Algorithm 13.3 Calibration using max-product BP in a Bethe-structured cluster graph

Procedure Generalized-MP-BP (Φ, // Set of factors
R, // Set of regions
\{κ_r\}_{r \in R}, \{κ_i\}_{X_i \in X} // Counting numbers
)

1. \(ρ_i \leftarrow \frac{1}{κ_i}\)
2. \(ρ_r \leftarrow \frac{1}{κ_r}\)
3. Initialize-CGraph
4. while region graph is not max-calibrated
5. Select \(C_r\) and \(X_i \in C_r\)
6. \(δ_{i \rightarrow r}(X_i) \leftarrow \left[\left(\prod_{r' \neq r} δ_{i \rightarrow r'}(X_i)\right)^{ρ_i}\left(\max_{X_r \in C_r} ψ_r(C_r)\left(\prod_{X_j \in C_r, j \neq i} δ_{j \rightarrow r}\right)^{ρ_r}\right]\right]^{-\frac{1}{ρ_i + ρ_r}}
7. for each region \(r \in R \cup \{1, \ldots, n\}\)
8. \(β_r(C_r) \leftarrow (ψ_r(C_r)\prod_{X_i \in C_r} δ_{i \rightarrow r}(X_i))^{ρ_r}\)
9. return \(\{β_r\}_{r \in R}\)

13.4.2.1 Max-Product with Counting Numbers

We begin with a reminder of the notion of belief propagation with counting numbers. For concreteness, we also provide the max-product variant of a message passing algorithm for this case, although (as we mentioned) the max-product variant can be obtained from the sum-product algorithm using a simple syntactic substitution.

In section 11.3.7, we defined a set of sum-product message passing algorithms; these algorithms were defined in terms of a set of counting numbers that specify the extent to which entropy terms for different subsets of variables are counted in the entropy approximation used in the energy functional. For a given set of counting numbers, one can derive a message passing algorithm by using the fixed point equations obtained by differentiating the Lagrangian for the energy functional, with its sum-product calibration constraints. The standard belief propagation algorithm is obtained from the Bethe energy approximation; other sets of counting numbers give rise to other message passing algorithms.

As we discussed, one can take these sum-product message passing algorithms (for example, those in exercise 11.17 and exercise 11.19) and convert them to produce a max-product variant by simply replacing each summation operation as maximization. For concreteness, in algorithm 13.3, we repeat the algorithm of exercise 11.17, instantiated to the max-product setting. Recall that this algorithm applies only to *Bethe cluster graphs*, that is, graphs that have two levels of regions: “large” regions \(r\) containing multiple variables with counting numbers \(κ_r\), and singleton regions containing individual variables \(X_i\) with counting numbers \(κ_i\); all factors in \(Φ\) are assigned only to large regions, so that \(ψ_i = 1\) for all \(i\).