# Causal Inference Theory with Information Algebras: Binary Relations, Alexandrov Topologies and Conditional Topological Separation

Benjamin Heymann, Michel De Lara, Jean-Philippe Chancelier

BH: Criteo AI Lab, Paris, France
MDL and JPC : Cermics, École des Ponts, Marne-la-Vallée, France

# Outline of the presentation

Background on binary relations, graphs and Alexandrov topologies

Conditional relations on a graph

Conditional topological separation

# Background on binary relations, graphs and Alexandrov topologies

# Background on binary relations, graphs and Alexandrov topologies

## Binary relations

# Binary relations (definition and exemples)

Let $\mathcal{V}$ be a nonempty set (finite or not)

- We recall that a (binary) relation $\mathcal{R}$ on $\mathcal{V}$ is a subset

$$\mathcal{R} \subset \mathcal{V} \times \mathcal{V}$$

  and that

$$\gamma \, \mathcal{R} \, \lambda \iff (\gamma, \lambda) \in \mathcal{R}$$

- For any subset $\Gamma \subset \mathcal{V}$, the (sub)diagonal relation is

$$\Delta_\Gamma = \left\{ (\gamma, \lambda) \in \mathcal{V} \times \mathcal{V} \,\middle|\, \gamma = \lambda \in \Gamma \right\}$$

  and the diagonal relation is $\Delta = \Delta_\mathcal{V}$

# Binary relations (follow up)

- A foreset of a relation $\mathcal{R}$ is any set of the form

$$\mathcal{R}\,\lambda = \left\{\gamma \in \mathcal{V} \,\middle|\, \gamma\,\mathcal{R}\,\lambda\right\}$$

- An afterset of a relation $\mathcal{R}$ is any set of the form

$$\gamma\,\mathcal{R} = \left\{\lambda \in \mathcal{V} \,\middle|\, \gamma\,\mathcal{R}\,\lambda\right\}$$

- The opposite or complementary $\mathcal{R}^{\mathsf{c}}$ of a binary relation $\mathcal{R}$ is the relation $\mathcal{R}^{\mathsf{c}} = \mathcal{V} \times \mathcal{V} \setminus \mathcal{R}$, that is, defined by

$$\gamma\,\mathcal{R}^{\mathsf{c}}\,\lambda \iff \neg(\gamma\,\mathcal{R}\,\lambda)$$

- The converse $\mathcal{R}^{-1}$ of a binary relation $\mathcal{R}$ is defined by

$$\gamma\,\mathcal{R}^{-1}\,\lambda \iff \lambda\,\mathcal{R}\,\gamma$$

  and a relation $\mathcal{R}$ is symmetric if $\mathcal{R}^{-1} = \mathcal{R}$

# Binary relations (composition)

- The composition $\mathcal{R}\mathcal{R}'$ of two binary relations $\mathcal{R}, \mathcal{R}'$ on $\mathcal{V}$ is the binary relation on $\mathcal{V}$ defined by

$$\gamma(\mathcal{R}\mathcal{R}')\lambda \iff \exists \delta \in \mathcal{V}\ ,\ \ \gamma\,\mathcal{R}\,\delta \text{ and } \delta\,\mathcal{R}'\,\lambda$$

  By induction we define $\mathcal{R}^{n+1} = \mathcal{R}\mathcal{R}^n$ for $n \in \mathbb{N}$, with the convention $\mathcal{R}^0 = \Delta$

- The transitive closure of a binary relation $\mathcal{R}$ is

$$\mathcal{R}^+ = \bigcup_{k=1}^{\infty} \mathcal{R}^k$$

  and $\mathcal{R}$ is transitive if $\mathcal{R}^+ = \mathcal{R}$

- The reflexive and transitive closure is

$$\mathcal{R}^* = \mathcal{R}^+ \cup \Delta = \bigcup_{k=0}^{\infty} \mathcal{R}^k$$

- A partial equivalence relation is a symmetric and transitive binary relation

# Background on binary relations, graphs and Alexandrov topologies

## Graphs defined by a binary relation

# Graphs defined by a binary relation

- Let $\mathcal{V}$ be a nonempty set (finite or not), whose elements are called vertices
- Let $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ be a relation on $\mathcal{V}$, whose elements are ordered pairs (that is, couples) of vertices called edges
  - the first element of an edge is the tail of the edge
  - whereas the second one is the head of the edge
  - both tail and head are called endpoints of the edge, and we say that the edge connects its endpoints
- We define a loop as an element of $\Delta \cap \mathcal{E}$, that is, a loop is an edge that connects a vertex to itself

**Definition**

A graph, as we use it, is a couple $(\mathcal{V}, \mathcal{E})$ where $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$

# Graphs (comments)

As we define a graph,

- it may hold a finite or infinite number of vertices
- there is at most one edge that has a couple of ordered vertices as single endpoints, hence a graph (in our sense) is not a multigraph (in graph theory)
- loops are not excluded (since we do not impose $\Delta \cap \mathcal{E} = \emptyset$)

Hence, what we call a graph would be called, in graph theory,
a directed simple graph permitting loops

# Background on binary relations, graphs and Alexandrov topologies

## Alexandrov topology induced by a binary relation

Let $(\mathcal{V}, \mathcal{E})$ be a graph, that is, $\mathcal{V}$ is a set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$

**Proposition**

*The following set*

$$\mathcal{T}_{\mathcal{E}} = \left\{ \mathcal{O} \subset \mathcal{V} \,\middle|\, \mathcal{O}\mathcal{E} \subset \mathcal{O} \right\}$$

*is an Alexandrov topology on* $\mathcal{V}$*,*
*with the property that open subsets are characterized by*

$$\mathcal{O} \in \mathcal{T}_{\mathcal{E}} \iff \mathcal{O}\mathcal{E} \subset \mathcal{O} \iff \mathcal{O}\mathcal{E}^{+} \subset \mathcal{O} \iff \mathcal{O}\mathcal{E}^{*} \subset \mathcal{O} \iff \mathcal{O}\mathcal{E}^{*} = \mathcal{O}$$

*In the Alexandrov topology* $\mathcal{T}_{\mathcal{E}}$*,*
*the topological closure* $\overline{\Gamma}^{\mathcal{E}}$ *of a subset* $\Gamma \subset \mathcal{V}$ *is given by*

$$\overline{\Gamma}^{\mathcal{E}} = \mathcal{E}^{*}\Gamma \,, \quad \forall \Gamma \subset \mathcal{V}$$

*that is, is the* $\mathcal{E}^{*}$*-foreset of* $\Gamma$

# Conditional relations on a graph

# Conditional parental relation

Let $(\mathcal{V}, \mathcal{E})$ be a graph, that is, $\mathcal{V}$ is a set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, and let $W \subset \mathcal{V}$ be a subset of (conditioning) vertices

**Definition**

We define the conditional parental relation $\mathcal{E}^W$ as

$$\mathcal{E}^W = \Delta_{W^c}\mathcal{E}\,, \text{ that is, } \gamma\mathcal{E}^W\lambda \iff \gamma \in W^c \text{ and } \gamma\mathcal{E}\lambda \qquad (\forall \gamma, \lambda \in \mathcal{V})$$

and the conditional ascendent relation $\mathcal{B}^W$ as

$$\mathcal{B}^W = \mathcal{E}(\Delta_{W^c}\mathcal{E})^* = \mathcal{E}\mathcal{E}^{W*} \text{ where } \mathcal{E}^{W*} = (\mathcal{E}^W)^*$$

which relates descendent with ascendent by means of elements in $W^c$

We define their converses $\mathcal{E}^{-W}$ and $\mathcal{B}^{-W}$ as

$$\mathcal{E}^{-W} = (\mathcal{E}^W)^{-1} = \mathcal{E}^{-1}\Delta_{W^c}$$

$$\mathcal{B}^{-W} = \left(\mathcal{B}^W\right)^{-1} = (\mathcal{E}^{-1}\Delta_{W^c})^*\mathcal{E}^{-1} = \mathcal{E}^{-W*}\mathcal{E}^{-1} \text{ where } \mathcal{E}^{-W*} = (\mathcal{E}^{-W})^*$$

# Conditional common cause, cousinhood and active relations

With these elementary binary relations, we define the conditional common cause relation $\mathcal{K}^w$ as the symmetric relation

$$\mathcal{K}^w = \mathcal{B}^{-w} \Delta_{W^c} \mathcal{B}^w = \mathcal{E}^{-W+} \mathcal{E}^{W+}$$

the conditional cousinhood relation $\mathcal{C}^w$ as the partial equivalence relation

$$\mathcal{C}^w = \left( \Delta_W \mathcal{K}^w \Delta_W \right)^+ \cup \Delta_W$$

and the conditional active relation $\mathcal{A}^w$ as the symmetric relation

$$\mathcal{A}^w = \Delta \cup \mathcal{B}^w \cup \mathcal{B}^{-w} \cup \mathcal{K}^w \cup \left( \mathcal{B}^w \cup \mathcal{K}^w \right) \mathcal{C}^w \left( \mathcal{B}^{-w} \cup \mathcal{K}^w \right)$$

# Conditional topological separation

# Conditional topological separation

## Definitions of d- and t-separation as binary relations

# d-separation between vertices

Let $(\mathcal{V}, \mathcal{E})$ be a graph, that is, $\mathcal{V}$ is a set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$, and let $W \subset \mathcal{V}$ be a subset of (conditioning) vertices

**Definition ("a la Pearl")**

The vertices $\gamma$ and $\lambda$ are (conditionally) directionally separated (w.r.t. the subset $W$)

$$\gamma \underset{d}{\perp\!\!\!\perp} \lambda \mid W \iff \underbrace{D_U\big[\{(\gamma,\lambda)\} \mid \mathcal{V},\mathcal{E}\big]}_{\text{all "paths"}} \subset \underbrace{U_b^W(\mathcal{V},\mathcal{E})}_{\text{blocked "paths"}} \qquad (\forall \gamma, \lambda \in \mathcal{V})$$

- The vertices $\gamma$ and $\lambda$ are (conditionally) directionally separated if and only if all the extended-oriented paths, having them as endpoints, are blocked
- This definition mimics Pearl's d-separation, but the separation is between vertices and not between disjoint subsets

# d-separation as a binary relation

**Theorem (d-separation as a binary relation)**

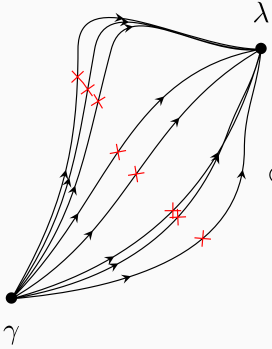
*For any vertices* $\gamma, \lambda \in \mathcal{V}$,

$$\gamma \perp\!\!\!\perp_d \lambda \mid W \iff \neg(\gamma \mathcal{A}^W \lambda)$$

- The proof of this theorem (d-separation as a binary relation) is quite technical, but involving simple mathematical objects, like paths in graphs and relations
- Pearl's d-separation between disjoint subsets is now expressed as

$$\Gamma \perp\!\!\!\perp_d \Lambda \mid W \iff \forall \gamma \in \Gamma\,,\ \forall \lambda \in \Lambda\,,\ \neg(\gamma \mathcal{A}^W \lambda)$$

$$\gamma \underset{d}{\nparallel} \lambda \mid W \iff \left\{ \vphantom{\int} \right.$$




$$\left. \subset U_b^W(\mathcal{V}, \mathcal{E}) \right\}$$

# Conditional topological separation between vertices, t-separation

We recall that $\overline{\Gamma}^{\mathcal{E}^W} = \mathcal{E}^{W*}\Gamma$ denotes
the $\mathcal{T}_{\mathcal{E}^W}$-topological closure of a subset $\Gamma \subset W$

**Definition**

We set

$$\mathfrak{S}^W = \Delta \cup \mathcal{C}^W\left(\mathcal{B}^{-W} \cup \mathcal{K}^W\right)$$

For any vertices $\gamma$, $\lambda \in \mathcal{V}$, we denote

$$\gamma \mathrel{\underset{t}{\perp\!\!\!\perp}} \lambda \mid W \iff \overline{\mathfrak{S}^W\gamma}^{\mathcal{E}^W} \cap \overline{\mathfrak{S}^W\lambda}^{\mathcal{E}^W} = \emptyset$$

and we say that the vertices $\gamma$ and $\lambda$ are
conditionally topologically separated (w.r.t. $W$)
or, shortly, t-separated

# Conditional topological separation

## Properties of d- and t-separation

# Topological separation is equivalent to d-separation

**Theorem**

*We have the equivalence*

$$\gamma \underset{t}{\perp\!\!\!\perp} \lambda \mid W \iff \gamma \underset{d}{\perp\!\!\!\perp} \lambda \mid W \qquad (\forall \gamma, \lambda \in W^{c})$$

**Proof** We have that

$$\begin{aligned}
&\Delta_{W^c}\big(\Delta \cup (\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}\big)\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\big(\Delta \cup \mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\big)\Delta_{W^c} \\
&\quad = \Delta_{W^c}\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\Delta_{W^c} && \text{(by developing)} \\
&\qquad \cup \Delta_{W^c}\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\big(\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\big)\Delta_{W^c} \\
&\qquad \cup \Delta_{W^c}\big((\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}\big)\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\Delta_{W^c} \\
&\qquad \cup \Delta_{W^c}\big((\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}\big)\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\big(\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\big)\Delta_{W^c} \\
&\quad = \Delta_{W^c}\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\Delta_{W^c} \\
&\qquad \cup \Delta_{W^c}\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\Delta_{W^c} && \text{(as } \mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\mathcal{C}^{W} = \mathcal{C}^{W} \text{ by (34c))} \\
&\qquad \cup \Delta_{W^c}(\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}\mathcal{E}^{-W*}\mathcal{E}^{W*}\Delta_{W^c} && \text{(also by (34c))} \\
&\qquad \cup \Delta_{W^c}(\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\Delta_{W^c} && \text{(also by (34c) applied twice)} \\
&\quad = \Delta_{W^c}\big(\mathcal{B}^{W} \cup \mathcal{K}^{W}\big)\mathcal{C}^{W}\big(\mathcal{B}^{-W} \cup \mathcal{K}^{W}\big)\Delta_{W^c} && \text{(by (34d) and (34e))} \\
&\qquad \cup \Delta_{W^c}\big(\mathcal{B}^{W} \cup \mathcal{K}^{W}\big)(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\Delta_{W^c} && \text{(by (34e))} \\
&\qquad \cup \Delta_{W^c}(\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}\big(\mathcal{B}^{-W} \cup \mathcal{K}^{W}\big)\Delta_{W^c} && \text{(by (34d))} \\
&\qquad \cup \Delta_{W^c}(\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\Delta_{W^c} \\
&\quad = \Delta_{W^c}(\mathcal{B}^{W} \cup \mathcal{K}^{W})\mathcal{C}^{W}(\mathcal{B}^{-W} \cup \mathcal{K}^{W})\Delta_{W^c}\,.
\end{aligned}$$

This ends the proof. ■

- We have started to check all the mathematical results with the help of the Coq proof assistant

# Topological separation (t-separation) between subsets

**Definition**

We define t-separation between subsets $\Gamma, \Lambda \subset W$ by

$$\Gamma \underset{t}{\perp\!\!\!\perp} \Lambda \mid W \iff \gamma \underset{t}{\perp\!\!\!\perp} \lambda \mid W\ ,\ \ \forall \gamma \in \Gamma\ ,\ \ \forall \lambda \in \Lambda$$

and we say that $\Gamma$ and $\Lambda$ are (conditionally) topologically separated (w.r.t.$W$)

**Theorem**

*The disjoint subsets $\Gamma, \Lambda \subset W$ are (conditionally) topologically separated (w.r.t.$W$) if and only if there exists $W_\Gamma, W_\Lambda \subset W$ such that*

$$W_\Gamma \sqcup W_\Lambda = W \ \ \textit{and} \qquad \underbrace{\overline{\Gamma \cup W_\Gamma}^{\mathcal{E}^W}}_{\mathcal{T}_{\mathcal{E}^W}\ \text{topological closure}} \quad \bigcap \overline{\Lambda \cup W_\Lambda}^{\mathcal{E}^W} = \emptyset$$

# Conditional topological separation

## Examples

Prove that $Y_1 \perp\!\!\!\perp_t Y_2 \mid W$ using $W = W_{Y_1} \sqcup \emptyset$

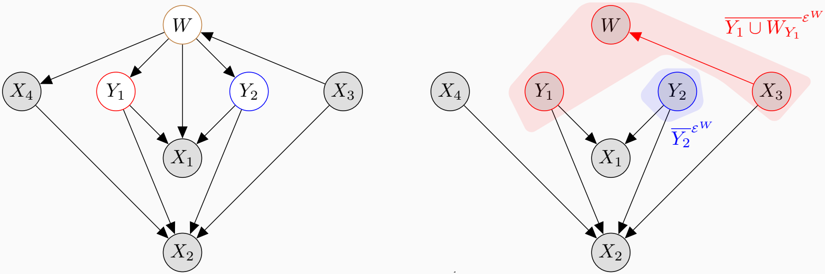
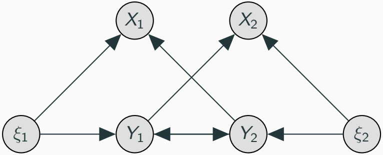
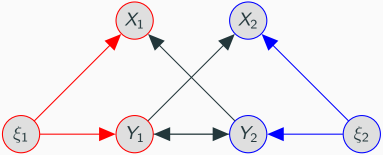



**Figure 1:** The split of $W$ is a piece of information that can be insightful




**(a)** Original graph.




**(b)** Let $W_{X_i} = Y_i$, for $i = 1, 2$. The closure of $X_1 \cup Y_1$ (resp. $X_2 \cup Y_2$), with the edges followed to build the closure, is in red (resp. blue).

**Figure 2:** Topological separation is easy to check: nonrecursive system

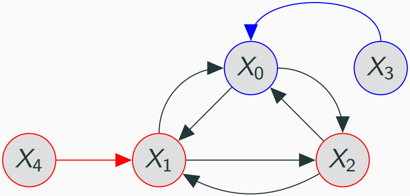



**Figure 3:** $X_3$ and $X_4$ are independent conditioned on $(X_0, X_1, X_2)$ but not independent if we only condition on $(X_0, X_1)$. The visual proof of topological separation is obtained by considering the splitting $W_{X_4} = \{X_1, X_2\}$ and $W_{X_3} = \{X_0\}$ and observing that the topological closure of $X_3 \cup W_{X_3}$ in blue does not intersect the topological closure of $X_4 \cup W_{X_4}$ in red

.

# References


Jean-Philippe Chancelier, Michel De Lara, and Benjamin Heymann, *Conditional separation as a binary relation*, 2021, Preprint.

Michel De Lara, Jean-Philippe Chancelier, and Benjamin Heymann, *Topological conditional separation*, 2021, Preprint.