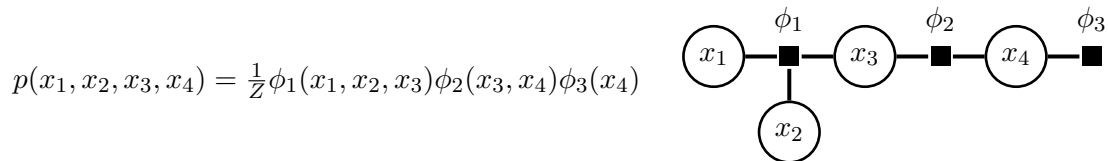


*These notes are intended to give a summary of relevant concepts from the lectures which are helpful to complete the tutorial sheet. It is not intended to cover the lectures thoroughly. Learning this content is not a replacement for working through the lecture material and the tutorial sheet.*

**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  $\phi_c$  are represented with factor nodes (squares). Edges connect each factor node  $\phi_c$  to all its variable nodes  $x_i \in \mathcal{X}_c$ .



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

We sum out the variable  $x^*$  by first finding all factors  $\phi_i(\mathcal{X}_i)$ , such that  $x^* \in \mathcal{X}_i$ . We combine these factors into a new factor  $\tilde{\phi}^*(\tilde{\mathcal{X}}^*) = \sum_{x^*} \prod_{i: x^* \in \mathcal{X}_i} \phi_i(\mathcal{X}_i)$ . Note that  $\tilde{\phi}^*$  does not depend on  $x^*$ , i.e.  $\tilde{\mathcal{X}}^* = \bigcup_{i: x^* \in \mathcal{X}_i} (\mathcal{X}_i \setminus x^*)$ . 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.

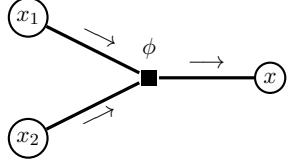
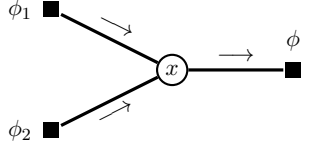
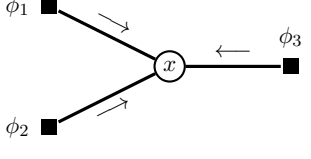
$$p(\mathcal{X} \setminus x^*) \propto \sum_{x^*} \prod_c \phi_c(\mathcal{X}_c) \propto \left[ \prod_{i: x^* \notin \mathcal{X}_i} \phi_i(\mathcal{X}_i) \right] \left[ \sum_{x^*} \prod_{i: x^* \in \mathcal{X}_i} \phi_i(\mathcal{X}_i) \right] \quad (1)$$

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

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  $\phi_c$  depend upon.

**Sum-product algorithm** — Variable elimination (for factor trees) reformulated with “messages” which allows for re-use of computations already done. See table on following page.

**Max-product algorithm** — Same as the sum-product algorithm, but *max* replaces  $\sum$ .

$\mu_{\phi \rightarrow x}(x)$	Factor to variable $\mu_{\phi \rightarrow x}(x) = \sum_{x_1, \dots, x_j} \phi(x_1, \dots, x_j, x) \prod_{i=1}^j \mu_{x_i \rightarrow \phi}(x_i)$	
$\mu_{x \rightarrow \phi}(x)$	Variable to factor $\mu_{x \rightarrow \phi}(x) = \prod_{i=1}^j \mu_{\phi_i \rightarrow x}(x)$	
$\tilde{p}(x)$	Univariate marginals – unnormalised $p(x) \propto \prod_{i=1}^j \mu_{\phi_i \rightarrow x}(x)$	
$\tilde{p}(x_1, \dots, x_j)$	Joint marginals of variables sharing a factor – unnormalised $p(x_1, \dots, x_j) \propto \phi(x_1, \dots, x_j) \prod_{i=1}^j \mu_{x_i \rightarrow \phi}(x_i)$	