Problem 1 – A graph of tree-width 2

The graph $G$ below has treewidth 2. The treewidth of a graph is one less than the size of the maximum clique in that graph.

![Graph G](image)

a. Find an orientation for the graph $G$, which produces no V-structures. Denote the resulting DAG by $G'$.

[b. NOT GRADED] Write the general factored form of a distribution $P$ for which the undirected graph is an I-map.

c. NOT GRADED Write the general factored form of a distribution $P'$ of which the DAG you found in question a. is an I-map.

[d. NOT GRADED] The two factorizations in b,c must be equal. Find a way to group the factors in $P'$ to obtain the factorization in $P$. The grouping may not be unique.

Using the grouping you found, show that the potentials of $P$ have a probabilistic interpretation. Since there will be many potentials, it is sufficient to find a probabilistic interpretation for one potential containing variable $A$ and for one containing the variable $E$.

e. Verify that $G$ is chordal by the Tarjan elimination algorithm.
f. Construct a junction tree \( J = (\mathcal{C}, \mathcal{S}) \) for the graph \( \mathcal{G} \). Is this tree unique? List its separators (with multiplicities).

g. Write the factorization of \( P_{ABCDEF} \) w.r.t the junction tree \( J \) in f. (i.e clique marginals over separator marginals).

h. Show step by step the variable elimination procedure for computing \( P_{A,C=0,D=0,E=0} \) using the factorization of \( P \) obtained in c. Choose an elimination ordering that does not create new factors (a.k.a potentials) involving more than 2 variables.

i. Each potential represents a (conditional) probability table; find the probabilistic semantics for each of the potentials created during this elimination. Some of the eliminations will lead to potentials equal to constants. Be sure to identify these (so, in practice, these eliminations can be skipped). Explain why and when this happens.

j. Now run the Sum-Product algorithm on the junction tree to obtain \( P_{A,C=0,D=0,E=0} \). More precisely, describe the sequence of messages, and give the expression of each message, in terms of the factorization of \( J \). Note that you must find a way to re-express \( P_V \) as a product only (no division).

It is OK to: (i) express \( P_V \) as a product of new potentials \( \phi_U \), giving the expression of each \( \phi_U \) in relation to the factors you obtained in g.; (ii) give the sequence of messages in terms of the factors \( \phi_U \)'s and new factors \( \psi_U \)'s.

k. Optional, not graded: Each message and factor appearing in the Sum-Product algorithm represents a probability. Find what is this probability (you can compare with the steps of VE to get to the answer).

l. Now show how to get \( P_{A|C=0,D=0,E=0} \) using the Sum-Product algorithm.

**Problem 3 – IPF in a decomposable model**

Prove that in a junction tree, i.e. in a MRF over a chordal graph, there is a sequence of clique updates so that IPF converges in a finite number of steps.

More precisely, use the rooted tree factorization of the junction tree, which has no explicit separator potentials. (The statement is true for other j.t. parametrizations as well, but one would have to appropriately re-define the IPF algorithm.)

Assume that the initialization is compatible with some \( P_V^0 > 0 \) joint distribution over the variables in \( V \). You must show that the algorithm finds the correct values after a finite number of steps from any such initialization. A “step” here means updating all the parameters of a single clique, i.e. updating \( \phi_C \) for a single \( C \). You can prescribe in the algorithm the order in which the updates should be
made.

[Optional, extra credit: What would happen, if the initialization was not compatible with a $P^0_\gamma$, but an initialization with arbitrary potential values $\phi_C(x_C) > 0$? Can you modify the IPF algorithm, so that it converges to the correct j.t. parametrization for this initialization?]