Counterfactuals and Updates In a Causal Setting

Alexander Bochman

Holon Institute of Technology (HIT)
Israel

BRA2015

Pearl's Causal Models

Causal model $M = \langle U, V, F \rangle$

(i) U is a set of *background* (*exogenous*) variables, V is a finite set of *endogenous* variables.

(ii) F is a set of functions $f_i : U \cup (V \backslash \{V_i\}) \mapsto V_i$ for each $V_i \in V$.

F is represented by equations $v_i = f_i(pa_i, u_i)$, where PA_i (parents) is the unique minimal set in $V \backslash \{V_i\}$ sufficient for representing f_i.

Every instantiation $U = u$ determines a "causal world" of the model.

Submodels

A submodel M_x of M is obtained by replacing F with the set:

$$F_x = \{f_i \mid V_i \notin X\} \cup \{X = x\}.$$

Submodels provide answers to counterfactual queries.

Propositional reformulation

Propositional atoms are partitioned into a set of *exogenous* atoms and a finite set of *endogenous* atoms.

- A *Boolean structural equation* is an expression of the form $A = F$, where A is an endogenous atom and F is a propositional formula in which A does not appear.
- A *Boolean causal model* is a set of Boolean structural equations $A = F$, one for each endogenous atom A.
- A *solution* (or a *causal world*) of a Boolean causal model M is any propositional interpretation satisfying $A \leftrightarrow F$ for all $A = F$ in M.

Submodels

If I is a truth-valued function on a set X of endogenous atoms, the *submodel* M_X^I of M is obtained from M by replacing every equation $A = F$, where $A \in X$, with $A = I(A)$.

Firing squad

U, C, A, B, D stand for “Court orders the execution”, “Captain gives a signal”, “Rifleman A shoots”, “Rifleman B shoots”, and “Prisoner dies.” The Boolean causal model $\{C = U,\ A = C,\ B = C,\ D = A \vee B\}$ has two solutions, which give us a *prediction* $\neg A \rightarrow \neg D$:

If rifleman A did not shoot, the prisoner is alive.

an *abduction* $\neg D \rightarrow \neg C$, and even a *transduction* $A \rightarrow B$:

If the prisoner is alive, the Captain did not signal.

If rifleman A shot, then B shot as well.

The submodel $\{C = U,\ A = \mathbf{t},\ B = C,\ D = A \vee B\}$ implies $\neg C \rightarrow (D \wedge \neg B)$, which justifies

If the captain gave no signal and rifleman A decides to shoot, the prisoner will die and B will not shoot.

First-order reformulation

Object constants are partitioned into *rigid*, *exogenous*, and a finite set of *endogenous* symbols.

- A *structural equation* is an expression $c = t$, where c is endogenous, and t a ground term in which c does not appear.
- A *causal model* is a first-order interpretation of rigid and function symbols, plus a set of structural equations $c = t$, one for each endogenous symbol c.
- A *causal world* of a causal model M is an extension of the interpretation of rigid and function symbols in M to the exogenous and endogenous symbols that satisfies all *equalities* $c = t$ in M.

Submodels

For a set X of endogenous symbols and a function I from X to the set of rigid constants, the *submodel* M^I_X of M is the causal model obtained from M by replacing every equation $c = t$, where $c \in X$, with $c = I(c)$.

An Ideal Gas model

The physical setup: a closed gas container with variable volume that can be heated. Pressure (P) and volume (V) are endogenous, while temperature (T) is exogenous.

$$P = c \cdot \frac{T}{V} \qquad V = c \cdot \frac{T}{P}$$

Fixing the volume V produces a submodel

$$P = c \cdot \frac{T}{V} \qquad V = v$$

that corresponds to the *Gay-Lussac's Law*: pressure is proportional to temperature (though the temperature is not *determined* by the pressure). Similarly, fixing the pressure P gives a submodel

$$P = p \qquad V = c \cdot \frac{T}{P}$$

that represents the *Charles's Law*: volume is proportional to temperature (though not vice versa).

Causal Calculus
Propositional case

Causal rules: $A \Rightarrow B$, where A, B are classical propositional formulas.

A *causal theory* Δ is a set of causal rules.

$$\Delta(u) = \{B \mid A \Rightarrow B \in \Delta, \text{ for some } A \in u\}$$

Nonmonotonic Semantics

A world α is an *exact model* of a causal theory Δ if it is a unique model of $\Delta(\alpha)$.

$$\alpha = \mathrm{Th}(\Delta(\alpha))$$

Exact world is closed wrt the causal rules, and any proposition in it is caused (explained).

Determinate causal theories and completion

Determinate causal theory: heads are literals or **f**. A determinate causal theory is *definite* if no literal is the head of infinitely many rules.

The *(literal) completion* of a definite causal theory Δ is the set of classical formulas

$$p \leftrightarrow \bigvee\{A \mid A \Rightarrow p \in \Delta\} \qquad \neg p \leftrightarrow \bigvee\{A \mid A \Rightarrow \neg p \in \Delta\},$$

for every atom p, plus the set $\{\neg A \mid A \Rightarrow \mathbf{f} \in \Delta\}$.

Proposition (McCain&Turner 1997)

The nonmonotonic semantics of a definite causal theory coincides with the classical semantics of its completion.

Causal Logic (Bochman 2003)

Causal inference relations:

(Strengthening) If $A \vDash B$ and $B \Rightarrow C$, then $A \Rightarrow C$;

(Weakening) If $A \Rightarrow B$ and $B \vDash C$, then $A \Rightarrow C$;

(And) If $A \Rightarrow B$ and $A \Rightarrow C$, then $A \Rightarrow B \wedge C$;

(Or) If $A \Rightarrow C$ and $B \Rightarrow C$, then $A \vee B \Rightarrow C$;

(Cut) If $A \Rightarrow B$ and $A \wedge B \Rightarrow C$, then $A \Rightarrow C$;

(Truth/Falsity) $\mathbf{t} \Rightarrow \mathbf{t}$; $\mathbf{f} \Rightarrow \mathbf{f}$.

Logical Semantics

$A \Rightarrow B$ is *valid* in a Kripke model (W, R, V) if, for any worlds α, β such that $R\alpha\beta$, if A holds in α, then B holds in β.

A modal representation of causal rules: $A \Rightarrow B \equiv A \rightarrow \Box B$.

Let $\Rightarrow_\Delta$ be the least causal inference relation that includes a causal theory Δ.

Adequacy

Exact models of Δ coincide with the exact models of $\Rightarrow_\Delta$.

Causal theories Δ and Γ are

- *strongly equivalent* if, for any set Φ of causal rules, $\Delta \cup \Phi$ has the same nonmonotonic semantics as $\Gamma \cup \Phi$;
- *causally equivalent* if $\Rightarrow_\Delta = \Rightarrow_\Gamma$.

Strong equivalence

Causal theories Δ and Γ are strongly equivalent if and only if they are causally equivalent.

First-order causal calculus (Lifschitz 1997)

Causal rules: $G \Rightarrow F$, where F and G are first-order formulas.

A *first-order causal theory* Δ is a finite set of causal rules and a list **c** of object, function and predicate constants - the *explainable* symbols of Δ.

$$\Delta(v\mathbf{c}) \equiv \bigwedge\{\forall \mathbf{x}(G \to F^{\mathbf{c}}_{v\mathbf{c}}) \mid G \Rightarrow F \in \Delta\},$$

where $F^{\mathbf{c}}_{v\mathbf{c}}$ is the result of substituting new variables $v\mathbf{c}$ for **c** in F.

The *nonmonotonic semantics* of the causal theory Δ is described by

$$\forall v\mathbf{c}(\Delta(v\mathbf{c}) \leftrightarrow (v\mathbf{c} = \mathbf{c})).$$

The interpretation of the explainable symbols is the only interpretation that is determined, or "causally explained," by the rules of Δ.

Functional completion

If every explainable symbol of Δ is an object constant, and Δ consists of rules of the form

$$G(x) \Rightarrow c = x,$$

one for each explainable symbol c, then the *(functional) completion* of Δ is the conjunction of the first-order sentences

$$\forall x(c = x \leftrightarrow G(x))$$

for all rules of Δ.

Two-level representation

Pearl's causal models

- Structural equations and their solutions
- Interventions/submodels

$\Rightarrow$

$\Rightarrow$

Causal calculus

- Nonmonotonic semantics
- Causal logic

The Representation
Propositional case

For a Boolean causal model M, Δ_M is the propositional causal theory consisting of the rules

$$F \Rightarrow A \qquad \neg F \Rightarrow \neg A$$

for all equations $A = F$ in M and the rules

$$A \Rightarrow A \qquad \neg A \Rightarrow \neg A$$

for all exogenous atoms A of M.

Theorem

The causal worlds of a Boolean causal model M are identical to the exact models of Δ_M.

Firing squad, continued

The causal theory Δ_M for the firing squad example:

$$U \Rightarrow C,\ \neg U \Rightarrow \neg C,$$
$$C \Rightarrow A,\ \neg C \Rightarrow \neg A, \quad C \Rightarrow B,\ \neg C \Rightarrow \neg B,$$
$$A \lor B \Rightarrow D,\ \neg(A \lor B) \Rightarrow \neg D,$$
$$U \Rightarrow U,\ \neg U \Rightarrow \neg U.$$

This causal theory has two exact models, identical to the solutions (causal worlds) of M.

The Representation
Subtheories

Given a set X of atoms and a truth-valued function I on X, the *subtheory* Δ^I_X of a determinate causal theory Δ is obtained from Δ by

- removing all rules $A \Rightarrow p$ and $A \Rightarrow \neg p$ with $p \in X$, and
- adding $\mathbf{t} \Rightarrow p$ for each $p \in X$ such that $I(p) = \mathbf{t}$,
- adding $\mathbf{t} \Rightarrow \neg p$ for each $p \in X$ such that $I(p) = \mathbf{f}$.

Example (Firing squad, continued)

The submodel $M^I_{\{A\}}$ with $I(A) = \mathbf{t}$ corresponds to the subtheory $\Delta^I_{\{A\}}$:

$$
\begin{gathered}
U \Rightarrow C,\ \neg U \Rightarrow \neg C, \\
\mathbf{t} \Rightarrow A, \\
C \Rightarrow B,\ \neg C \Rightarrow \neg B,\ A \vee B \Rightarrow D,\ \neg(A \vee B) \Rightarrow \neg D, \\
U \Rightarrow U,\ \neg U \Rightarrow \neg U.
\end{gathered}
$$

First-order Representation

For a first-order causal model M, Δ_M is the first-order causal theory whose explainable constants are the endogenous symbols of M, and whose rules are

$$x = t \Rightarrow x = c,$$

for every structural equation $c = t$ from M.

Theorem

An extension of the interpretation of rigid and function symbols in M to the exogenous and endogenous symbols on a universe of cardinality > 1 is a solution of M iff it is a nonmonotonic model of Δ_M.

Summary

- The causal calculus provides an adequate logical framework for representing and computing updates and counterfactuals in a causal setting.