



Figure 3.7: A directed graphical model over random variables  $a$ ,  $b$ ,  $c$ ,  $d$  and  $e$ . This graph corresponds to probability distributions that can be factored as

$$p(a, b, c, d, e) = p(a)p(b | a)p(c | a, b)p(d | b)p(e | c). \quad (3.54)$$

This graph allows us to quickly see some properties of the distribution. For example,  $a$  and  $c$  interact directly, but  $a$  and  $e$  interact only indirectly via  $c$ .

are usually not probability distributions of any kind. Any set of nodes that are all connected to each other in  $\mathcal{G}$  is called a clique. Each clique  $\mathcal{C}^{(i)}$  in an undirected model is associated with a factor  $\phi^{(i)}(\mathcal{C}^{(i)})$ . These factors are just functions, not probability distributions. The output of each factor must be non-negative, but there is no constraint that the factor must sum or integrate to 1 like a probability distribution.

The probability of a configuration of random variables is **proportional** to the product of all of these factors—assignments that result in larger factor values are more likely. Of course, there is no guarantee that this product will sum to 1. We therefore divide by a normalizing constant  $Z$ , defined to be the sum or integral over all states of the product of the  $\phi$  functions, in order to obtain a normalized probability distribution:

$$p(\mathbf{x}) = \frac{1}{Z} \prod_i \phi^{(i)}(\mathcal{C}^{(i)}). \quad (3.55)$$

See figure 3.8 for an example of an undirected graph and the factorization of probability distributions it represents.

Keep in mind that these graphical representations of factorizations are a language for describing probability distributions. They are not mutually exclusive families of probability distributions. Being directed or undirected is not a property of a probability distribution; it is a property of a particular **description** of a