6.892 Machine learning seminar

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

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

March 1, 1999

This assignment is due **March 8**. You can either email the assignment to me as plain text (use subject "HOMEWORK 4") or give it to me in class.

1. a) Show that triangulation implies decomposability [1].
   b) Why are triangulated graphs so important in probabilistic calculations? (2 sentence limit :-)

2. a) We can construct a junction tree from the cliques $\{C_i\}$ of a decomposable graph by applying a maximum weight spanning tree algorithm with weights $w_{ij} = |C_i \cap C_j|$ (the number of nodes in the intersection). Show that any tree of cliques that is not a maximum weight spanning tree cannot be a junction tree.
   b) Often the max weight spanning tree is not unique (e.g. when the weights are all equal). How would you select among the alternative trees?

3. **Research level:** We can use the MWST algorithm to find the ML structure of simple tree models. Formulate an analogous search method for the ML junction tree structure subject to e.g. specific limit on the maximum allowable clique size.

   More generally, assigning variables to clusters rather than clustering data points (vectors of all the variables) is the dual clustering problem. Can you exploit the duality by adapting existing clustering algorithms to generate meaningful hierarchical clusters of variables?

[1] Hint: use induction on the size of the graph