Argumentation-based Probabilistic Causal Reasoning

Proofs of technical results

Lars Bengel[1], Lydia Blümel[1], Tjitze Rienstra[2], and Matthias Thimm[1]

[1] Artificial Intelligence Group, University in Hagen, Germany
{firstname.lastname}@fernuni-hagen.de
[2] Dep. of Advanced Computing Sciences, Maastricht University, The Netherlands
t.rienstra@maastrichtuniversity.nl

Proposition 2. *For any causal model $\mathcal{C} = (K, \mathsf{P})$, the probability distribution $P_{\mathcal{C}}$ sums up to 1.*

Proof. *A causal state $C \in 2^{U(K)}$ is a configuration of the background atoms. The probabilities of the background atoms are independent of each other. The probability of the causal state C is then per definition the product of the probabilities that each background atom holds true (or false, if $u \notin C$). Clearly, since the causal states span the power set of background atoms, there exists exactly one causal state per configuration of background atoms and it follows directly that $\sum_{C \in 2^{U(K)}} P_{\mathcal{C}}(C) = 1$.*

Proposition 3. *For any causal model $\mathcal{C} = (K, \mathsf{P})$ and observation ϕ, the probability distribution P_{AF} sums up to 1.*

Proof. *From Proposition 2 we know that $P_{\mathcal{C}}$ is well defined. Now, per Definition 9 for some framework state S $P_{AF}(S)$ is the sum of probabilities $P_{\mathcal{C}}(C)$ over all causal states C that induce exactly the set of arguments corresponding to the framework state S. Recall the definition of $\mathsf{Arg}_{\mathcal{C}}(C)$:*

$$\mathsf{Arg}_{\mathcal{C}}(C) = \{(\Phi, \psi) \in \mathsf{Arg}_{\mathcal{C}} \mid \Phi \cup C \cup \{\neg u \mid u \notin C\} \nvdash \bot\}.$$

Clearly, every causal state corresponds exactly to set of arguments. Since the set framework states spans the whole power set of Arg, it follows that every causal state must correspond to exactly one framework state. If a framework state has no corresponding causal state its probability is per definition 0. Thus, since $P_{\mathcal{C}}$ sums up to 1 it follows directly that P_{AF} also sums up to 1.

Theorem 1. *Let $\mathcal{C} = (K, \mathsf{P})$ be a probabilistic causal model and $\phi \mathrel{|\!\sim}_{\mathcal{C}} \psi$ is a causal statement. Then $P(\phi \mathrel{|\!\sim}_{\mathcal{C}} \psi) = P_{\mathcal{C}}(\psi | \phi)$.*

Proof. *Consider the probabilistic causal model $\mathcal{C} = (K, \mathsf{P})$. The probability $P(\phi \mathrel{|\!\sim}_{\mathcal{C}} \psi)$ that the statement holds is computed according to (2) as*

$$P(\phi \mathrel{|\!\sim}_{\mathcal{C}} \psi) = \frac{\sum_{S \in \mathsf{S}_{[\psi = true]}} P_{AF}(S)}{\sum_{C \in \mathsf{C}(\phi)} P_{\mathcal{C}}(C)}.$$

On the other hand, the conditional probability is computed as $P_{\mathcal{C}}(\psi|\phi) = \frac{P_{\mathcal{C}}(\psi \wedge \phi)}{P_{\mathcal{C}}(\phi)}$. *We show*

1. $P_{\mathcal{C}}(\phi) = \sum_{C \in \mathsf{C}(\phi)} P_{\mathcal{C}}(C)$, *and*
2. $P_{\mathcal{C}}(\psi \wedge \phi) = \sum_{S \in \mathsf{S}_{[\psi = true]}} P_{AF}(S)$.

to 1.) *This follows directly from Definition 5 and Equation (1), since causal states correspond directly to valuations of the background variables. The probability is computed as the sum of probabilities of all valuations (causal states) for which* ϕ *is true, which is exactly how the probability* $P_{\mathcal{C}}(\phi)$ *is computed.*

to 2.) *The probability* $P_{\mathcal{C}}(\phi \wedge \psi)$ *is defined as the sum of probabilities over all valuations of the background atoms under which* $K \vdash \phi \wedge \psi$.

$$P_{\mathcal{C}}(\phi \wedge \psi) = \sum_{V \in \mathcal{V}_{\phi \wedge \psi}} P_{\mathcal{C}}(V(U(K)))$$

$$\text{where } \mathcal{V}_{\phi \wedge \psi} = \{u : U(K) \to \{true, false\} \mid val(K \vdash \phi \wedge \psi, u) = true\}$$

On the other hand,

$$\sum_{S \in \mathsf{S}_{[\psi = true]}} P_{\mathsf{AF}}(S)$$

amounts to the sum of probabilities over all framework states of the induced PAF that entails the conclusion ψ. *Important to note is that the induced PAF contains only arguments that are a subset of some causal states* $C \in \mathsf{C}(\phi)$, *i. e., causal states in which* ϕ *is evaluated to true. A framework state* S *is a set of arguments* $S \subseteq \mathsf{Arg}_{\mathcal{C}}$ *to which we assign, according to Def. 9, the probability*

$$P_{\mathsf{AF}}(S) = \sum_{C \in \mathsf{C}(\phi, S)} P_{\mathcal{C}}(C).$$

In other words, the probability $P_{\mathsf{AF}}(S)$ *of a framework state* S *corresponds to the sum of probabilities over the set of causal states* $\mathsf{C}(\phi, S)$ *for which* ϕ *is evaluated to true and that induce the same set of arguments* S. *Per Equation (1), the probability of a causal state* C *is equal to the probability of the valuation that evaluates exactly those background atoms to true, that are in the causal state* C.

We defined the probability $P(\phi \mid\!\sim_{\mathcal{C}} \psi)$ *via* $\sum_{S \in \mathsf{S}_{[\psi = true]}} P_{\mathsf{AF}}(S)$, *which is the probability over those framework states for which every stable extension contains at least one argument with the conclusion* ψ. *This is then the sum of probabilities over exactly those valuations for which* ϕ *and* ψ *are true. Thus, it follows that*

$$P_{\mathcal{C}}(\psi \wedge \phi) = \sum_{V \in \mathcal{V}_{\phi \wedge \psi}} P_{\mathcal{C}}(V(U(K))) = \sum_{S \in \mathsf{S}_{[\psi = true]}} P_{\mathsf{AF}}(S).$$

Therefore, we have that $P(\phi \mid\!\sim_{\mathcal{C}} \psi) = P_{\mathcal{C}}(\psi|\phi)$ *and we have shown that the argumentation-based approach to probabilistic causal reasoning is equivalent to that introduced by Pearl.* □

Theorem 2. *Let $\mathcal{C} = (K, \mathsf{P})$ be a probabilistic causal model. Given a counterfactual statement $\phi \mid\!\sim_{\mathcal{C}^*_{[v^*=x]}} \psi^*$, we have that $P(\phi \mid\!\sim_{\mathcal{C}^*_{[v^*=x]}} \psi^*) = P_{\mathcal{C}}(\psi \mid \phi)$.*

Proof. We prove the equivalence by showing the following two things: (1) the equivalence between our argumentation-based probabilistic twin network approach and the standard twin network approach and (2) the equivalence between the standard three-step procedure and the twin network approach.

For (1), we have to show that our approach to computing the probability $P(\phi \mid\!\sim_{\mathcal{C}^*_{[v^*=x]}} \psi^*)$ in the twin model via the induced probabilistic argumentation framework is equivalent to computing the conditional probability $P_{\mathcal{C}}(\psi^* \mid \phi)$ via Pearl's twin network method. Per Definition 11, the twin model of some causal model is also a causal model itself with additional atoms representing the twin world. Thus, it follows directly from the proof of Theorem 1 that computing the probability via the induced probabilistic argumentation framework of the twin model is equivalent to the standard way of computing the conditional probability in that twin causal model.

For (2), the equivalence between the three-step procedure and the twin network method for evaluating counterfactuals in probabilistic causal models has already been shown by Pearl.

Hence, it follows that our argumentation-based twin network approach is equivalent to the three-step procedure for probabilistic counterfactual reasoning.

□