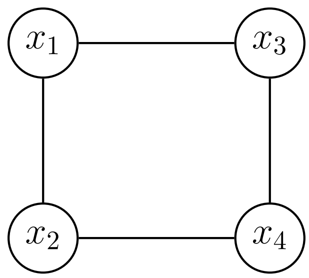
University of Illinois Spring 2018

CS 446: Machine Learning
Homework 7

Due on Tuesday, March 6, 2018, 11:59 a.m. Central Time

1. **[4 points]** Inference Methods for Discrete Markov Random Fields

For this problem, consider the following Markov Random Field, where each node can be assigned one of the values in $\{1, 2, 3, 4, 5\}$:



(a) To conduct MAP inference on this graph using exhaustive search, how many configurations must be checked?

Solution: $5^4 = 625$ (1 pt)

(b) Can MAP inference be run on this graph using a dynamic programming algorithm? Why or why not?

Solution: No. The graph is not a tree. (2 pts, one for "No" and one for explanation)

(c) To run MAP inference on this graph using loopy belief propagation, how many messages must be computed?

Solution: # Messages = Number of nodes $\times$ neighbors per node $\times$ number of node assignments
$= 4 \times 2 \times 5$
$= 40$ (1 pt)

2. **[7 points]** ILP Inference formulation in Discrete Markov Random Fields

(a) Suppose we have two variables $x_1 \in \{0, 1\}$ and $x_2 \in \{0, 1\}$ and their local evidence functions $\theta_1(x_1)$ and $\theta_2(x_2)$ as well as a pairwise function $\theta_{1,2}(x_1, x_2)$. Using this setup, inference solves $\arg\max_{x_1,x_2} \theta_1(x_1) + \theta_2(x_2) + \theta_{1,2}(x_1, x_2)$. Using

$$\theta_1(x_1) = \begin{cases} 1 & \text{if } x_1 = 0 \\ 2 & \text{otherwise} \end{cases} \qquad \theta_2(x_2) = \begin{cases} 1 & \text{if } x_2 = 0 \\ 2 & \text{otherwise} \end{cases}$$

$$\theta_{1,2}(x_1, x_2) = \begin{cases} -1 & \text{otherwise} \\ 2 & \text{if } x_1 = 0 \ \&\ x_2 = 1 \end{cases}$$

what is the integer linear programming formulation of the inference task? Make the cost function and constraints explicit for the given problem, i.e., do not use a general formulation.

Solution:

$$\max_b \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \\ -1 \\ -1 \\ 2 \\ -1 \end{bmatrix}^\top \begin{bmatrix} b_1(0) \\ b_1(1) \\ b_2(0) \\ b_2(1) \\ b_{1,2}(0,0) \\ b_{1,2}(1,0) \\ b_{1,2}(0,1) \\ b_{1,2}(1,1) \end{bmatrix}$$

$$\text{s.t.} \begin{cases} b_1(0), b_1(1), b_2(0), b_2(1), b_{1,2}(0,0), b_{1,2}(1,0), b_{1,2}(0,1), b_{1,2}(1,1) \in \{0,1\} \\ b_1(0) + b_1(1) = 1, b_2(0) + b_2(1) = 1 \\ b_{1,2}(0,0) + b_{1,2}(1,0) + b_{1,2}(0,1) + b_{1,2}(1,1) = 1 \\ b_1(0) = b_{1,2}(0,0) + b_{1,2}(0,1) \\ b_1(1) = b_{1,2}(1,0) + b_{1,2}(1,1) \\ b_2(0) = b_{1,2}(0,0) + b_{1,2}(1,0) \\ b_2(1) = b_{1,2}(0,1) + b_{1,2}(1,1) \end{cases}$$

(4 pts - one for max, one for $b_r \in \{0,1\}$, one for intra-region marginalization constraints, one for inter-region marginalization constraints)

(b) What is the solution (value and argument) to the program in part (a).

Solution:
argument: $b_1(0) = 1; b_1(1) = 0; b_2(0) = 0; b_2(1) = 1; b_{1,2}(0,0) = 0; b_{1,2}(1,0) = 0; b_{1,2}(0,1) = 1; b_{1,2}(1,1) = 0$; value: 5 (2 pts - one for value, one for argument)

(c) Why do we typically not use the integer linear program for reasonably sized MRFs?

Solution:
It can be very slow (ILP is NP-Hard). (1 pt)