T-61.5140 Machine Learning: Advanced Probabilistic Methods
Hollmén, Kallasjoki
Examination, 17th of May, 2011 from 13 to 16 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.

You should complete both questions 2a and 2b. You can choose one of the two alternatives in question 5, either the *(option 1)* or *(option 2)*, but not both.

1. Define the following terms shortly:
a) conditional independence
b) Bayesian network
c) treewidth
d) likelihood function
e) Hammersley-Clifford theorem
f) variational approximation

2a. Derive the Bayes's rule from the definition of conditional independence of discrete random variables X and C and name the parts of the Bayes's rule.

2b. What is the difference between rejection sampling and importance sampling? Write down the rejection rules and explain the difference.

3. 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: $p(\mathbf{x} \mid \theta) = \prod_{i=1}^{d} \theta_i^{x_i}(1-\theta_i)^{1-x_i}$.

4. 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 $x_i \in \{0, 1, 2, 3\}$, give the sizes of the tables representing the probabilities for the conditional probability distributions. Moreover, derive the junction tree representation (and name the steps). Draw the resulting junction tree.

Choose one of the following questions, either (*option 1*) or (*option 2*). The (*option 2*) is on the reverse side of this paper.

5. *(option 1)*: How can you generate samples from a Laplace distribution. Write a sketch of a program generating the samples. You have access to a computer that is able to generate samples from the uniform distribution. Hint: The probability density function for an Laplacian random variable is here: $p(x \mid \mu, b) = \frac{1}{2b}exp^{(-\frac{x-\mu}{b})}$.

5. *(option 2)*: In a TV game show, the contestant has the chance to win the price if he chooses a right door out of three doors behind which the price is waiting. First, the contestant is asked to select one of the doors. Then, the game show keeper opens another door which does not contain the prize. After that, the contestant is asked if he want to change his selection of the door (out of the two doors that remain unopened). Finally, the contestant wins the prize if it is found behind the selected door.

Model the domain with random variables *Price, First Selection, Game show keeper opens* with a Bayesian network and write the probability table(s) for the network according to the rules of the game. Calculate the probability of winning the prize if the contestant (a) does NOT change his original selection of the door, and (b) changes his original selection.