Causal Inference Theory with Information Algebras: Conditional Topological Separation with Localization

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

Conditional topological separation with localization

H-Precedence as a way to express (functional) dependence

Definition

For given subset $H \subset \mathbb{H}$ of configurations (localization),
and focal agent $a \in A$,
the H-predecessor set $\mathcal{P}^H a \subset A$
is the smallest subset $B \subset A$ of agents such that

$$\mathcal{I}_a \cap H \subset \left(\bigotimes_{c \notin B} \{\emptyset, \mathbb{U}_c\} \otimes \bigotimes_{b \in B} \mathcal{U}_b \otimes \mathcal{F} \right) \cap H$$

Thus, $\mathcal{P}^H$ defines a binary relation on the agents set A such that

$b \mathcal{P}^H a \iff b \in \mathcal{P}^H a$

$\iff$ agent b's actions indeed affect agent a's information (on H)

$\iff$ agent b's actions are arguments of agent a's W-strategies (on H)

$$\lambda_a(\omega, \{u_b\}_{b \in A}) = \lambda_a(\{u_b\}_{b \in \mathcal{P}^H a}, \cancel{\{u_b\}_{b \notin \mathcal{P}^H a}}, \omega)$$

(W, H)-conditional precedence

Definition

For given subset $H \subset \mathbb{H}$ of configurations (localization),
subset $W \subset A$ of agents (conditioning),
and focal agent $a \in A$,
the (W, H)-conditional predecessor set $\mathcal{P}^{W,H}a \subset A$
is the smallest subset $B \subset A$ of agents such that

$$\mathfrak{I}_a \cap H \subset \left(\bigotimes_{c \notin B \cup W} \{\emptyset, \mathbb{U}_c\} \otimes \bigotimes_{b \in B \cup W} \mathfrak{U}_b \otimes \mathfrak{F} \right) \cap H$$

(W, H)-conditional precedence

Proposition

$$\mathcal{P}^{W,H} = \underbrace{\Delta_{W^c}}_{\substack{\textit{subdiagonal} \\ \textit{relation}}} \quad \underbrace{\mathcal{P}^{H}}_{\substack{\textit{precedence} \\ \textit{relation } \mathcal{P}^{\emptyset,H}}}$$

$$b \in \mathcal{P}^{W,H} a \iff b \notin W \text{ and } b \in \mathcal{P}^{H} a$$

- The graph $(\mathcal{V}, \mathcal{P}^{\emptyset,\mathbb{H}})$ is the original precedence graph
- The graph $(\mathcal{V}, \mathcal{P}^{\emptyset,H})$ is a subgraph of $(\mathcal{V}, \mathcal{P}^{\emptyset,\mathbb{H}})$, where some edges have been removed (by an appropriate choice of H, do-variables may be introduced)

Discussion

Wrapping-up

- Information dependency models ≈
  causality with information σ-algebras
  (from Witsenhausen's 1971 "intrinsic model")
- An interesting perspective from decision theory
  - captures causality without reference to functional dependencies,
    but with information σ-algebras
    (Witsenhausen has his own definition of causality/causal ordering,
    which implies solvability by means of
    a recursively computable solution map)
  - elegant style of expression and proof: equational reasoning

→ In this presentation, we have focused on the topological separation,
which we have proved to be and alternative definition of d-separation
(but there is more)

Making the case for Information Dependency Model

- The three rules of do-calculus reduce to a unique sufficient condition for conditional independence

$$Y \underset{t}{\perp\!\!\!\perp} Z \mid (W, H) \Longrightarrow \Pr\left(U_Y \mid U_W, U_{\bar{Z}}, H\right) = \Pr\left(U_Y \mid U_W, H\right)$$

- Another axiomatization of causality
  - without functions — but with σ-algebras
  - without probabilities
  - without (finite) graphs — but with binary relations (on possibly infinite sets) and induced (Alexandrov) topologies
- A bridge between causality, game theory, control theory and reinforcement learning?

References

H. S. Witsenhausen, *On information structures, feedback and causality*, SIAM J. Control **9** (1971), no. 2, 149–160.