Griffith UNIVERSITY
Centre for Quantum Dynamics

A semantics for counterfactuals in quantum causal models

Eric G. Cavalcanti

Joint work with Ardra Kooderi Suresh and Markus Frembs

Workshop on Quantum Physics and Statistical Causal Models
Simons Institute, Berkeley, April 27, 2022

Australian Government
Australian Research Council

FQXi
FOUNDATIONAL QUESTIONS INSTITUTE

What if?

- What if I had jumped?
- What if Oswald had not killed Kennedy?
- Given that the patient died, what if they had not been given the treatment?
- What if we stopped emitting greenhouse gases?

Similarity analysis of counterfactuals

- Lewis [2] "Analysis 2" (see also Stalnaker [1]):

"A counterfactual 'If it were that A, then it would be that C' is (non-vacuously) true if and only if some (accessible) world where both A and C are true is more similar to our actual world, overall, than is any world where A is true but C is false".

"Counterfactuals are infected with vagueness… We ordinarily resolve the vagueness of counterfactuals in such a way that counterfactual dependence is asymmetric".

→ System of priorities to resolve vagueness.

[1] R. C. Stalnaker, "A Theory of Conditionals", in *Studies in Logical Theory*, N. Rescher (ed.), 98–112 (1968).
[2] D. K. Lewis, "Counterfactuals and Comparative Possibility", *Journal of Philosophical Logic*, 2(4): 418–446 (1973).
[3] D. K. Lewis, "Counterfactual Dependence and Time's Arrow", *Noûs*, 13(4): 455–476 (1979)

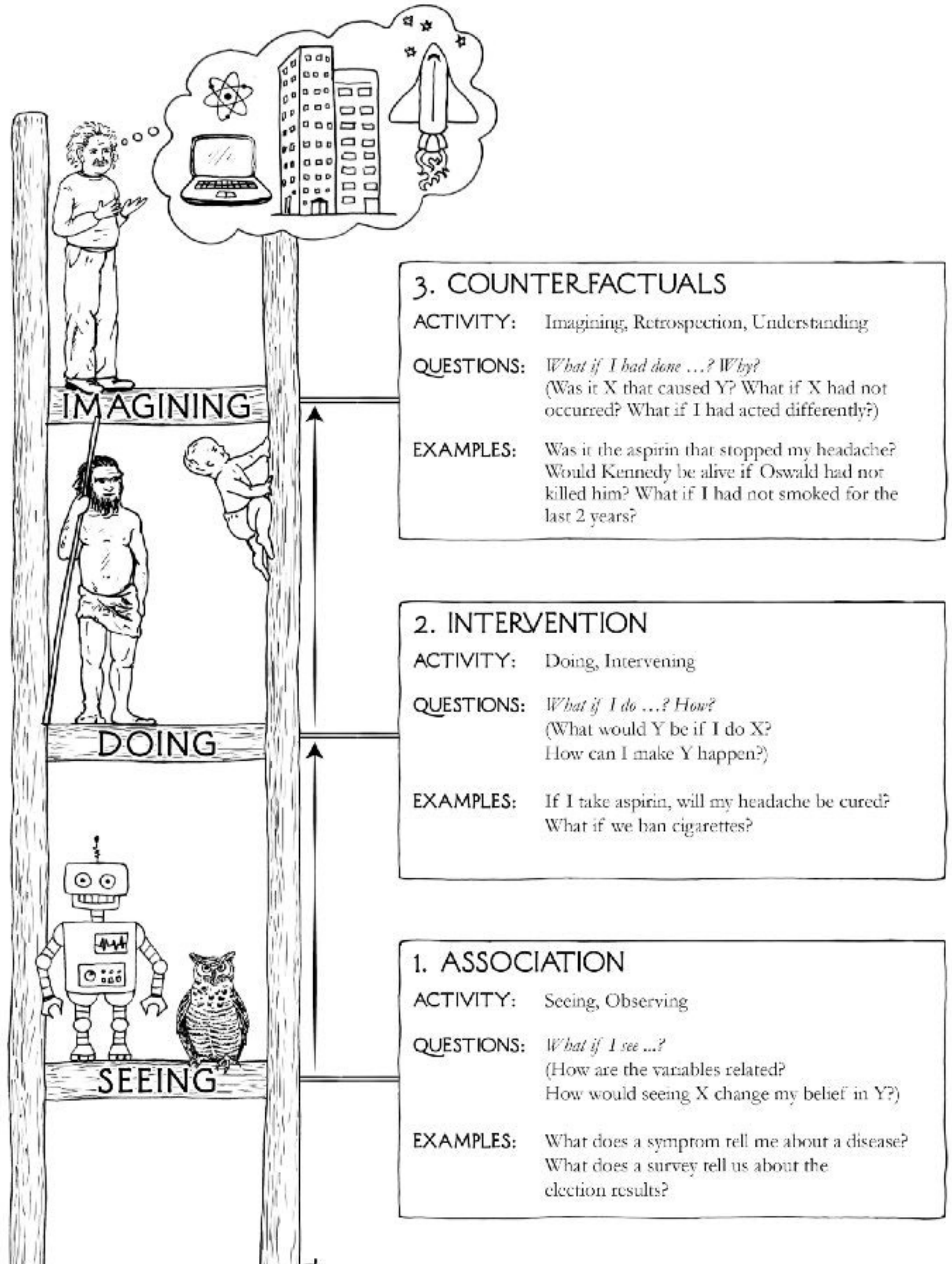
Counterfactuals and causation

- Hume [4]: “We may define a cause to be an object followed by another, and where all the objects similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed.”

- Lewis [5]: counterfactual analysis of causation
  - “I take Hume’s second definition as my definition not of causation itself, but of causal dependence among actual events”
  - assumes determinism, as does his theory of counterfactuals

[4] D. Hume, An Enquiry concerning Human Understanding, Section VII (1748).
[5] D.K. Lewis, “Causation”, *Journal of Philosophy*, 70(17): 556–567 (1973)

Pearl's ladder of causation



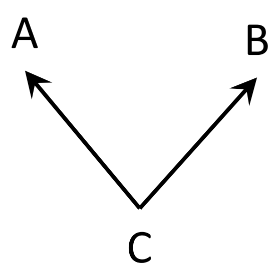
Level 1 - Association

- Bayesian networks
  - Directed Acyclic Graphs (DAGs)

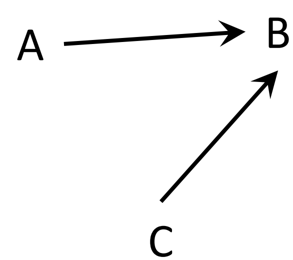
Markov condition: $P(X_1, \ldots, X_n) = \prod_i P(X_i | Pa(X_i))$

  - E.g.:

P(A,B,C) = P(A|C)P(B|C)P(C)



P(A,B,C) = P(B|A,C)P(A)P(C)



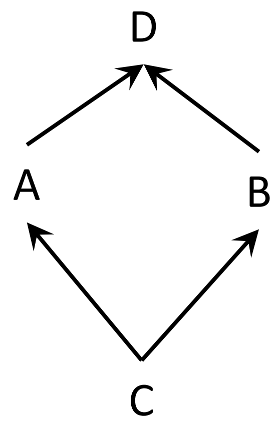
Level 2 - Intervention

- Causal Bayesian networks
  - Oracle for interventions

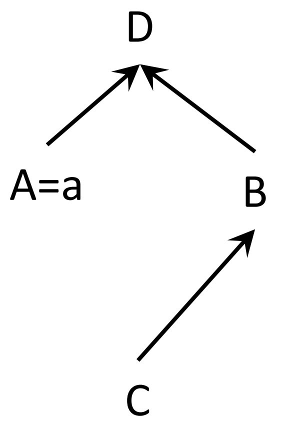
    do-intervention:

    $P_{do(A=a)}(A, B, C, D) = P(D|A = a, B)P(B|C)P(C)$

G:

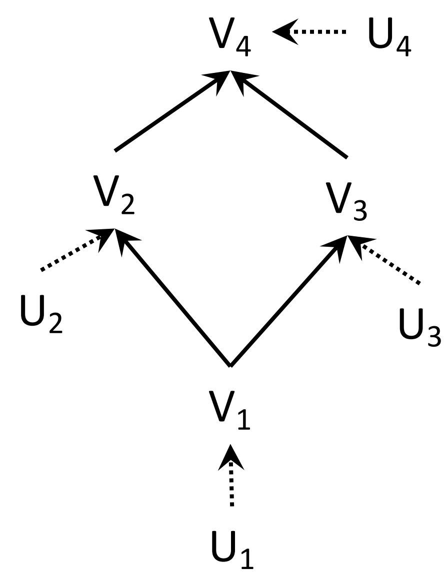


G(do (A=a)):



Level 3 - Counterfactuals

- (Probabilistic) Structural Causal Model
  - DAG G
  - Endogenous variables **V**=$\{V_1,\ldots,V_n\}$
  - Exogenous variables **U**
  - Functions $\{v_i = f_i(pa_i, u_i)\}_i$
  - Probability P(u)



Pearl's three-step algorithm for counterfactuals

Given evidence e, what is the probability that Y would have been y had X been x?

1. Abduction
   – Update P(u) by the evidence e to obtain P(u|e)
2. Action
   – Replace the equations for the variables in X by X=x. (do(X=x))
3. Prediction
   – Compute the probability P(Y=y) using the modified model.

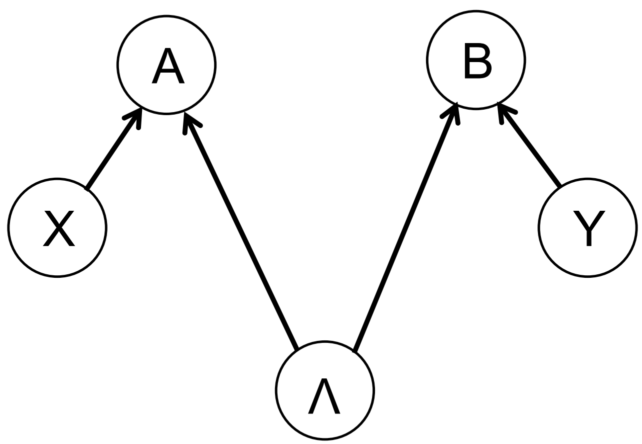
J. Pearl, *Causality: Models, Reasoning and Inference,* Cambridge Univ. Press (2009)

Quantum violations of CCMs

– Bell's theorem

- Classical causal models + relativistic causal structure (+ exogenous interventions)

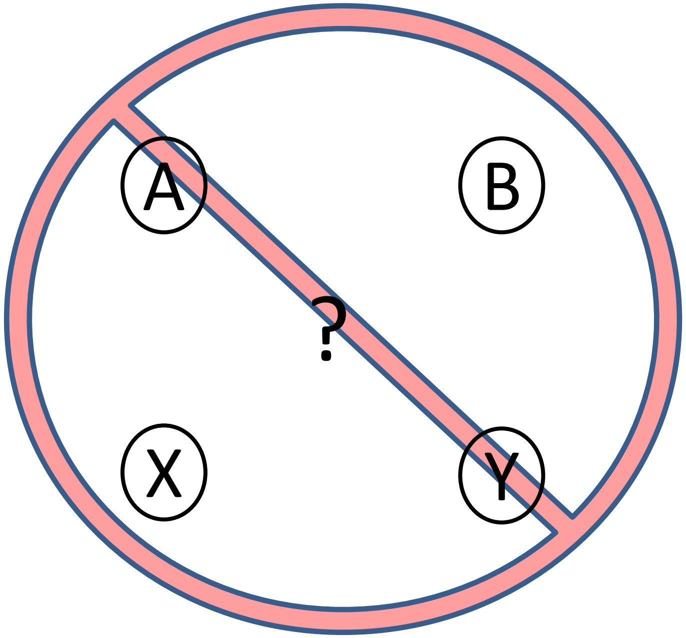
--> contradiction with quantum correlations



→ Bell inequalities

Finely-tuned Bells

No fine-tuning (Faithfulness)**:** Every conditional independence between variables must arise as a consequence of the causal graph and not due to special choices of causal-statistical parameters.



Wood and Spekkens, NJP **17**, 33002 (2015):

No classical causal model can explain all instances of bipartite Bell nonlocality without fine-tuning.

Finely-tuned Bells

EGC, Phys. Rev. X 8, 021018 (2018):

No faithful classical causal model can explain bipartite Bell nonlocality (or Kochen-Specker contextuality in scenarios with two measurements per context).

J.C Pearl and EGC, Quantum 5, 518 (2021), arXiv:1909.05434:

No faithful classical causal model can explain Bell nonlocality (or Kochen-Specker contextuality) in arbitrary scenarios.

Quantum causal models

Leifer and Spekkens, Phys. Rev. A 88, 052130 (2013), arXiv:1107.5849

Wood and Spekkens, New Journal of Physics, 17 (2015), arXiv:1208.4119

Cavalcanti and Lal, J. Phys. A 47, 424018 (2014), arXiv:1311.6852

Fritz, T. Comm. Math. Phys., 341, 391–434 (2016), arXiv:1404.4812

Henson, Lal and Pusey (HLP), New J. Phys. 16, 113043 (2014), arXiv:1405.2572

Pienaar and Brukner (PB), New J. Phys. 17, 073020 (2015), arXiv:1406.0430

Chaves, Majenz and Gross, Nat. Commun. 6, 5766 (2015), arXiv:1407.3800

Costa, Shrapnel, New J. Phys. 18 063032 (2016), arXiv:1512.07106

Allen, Barrett, Horsman, Lee and Spekkens, Phys. Rev. X 7, 031021 (2017), arXiv:1609.09487

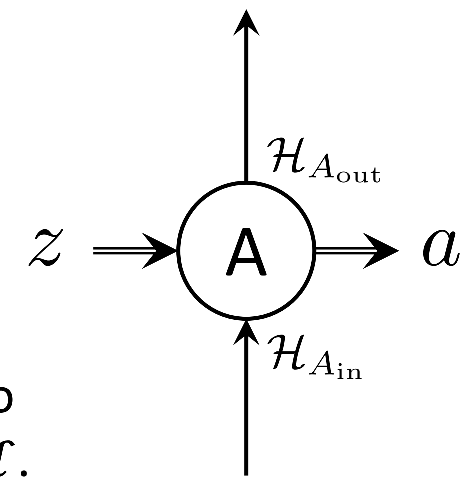
Giarmatzi and Costa. npj Quantum Information 4, 17 (2018), arxiv:1704.00800

Barrett, Lorenz, Oreshkov, arXiv:1906.10726

- Compatible with Relativistic Causality
- Allow for causal discovery / faithfulness
- Reproduce classical causal models as a special case

Quantum causal models

- Quantum node: locus of *potential interventions*
  - Input/output Hilbert spaces, with $d_{A_{in}} = d_{A_{out}}$



- Quantum instrument:
  - Set of completely positive (CP) maps $\mathcal{M}^a$ summing to a completely positive trace-preserving (CPTP) map $\mathcal{M}$.

$$z = \{\mathcal{M}^a : \mathcal{L}(\mathcal{H}_{A_\mathrm{in}}) \to \mathcal{L}(\mathcal{H}_{A_\mathrm{out}})\}_a$$

$$\mathcal{M} = \sum_a \mathcal{M}^a$$

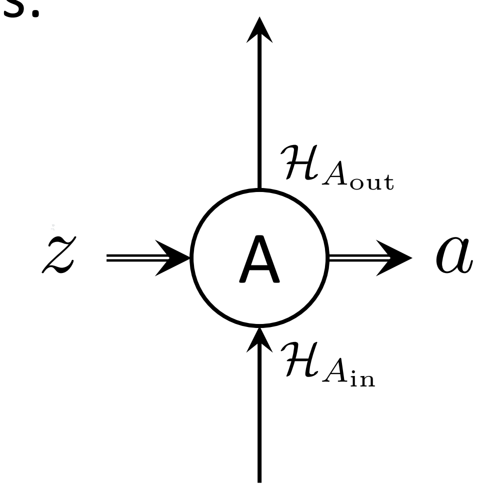
Quantum causal models

- Choi-Jamiolkowski isomorphism for instruments:

$$\tau_A^{a|z} = \sum_{i,j} \mathcal{M}^{a|z}(|i\rangle\langle j|)^T_{A_{out}} \otimes |j\rangle_{A_{in}}\langle i|$$

$$\tau_A^{z} = \sum_{a} \tau_A^{a|z}$$

$$\mathrm{Tr}_{A_{out}}[\tau_A^z] = \mathbb{I}_{A_{in}}$$



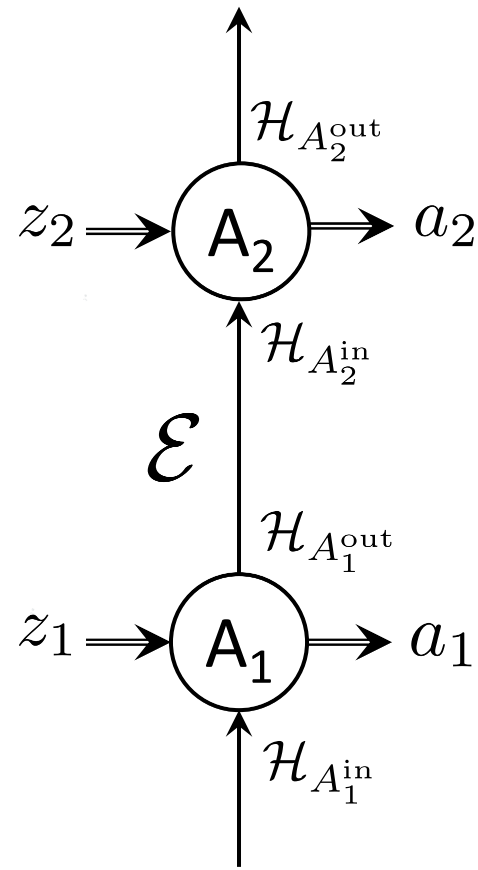
Quantum causal models

- CJ isomorphism for channels (CPTP maps) between nodes:

$$\mathcal{E} : \mathcal{L}(\mathcal{H}_{A_1^{\mathrm{out}}}) \to \mathcal{L}(\mathcal{H}_{A_2^{\mathrm{in}}})$$

$$\rho_{A_2|A_1} = \sum_{i,j} \mathcal{E}(|i\rangle\langle j|)_{A_2^{in}} \otimes |i\rangle_{A_1^{out}} \langle j|$$

(Compare to $\tau_A^{a|z} = \sum_{i,j} \mathcal{M}^{a|z}(|i\rangle\langle j|)^T_{A_{out}} \otimes |j\rangle_{A_{in}} \langle i|$)



Quantum causal models

A ***quantum process operator*** $\boldsymbol{\sigma_{A_1,\ldots,A_n}}$ over nodes A_1 , … , A_n is a positive semi-definite operator:

$$\sigma_{A_1,\cdots,A_n} \in \mathcal{L}(\bigotimes_i \mathcal{H}_{A_i^{in}} \otimes \mathcal{H}_{A_i^{out}})$$

such that for all choices of instruments z = (z1 , … , zn) at the nodes,

$$\mathrm{Tr}_{A_1\cdots A_n}[\sigma_{A_1,\cdots,A_n} \tau_{A_1}^{z_1} \otimes \cdots \otimes \tau_{A_n}^{z_n}] = 1$$

Probabilities are given by the generalized Born rule:

$$P(a_1,\cdots,a_n|z_1,\cdots,z_n) = \mathrm{Tr}_{A_1\cdots A_n}[\sigma_{A_1,\cdots,A_n} \tau_{A_1}^{a_1|z_1} \otimes \cdots \otimes \tau_{A_n}^{a_n|z_n}]$$

Where we use the shorthand notation:

$$\mathrm{Tr}_A[\cdots] = \mathrm{Tr}_{A_{in},A_{out}}[\cdots]$$

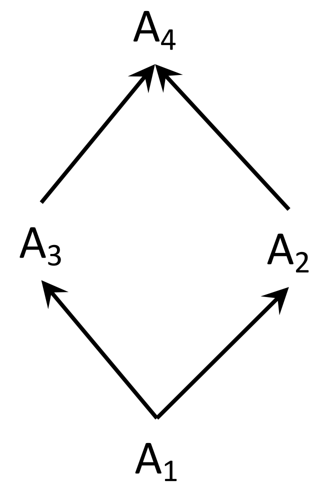
Quantum causal models

A ***quantum causal model*** is specified by

1) a DAG *G* with quantum nodes A_1 , … , A_n as vertices

2) A quantum channel $\rho_{A_i|Pa(A_i)}$ for each node such that

$$[\rho_{A_i|Pa(A_i)}, \rho_{A_j|Pa(A_j)}] = 0 \text{ for all } i, j.$$



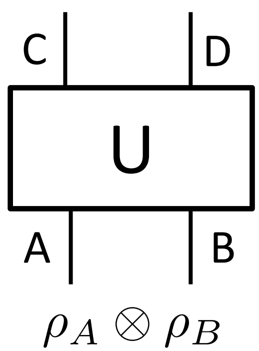
The process operator is then given by

$$\sigma_{A_1,\cdots,A_n} = \prod_i \rho_{A_i|Pa(A_i)}.$$

This is the *quantum causal Markov condition.*

Signalling in unitary processes

Consider the CJ operator for a unitary quantum channel $\rho^{U}_{CD|AB}$



We say that *A does not influence D* ($A \nrightarrow D$) iff for any state ρ_B it is not possible to send signals from A to D by varying ρ_A.

Quantum structural models

- A ***quantum structural model*** is specified by:

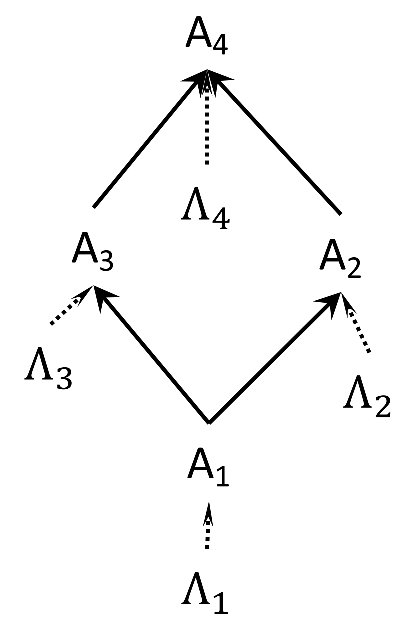
1) A DAG *G* over a set of quantum nodes **A** = { A_1 , … , A_n}
2) A set of exogenous nodes $\boldsymbol{\Lambda} = \{\Lambda_1, \ldots, \Lambda_n\}$
3) A final node *F (not shown in figure)*
4) A $\rho^U_{\boldsymbol{AF}|\boldsymbol{A\Lambda}}$ that satisfies the no-influence conditions

$$\{A_j \nrightarrow A_i\}_{A_j \notin Pa(A_i)} \text{ and } \{\Lambda_j \nrightarrow A_i\}_{j \neq i}$$

5) A set of preparation instruments

$$\{\tau^{\lambda_i}_{\Lambda_i}\}_{\lambda_i} \equiv \{p(\lambda_i)(\rho^{\lambda_i}_{\Lambda_i^{out}})^T \otimes \mathbb{I}_{\Lambda_i^{in}}\}_{\lambda_i}$$

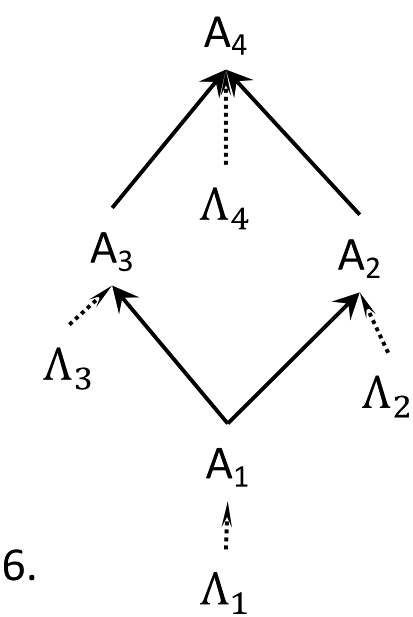
$\tau^\rho_\Lambda = \sum_\lambda \tau^\lambda_\Lambda$ prepares state $\rho = \sum_\lambda p(\lambda)\rho^\lambda$



Quantum structural models

We say that a process operator $\sigma_{\boldsymbol{A}}$ is ***compatible**** with a DAG G iff there exists a quantum structural model associated with G that recovers $\sigma_{\boldsymbol{A}}$ as a marginal:

$$\sigma_{\mathbf{A}} = \mathrm{Tr}_{F_{in}\mathbf{\Lambda}}[\rho^{U}_{\mathbf{A}F|\mathbf{A}\mathbf{\Lambda}}\tau^{\rho_1}_{\Lambda_1} \otimes \cdots \otimes \tau^{\rho_n}_{\Lambda_n}]$$



It can be shown** that:

$\sigma_{\boldsymbol{A}}$ is compatible with $G \Leftrightarrow \sigma_{\boldsymbol{A}}$ is Markov for G

* See also slightly different definition in Barrett *et al*, arXiv:1906.10726.

** Follows straightforwardly from proof in Barrett et al.

Quantum structural models

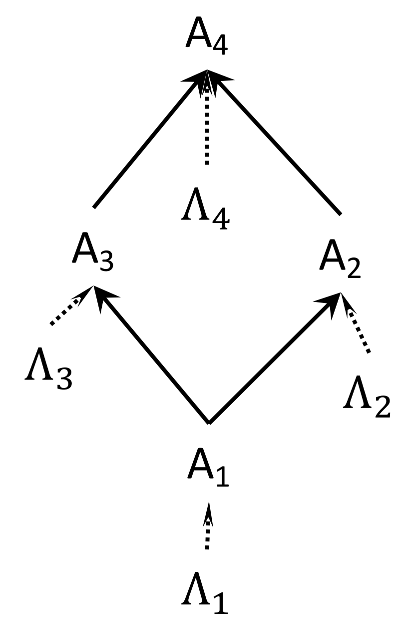
Given information about a particular set of outcomes **λ** = ($\lambda_1, \cdots, \lambda_n$) for the exogenous instruments, we can define a *conditional process operator* *

$$\sigma_{\mathbf{A}}^{\vec{\lambda}} = \frac{\mathrm{Tr}_{F_{in}\mathbf{\Lambda}}\left[\rho^{U}_{\mathbf{A}F|\mathbf{A}\mathbf{\Lambda}}\tau^{\lambda_1}_{\Lambda_1}\otimes\cdots\otimes\tau^{\lambda_n}_{\Lambda_n}\right]}{p(\lambda_1,\cdots,\lambda_n)}$$

This can be used to calculate the conditional probability for a set of outcomes **a** = (a_1, …, a_n) for instruments **z** = (z_1, …, z_n), given **λ:**

$$p_{\mathbf{z}}(\mathbf{a}|\vec{\lambda}) = \mathrm{Tr}_{\mathbf{A}}[\sigma_{\mathbf{A}}^{\vec{\lambda}}\tau_{\mathbf{a}}^{\mathbf{a}|\mathbf{z}}]$$

where $\tau_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}} = \tau_{A_1}^{a_1|z_1}\otimes\cdots\otimes\tau_{A_n}^{a_n|z_n}$



* Following Costa and Shrapnel (2016).

Evaluation of counterfactuals

- Standard counterfactual query:

Given that the set of instruments $\boldsymbol{z} = (z_1, \ldots, z_n)$ has been implemented and outcomes ***a*** = (a_1, …, a_n) obtained, what is the probability $p_{\boldsymbol{z}'}^{\boldsymbol{a}|z}(\boldsymbol{c}|\boldsymbol{b})$ that outcomes ***c*** would have obtained for a subset of nodes ***C***, had instruments $\boldsymbol{z}' = (z'_1, \ldots, z'_n)$ been implemented and outcomes ***b*** obtained at nodes ***B***?

1. ***Abduction***
2. ***Action***
3. ***Prediction***.

Evaluation of counterfactuals

Given that the set of instruments $\boldsymbol{z} = (z_1, \ldots, z_n)$ have been implemented and outcomes ***a*** obtained, what is the probability $p_{\boldsymbol{z'}}^{\boldsymbol{a}|\boldsymbol{z}}(\boldsymbol{c}|\boldsymbol{b})$ that outcomes ***c*** would have obtained for a subset of nodes ***C***, had instruments $\boldsymbol{z}' = (z'_1, \ldots, z'_n)$ been implemented and outcomes ***b*** obtained at nodes ***B***?

1. Abduction

- Infer “stable facts”*: $\{\tau_{\Lambda_i}^{\lambda_i}\}_i$
- Given observed outcomes: $\{\tau_{A_i}^{a_i|z_i}\}_i$
- Bayesian update:

$$p_{\mathbf{z}}(\vec{\lambda}|\mathbf{a}) = \frac{p_{\mathbf{z}}(\mathbf{a}|\vec{\lambda})p_{\mathbf{z}}(\vec{\lambda})}{p_{\mathbf{z}}(\mathbf{a})} = \frac{\mathrm{Tr}_{\mathbf{A}}\left[\sigma_{\mathbf{A}}^{\vec{\lambda}}\tau_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}}\right]p(\lambda_1, \cdots, \lambda_n)}{\mathrm{Tr}_{\mathbf{A}}\left[\sigma_{\mathbf{A}}\tau_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}}\right]}$$

- *Updated process operator*: $\sigma_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}} = \sum_{\lambda_1, \cdots, \lambda_n} p_{\mathbf{z}}(\vec{\lambda}|\mathbf{a})\sigma_{\mathbf{A}}^{\vec{\lambda}}$

* Di Biagio and Rovelli, Foundations of Physics 51, 30 (2021)

Evaluation of counterfactuals

Given that the set of instruments $\boldsymbol{z} = (z_1, \ldots, z_n)$ have been implemented and outcomes ***a*** obtained, what is the probability $p_{\boldsymbol{z}'}^{\boldsymbol{a}|\boldsymbol{z}}(\boldsymbol{c}|\boldsymbol{b})$ that outcomes ***c*** would have obtained for a subset of nodes ***C***, had instruments $\boldsymbol{z}' = (z'_1, \ldots, z'_n)$ been implemented and outcomes ***b*** obtained at nodes ***B***?

2. Action

- Change instruments to $\{\tau_{A_i}^{z'_i}\}_i$
- Some may be do-interventions: $\tau_A^{\mathrm{do}(\rho)} \equiv (\rho_{A_{out}})^T \otimes \mathbb{I}_{A_{in}}$

Evaluation of counterfactuals

Given that the set of instruments $\boldsymbol{z} = (z_1, \dots, z_n)$ have been implemented and outcomes $\boldsymbol{a}$ obtained, what is the probability $p_{\boldsymbol{z}'}^{\boldsymbol{a}|\boldsymbol{z}}(\boldsymbol{c}|\boldsymbol{b})$ that outcomes $\boldsymbol{c}$ would have obtained for a subset of nodes ***C***, had instruments $\boldsymbol{z}' = (z'_1, \dots, z'_n)$ been implemented and outcomes $\boldsymbol{b}$ obtained at nodes ***B***?

2. Prediction

$$p_{\mathbf{z}'}^{\mathbf{a}|\mathbf{z}}(\mathbf{c}|\mathbf{b}) = \frac{p_{\mathbf{z}'}^{\mathbf{a}|\mathbf{z}}(\mathbf{c},\mathbf{b})}{p_{\mathbf{z}'}^{\mathbf{a}|\mathbf{z}}(\mathbf{b})} = \frac{\mathrm{Tr}_{\mathbf{A}}\left[\sigma_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}} \tau_{\mathbf{B}}^{\mathbf{b}|\mathbf{z}'_{\mathbf{B}}} \tau_{\mathbf{C}}^{\mathbf{c}|\mathbf{z}'_{\mathbf{C}}} \tau_{\mathbf{A}\setminus\mathbf{B}\cup\mathbf{C}}^{\mathbf{z}'}\right]}{\mathrm{Tr}_{\mathbf{A}}\left[\sigma_{\mathbf{A}}^{\mathbf{a}|\mathbf{z}} \tau_{\mathbf{B}}^{\mathbf{b}|\mathbf{z}'} \tau_{\mathbf{A}\setminus\mathbf{B}}^{\mathbf{z}'}\right]}$$

where $\tau_{\mathbf{B}}^{\mathbf{b}|\mathbf{z}'_B} = \bigotimes_{B_j \in \mathbf{B}} \tau_{B_j}^{b_j|z'_j}\,,\ \tau_{\mathbf{C}}^{\mathbf{c}|\mathbf{z}'_C} = \bigotimes_{C_k \in \mathbf{C}} \tau_{C_k}^{c_k|z'_k}\,,\ \tau_{\mathbf{A}\setminus\mathbf{B}\cup\mathbf{C}}^{\mathbf{z}'} = \bigotimes_{A_i \notin \mathbf{B}\cup\mathbf{C}} \tau_{A_i}^{z'_i}\,, \mathrm{etc.}$

Interpretation of counterfactual queries

- Classical structural model:
  - All endogenous variables determined by background variables ***u.***

$\Longrightarrow$ *The only way the antecedent could have been different, while keeping **u** fixed, is if some intervention had occurred.*

"To model an action *do(X = x)* one performs a "minisurgery" on the causal model, that is, a minimal change necessary for establishing the antecedent *X = x*, while leaving the rest of the model intact… This mini-surgery (not unlike Lewis's "little miracle"), makes precise the idea of using a "minimal deviation from actuality" to define counterfactuals." [J. Pearl, Causal and Counterfactual Inference, TECHNICAL REPORT R-485 (2019)]

- Quantum structural model:
  - Complete knowledge of the exogenous variables λ does not in general determine all instrument outcomes.

$\Longrightarrow$ *The antecedent could have been different, even keeping all $\lambda's$ fixed and the rest of the model intact, if **z = z'** and*

$$p_{\mathbf{z}}(\mathbf{b}|\vec{\lambda}) > 0 \quad \forall\vec{\lambda}\,\text{s.t.}\;\; p_{\mathbf{z}}(\vec{\lambda}|\mathbf{a}) > 0$$

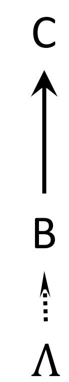
Interpretation of counterfactual queries

- ***Passive*** counterfactuals

  The antecedent could have been different, even keeping all λ's fixed and the rest of the model intact, if ***z = z'*** *and*

  $$p_{\mathbf{z}}(\mathbf{b}|\vec{\lambda}) > 0 \quad \forall\vec{\lambda}\,\text{s.t.}\;\; p_{\mathbf{z}}(\vec{\lambda}|\mathbf{a}) > 0$$

  – E.g.
    - $\tau_\Lambda \;\; = \{\frac{1}{2}[0]_{\Lambda^{out}} \otimes \mathbb{I}_{\Lambda^{in}}, \frac{1}{2}[1]_{\Lambda^{out}} \otimes \mathbb{I}_{\Lambda^{in}}\}$
    - $\tau_{\mathrm{B}} \;\; = \{[+]_{\mathrm{B}^{out}} \otimes [+]_{\mathrm{B}^{in}}, [-]_{\mathrm{B}^{out}} \otimes [-]_{\mathrm{B}^{in}}\}$
    - $\tau_{\mathrm{C}} \;\; = \{[+]_{\mathrm{C}^{out}} \otimes [+]_{\mathrm{C}^{in}}, [-]_{\mathrm{C}^{out}} \otimes [-]_{\mathrm{C}^{in}}\}$
    - $\rho^U_{BC|B\Lambda} = \rho^{id}_{C|B}\rho^{id}_{B|\Lambda}$



- Q: Given that *b = +*, what's the probability that *c = +* had it been that *b = -* ?

  – $p_{\boldsymbol{z}}(\lambda|b=+) > 0$ for all λ, $p_{\boldsymbol{z}}(b=+|\lambda) > 0$ for all λ.

  – Answer to passive CF query: $p_{\boldsymbol{z}}^{b=+|\boldsymbol{z}}(c=+|b=-) = 0$.

Interpretation of counterfactual queries

- ***Active*** counterfactuals

Not $p_{\mathbf{z}}(\mathbf{b}|\vec{\lambda}) > 0 \quad \forall\vec{\lambda} \text{ s.t. } p_{\mathbf{z}}(\vec{\lambda}|\mathbf{a}) > 0$

E.g.:

- $\tau_\Lambda \;= \{\frac{1}{2}[+]_{\Lambda^{out}} \otimes \mathbb{I}_{\Lambda^{in}}, \frac{1}{2}[-]_{\Lambda^{out}} \otimes \mathbb{I}_{\Lambda^{in}}\}$
- $\tau_{\mathrm{B}} \;= \{[+]_{\mathrm{B}^{out}} \otimes [+]_{\mathrm{B}^{in}}, [-]_{\mathrm{B}^{out}} \otimes [-]_{\mathrm{B}^{in}}\}$
- $\tau_{\mathrm{C}} \;= \{[+]_{\mathrm{C}^{out}} \otimes [+]_{\mathrm{C}^{in}}, [-]_{\mathrm{C}^{out}} \otimes [-]_{\mathrm{C}^{in}}\}$
- $\rho^{U}_{BC|B\Lambda} = \rho^{id}_{C|B}\rho^{id}_{B|\Lambda}$

- Q: Given that *b = +*, what's the probability that *c = +* had it been that *b* = -?

  $p_{\mathbf{z}}(\lambda = -|b = +) = 0\,,\; p_{\mathbf{z}}(b = -|\lambda = +) = 0.$

  – The only way to accommodate the antecedent is via an intervention. We use the do-intervention:

  $$\tau_B^{\mathrm{do}(b=-)} = ([-]_{B_{out}})^T \otimes \mathbb{I}_{B_{in}}$$

  – Answer to active CF query: $p^{b=+}_{do(b=-)}(c = +) = 0.$

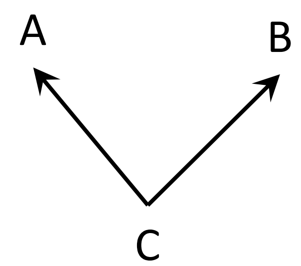
Passive vs Active?

- *Principle of Similarity:*

  If it is ambiguous whether a CF query is intended as a passive or active CF and it can be interpreted passively, then it should be interpreted as a passive CF query.

    - “minimal modification”, “closest possible world”

• Example: Bell scenario



- $\tau_{\mathrm{C}} \quad = [\Phi^+]_{\mathrm{C}^{out}} \otimes \mathbb{I}_{\mathrm{C}^{in}}$, where $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$
- $\rho^{U}_{AB|C} = \rho^{id}_{A|C}\rho^{id}_{B|C}$
- $\tau_{\mathrm{A}} \quad = \{[0]_{\mathrm{A}^{out}} \otimes [0]_{\mathrm{A}^{in}}, [1]_{\mathrm{A}^{out}} \otimes [1]_{\mathrm{A}^{in}}\}$
- $\tau_{\mathrm{B}} \quad = \{[0]_{\mathrm{B}^{out}} \otimes [0]_{\mathrm{B}^{in}}, [1]_{\mathrm{B}^{out}} \otimes [1]_{\mathrm{B}^{in}}\}$

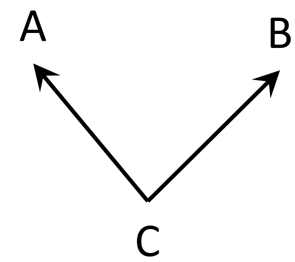
Q: Given that *a = 0*, what's the probability that *b = 1* had it been that *a = 1*?

- Principle of similarity → interpret as a passive CF.
  - Ans.: $p^{a=0}_{\boldsymbol{z}}(b = 1|a = 1) = 1$

→ CF dependence does not in general imply causation

- If we were to interpret as an active CF: "Given that a = 0, what's the probability that b = 1 *had we set* a = 1"?
  - Ans.: $p^{a=0}_{do(a=1)}(b = 1) = \frac{1}{2} = p_z(b = 1)$

- “Active-observational” CF



- $\tau_{\mathrm{C}} = [\Phi^+]_{\mathrm{C}^{out}} \otimes \mathbb{I}_{\mathrm{C}^{in}}$, where $|\Phi^+\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$
- $\rho^{U}_{AB|C} = \rho^{id}_{A|C}\rho^{id}_{B|C}$
- $\tau_{\mathrm{A}} = \{[0]_{\mathrm{A}^{out}} \otimes [0]_{\mathrm{A}^{in}}, [1]_{\mathrm{A}^{out}} \otimes [1]_{\mathrm{A}^{in}}\}$
- $\tau_{\mathrm{B}} = \{[0]_{\mathrm{B}^{out}} \otimes [0]_{\mathrm{B}^{in}}, [1]_{\mathrm{B}^{out}} \otimes [1]_{\mathrm{B}^{in}}\}$
- $\tau'_{\mathrm{A}} = \{[+]_{\mathrm{A}^{out}} \otimes [+]_{\mathrm{A}^{in}}, [-]_{\mathrm{A}^{out}} \otimes [-]_{\mathrm{A}^{in}}\}$

Q: What’s the probability that *b = 1* had it been that *a = 1*?

- Ans.: $p_{\boldsymbol{z}}\ (b = 1|a = 1) = 1$

Q: What’s the probability that *b = 1* had it been that *a = +*?

- Ans.: $p_{\boldsymbol{z}'}(b = 1|a = +) = \frac{1}{2}$

Q: What’s the probability that *b = 1* had instrument τ'_A been used?

- Ans.: $p_{\boldsymbol{z}'}(b = 1) = \frac{1}{2} = p_{\boldsymbol{z}}\ \ (b = 1)$

→ Conditioning on outcomes can create CF dependence without causal dependence, if the CF instrument does not block the input to the node (as in a do-intervention).

Summary

- Defined a quantum structural model
- Semantics for counterfactuals
- CF dependence and causation
  - CF dependence may occur without causation
  - CF dependence is associated with causal dependence only for active CFs

Limitations

- Not intended as a general theory of counterfactuals
- Different interpretations of QM would give different answers
  - E.g. Bohmian mechanics
- QCMs do not resolve the measurement problem*
  - Classical and quantum causal models are both *effective* theories, *pragmatically useful* at different levels of description:

    Classical causal models → Classical limit

    Quantum causal models → Fixed Heisenberg cut

* EGC, J. Phys.: Conf. Ser. 701 012002, arXiv:1602.07404
Bong *et al, A strong no-go theorem on the Wigner's friend paradox*, Nature Physics 16, 1199–1205 (2020)], arXiv:1907.05607
EGC, H. Wiseman*, Implications of Local Friendliness violation for quantum causality*, Entropy 2021, 23(8), 925 arXiv:2106.04065
https://www.youtube.com/watch?v=4KKZzqTmqBE

Further questions

- Violations of “Counterfactual definiteness”

- More interesting applications?
  - E.g. cryptographic scenarios: “Would I have observed something different had there been an eavesdropper?”
  - More general causal networks

Further questions

- Violations of "Counterfactual definiteness"

- More interesting applications?
  - E.g. cryptographic scenarios: "Would I have observed something different had there been an eavesdropper?"
  - More general causal networks

Thank you!