Refutation of probabilistic reasoning across the causal hierarchy

Abstract: We evaluate equations for axioms, propositions, and theorems as not tautologous. The last as \( P(X \land Z)Y \mid X \land Z \mid W) \equiv P(X \mid Y \mid X \mid W) \) is the briefest refutation of the conjecture of probabilistic reasoning across the causal hierarchy. These results form a non tautologous fragment of the universal logic VŁ4.

We assume the method and apparatus of Meth8/VŁ4 with Tautology as the designated proof value, F as contradiction, N as truthity (non-contingency), and C as falsity (contingency). The 16-valued truth table is row-major and horizontal, or repeating fragments of 128-tables, sometimes with table counts, for more variables. (See ersatz-systems.com.)

Note for clarity, we usually distribute quantifiers onto each designated variable.


Abstract We propose a formalization of the three-tier causal hierarchy of association, intervention, and counterfactuals as a series of probabilistic logical languages. Our languages are of strictly increasing expressivity, the first capable of expressing quantitative probabilistic reasoning—including conditional independence and Bayesian inference—the second encoding do-calculus reasoning for causal effects, and the third capturing a fully expressive do-calculus for arbitrary counterfactual queries. We give a corresponding series of finitary axiomatizations complete over both structural causal models and probabilistic programs, and show that satisfiability and validity for each language are decidable in polynomial space.

Introduction and summary … The aim of the present article is to gain conceptual as well as technical insight into this hierarchy by employing tools from logic. … The languages differ in how much they can express about these models. L₁, the language of association, expresses only “pure” probabilistic facts and relationships; L₂, the language of probabilistic intervention, allows expressing probabilities of basic conditional “if. . . then. . . ” statements; L₃, the language of probabilistic counterfactuals, encodes probabilities for arbitrary boolean combinations of such conditional statements. Using standard ideas from logic and existing results, we can address questions about definability and expressiveness. For instance, it is easy to prove in our framework that each language is strictly more expressive than those below it in the hierarchy (Prop. 1, 2 below). We can also interpret well-known insights from the graphical models and causal learning literatures as graph definability results for appropriate probabilistic logical languages, analogously to correspondence theory in modal logic . . .

Comparing expressivity With a precise semantic interpretation of our three languages in hand, we can now show rigorously that they form a strict hierarchy, in the sense that models ma be distinguishable only by moving up to higher levels of the hierarchy.
**Proposition 1.** \(L_2\) is strictly more expressive than \(L_1\). Proof. … It is easy to see by an induction on terms in \(L_1\) that \([t]M_1 = [t]M_2\), and thus \(M_1\) and \(M_2\) validate the same \(L_1\) formulas. Yet, \(M_1 \vDash P(X = 1) Y = 1 \equiv 1\), while \(M_2 \nvdash P(X = 1) Y = 1 \equiv 1\). In particular the schema \(P(\beta | \alpha) \equiv P(\alpha | \beta)\) is also falsified by \(M_2\), a reflection of the distinction between observation and intervention. (1.4.1)

**Remark 1.4.1:** We evaluate the particular schema.

\[
\text{LET} \quad p, q, r, s, t, u, v, x, y, z: \quad P, M_1, M_2, \ s, G, \alpha, \beta, \ x, \ y, \ z.
\]

\[
\sim((p\land((x=(s>s))\land y)=(s>s)))>r)\lor((p\land(v\land u))=(p\land(u\land v))) ;
\]

\[
\begin{align*}
\text{TNTN} & \quad \text{TNTT} & \quad \text{TNTN} & \quad \text{TNTT} (32) \\
\text{TTFT} & \quad \text{TTTT} & \quad \text{TTTT} & \quad \text{TTTT} (16)
\end{align*}
\]

(1.4.2)

**Proposition 2.** \(L_3\) is strictly more expressive than \(L_2\). Proof. … It is then easy to check (by induction) that \(M_1\) and \(M_2\) validate all the same \(L_2\) formulas, whereas …, \([p]M_1\neq [p]M_2\), with \(p\) the term denoting the probability of necessity and sufficiency \(P([X=0]Y=0 \land [X=1]Y=1)\). (2.2.1)

\[
(((x=(s>s))\land(y=(s>s)))\land((x=(s>s))\land(y=(s>s))))\land((p\land((x=(s>s))\land y)=(s>s)))>q)) @ (((x=(s>s))\land(y=(s>s)))\land((x=(s>s))\land(y=(s>s)))) \sim((p\land((x=(s>s))\land y)=(s>s)))>r) ;
\]

\[
\begin{align*}
\text{CCCF} & \quad \text{CFCC} & \quad \text{CCCF} & \quad \text{CFCC} (16) \\
\text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF} (48)
\end{align*}
\]

(2.2.2)

**Graph definability and do-calculus … Proposition 3.** \(G \vDash P(X \land Y | Z) \equiv P(X | Z) P(Y | Z)\) if and only if \((X \perp \perp Y | Z)\). In other words, the graph property of d-separation is definable in \(L_1\)… We leave further exploration of questions about graph definability in these languages for a future occasion. (3.1.1)

\[
(((x@y)\land(z\land t))>((p\land((x@y)\land z))=((p\land(x\land z))\land(p\land(y\land z))))>t ;
\]

\[
\begin{align*}
\text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF} \times 40 \\
\text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF} \times 16 \\
\text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF} \times 8 \\
\text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF} & \quad \text{TFFF}
\end{align*}
\]

(3.1.2)

**Axiomatizations …** The following 15 axioms constitute Poly.

\[
\text{OrdTot.} \quad t_1 > t_2 \lor t_2 > t_1
\]

\[
\text{LET} \quad p, q, r, s: \quad t \lor t_1, t_2, t_3, s; \sim \sim.
\]

\[
\neg(q>p)+\neg(p>q) ;
\]

\[
\begin{align*}
\text{TTTF} & \quad \text{FTTF} & \quad \text{FTTF} & \quad \text{FTTF}
\end{align*}
\]

(6.1.2)

**One.** \(t \cdot 1 = t\) (6.6.1)

\[
(p\land( s>s))= p ;
\]

\[
\begin{align*}
\text{TNTN} & \quad \text{TNTN} & \quad \text{TNTN} & \quad \text{TNTN}
\end{align*}
\]

(6.6.2)

**Neg.** \(t + (−t) = 0\) (6.15.1)

\[
(p+~p)=(s@s) ;
\]

\[
\begin{align*}
\text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF} & \quad \text{FFFF}
\end{align*}
\]

(6.15.2)
We use the principles below, all derivable from Poly: ...

\[
\text{NegAdd. } - (t_1 + t_2) \equiv (-t_1) + (-t_2) \tag{7.3.1}
\]

\[
\text{LET } p, q: t_1, t_2.
\]

\[
\neg(p+q) = \neg p \neg q; \quad \text{TFF\ TFF\ TFF\ TFF} \tag{7.3.2}
\]

**Conclusion** We noted two seminal formula schemas from L2 that feature centrally in the do-calculus:

\[
P([X \land Z]Y \mid [X \land Z]W) \equiv P(X)Y \mid [X](Z \land W) \tag{10.1}
\]

\[
P([X \land Z]Y \mid [X \land Z]W) \equiv P([X]Y \mid [X]W) \tag{11.1}
\]

\[
(p \& (((x \& z) \& y) \neg (x \& z) \& w)) = (p \& ((x \& y) \neg (x \& w))) ;
\]

\[
\text{TFF\ TFF\ TFF\ TFF (56)}
\]

\[
\text{TFF\ TFF\ TFF\ TFF (8)}
\]

\[
\text{TFF\ TFF\ TFF\ TFF (64)}
\]

**Remark 11.2:** Eq. 11.2 is not tautologous. Because the antecedents of 10.1 and 11.1 are identical, and 10.1 is a trivial tautology, this means further that the consequent of 10.1 cannot be identical to the consequent of 11.1. In one formula, this is the briefest refutation of the conjecture of probabilistic reasoning across the causal hierarchy.

The eight Eqs. 1.4.2, 2.2.2, 3.1.2, 6.1.2, 6.6.2, 6.15.2, 7.3.2, and 11.2 as rendered are not tautologous. These refute the claimed probabilistic reasoning across the causal hierarchy.