



Figure 1: A length-4 example of our generative model fac-torized as in Eq. (1) and overlaid with example values of the random variables in the sequence. We highlight that all the conditionals in the Bayesian network are recurrent neural networks, e.g., we note that  $m_i$  depends on  $m_{<i}$  because we employ a recurrent neural network to model the morphological tag sequence.