A Short Primer on Causals Models

Mathias Winther Madsen

November 3, 2016

Abstract

This brief note introduces the concepts relavant to causal modeling, assuming familiarity with graphical models. Its main purpose is to list the mathematical objects used in causal models, and to pair these objects up with the real-world phenomena they are supposed to represent.

In a conventional probabilistic model, a joint distribution over n variables can be factorized into a product of n conditional distributions in $n!$ different ways. For instance, a joint distribution $P(X, Y)$ can be expanded into either

$$P(X)\,P(Y \,|\, X) \qquad \text{or} \qquad P(Y)\,P(X \,|\, Y).$$

From a probabilistic perspective, there is no difference between these two factorizations. One might be more convenient than the other, but they define the same joint distribution.

This is not the case in the world of causal models. A **causal model** is a joint probability distribution with a **canonical factorization**. For instance, a joint distribution $P(X, Y, Z)$ can turned into a causal model by canonizing one of its $3! = 6$ factorizations, such as

$$P(X)\,P(Y \,|\, X)\,P(Z \,|\, X, Y).$$

The canonical factorization is special because it expresses an assumption about how the system will respond to interventions. The other $n! - 1$ factorizations are still valid tools for observational probabilistic inference, but they do not reflect the causal structure of the system.

On the purely formal level, an **intervention** is an operation that derives a probabilistic model from another. It does so by replacing one of the factors in a canonical factorization by a given marginal distribution. For instance, the factorized probabilistic model

$$P(X)\,P(Y \,|\, X)\,P(Z \,|\, X, Y)$$

produces the probabilistic model

$$P(X)\,Q(Y)\,P(Z \,|\, X, Y),$$

under the intervention $Y \sim Q$.

This concept of intervention is purely formal operation, but it is intended to model the effects of a real-world manipulation of a system. We interpret the imposition of a certain marginal distribution Q as the equivalent of physically imposing a specific behavior on the system.

The order of factorization can change the effect of an intervention: a single intervention on a single joint distribution can produce two different distributions under two different canonical factorizations. For instance, depending on which of the factorizations

$$P(X)\,P(Y\,|\,X) \qquad \text{or} \qquad P(Y)\,P(X\,|\,Y)$$

we assume, the intervention $X \sim Q$ will produce either of two distributions:

$$Q(X)\,P(Y\,|\,X) \qquad \text{or} \qquad P(Y)\,Q(X).$$

We can thus use different canonical factorizations to represent different hypotheses about the effect of interventions. Each canonical factorization of a joint distribution defines a behavior for the system under any possible intervention.

These hypothetical behaviors can be compared to the behavior of the actual system under real-world intervention. We can thus select good factorizations and reject bad ones by performing real-world interventions. When arbitrary interventions are a practical option, this reduces causal inference to a matter of hypothesis testing.

Thus, while a probabilistic model describes a system in one specific state, a causal model thus describes a system in a variety of states, one for each intervention. Table 1 summarizes how these relationships inside of the model correspond to relationships between the putative real-world system.

Descriptive	Mathematical
System	Joint distribution
Causal structure	Factorization
Intervention	New marginal distribution

Table 1: Causal concepts and their model counterparts.