

## 0.1 Parametrization of the transition matrix

The square transition matrix  $\mathbf{A}$  has  $S = 2^{2V}$  entries and therefore  $S(S - 1) = 2^{2V} - 2^V$  degrees of freedom. Although searching the full space of viable transition matrices is possible via unparametrized sampling techniques, it is computationally challenging and hard to interpret. To achieve this reduction in degrees of freedom, and also preserve the anatomically and medically motivated structure of the Bayesian network from ??, we can represent the transition probability from one state  $\mathbf{x}[t]$  to another state  $\mathbf{x}[t + 1]$  using the conditional probabilities defined for the Bayesian network (BN). The difference is that the probability of observing a certain state of lymph node level (LNL)  $v$  now depends on the state of the patient one time-step before. Note that from here on, we will mostly drop the probabilistically correct notation  $P(X = x)$  and just write  $P(x)$  for brevity

$$P_{HMM}(\mathbf{x}[t + 1] | \mathbf{x}[t]) = \prod_{v \leq V} Q(x_v[t + 1]; x_v[t]) \times \left[ P_{BN}(x_v[t + 1] | \{x_r[t], \tilde{t}_{rv}\}_{r \in \text{pa}(v)}, \tilde{b}_v) \right]^{1-x_v[t]} \quad (1)$$

Here we have reused the conditional probability from the BN for each LNL, but we take it to the power of one minus that node's previous value. This ensures that an involved node stays involved with probability 1. The parameters  $\tilde{t}_{\text{pa}(v)v}$  and  $\tilde{b}_v$  take the same role as in the BN, but they are now probability *rates*, since they act per time-step. Lastly, the first term  $Q$  in the product formalizes the fact that a metastatic lymph node level cannot become healthy again once it was involved. This also means that several entries in the transition matrix  $\mathbf{A}$  must be zero. In a table the values of  $Q(x_v[t + 1]; x_v[t])$  can be written like this:

$$\begin{aligned} Q(X_v[t + 1] = 0; X_v[t] = 0) &= 1 \\ Q(X_v[t + 1] = 0; X_v[t] = 1) &= 0 \\ Q(X_v[t + 1] = 1; X_v[t] = 0) &= 1 \\ Q(X_v[t + 1] = 1; X_v[t] = 1) &= 1 \end{aligned} \quad (2)$$

which gives rise to a "mask" for  $\mathbf{A}$  which can be seen in ??.

To illustrate eq. (1), it helps to look at a specific example. E.g., the transition probability from state  $\xi_5 = (0 \ 1 \ 0 \ 0)$  to state  $\xi_7 = (0 \ 1 \ 1 \ 0)$ , which represents starting with involvement only in LNL II and asking for the probability that LNL III becomes involved as well over the next time-step:

$$\begin{aligned} P_{HMM}(\mathbf{X}[t + 1] = \xi_7 | \mathbf{X}[t] = \xi_5) \\ = & Q(X_1[t + 1] = 0; X_1[t] = 0) P_{BN}(X_1[t + 1] = 0 | \tilde{b}_1)^1 \\ & \times Q(X_2[t + 1] = 1; X_2[t] = 1) P_{BN}(X_2[t + 1] = 1 | X_1[t] = 0, \tilde{t}_{12}, \tilde{b}_2)^0 \\ & \times Q(X_3[t + 1] = 1; X_3[t] = 0) P_{BN}(X_3[t + 1] = 1 | X_2[t] = 1, \tilde{t}_{23}, \tilde{b}_3)^1 \\ & \times Q(X_4[t + 1] = 0; X_4[t] = 0) P_{BN}(X_4[t + 1] = 0 | X_3[t] = 0, \tilde{t}_{34}, \tilde{b}_4)^1 \\ = & (1 - \tilde{b}_1) \cdot 1 \cdot (\tilde{b}_3 + \tilde{t}_{23} - \tilde{b}_3 \tilde{t}_{23}) \cdot (1 - \tilde{b}_4) \end{aligned} \quad (3)$$

The interpretation of the last line is that this is the probability that LNL I and IV do not become involved, while LNL III gets infected through lymphatic drainage from either the main tumor or LNL II. The probability of LNL II remaining involved is 1, of course, which is why we take the respective term to the power of 0.