CLIMBING THE LATTER OF CAUSATION WITH tnreason

RESEARCH NOTES IN THE ENEXA AND QROM PROJECTS

January 27, 2026

ABSTRACT

We here describe how to represent Causal AI models in the tnreason formalism.

We find tensor network models for the different rungs of the *ladder of causation* Pearl and Mackenzie (2018). The notes are organized to follow the three rungs of the ladder:

- **Rung 1: Association** – captured by Bayesian Networks
- **Rung 2: Intervention** – captured by do-calculus, for which we introduce intervention variables
- **Rung 3: Counterfactuals** – captured by functional causal models (see Chapter 21 in Koller and Friedman (2009)) involving response variables

Following the formalism of Chapter 21 in Koller and Friedman (2009):

- Mutilation of the Bayesian network is captured by contractions with one-hot encodings to do-variables D_k.
- Intervention queries are formalized by contractions with one-hot encodings to all do-variables
- Response variables are formalized by selection variables
- Counterfactual queries are modeled as intervention queries on counterfactual twin networks (see (Koller and Friedman, 2009, Chapter 21))

1 Rung 1: Association

See main report: Bayesian Networks are tensor networks and probabilistic queries can be answered by their contractions.

2 Rung 2: Intervention

So far, there are several hypergraphs to model a probability distributions by a Bayesian Network. We now want to understand the directed hyperedges as causal relationships between the variables, beyond encoding of conditional independencies (see Example 1). To this end, we model the effects of interventions, which will distinguish between multiple Bayesian Networks modeling the same joint distribution.

Example 1 (Patient treatment outcome)**.** *We model a relationship about the treatment of a sick patient and the outcome of the treatment. Whether a treatment is conducted is modelled by the binary variable X_T and whether the patient improves after the treatment by the variable X_0.*

There are in general two possibilities of representing the joint distribution of X_T and X_O by a Bayesian Network:

- *Treatment directs to outcome:*

$$\mathcal{G} = (\{X_T, X_O\}, \{(\varnothing, X_T), (\{X_T\}, \{X_O\})\})$$

- *Outcome directs to treatment:*

$$\mathcal{G} = (\{X_T, X_O\}, \{(\varnothing, X_T), (\{X_O\}, \{X_T\})\})$$

In the typical intuition, we expect the treatment to have a causal effect on the outcome and not the other way round. From a Bayesian Network perspective, both models are however equivalent and represent the same joint distribution.

Interventions on Bayesian Networks are modeled by the do-calculus Pearl (2009), which we capture here based on do-variables.

2.1 Do calculus

We introduce the do-variables D_k taking values in $[m_k + 1]$ for each variable X_k of dimension m_k, which model the effects of interventions. The states are interpreted as

- $d_k \in [m_k]$: An intervention setting X_k to the value d_k
- $d_k = m_k$: No intervention on X_k

Definition 1. *The causally augmented conditional probability core is defined as*

$$\begin{aligned}\mathbb{P}^c\left[X_k \middle| D_k, X_{\mathrm{Pa}(k)}\right] =& \mathbb{P}\left[X_k \middle| X_{\mathrm{Pa}(k)}\right] \otimes \epsilon_{m_k}\left[D_k\right] \\ &+ \sum_{d_k \in [m_k]} \epsilon_{d_k}\left[X_k\right] \otimes \epsilon_{d_k}\left[D_k\right] \otimes \mathbb{I}\left[X_{\mathrm{Pa}(k)}\right] .\end{aligned}$$

The causally augmented Bayesian Network is the tensor network

$$\left\langle \{\mathbb{P}^c\left[X_k \middle| D_k, X_{\mathrm{Pa}(k)}\right] \,:\, k \in [d]\} \right\rangle_{\left[X_{[d]}, D_{[d]}\right]} .$$

The Bayesian Networks investigated so far are retrieved by contractions with $\epsilon_{m_k}\left[D_k\right]$. This corresponds with the situation, where no interventions are performed on any variable.

Lemma 1. *For any Bayesian Network $\mathbb{P}\left[X_{[d]}\right]$ on a directed acyclic hypergraph $\mathcal{G} = ([d], \mathcal{E})$ we have*

$$\mathbb{P}\left[X_{[d]}\right] = \left\langle \{\mathbb{P}^c\left[X_k \middle| D_k, X_{\mathrm{Pa}(k)}\right] \,:\, k \in [d]\} \cup \{\epsilon_{m_k}\left[D_k\right] \,:\, k \in [d]\} \right\rangle_{\left[X_{[d]}\right]} .$$

Proof. By construction we have

$$\mathbb{P}^c\left[X_k \middle| D_k = m_k, X_{\mathrm{Pa}(k)}\right] = \mathbb{P}\left[X_k \middle| X_{\mathrm{Pa}(k)}\right] .$$

Performing this on each causally augmented conditional distribution retrieves thus the original Bayesian Network. □

2.2 Intervention Queries

In a Bayesian Network, when intervening on a variable X_v, the corresponding conditional probability gets trivialized. The probability tensor is then captured by

$$\mathbb{P}\left[X_{\mathcal{U}} \middle| \mathrm{do}\left(X_{\mathcal{W}} = x_{\mathcal{W}}\right)\right] = \left\langle \{\mathbb{P}\left[X_v \middle| X_{\mathrm{Pa}(v)}\right] \,:\, v \notin \mathcal{W}\} \cup \{\frac{1}{m_v}\mathbb{I}\left[X_v, X_{\mathrm{Pa}(v)}\right] \,:\, v \in \mathcal{W}\} \cup \{\epsilon_{x_v}\left[X_v\right] \,:\, v \in \mathcal{W}\} \right\rangle_{\left[X_{\mathcal{U}}\right]} .$$

Since $\frac{1}{m_v}\mathbb{I}\left[X_v, X_{\mathrm{Pa}(v)}\right]$ is directed with $X_{\mathrm{Pa}(v)}$ incoming and X_v outgoing, the partition function of the network stays 1 and the normalization is captured by the contraction.

2.3 Simplifications

In general, when all variables are observable, intervention queries can be expressed by a collection of conditional queries (since any cpd of a Bayesian network is a conditional query, and we contract all except those with interventions on the head variables). In realistic scenarios, however, not all variables are observable (for example, the smoking gene has been only assumed, but no observations have been made). We are interested in situations, where intervention queries can be answered by conditional queries of observable variables. We can proof them in the tensor network formalism based on network separations, which contribution to contractions are scalar multiplications, which therefore are dropped in normalizations.

An example for a simplification is given by the randomized controlled trial.

Example 2 (Randomized controlled trial)**.** *The randomized controlled trial is a joint distribution of variables*

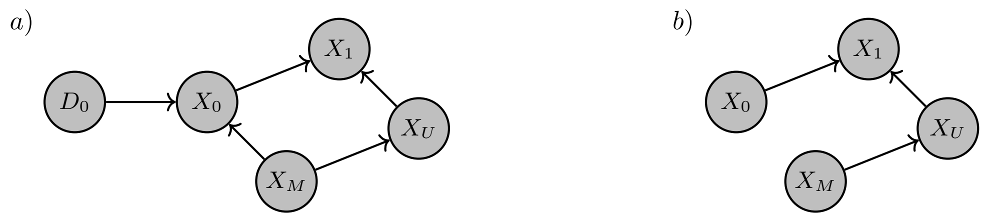

Figure 1: Causal graph for the Backdoor Criterion. a) The mediator variable X_M blocks any backdoor path from X_0 to X_1, potentially through a confounding variable X_U. b) The effect of the intervention on X_0 reduces the Bayesian network effectively to the shown subgraph.

- *Confounding variable X_C summarizing other factors affecting the outcome.*
- *Treatment assignment X_T with values in $\{0,1\}$. Here, 1 indicates assignment to the treatment group and 0 to the control group. The experiment is conducted such that the treatment variable is independent of the confounding variable X_C.*
- *Outcome X_O .*

We want to estimate the causal effect of the treatment on the outcome, namely calculate the intervention query

$$\mathbb{P}\left[X_O \middle| D_T = 1\right].$$

By design of the randomized control, the confounding variable has no causal effect on X_T. Thus, we have

$$\mathbb{P}\left[X_O \middle| D_T = 1\right] = \mathbb{P}\left[X_O \middle| X_T = 1\right]$$

that is, the intervention query is equal to the conditional probability.

2.3.1 Backdoor Criterion

More general than in the randomized controlled trial we have to deal with confounding variables. The Backdoor Criterion and the Frontdoor Criterion exploit mediator variables X_M, which are also observable and block certain causal paths to unobserved confounding variables X_U.

Lemma 2 (Backdoor Criterion)**.** *When X_M blocks any backdoor path from X_0 to X_1 and no node in X_M is a descendant of X_0, then*

$$\mathbb{P}^c\left[X_1 \middle| D_0 = x_0\right] = \left\langle \mathbb{P}\left[X_1 \middle| X_0 = x_0, X_M\right], \mathbb{P}\left[X_M\right] \right\rangle_{[X_1]}.$$

Proof. By assumption, we can contract all variables beside $X_{[2]}, X_M$ into cpds, and have a simplified causal model (see Figure 1a). We have

$$\mathbb{P}^c\left[X_0 \middle| D_0 = x_0, X_M\right] = \epsilon_{x_0}\left[X_0\right] \otimes \mathbb{I}\left[X_M\right]$$

and thus

$$\mathbb{P}^c\left[X_1 \middle| D_0 = x_0\right] = \left\langle \mathbb{P}\left[X_1 \middle| X_0 = x_0, X_M\right], \mathbb{P}\left[X_M\right] \right\rangle_{[X_1]}.$$

□

Lemma 3 (Backdoor Criterion)**.** *When X_M blocks any front door path from X_0 to X_1 and there is no confounder between X_0 and X_M, as well as between X_M and X_1, then*

$$\mathbb{P}^c\left[X_1 \middle| D_0 = x_0\right] = \left\langle \mathbb{P}\left[X_M \middle| X_0 = x_0\right], \left\langle \mathbb{P}\left[X_1 \middle| X_M, X_0\right], \mathbb{P}\left[X_0\right] \right\rangle_{[X_1, X_M]} \right\rangle_{[X_1]}.$$

Note that the inner contraction has X_0 as a closed variable.

Proof. We show the claim for the situation in Figure 2

$$\begin{aligned}\mathbb{P}^c\left[X_1 \middle| D_0 = x_0\right] &= \left\langle \mathbb{P}\left[X_M \middle| X_0 = x_0\right], \mathbb{P}\left[X_1 \middle| X_M, X_U\right], \mathbb{P}\left[X_U\right] \right\rangle_{[X_1]} \\ &= \left\langle \mathbb{P}\left[X_M \middle| X_0 = x_0\right], \left\langle \mathbb{P}\left[X_1 \middle| X_M, X_U\right], \mathbb{P}\left[X_U\right] \right\rangle_{[X_1, X_M]} \right\rangle_{[X_1]}.\end{aligned}$$

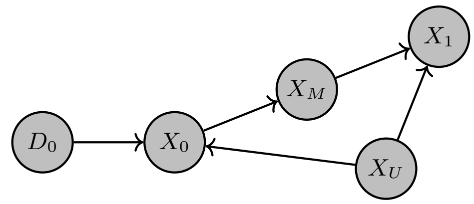

Figure 2: Causal graph for the Frontdoor Criterion.

We note that

$$\begin{aligned}\left\langle \mathbb{P}\left[X_1\middle|X_M, X_U\right], \mathbb{P}\left[X_U\right]\right\rangle_{[X_1, X_M = x_M]} &= \mathbb{P}^c\left[X_1\middle|D_M = x_M\right] \\ &= \left\langle \mathbb{P}\left[X_1\middle|X_M, X_0\right], \mathbb{P}\left[X_0\right]\right\rangle_{[X_1, X_M = x_M]}\end{aligned}$$

Here we used in the second equation that the variable X_0 blocks all backdoor paths from X_M to X_1 (i.e. the backdoor criterion). Combining both equations proves the claim. □

3 Rung 3: Counterfactuals

Following the formalism of Chapter 21 in Koller and Friedman (2009), counterfactual queries are modeled as intervention queries on counterfactual twin networks. Counterfactual twin networks are two copies of the original causal model, one representing the factual world and one representing the counterfactual world. They further share response variables, which capture the functional relationships between variables in the causal model. These response variables are modeled as selection variables in the $\mathrm{tnreason}$ formalism.

Example 3 (Patient counterfactual query)**.** *Let us assume, that we know whether a patient did not get assigned to the treatment group, and we know whether he did improve. We want to ask, whether he would have improved, if he had been assigned to the treatment group. This is not an intervention query, but a counterfactual query, since we reason about a different outcome, although we already know that the patient has not been treated.*

The causally augmented Bayesian Network is not rich enough to answer such queries. Let us consider a situation where X_T is chosen uniformly from $[2]$. There might be situations, where the probabilistic (and intervention) queries are the same, but the counterfactual queries differ:

(i) The outcome is independent of the treatment.

(ii) With probability 0.5, the treatment always leads to improvement and missing the treatment not. And also with probability 0.5, the treatment never leads to improvement, but missing the treatment does.

Intuitively, in the situation (i), the counterfactual query should return probability $1/2$ for improvement under treatment. In the situation (ii), the counterfactual query returns the same value with probability 1 as observed in the real world. Both cases are however indistinguishable by causal models itself, since in both cases the causally augmented conditional probabilities are

$$\mathbb{P}^c\left[X_T\middle|D_T\right] = \frac{1}{2}\mathbb{I}\left[X_T\right] \otimes \epsilon_2\left[D_T\right] + \left(\sum_{d_T \in [2]} \epsilon_{d_T}\left[D_T\right] \otimes \epsilon_{d_T}\left[X_T\right]\right)$$

and

$$\mathbb{P}^c\left[X_O\middle|D_O, X_T\right] = \frac{1}{2}\mathbb{I}\left[X_O, X_T\right] \otimes \epsilon_2\left[D_O\right] + \left(\sum_{d_O \in [2]} \epsilon_{d_O}\left[D_O\right] \otimes \epsilon_{d_O}\left[X_O\right] \otimes \mathbb{I}\left[X_T\right]\right).$$

To answer counterfactual queries, we therefore have to also reason about the mechanisms behind the causal relationships, which we do by introducing response variables in the following.

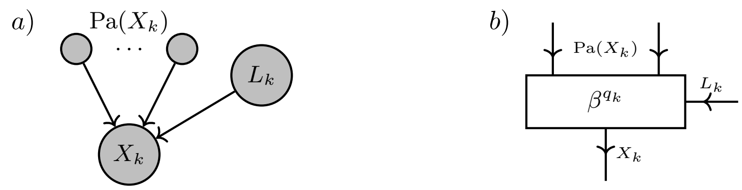

Figure 3: Introduction of response variables: a) as a Bayesian network, b) decoration by the basis encoding of the corresponding function.

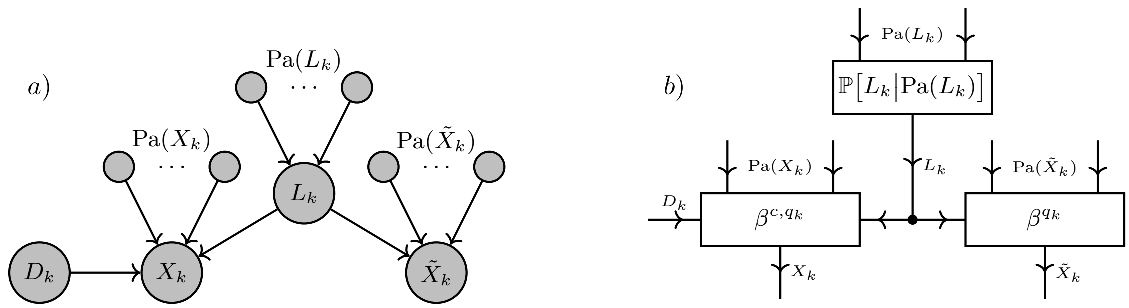

Figure 4: Counterfactual twinned network: a) as a Bayesian network. b) Decorations by causal augmented basis encodings by a do-variable for X_k, a conditional probability distribution for L_k and a basis encoding for the counterfactual variable $\tilde{X}_k$.

3.1 Response Variables

We introduce response variables L_k, such that X_k has a deterministic dependence on L_k and the parents of a variable to the variable itself. The dependence is captured by a function

$$q_k \,:\, [p_k] \times \left(\bigtimes_{\tilde{k}\in\mathrm{Pa}(k)} [m_{\tilde{k}}] \right) \to [m_k] \,.$$

The response augmented conditional probability core is defined as (see Figure 3):

$$\mathbb{P}^f\left[X_k \middle| L_k, X_{\mathrm{Pa}(k)}\right] = \beta^{q_k}\left[X_k, L_k, X_{\mathrm{Pa}(k)}\right]$$

3.2 Counterfactual Twinned Network

To answer counterfactual queries, we now construct the counterfactual twinned network (see Figure 4).

Definition 2. *Let there be a response augmented Bayesian Network on a directed acyclic hypergraph $\mathcal{G} = ([d], \mathcal{E})$ with response variables $L_{[d]}$. Further another hypergraph $\tilde{\mathcal{G}}$ and a Bayesian Network on $\tilde{\mathcal{G}}$ modelling the joint distribution of the response variables $L_{[d]}$. Be build counterfactual copies $\tilde{X}_k$ of each variable X_k, which are controlled by the same response variables L_k as the factual variables X_k. The counterfactual twinned network is then the tensor network consists of the causal and response augmented conditional probability cores to $X_{[d]}$, the conditional probability cores of the response variables $L_{[d]}$ and the response augmented conditional probability cores to the counterfactual variables $\tilde{X}_{[d]}$.*

4 Learning

Given a parametrization of hypothesis distribution by the states of selection variables, we can include the intervention variables. The cores to be trivialized by intervention can depend on the state of the selection variables. For example, a selection variable might select whether the cores direction is from X_0 to X_1 or in the other direction. In both cases only one of the to intervention variables D_0 and D_1 influences the core.

References

Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, Cambridge, Mass., 1. edition edition, July 2009. ISBN 978-0-262-01319-2.

Judea Pearl. *Causality: Models, Reasoning and Inference. Ausgezeichnet: ACM Turing Award for Transforming Artificial Intelligence 2011*. Cambridge University Press, Cambridge New York, NY Port Melbourne New Delhi Singapore, 2 edition, November 2009. ISBN 978-0-521-89560-6.

Judea Pearl and Dana Mackenzie. *The Book of Why: The New Science of Cause and Effect*. Basic Books, New York, May 2018. ISBN 978-0-465-09760-9.