



Figure 3.8: Highlighted nodes belong to the Markov blanket of node a in directed graph (a) and an undirected graph (b).

in a cutset, the nodes in one connected component are independent of those in another component. In Figure 2.8, the graph is partitioned into two cutsets of nodes that partitions the graph into two independent components. Markov Blankets (Definition 2.1) can be used to show that the nodes in a cutset are independent of the nodes in another cutset.

Finally, we can revisit Markov blankets highlighted in Figure 3.8, the Markov blanket for a node can differ from its Markov blanket in the core skeleton.

Definition 3.8 (Markov blankets in undirected G) Let $G = (V, E)$ be a graph and node of interest, $X \in V$. The Markov blanket for X in an undirected graph, the Markov blanket for X , is the set of nodes $Y \in V$ such that Y is a neighbor of X , i.e., $Y \in \text{ne}(X)$.

Thus, X is independent from the remaining nodes in the Markov blanket is observed. The utility of X is clear when we explore approximate inference

Comparison to Bayesian Networks

As with Bayesian networks, MRFs are poor representation. In the directed case, we find sparse distributions p whose independent components are the undirected case,

So, what independencies cannot be described by \mathcal{G} ? In the undirected graph, a probability distribution described by a v -structure (Figure 3.9). Neither

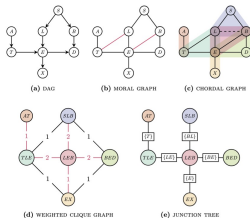
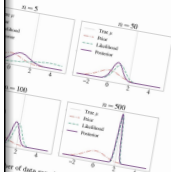


Figure 4.5: Transforming the ASIA DAG (Lauritzen and Spiegelhalter, 1988) into a junction tree, per Algorithm 2. We begin with the original DAG (a) and obtain the undirected moral graph (b). We then chordalize the moral graph to identify maximal cliques (c, chord denoted by a dashed edge; cliques in highlights). We transform the chordal graph into a weighted clique graph, where each node represents a maximal elimination clique and edge weights are assigned according to the cardinality of the separator (d). From this, we can obtain the maximum weight spanning tree (pink edges). Finally, we have our junction tree (e).

Proposition 4.1. Any chordal graph has a corresponding junction tree. Furthermore, any graph with a junction tree must be chordal.

Proposition 4.2. Any chordal graph with n nodes has at most n maximal cliques. Further, the chordal graph with n nodes and n maximal cliques is the graph with no edges, and a connected chordal graph has at most $n - 1$ maximal cliques.

For proof of Propositions 4.1 and 4.2, see Chapter 3 of Vandenberghe and Andersen (2015). Following from these observations, we can see



er of data samples n increases, the influence of the prior decreases. Meanwhile, the influence of the likelihood comes to our data are drawn from a Gaussian distribution with $\mu = 2$.

about θ , we can choose a weak prior or uniform prior. If we are very confident in our domain expertise, we can choose a more informative prior. Further, the prior $p(\theta)$ is a function of the sample size n . As n is small, the posterior is heavily influenced by the prior and is more concentrated around the true value of θ . As n increases, the influence of the prior on the posterior decreases. In other words, as the likelihood function incorporates more data, the influence of the prior on the posterior decreases.

distribution using Bayes' rule, it should calculate the numerator. However, calcu-