T-61.5140 Machine Learning: Advanced Probabilistic Methods
Hollmén, Raiko
Examination, 15th of May, 2008 from 9 to 12 o'clock.

In order to pass the course and earn 5 ECTS credit points, you must also pass the term project. Results of this examination are valid for one year after the examination date. Information for Finnish speakers: Voit vastata kysymyksiin myös suomeksi, kysymykset on ainoastaan englannin kielellä. Information for Swedish speakers: Du får också svara på svenska, frågorna finns dock endast på engelska.

1. Define the following terms shortly:
a) conditional independence
b) treewidth
c) d-separation
d) Markov Blanket
e) complete-data likelihood
f) proposal distribution

2. Given a Hidden Markov Model (HMM) for a sequence of observations $Y = (y_1, \ldots, y_t)$, show that the predictive distribution of the observations y_t follows a mixture distribution.

3. Write the algorithm for Gibbs sampling and write the distributions to sample from in the case of $p(x_1, x_2, x_3, x_4)$.

4. Write the probability $p(\mathbf{x})$ for the finite mixture model of multivariate Bernoulli distributions, name the parts of the mixture model, and derive the E-step and the M-step of the Expectation-Maximization (EM) algorithm.
Hint: The probability for a d-dimensional vector of 0-1 data can be calculated with the following equation: $p(\mathbf{x} \mid \theta) = \prod_{i=1}^{d} \theta_i^{x_i} (1 - \theta_i)^{1-x_i}$.

5. For the Bayesian network that decomposes the joint probability as in $p(x_1, \ldots, x_5) = p(x_1)p(x_2|x_1)p(x_3|x_1)p(x_4|x_2, x_3)p(x_5|x_4)$, draw the corresponding graphical representation. Assuming all the variables have discrete values $x_i \in \{0, 1, 2\}$, give the sizes of the tables representing the probabilities for the conditional distributions. Moreover, derive the junction tree representation (and name the steps). Draw the resulting junction tree.