applications is that various types of restrictions might be made on some of the parameters, for example, some parameters might be required to be equal to (functions of) other parameters in the model.

Example 13.4.1 (Markov chain)**.** Let $X_1, X_2, \ldots, X_m$ be a sequence of discrete random variables, each with state space $[r]$. Consider the directed path graph P_m, with edges $i \to i+1$ for $i \in [m-1]$. According to the recursive factorization theorem, the joint probability distribution for the Bayesian network associated to this DAG has the form

$$P(X_1 = x_1, \ldots, X_m = x_m) = P_1(X_1 = x_1) \prod_{i=2}^{m} P_i(X_i = x_i | X_{i-1} = x_{i-1}).$$

Note that the conditional independence structure in this model will have the form $X_{i+1} \perp\!\!\!\perp X_{[i-1]} | X_i$ for $i = 2, \ldots, m-1$, so that the sequence of random variables forms a Markov chain, as discussed in Chapter 1.

If, in addition, we require that each of the conditional distributions P_i are equal,

$$P_i(X_i = x | X_{i-1} = y) = P_2(X_2 = x | X_1 = y) \text{ for all } i, x, y,$$

then we have a *homogeneous Markov chain*, which is the way that discrete Markov chains are often discussed, for example in [**Nor98**]. Letting $A = (a_{i_1 i_2})_{i_1, i_2 \in [r]}$ denote the matrix representing the conditional distribution, that is, $a_{i_1 i_2} = P(X_{t+1} = i_2 | X_t = i_1)$, then A is the transition matrix of the homogeneous Markov chain. Thus, the example of a Markov chain from Chapter 1 is an example of a Bayesian network.

As we saw in Example 13.4.1, models that can be expressed by a sequence of simple Markov processes connecting random variables can usually be thought of as graphical models on a directed acyclic graph, with the underly DAG being a directed tree. Another important example is the hidden Markov model.

Example 13.4.2 (Hidden Markov model)**.** Let $Y_1, \ldots, Y_m$ be a sequence of discrete random variables each with state space $[r]$, and let $X_1, \ldots, X_m$ be a sequence of discrete random variables, each with state space $[s]$. Consider the directed graph with edges $Y_i \to Y_{i+1}$ for $i \in [m-1]$ and $Y_i \to X_i$ for $i \in [m]$. This DAG is illustrated in Figure 13.4.1, a type of tree called a caterpillar.

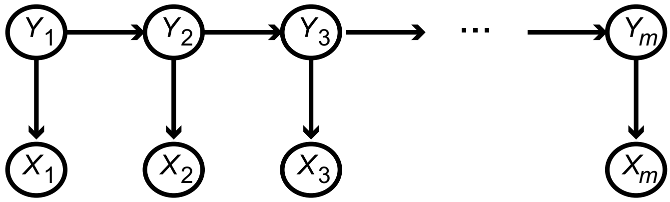

Figure 13.4.1. The caterpillar tree, the underlying graph of a hidden Markov model.

The recursive factorization property for this graph yields the joint distribution

$$P(X_1 = x_1, \ldots, X_m = x_m, Y_1 = y_1, \ldots, Y_m = y_m)$$
$$= P_1(Y_1 = Y_1) \prod_{i=2}^{m} P_i(Y_i = y_i | Y_{i-1} = y_{i-1}) \prod_{i=1}^{m} Q_i(X_i = x_i | Y_i = y_i).$$

In this model, we typically require that each of the conditional distributions of Y_i given Y_{i-1} are the same, and that each of the conditional distributions of X_i given Y_i are the same. Furthermore, we assume that the random variables $Y_1, \ldots, Y_m$ are all hidden random variables, that is, they are unobserved (we discuss this point in more detail in later chapters).

In applications, one should think of the hidden Markov model in the following way. The sequence of random variables $Y_1, \ldots, Y_m$ evolves according to a homogeneous Markov chain model. However, we are not able to observe Y, we only observe a corrupted or noisy version of it or some information X computed from Y in a possibly random way. Usually we want to recover Y from X. The hidden Markov model is commonly used in computational biology where it is used to align DNA sequences and annotate DNA for genes (see, e.g., [**DEKM98**]). We will discuss it in more detail in Chapter 18. □

The Ising model is an important interpretation of undirected graphical models. Although we usually associate the Ising model with statistical physics, this model and variations on it are also used in spatial statistical and agriculture.

Example 13.4.3 (Ising model)**.** Let $G = (V, E)$ be an undirected graph, and let X be a discrete random vector where all random variables are binary with state space $\{-1, 1\}$. Introduce a probability distribution

$$P(x_1, \ldots, x_n) = \frac{1}{Z} \exp\left(\sum_{(i,j) \in E} \alpha_{ij} x_i x_j + \sum_{i \in V} \beta_i x_i \right).$$