6.892 Machine learning seminar

http://www.ai.mit.edu/courses/6.892-ml

Prof. Tommi S. Jaakkola
tommi@ai.mit.edu

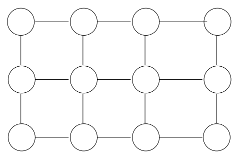
March 8, 1999

This assignment is due **March 29**. You can either email the assignment to me as plain text (use subject "HOMEWORK 6") or give it to me in class.

1. Very briefly, identify an actual problem for which you think maximum entropy principle gives the right solution.

2. a) Assume $P_0(x_1, \ldots, x_m)$ is a zero mean Gaussian with a covariance matrix C where $C_{ij} = e^{-|i-j|}$. What is the minimum attainable relative entropy between P_0 and P that satisfies the constraints $x_i \geq \epsilon$ for all $i = 1, \ldots, m$?
   b) Evaluate the minimun relative entropy when $m = 5$ and $\epsilon = 1$.
   c) Explain your calculation as a large deviation bound for markov chains.

3. Provide at least four qualitatively different variational lower bound approximations for the following graph assuming all the variables in the graph are hidden variables (the set of visible variables is the empty set). Explain how the approximations scale as a function of m where m is the number of values each variable can take.



4. **Research level:** In the variational framework, the variational distribution need not be completely factorized but can be any tractable directed model. This naturally includes trees as well. The structure of the variational distribution is typically assumed to be given by the user which, in many cases, is a serious limitation. Assuming we are dealing with trees, formulate the variational inference problem so that it includes a feasible search over the tree structures. Hint: use a maximum weight spanning tree algorithm with weights derived e.g. from the effects that any particular pairwise dependence would have on the lower bound beyond the completely factorized mean field approximation.

   The utility of this algorithm would extend beyond the inference problem. For example, it would allow us to extract succint context dependent explanations from complicated dependency structures.