```
function GIBBS-ASK(X, e, bn, N) returns an estimate of P(X | e)
  local variables: C, a vector of counts for each value of X, initially zero
                   Z, the nonevidence variables in bn
                   x, the current state of the network, initialized from e

  initialize x with random values for the variables in Z
  for k = 1 to N do
     choose any variable Z_i from Z according to any distribution ρ(i)
     set the value of Z_i in x by sampling from P(Z_i | mb(Z_i))
     C[j] ← C[j] + 1 where x_j is the value of X in x
  return NORMALIZE(C)
```

Figure 13.20 The Gibbs sampling algorithm for approximate inference in Bayes nets; this version chooses variables at random, but cycling through the variables but also works.

2. *Rain* is chosen and then sampled, given the current values of its Markov blanket: in this case, we sample from $\mathbf{P}(Rain \mid Cloudy = false, Sprinkler = true, WetGrass = true)$. Suppose this yields $Rain = true$. The new current state is $[false, \mathbf{true}, true, \mathbf{true}]$.

The one remaining detail concerns the method of calculating the Markov blanket distribution $\mathbf{P}(X_i \mid mb(X_i))$, where $mb(X_i)$ denotes the values of the variables in X_i's Markov blanket, $MB(X_i)$. Fortunately, this does not involve any complex inference. As shown in Exercise 13.MARB, the distribution is given by

$$P(x_i \mid mb(X_i)) = \alpha P(x_i \mid parents(X_i)) \prod_{Y_j \in Children(X_i)} P(y_j \mid parents(Y_j))\,. \qquad (13.10)$$

In other words, for each value x_i, the probability is given by multiplying probabilities from the CPTs of X_i and its children. For example, in the first sampling step shown above, we sampled from $\mathbf{P}(Cloudy \mid Sprinkler = true, Rain = false)$. By Equation (13.10), and abbreviating the variable names, we have

$$\begin{aligned} P(c \mid s, \neg r) &= \alpha P(c)P(s \mid c)P(\neg r \mid c) = \alpha\, 0.5 \cdot 0.1 \cdot 0.2 \\ P(\neg c \mid s, \neg r) &= \alpha P(\neg c)P(s \mid \neg c)P(\neg r \mid \neg c) = \alpha\, 0.5 \cdot 0.5 \cdot 0.8\,, \end{aligned}$$

so the sampling distribution is $\alpha\langle 0.001, 0.020\rangle \approx \langle 0.048, 0.952\rangle$.

Figure 13.21(a) shows the complete Markov chain for the case where variables are chosen uniformly, i.e., $\rho(Cloudy) = \rho(Rain) = 0.5$. The algorithm is simply wandering around in this graph, following links with the stated probabilities. Each state visited during this process is a sample that contributes to the estimate for the query variable *Rain*. If the process visits 20 states where *Rain* is true and 60 states where *Rain* is false, then the answer to the query is NORMALIZE$(\langle 20, 60\rangle) = \langle 0.25, 0.75\rangle$.

Analysis of Markov chains

We have said that Gibbs sampling works by wandering randomly around the state space to generate samples. To explain why Gibbs sampling works *correctly*—that is, why its estimates converge to correct values in the limit—we will need some careful analysis. (This section is somewhat mathematical and can be skipped on first reading.)