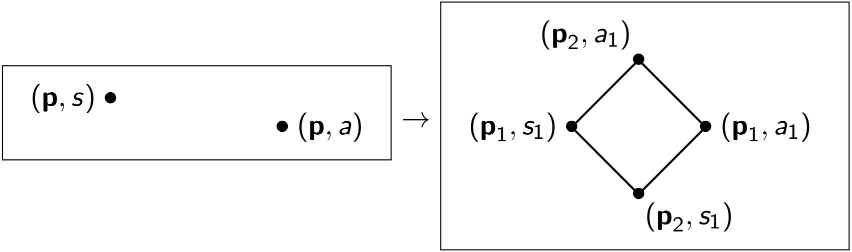
Winograd Schema Challenge

- Sheaf-theoretic measurement scenario:
  - $X = \{(\mathbf{p}, s), (\mathbf{p}, a)\}$,
  - $\mathcal{M} = \{\{(\mathbf{p}, s)\}, \{(\mathbf{p}, a)\}\}$,
  - $O = \{A, B\}$.

$(\mathbf{p}, s) \bullet$

$\bullet (\mathbf{p}, a)$

Generalised Winograd Schema



- 1 pronoun → 2 pronouns
- 1 special-alternate pair → 2 special-alternate pairs

An example

Example

A and B belong to the same [*cannibalistic*/*herbivorous*]$_1$ species of animal. On a hot afternoon in the south Sahara, **one of them**$_1$ was very hungry. They noticed each other when they were roaming in the field. After a while, **one of them**$_2$ is no longer [*hungry*/*alive*]$_2$.

Generated on Stable Diffusion Online with prompts: *In the hot desert, cute but hungry cats are in a tough spot, with one eating the other out of desperation.*

An example - scenario

Instruction: Please read the following short story which contains some ambiguities, then select the interpretations you think are the most appropriate.

Story: A and B belong to the same ${word1} species of animals. In a hot afternoon in south Sahara, one of them was very hungry. They notice each other when they were roaming in the field. In a while, one of them is no longer ${word2}.

Question: The following are 4 different interpretations of the story. Please select the **2** most appropriate interpretations.

- ☐ A was the very hungry ${word1} animal. A is no longer ${word2}.
- ☐ A was the very hungry ${word1} animal. B is no longer ${word2}.
- ☐ B was the very hungry ${word1} animal. A is no longer ${word2}.
- ☐ B was the very hungry ${word1} animal. B is no longer ${word2}.

- $\$\{\text{word1}\} \in \{\text{cannibalistic}, \text{herbivorous}\}$
- $\$\{\text{word2}\} \in \{\text{hungry}, \text{alive}\}$
- Human judgements collected on Amazon Mechanical Turk

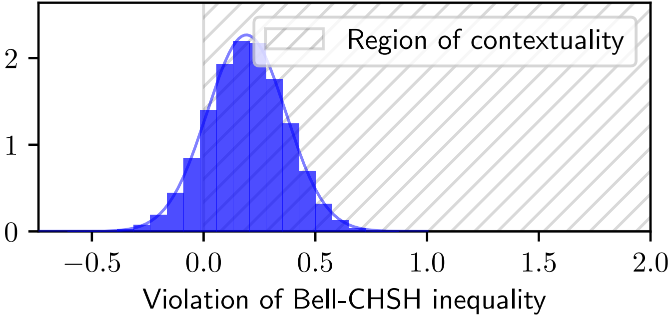
Human judgements - Results

(word1,	word2)	(A, A)	(A, B)	(B, A)	(B, B)
(*canni*,	*hungry*)	0.402	0.097	0.097	0.402
(*canni*,	*alive*)	0.044	0.455	0.455	0.044
(*herbi*,	*hungry*)	0.345	0.154	0.154	0.345
(*herbi*,	*alive*)	0.344	0.155	0.155	0.344

- Gives 0.192 violation to Bell-CHSH inequality
- $\implies$ Our example of generalised Winograd schema is contextual

Human judgements - Bootstrapping

- Bootstrap resampled to establish statistical significance
- 87% resampled datasets lie in the contextual region
- Standard deviation is 0.176



Conclusion

- Quantum **contextuality** has been connected to the computational power of quantum computers
- Disambiguation in language requires **context**
- We constructed two families of linguistic schemas that are able to host contextual models
- In particular, we generalised the Winograd Schema Challenge such that it is able to host contextual models

Future work

- Construct a set of examples of the proposed Generalised Winograd Schema
- Detect contextuality in corpus
- Connect contextuality to established phenomena in psycholinguistics, e.g. reading time
- It is not clear how to use the results to develop new quantum algorithms for NLP...