Exercises for the lecture:
Bayesian networks: inference and prediction.
D-separation.

December 3, 2015

1 Hidden Markov models

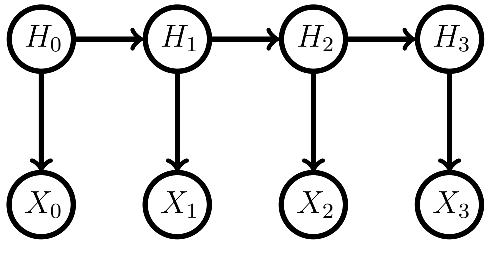
The graph below depicts the Bayesian network representation of a *Hidden Markov model* (HMM). HMMs are popular tools for time series modelling, e.g. in speech recognition or for the generation of trajectories for robotic limbs. The X_i represent a time-series of observations (e.g. wavelet coefficients in speech modelling), whereas the H_i are a sequence of hidden variables which describe the dependency structure of the X_i across time. In real-world applications, the X_i are often continuous random variables, and each H_i can typically take on a large number of possible values. For simplicity, we assume here that the H_i and the X_i are binary. Furthermore, let

$$P(X_i|H_i) = 0.8 \text{ if and only if } X_i = H_i$$

$$P(H_{i+1}|H_i) = 0.7 \text{ if and only if } H_i = H_{i+1}$$

$$P(H_0 = T) = 0.5$$

$P(X_i|H_i)$ is often called the *emission model*, and $P(H_{i+1}|H_i)$ is the *transition matrix*.



E1. (**1p**) draw the factor graph corresponding to this Bayesian network. Explain why the sum-product algorithm can be used to compute marginals in this network.

Use the sum-product algorithm, where applicable, to solve the exercises below. Compute numerical probabilities to 3 significant digits in exponential notation, e.g. write $1.73 \cdot 10^{-4}$ for $\frac{1}{5772}$.

E2. (**4p**) what are the marginal distributions $P(X_0), \ldots, P(X_3)$?

Assume you observe the sequence $S = (X_0 = T, X_1 = F, X_2 = F, X_3 = T)$.

E3. (**3p**) **optional**: what is the probability distribution $P(H_3|S)$? This is an example of *inference* of the hidden state.

E4. (**3p**) **optional**: what is the probability distribution $P(H_0|S)$?

E5. (**3p**) **optional**: what is the most probable hidden state sequence H_0, H_1, H_2, H_3 given these observations ? Answering this question is typically important when the hidden states are used as input into classification systems, e.g. to identify a spoken word.

Assume you observe the sequence $(X_0 = T, X_1 = F, X_2 = F)$.

E6. (**3p**) what is the probability distribution of X_3 given this sequence? This is an example of *time-series prediction*.

E7. (**2p**) assume all variables were unobserved. Are X_0 and X_3 independent? Assume you knew $H_2 = F$. Are X_0 and X_3 independent given this knowledge?