THE UNIVERSITY *of* EDINBURGH informatics

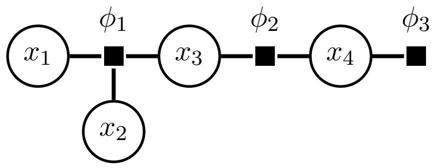
Probabilistic Modelling and Reasoning
Exercises 3 — Notes

Spring 2023
Hodari & Gutmann

These notes are intended to give a summary of relevant concepts from the lectures which are helpful to complete the exercises. It is not intended to cover the lectures thoroughly. Learning this content is not a replacement for working through the lecture material and the exercises.

Factor graph — A factor graph represents an arbitrary function in terms of factors and their connections with variables. For example, a factor graph can represent a distribution written as a Gibbs distribution – $p(\mathbf{x}) = \frac{1}{Z}\prod_c \phi_c(\mathcal{X}_c)$ – where variables $x_i \in \mathbf{x}$ are represented with variable nodes (circles) and potentials ϕ_c are represented with factor nodes (squares). Edges connect each factor node ϕ_c to all its variable nodes $x_i \in \mathcal{X}_c$.

$$p(x_1, x_2, x_3, x_4) = \frac{1}{Z}\phi_1(x_1, x_2, x_3)\phi_2(x_3, x_4)\phi_3(x_4)$$



Variable elimination — Given $p(\mathcal{X}) \propto \prod_c \phi_c(\mathcal{X}_c)$, we compute the marginal $p(\mathcal{X} \setminus x^*)$ via the sum rule by exploiting the factorisation by means of the distributive law.

We sum out the variable x^* by first finding all factors $\phi_i(\mathcal{X}_i)$ such that $x^* \in \mathcal{X}_i$, and forming the compound factor $\phi^*(\mathcal{X}^*) = \prod_{i:x^* \in \mathcal{X}_i} \phi_i(\mathcal{X}_i)$, with $\mathcal{X}^* = \bigcup_{i:x^* \in \mathcal{X}_i} \mathcal{X}_i$. Summing out x^* then produces a new factor $\tilde{\phi}^*(\tilde{\mathcal{X}}^*) = \sum_{x^*} \phi^*(\mathcal{X}^*)$ that does not depend on x^*, i.e. $\tilde{\mathcal{X}}^* = \mathcal{X}^* \setminus x^*$. This is possible as products are commutative, and a sum can be distributed within a product as long as all terms depending on the variable(s) being summed come to the right of the sum.

$$p(\mathcal{X} \setminus x^*) \propto \sum_{x^*} \prod_c \phi_c(\mathcal{X}_c) \propto \left[\prod_{i:x^* \notin \mathcal{X}_i} \phi_i(\mathcal{X}_i)\right] \left[\sum_{x^*} \prod_{i:x^* \in \mathcal{X}_i} \phi_i(\mathcal{X}_i)\right] \tag{1}$$

$$\propto \left[\prod_{i:x^* \notin \mathcal{X}_i} \phi_i(\mathcal{X}_i)\right] \tilde{\phi}^*(\tilde{\mathcal{X}}^*) \tag{2}$$

When eliminating variables, order of elimination matters. However, optimal choice of elimination order is difficult. Picking variables greedily is a common heuristic, where the "best" x^* is the one that fewest factors ϕ_c depend upon.