7.2 Inference in HMMs

Consider a discrete HMM with hidden states S_t, observations O_t, transition matrix $a_{ij} = P(S_{t+1} = j | S_t = i)$ and emission matrix $b_{ik} = P(O_t = k | S_t = i)$. In class, we defined the forward-backward probabilities:

$$\begin{aligned} \alpha_{it} &= P(o_1, o_2, \ldots, o_t, S_t = i), \\ \beta_{it} &= P(o_{t+1}, o_{t+2}, \ldots, o_T | S_t = i), \end{aligned}$$

for a particular observation sequence $\{o_1, o_2, \ldots, o_T\}$ of length T. This problem will increase your familiarity with these definitions. In terms of these probabilities, which you may assume to be given, as well as the transition and emission matrices of the HMM, show how to (efficiently) compute the following probabilities:

(a) $P(S_{t+1} = j | S_t = i, o_1, o_2, \ldots, o_T)$

(b) $P(S_t = i | S_{t+1} = j, o_1, o_2, \ldots, o_T)$

(c) $P(S_{t-1} = i, S_t = k, S_{t+1} = j | o_1, o_2, \ldots, o_T)$

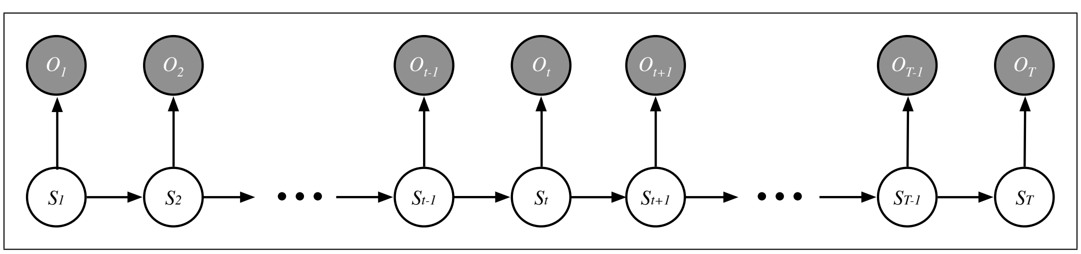
(d) $P(S_{t+1} = j | S_{t-1} = i, o_1, o_2, \ldots, o_T)$

In all these problems, you may assume that $t > 1$ and $t < T$; in particular, you are *not* asked to consider the boundary cases.

7.3 Conditional independence

Consider the hidden Markov model (HMM) shown below, with hidden states S_t and observations O_t for times $t \in \{1, 2, \ldots, T\}$. State whether the following statements of conditional independence are true or false.

$$\begin{aligned}
________ \quad P(S_t|S_{t-1}) &= P(S_t|S_{t-1}, S_{t+1}) \\
________ \quad P(S_t|S_{t-1}) &= P(S_t|S_{t-1}, O_{t-1}) \\
________ \quad P(S_t|S_{t-1}) &= P(S_t|S_{t-1}, O_t) \\
________ \quad P(S_t|O_{t-1}) &= P(S_t|O_1, O_2, \ldots, O_{t-1}) \\
________ \quad P(O_t|S_{t-1}) &= P(O_t|S_{t-1}, O_{t-1}) \\
________ \quad P(O_t|O_{t-1}) &= P(O_t|O_1, O_2, \ldots, O_{t-1}) \\
________ \quad P(S_2, S_3, \ldots, S_T|S_1) &= \textstyle\prod_{t=2}^{T} P(S_t|S_{t-1}) \\
________ \quad P(S_1, S_2, \ldots, S_{T-1}|S_T) &= \textstyle\prod_{t=1}^{T-1} P(S_t|S_{t+1}) \\
________ \quad P(S_1, S_2, \ldots, S_T|O_1, O_2, \ldots, O_T) &= \textstyle\prod_{t=1}^{T} P(S_t|O_t) \\
________ \quad P(S_1, S_2, \ldots, S_T, O_1, O_2, \ldots, O_T) &= \textstyle\prod_{t=1}^{T} P(S_t, O_t) \\
________ \quad P(O_1, O_2, \ldots, O_T|S_1, S_2, \ldots, S_T) &= \textstyle\prod_{t=1}^{T} P(O_t|S_t) \\
________ \quad P(O_1, O_2, \ldots, O_T) &= \textstyle\prod_{t=1}^{T} P(O_t|O_1, \ldots, O_{t-1})
\end{aligned}$$



7.4 Belief updating

In this problem, you will derive recursion relations for real-time updating of beliefs based on incoming evidence. These relations are useful for situated agents that must monitor their environments in real-time.

(a) Consider the discrete hidden Markov model (HMM) with hidden states S_t, observations O_t, transition matrix a_{ij} and emission matrix b_{ik}. Let

$$q_{it} \;=\; P(S_t\!=\!i|o_1, o_2, \ldots, o_t)$$

denote the conditional probability that S_t is in the i^{th} state of the HMM based on the evidence up to and including time t. Derive the recursion relation:

$$q_{jt} = \frac{1}{Z_t} b_j(o_t) \sum_i a_{ij} q_{it-1} \quad \text{where} \quad Z_t = \sum_{ij} b_j(o_t) a_{ij} q_{it-1}.$$

Justify each step in your derivation—for example, by appealing to Bayes rule or properties of conditional independence.

(b) Consider the dynamical system with *continuous, real-valued* hidden states X_t and observations Y_t, represented by the belief network shown below. By analogy to the previous problem (replacing sums by integrals), derive the recursion relation:

$$P(x_t|y_1, y_2, \ldots, y_t) \;=\; \frac{1}{Z_t} P(y_t|x_t) \int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1, y_2, \ldots, y_{t-1}),$$

where Z_t is the appropriate normalization factor,

$$Z_t = \int dx_t\, P(y_t|x_t) \int dx_{t-1}\, P(x_t|x_{t-1})\, P(x_{t-1}|y_1, y_2, \ldots, y_{t-1}).$$

In principle, an agent could use this recursion for real-time updating of beliefs in arbitrarily complicated continuous worlds. In practice, why is this difficult for all but Gaussian random variables?