T.61.5140 Machine Learning: Advanced Probablistic Methods
Hollmén, Raiko (Spring 2008)
Problem session, 22nd of February, 2008
http://www.cis.hut.fi/Opinnot/T-61.5140/

1. Draw the graphical model and factor graph (page 399), and run the sum-product algorithm (page 402, also known as belief propagation) on pencil and paper for the model below. Variables are S for sprinkler being on, R for raining, and W for the grass being wet. Compute $P(W = t)$.

P(S)	
S=f	0.7
S=t	0.3

P(R)	
R=f	0.8
R=t	0.2

$P(W \mid S, R)$	S=f, R=f	S=f, R=t	S=t, R=f	S=t, R=t
W=f	1.0	0.0	0.0	0.0
W=t	0.0	1.0	1.0	1.0

2. The model in Problem 1 is extended with a variable C for the sky being cloudy. (a) Draw the graphical model and the factor graph. (b) Why cannot you run the sum-product algorithm? (c) Where does the inference stop if you try anyway? (d) One could avoid the problem by using loopy belief propagation (page 417), explain shortly how.

P(C)	
C=f	0.5
C=t	0.5

$P(S \mid C)$	C=f	C=t
S=f	0.5	0.9
S=t	0.5	0.1

$P(R \mid C)$	C=f	C=t
R=f	1.0	0.4
R=t	0.0	0.6

$P(W \mid S, R)$	S=f, R=f	S=f, R=t	S=t, R=f	S=t, R=t
W=f	1.0	0.0	0.0	0.0
W=t	0.0	1.0	1.0	1.0

3. The junction tree algorithm is the generalization of the sum-product algorithm to general graphs. When applied to the model in Problem 2, it would effectively cluster variables S and R into a single variable, say S_R, that has four possible values, f_f, f_t, t_f, and t_t to represent the different combinations of S and R being false or true. (a) Write the joint probability $P(C, S_R, W)$. (b) Draw the graphical model and a factor graph. (c) Compute the conditional probability tables for the modified model.

4. For the model in Figure 1, construct a junction tree (also known as join tree). This is formed by moralizing the graph (connecting co-parents), forgetting the direction of the edges, triangulating (adding a chord to each chordless cycle of four or more nodes), and finally creating a junction tree based on the resulting graph. The junction tree has a node for each maximal clique of the previous graph. Temporarily all the nodes are connected with weighted edges, the weight being the number of shared variables in the two cliques. A maximum weight spanning tree is formed, and other edges and all weight can be forgotten. These nodes (cliques) correspond to the small square nodes of the factor graph. One still needs to add the variables in the graph as round nodes. Variables in the intersection of neighbouring cliques need to be clustered into a single variable, just like S and R in Problem 3. (Note that a variable may appear many times. Note also that the junction tree is not unique, there can be many ways to triangulate and to choose a maximum spanning tree.)

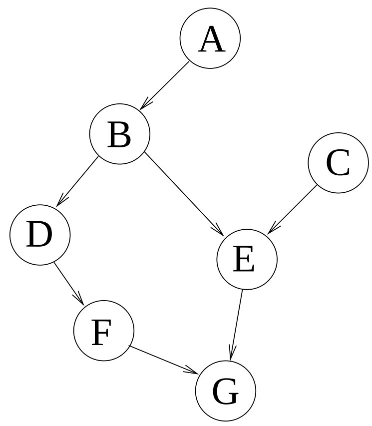

Figure 1: Problem 4.