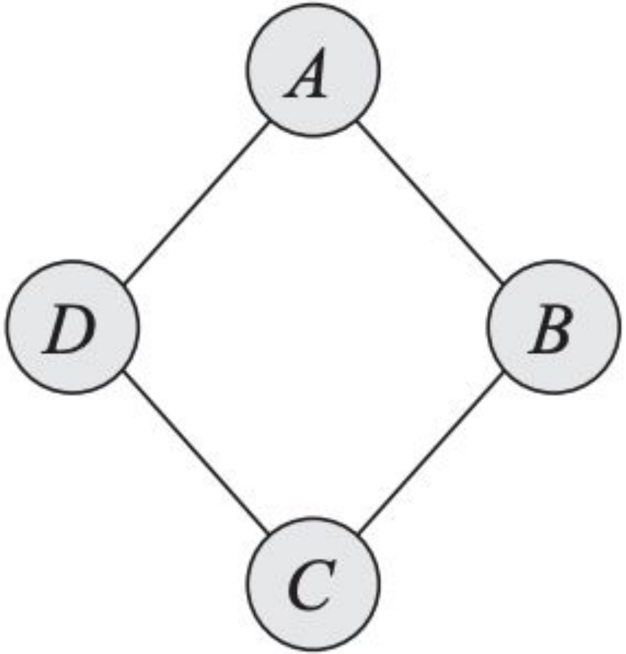
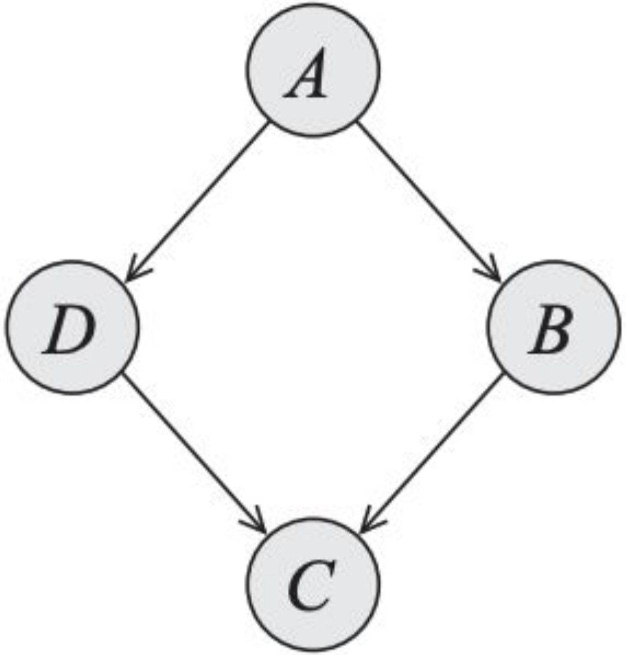
Advanced Machine Learning

Likhit Nayak

Undirected Graphs





Factors

Let $\boldsymbol{D}$ *be a set of random variables. We define a* factor ϕ *to be a function from* $Val(\boldsymbol{D})$ *to* $\mathbb{R}$. *A factor is* nonnegative *if all its entries are nonnegative. The set of variables* $\boldsymbol{D}$ *is called the* scope *of the factor and denoted* $Scope[\phi]$.

$\phi_1(A, B)$

a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$\phi_2(B, C)$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

$\phi_3(C, D)$

c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_4(D, A)$

d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

Factors

A distribution P_Φ *is a* Gibbs distribution *parameterized by a set of factors* $\Phi = \{\phi_1(\boldsymbol{D}_1), \ldots, \phi_K(\boldsymbol{D}_K)\}$ *if it is defined as follows:*

$$P_\Phi(X_1, \ldots, X_n) = \frac{1}{Z}\tilde{P}_\Phi(X_1, \ldots, X_n),$$

where

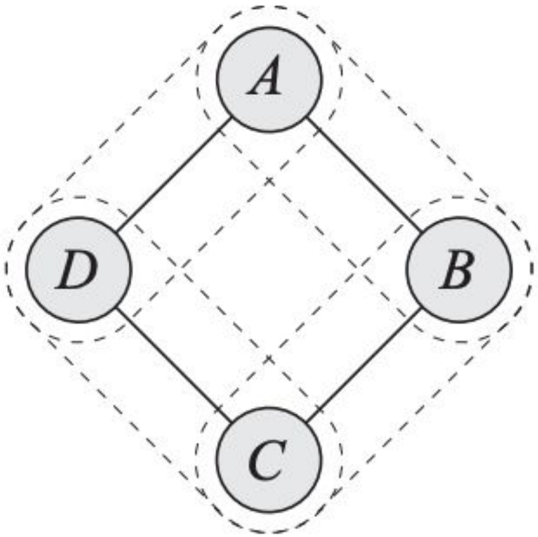
$$\tilde{P}_\Phi(X_1, \ldots, X_n) = \phi_1(\boldsymbol{D}_1) \times \phi_2(\boldsymbol{D}_2) \times \cdots \times \phi_m(\boldsymbol{D}_m)$$

$$Z = \sum_{X_1, \ldots, X_n} \tilde{P}_\Phi(X_1, \ldots, X_n)$$

is a normalizing constant called the partition function.

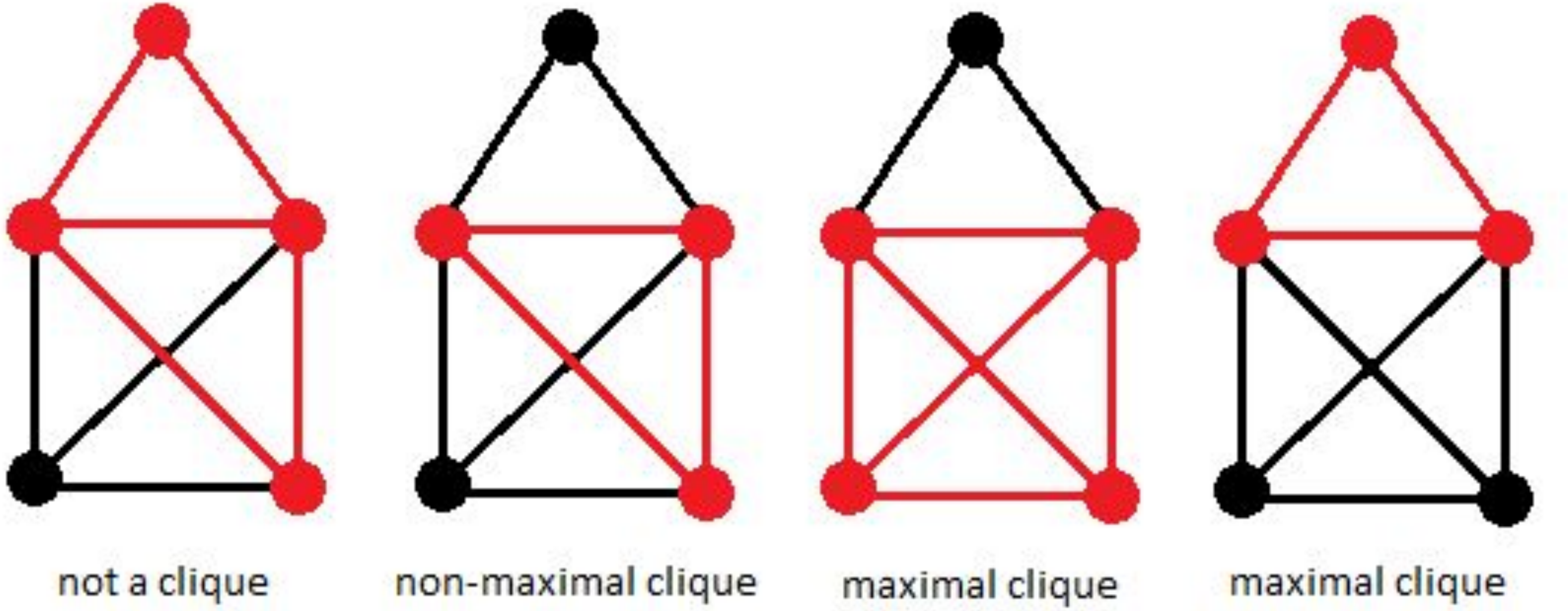
Markov Network

We say that a distribution P_Φ *with* $\Phi = \{\phi_1(\boldsymbol{D}_1), \ldots, \phi_K(\boldsymbol{D}_K)\}$ factorizes *over a Markov network* $\mathcal{H}$ *if each* $\boldsymbol{D}_k$ *(*$k = 1, \ldots, K$*) is a* $\boxed{\textit{complete subgraph}}$ *of* $\mathcal{H}$.

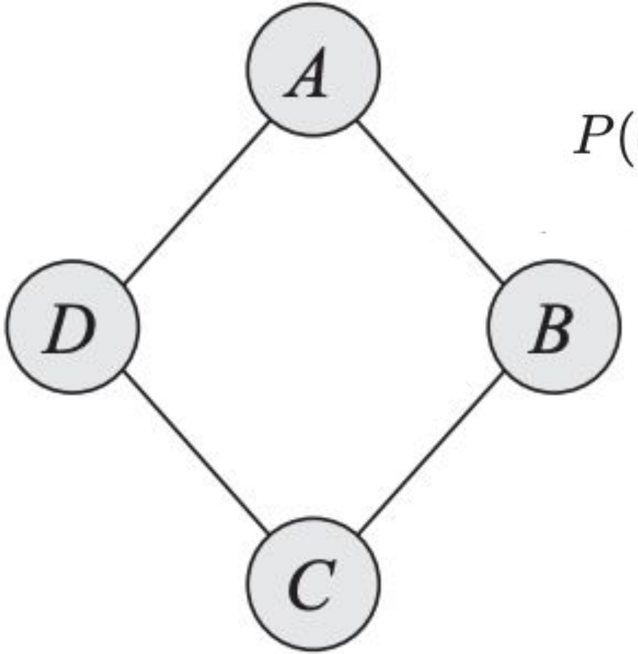


Clique

Cliques and maximal cliques



Markov Network



$$P(a,b,c,d) = \frac{1}{Z}\phi_1(a,b)\cdot\phi_2(b,c)\cdot\phi_3(c,d)\cdot\phi_4(d,a)$$

$$Z = \sum_{a,b,c,d}\phi_1(a,b)\cdot\phi_2(b,c)\cdot\phi_3(c,d)\cdot\phi_4(d,a)$$

Factor Product

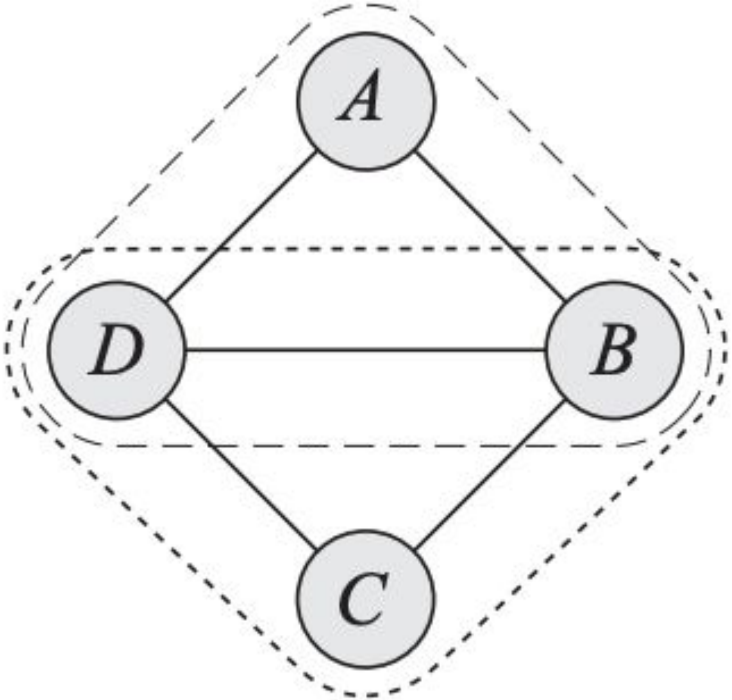
Let $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$ be three disjoint sets of variables, and let $\phi_1(\boldsymbol{X}, \boldsymbol{Y})$ and $\phi_2(\boldsymbol{Y}, \boldsymbol{Z})$ be two factors. We define the factor product *$\phi_1 \times \phi_2$ to be a factor $\psi : Val(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) \mapsto \mathbb{R}$ as follows:*

$$\psi(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) = \phi_1(\boldsymbol{X}, \boldsymbol{Y}) \cdot \phi_2(\boldsymbol{Y}, \boldsymbol{Z}).$$

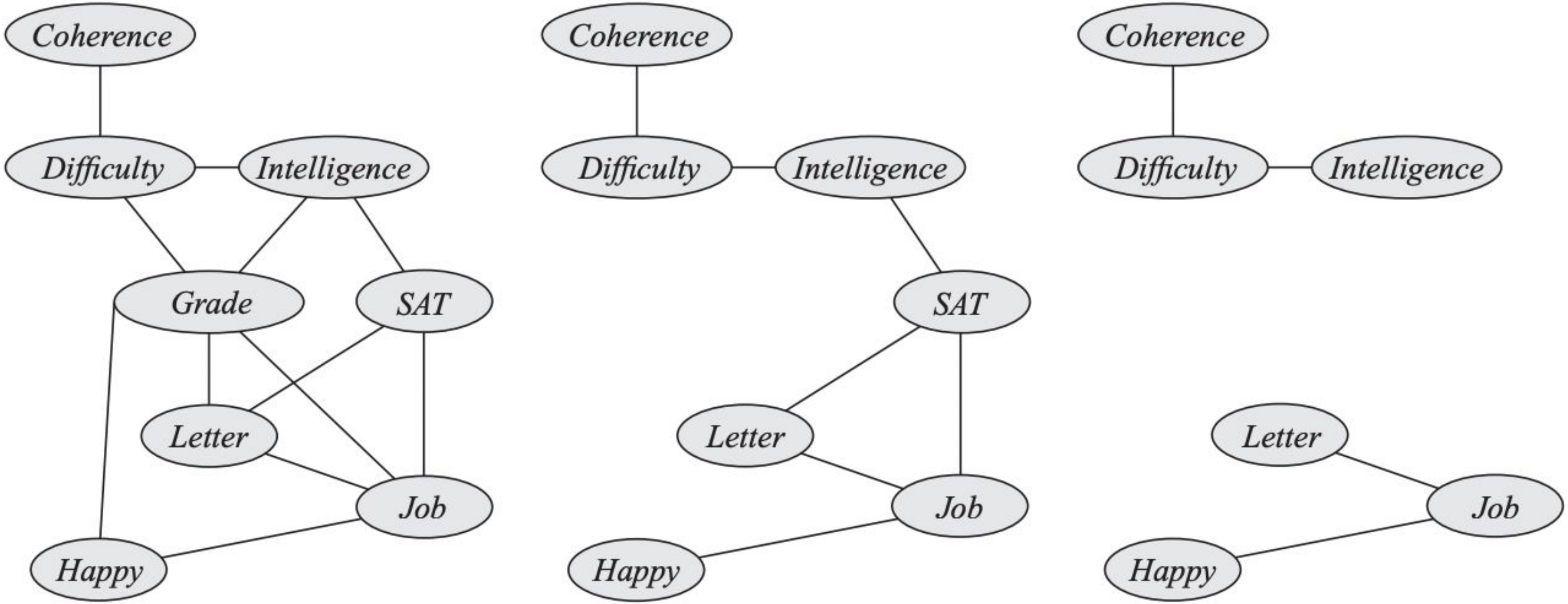
a^1	b^1	0.5
a^1	b^2	0.8
a^2	b^1	0.1
a^2	b^2	0
a^3	b^1	0.3
a^3	b^2	0.9

b^1	c^1	0.5
b^1	c^2	0.7
b^2	c^1	0.1
b^2	c^2	0.2

Markov Network



Reduced Markov Networks



Markov Networks - Image Segmentation