

equal. Finally, Carol’s finishing time depends on both her teammates. If Alice is slow, Bob will probably finish late too. As a consequence, Carol will have quite a late starting time and thus is likely to have a late finishing time as well. However, Carol’s finishing time depends only *indirectly* on Alice’s finishing time via Bob’s. If we already know Bob’s finishing time, we will not be able to estimate Carol’s finishing time better by finding out what Alice’s finishing time was. This means we can model the relay race using only two interactions: Alice’s effect on Bob and Bob’s effect on Carol. We can omit the third, indirect interaction between Alice and Carol from our model.

Structured probabilistic models provide a formal framework for modeling only direct interactions between random variables. This allows the models to have significantly fewer parameters and therefore be estimated reliably from less data. These smaller models also have dramatically reduced computational cost in terms of storing the model, performing inference in the model, and drawing samples from the model.

16.2 Using Graphs to Describe Model Structure

Structured probabilistic models use graphs (in the graph theory sense of “nodes” or “vertices” connected by edges) to represent interactions between random variables. Each node represents a random variable. Each edge represents a direct interaction. These direct interactions imply other, indirect interactions, but only the direct interactions need to be explicitly modeled.

There is more than one way to describe the interactions in a probability distribution using a graph. In the following sections we describe some of the most popular and useful approaches. Graphical models can be largely divided into two categories: models based on directed acyclic graphs, and models based on undirected graphs.

16.2.1 Directed Models

One kind of structured probabilistic model is the **directed graphical model**, otherwise known as the **belief network** or **Bayesian network**² (Pearl, 1985).

Directed graphical models are called “directed” because their edges are directed,

² Judea Pearl suggested using the term “Bayesian network” when one wishes to “emphasize the judgmental” nature of the values computed by the network, i.e. to highlight that they usually represent degrees of belief rather than frequencies of events.