Figure 1: A length-4 example of our generative model factorized as in Eq. (1) and overlayed 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 $\boldsymbol{m}_{<i}$ because we employ a recurrent neural network to model the morphological tag sequence.