Neural networks

Conditional random fields - Markov network

LINEAR CHAIN CRF

Topics: factor, sufficient statistic

- With hidden units, the CRF factors could be:

$$\Psi_f(\mathbf{y}, \mathbf{X}) = \begin{cases} \phi_f(y_k, \mathbf{x}_{k-1}) = \exp\left(a^{(L+1,-1)}(\mathbf{x}_{k-1})_{y_k}\right) \\ \phi_f(y_k, \mathbf{x}_k) = \exp\left(a^{(L+1,0)}(\mathbf{x}_k)_{y_k}\right) \\ \phi_f(y_k, \mathbf{x}_{k+1}) = \exp\left(a^{(L+1,+1)}(\mathbf{x}_{k+1})_{y_k}\right) \\ \phi_f(y_k, y_{k+1}) = \exp\left(V_{y_k, y_{k+1}}\right) \end{cases}$$

- There is no simple form for the sufficient statistics

LINEAR CHAIN CRF

Topics: factor, sufficient statistic

- With hidden units, the CRF factors could be:

unary factors

$$\Psi_f(\mathbf{y},\mathbf{X}) = \begin{cases} \boxed{\phi_f(y_k,\mathbf{x}_{k-1})} = \exp\left(a^{(L+1,-1)}(\mathbf{x}_{k-1})_{y_k}\right) \\ \boxed{\phi_f(y_k,\mathbf{x}_{k})} = \exp\left(a^{(L+1,0)}(\mathbf{x}_{k})_{y_k}\right) \\ \boxed{\phi_f(y_k,\mathbf{x}_{k+1})} = \exp\left(a^{(L+1,+1)}(\mathbf{x}_{k+1})_{y_k}\right) \\ \phi_f(y_k,y_{k+1}) = \exp\left(V_{y_k,y_{k+1}}\right) \end{cases}$$

- There is no simple form for the sufficient statistics

LINEAR CHAIN CRF

Topics: factor, sufficient statistic

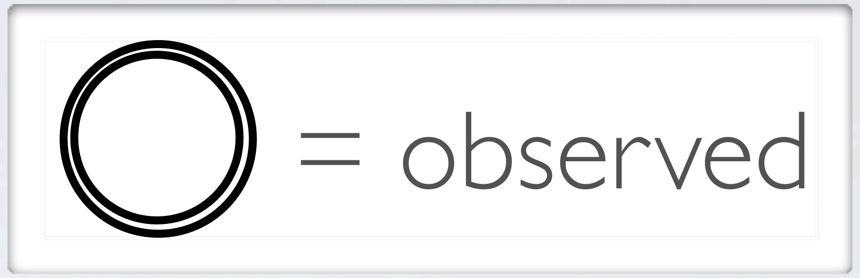
- With hidden units, the CRF factors could be:

$$\Psi_f(\mathbf{y},\mathbf{X}) = \begin{cases} \underbrace{\boxed{\phi_f(y_k,\mathbf{x}_{k-1})}}_{\text{unary factors}} = \exp\left(a^{(L+1,-1)}(\mathbf{x}_{k-1})_{y_k}\right) \\ \boxed{\phi_f(y_k,\mathbf{x}_k)} = \exp\left(a^{(L+1,0)}(\mathbf{x}_k)_{y_k}\right) \\ \boxed{\phi_f(y_k,\mathbf{x}_{k+1})} = \exp\left(a^{(L+1,+1)}(\mathbf{x}_{k+1})_{y_k}\right) \\ \underbrace{\boxed{\phi_f(y_k,y_{k+1})}}_{\text{pairwise factors}} = \exp\left(V_{y_k,y_{k+1}}\right) \end{cases}$$

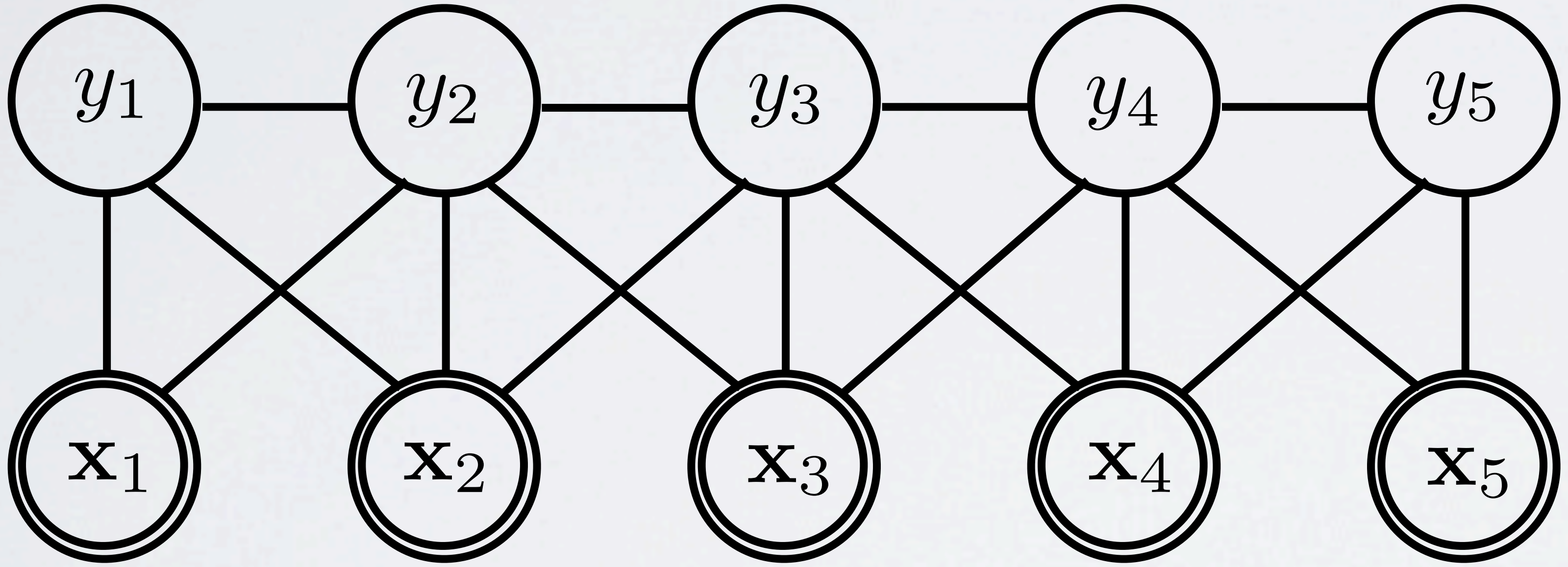
- There is no simple form for the sufficient statistics

MARKOV NETWORK VISUALIZATION

Topics: Markov network



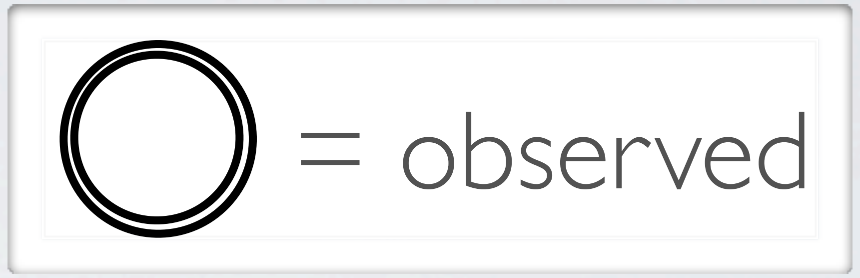
- Illustration for $K{=}5$



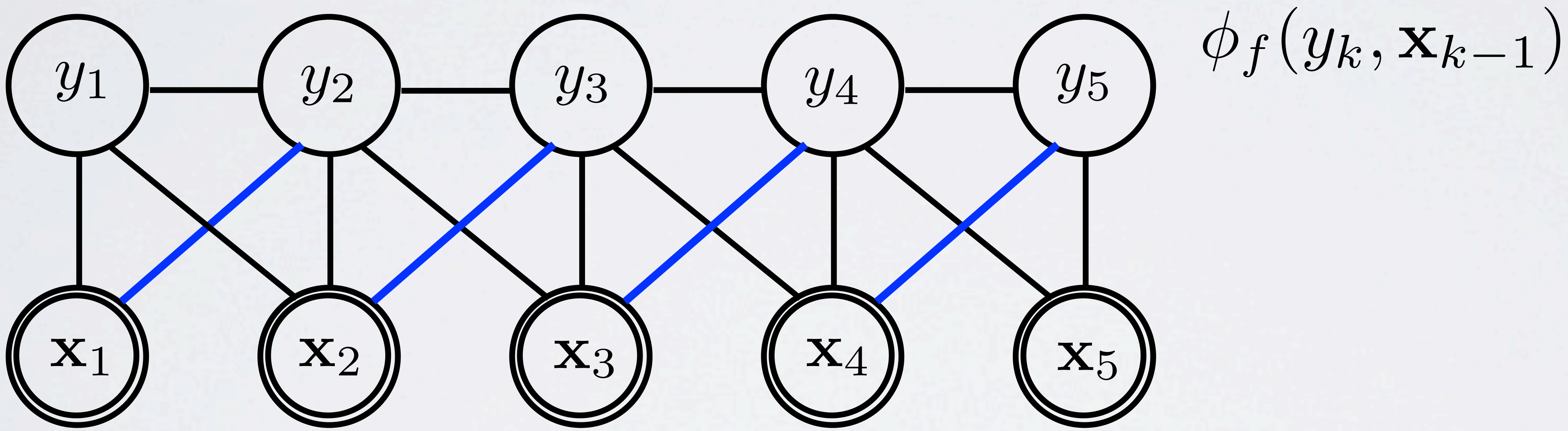
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



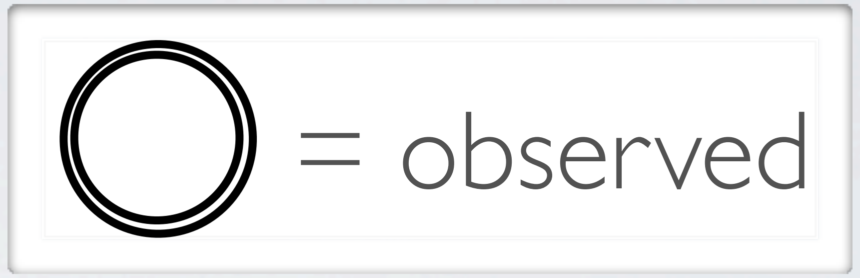
- Illustration for $K{=}5$



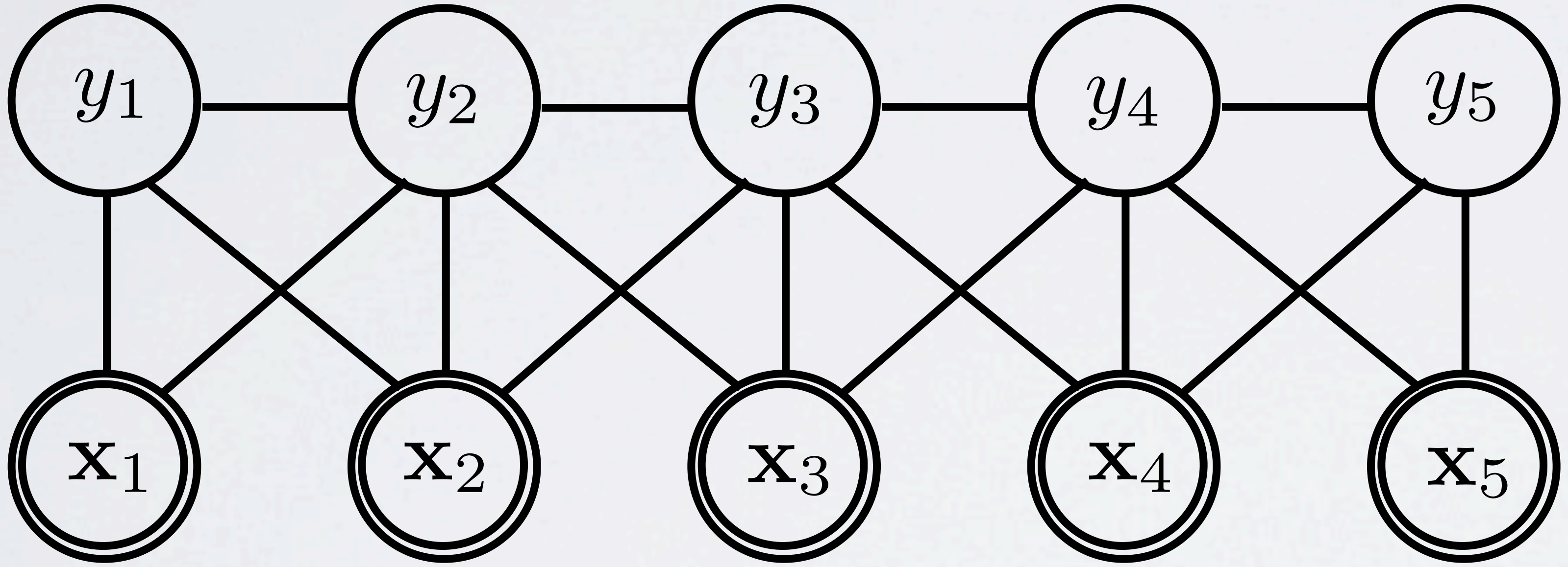
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



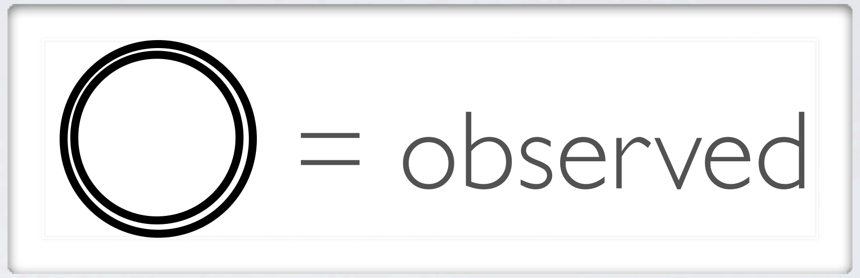
- Illustration for $K=5$



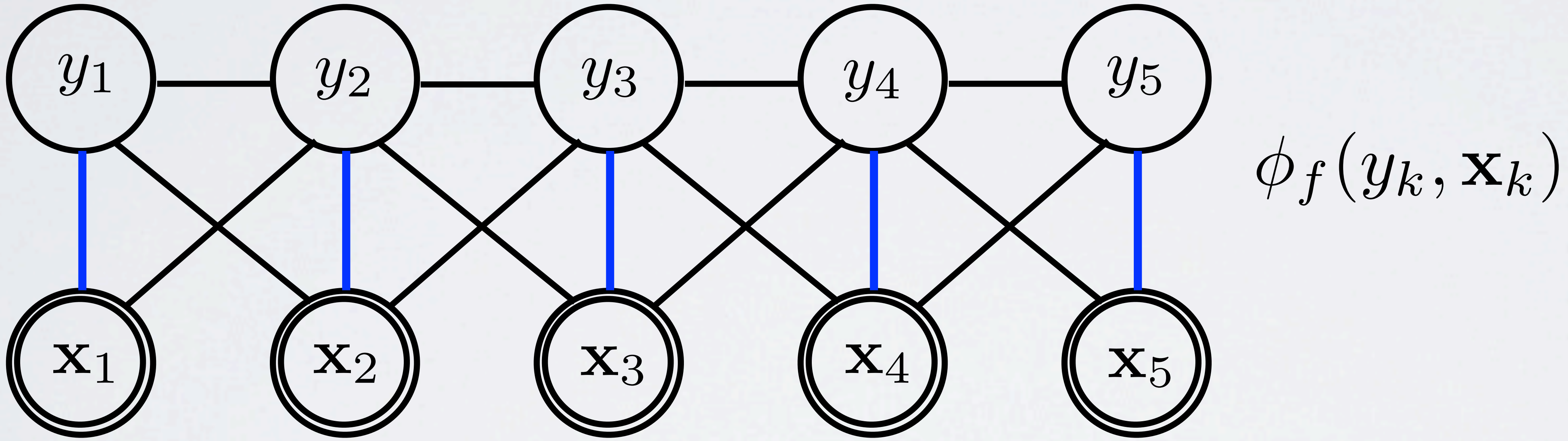
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



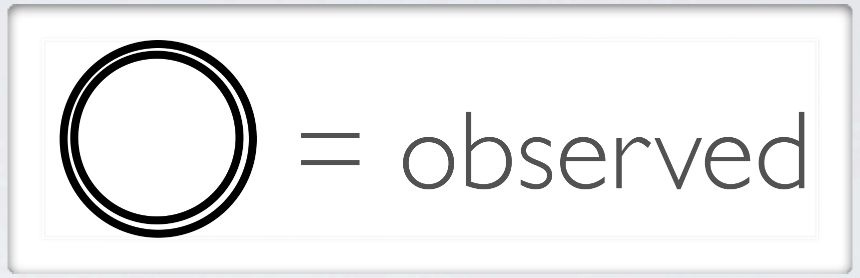
- Illustration for K=5



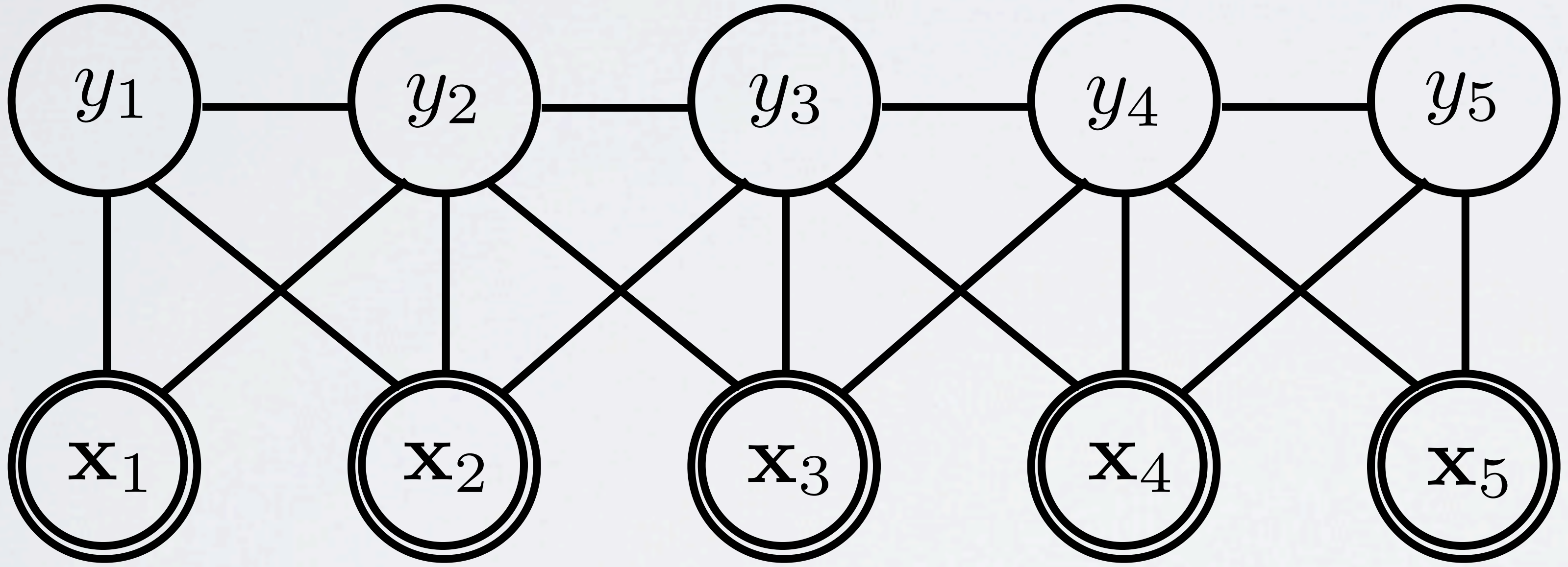
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



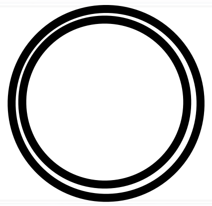
- Illustration for K=5



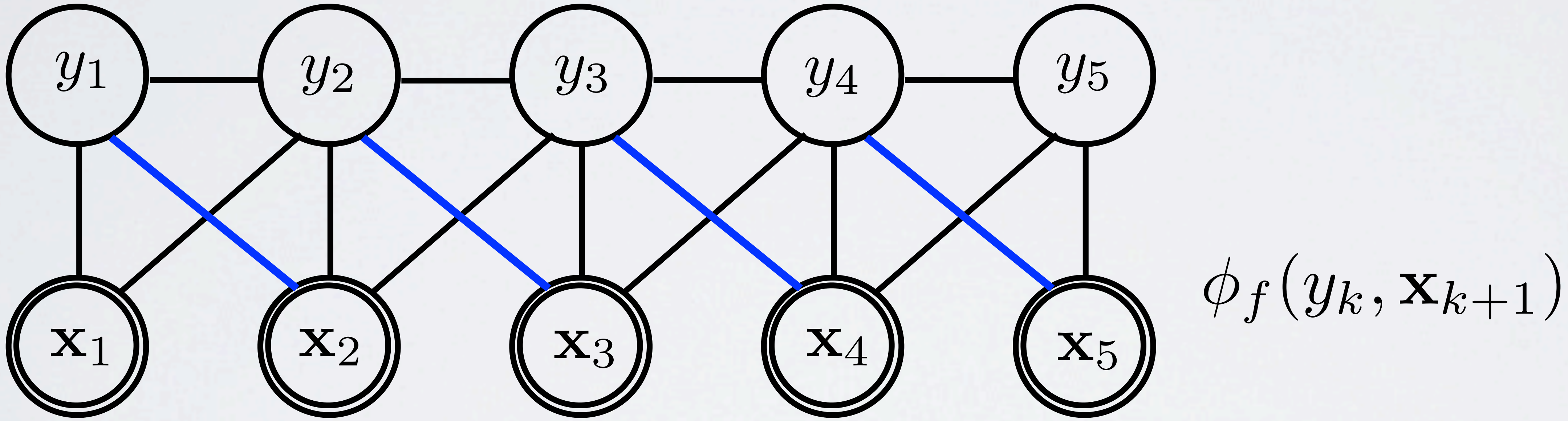
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network

 = observed

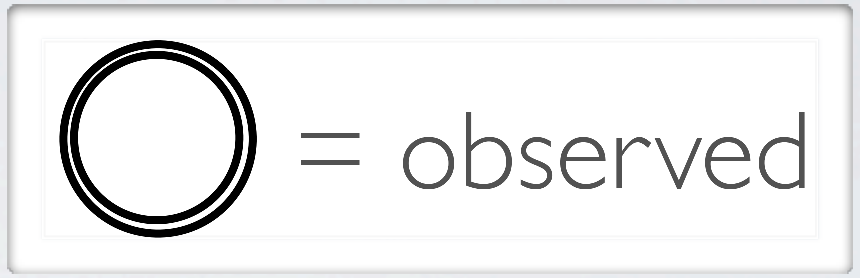
- Illustration for $K{=}5$



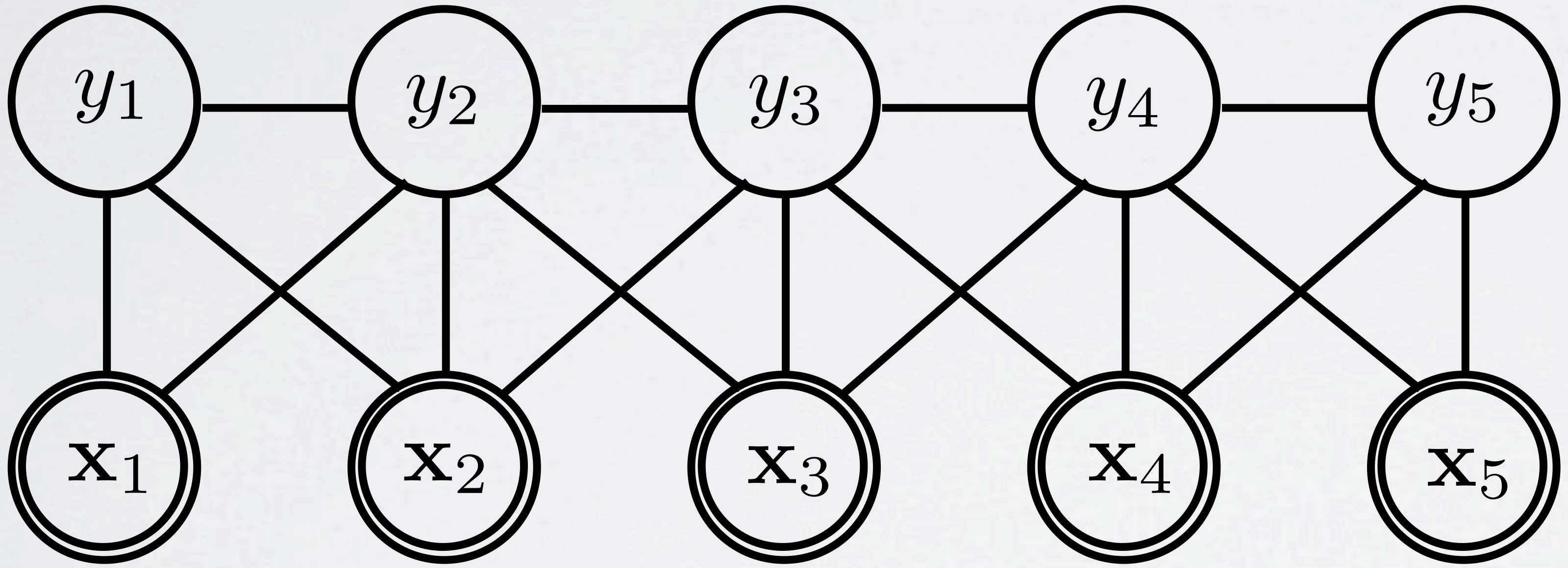
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



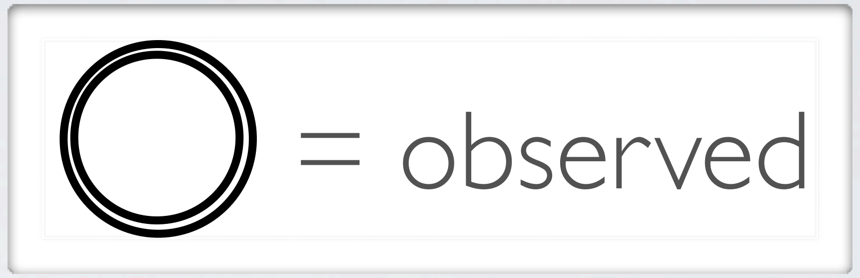
- Illustration for $K=5$



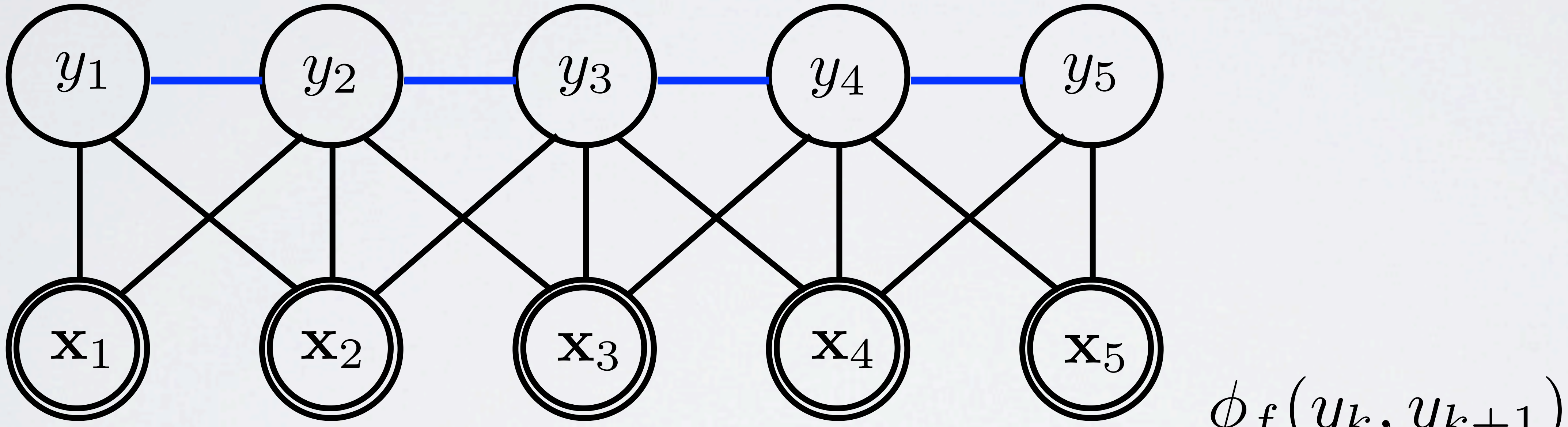
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



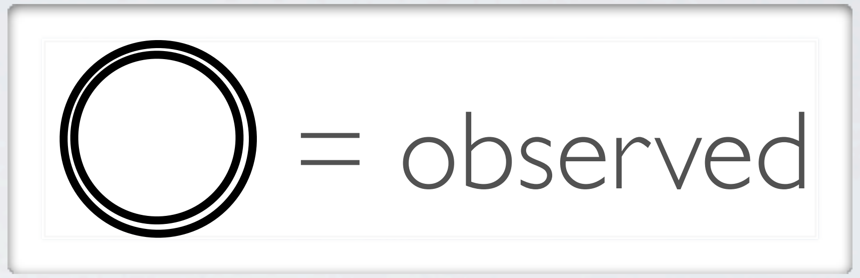
- Illustration for $K=5$



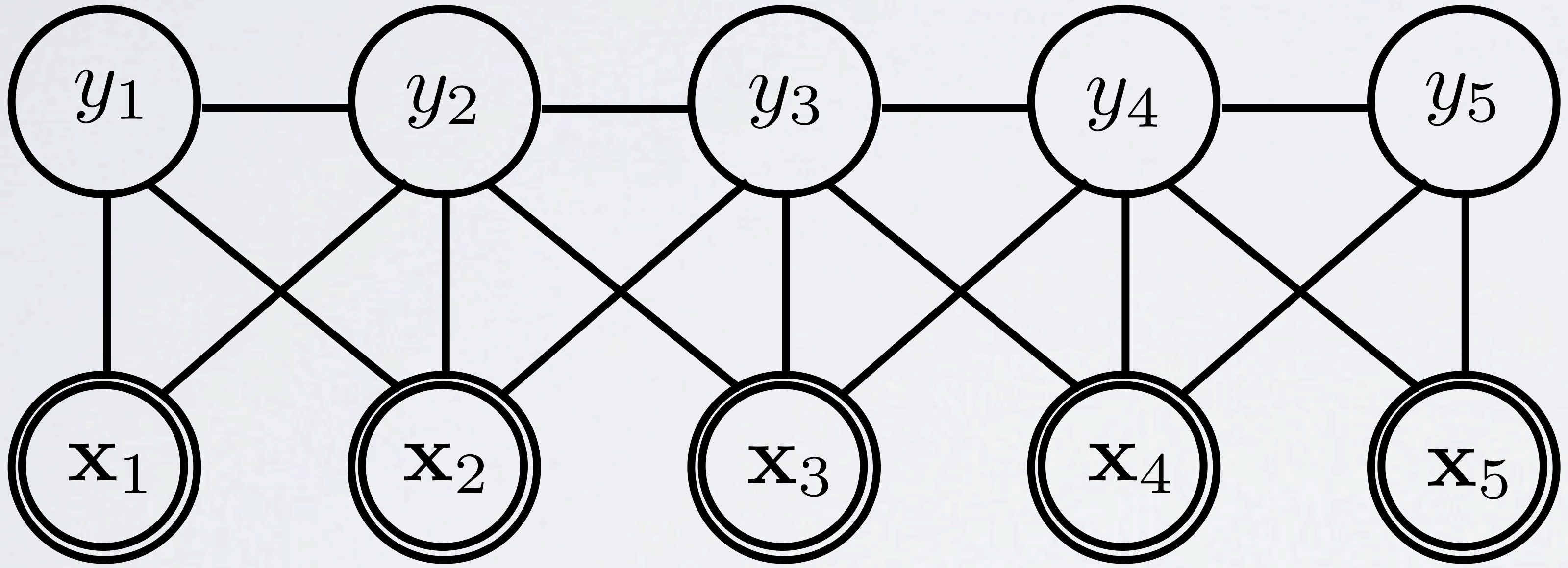
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



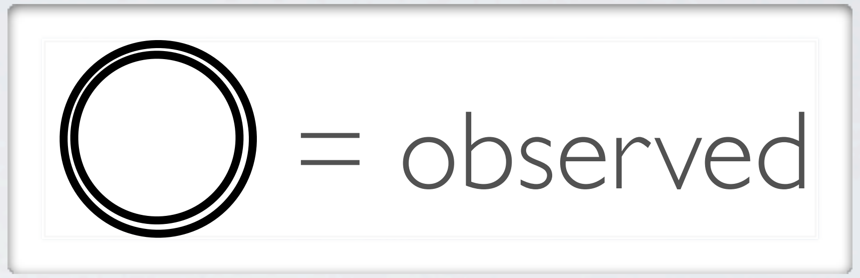
- Illustration for K=5



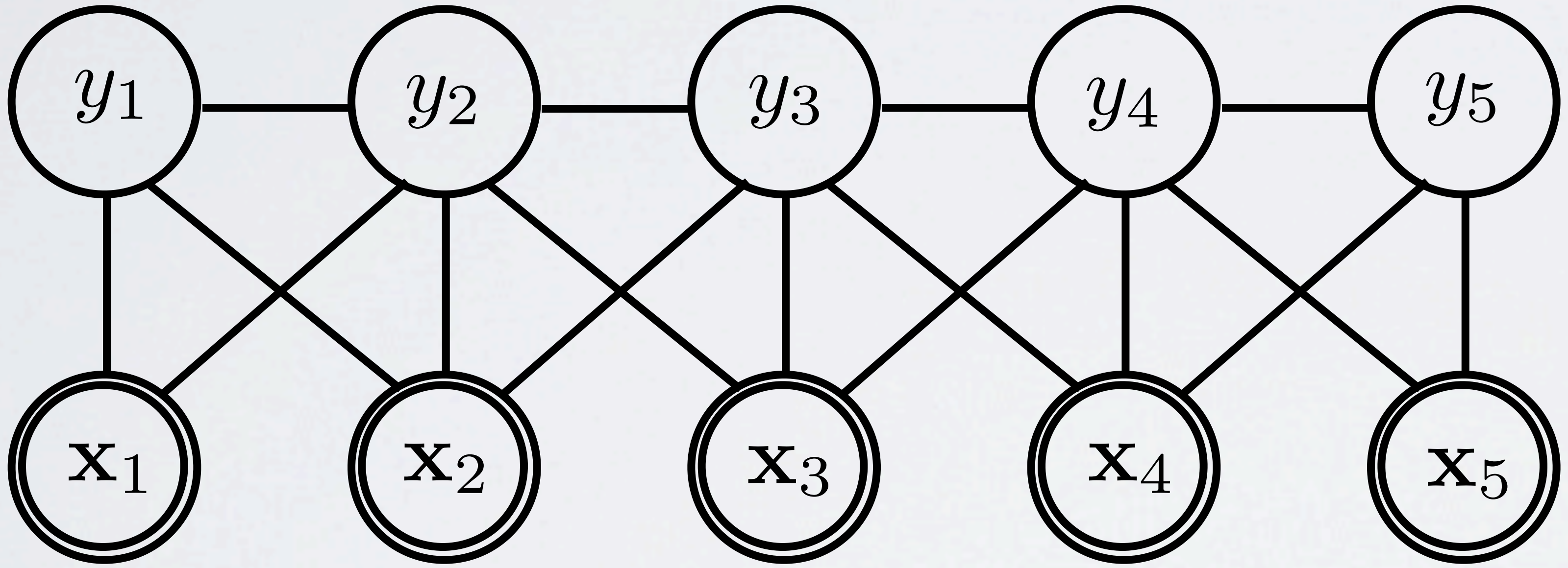
- Network with an edge between each node (random variable) that shares a factor

MARKOV NETWORK VISUALIZATION

Topics: Markov network



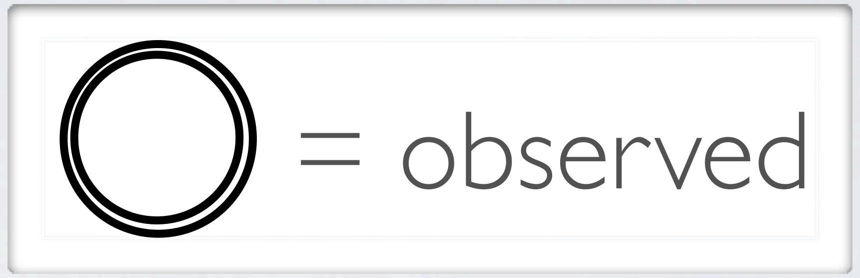
- Illustration for $K=5$



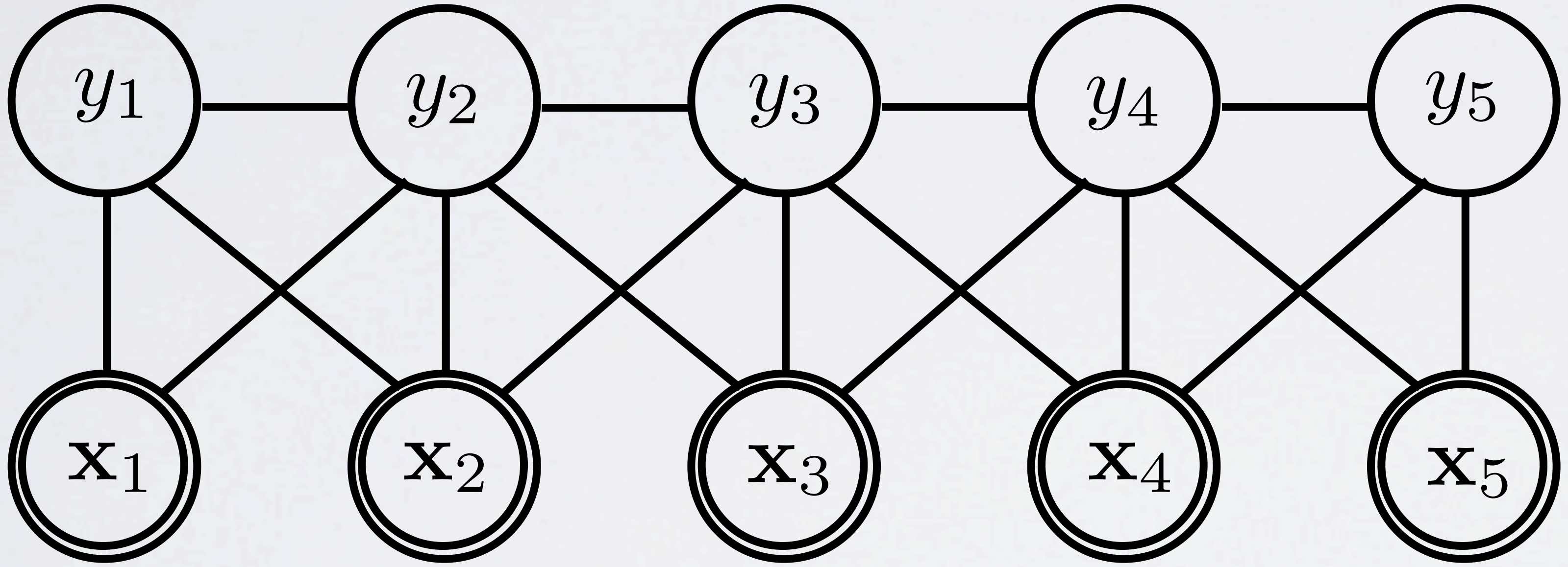
- Local Markov property
  - each node is independent of other nodes **given** its neighbors

MARKOV NETWORK VISUALIZATION

Topics: Markov network



- Illustration for $K=5$

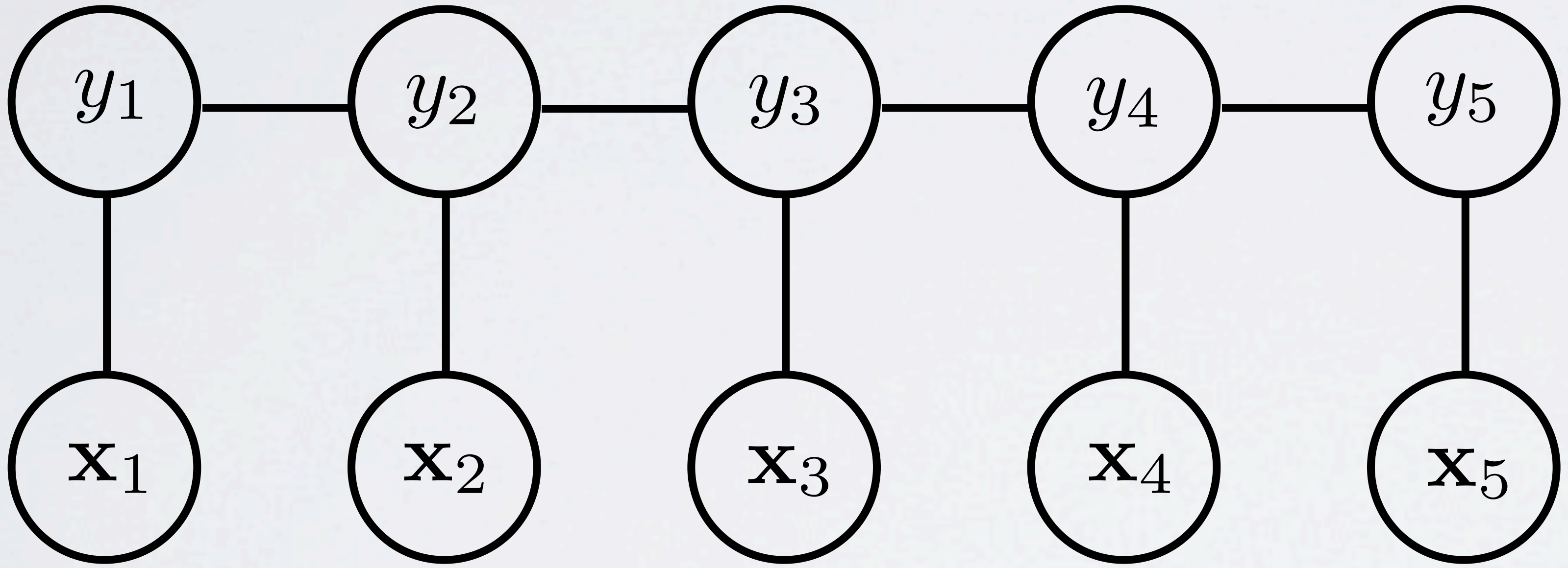


- General conditional independence
  - two nodes are conditionally independent if all paths between them contain at least one of the conditioning node

DIRECTED VS. UNDIRECTED

Topics: undirected graphical model, directed graphical model

- CRF is an undirected graphical model

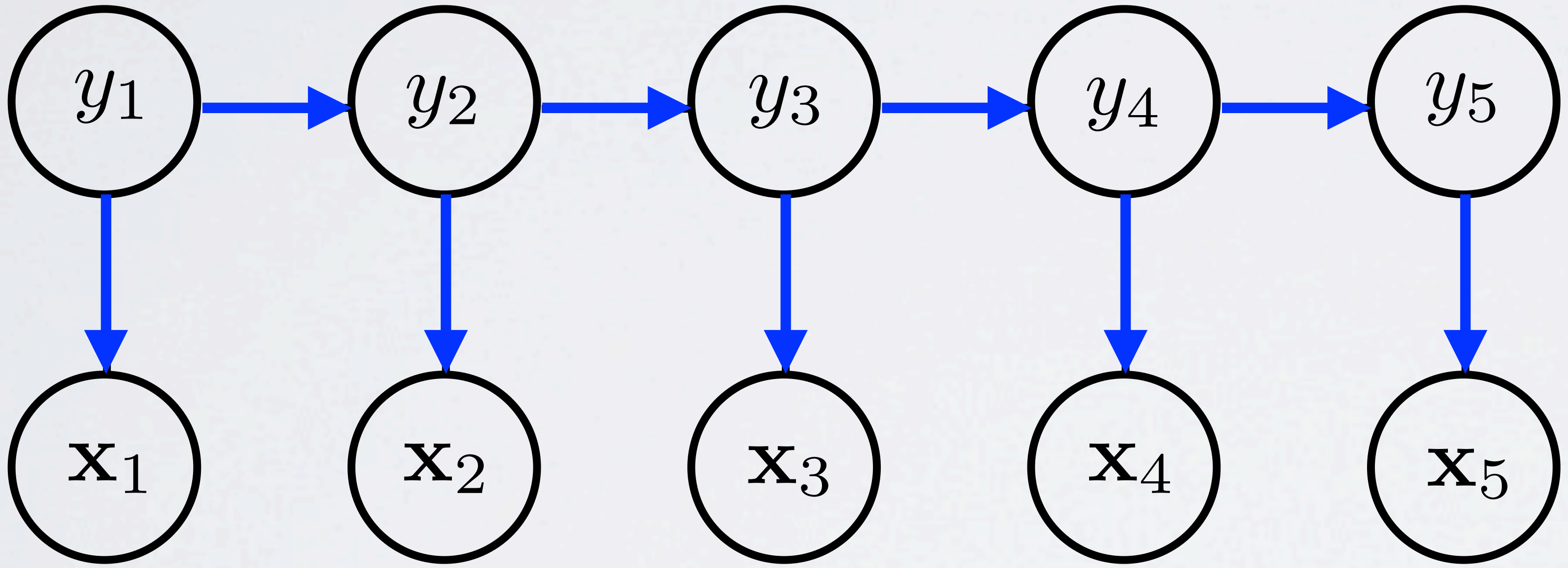


$$\left.\begin{array}{l} \phi_f(y_k, \mathbf{x}_k) \\ \\ \phi_f(y_k, y_{k+1}) \end{array}\right\} \text{only need to be non-negative}$$

DIRECTED VS. UNDIRECTED

Topics: undirected graphical model, directed graphical model

- HMM is a directed graphical model



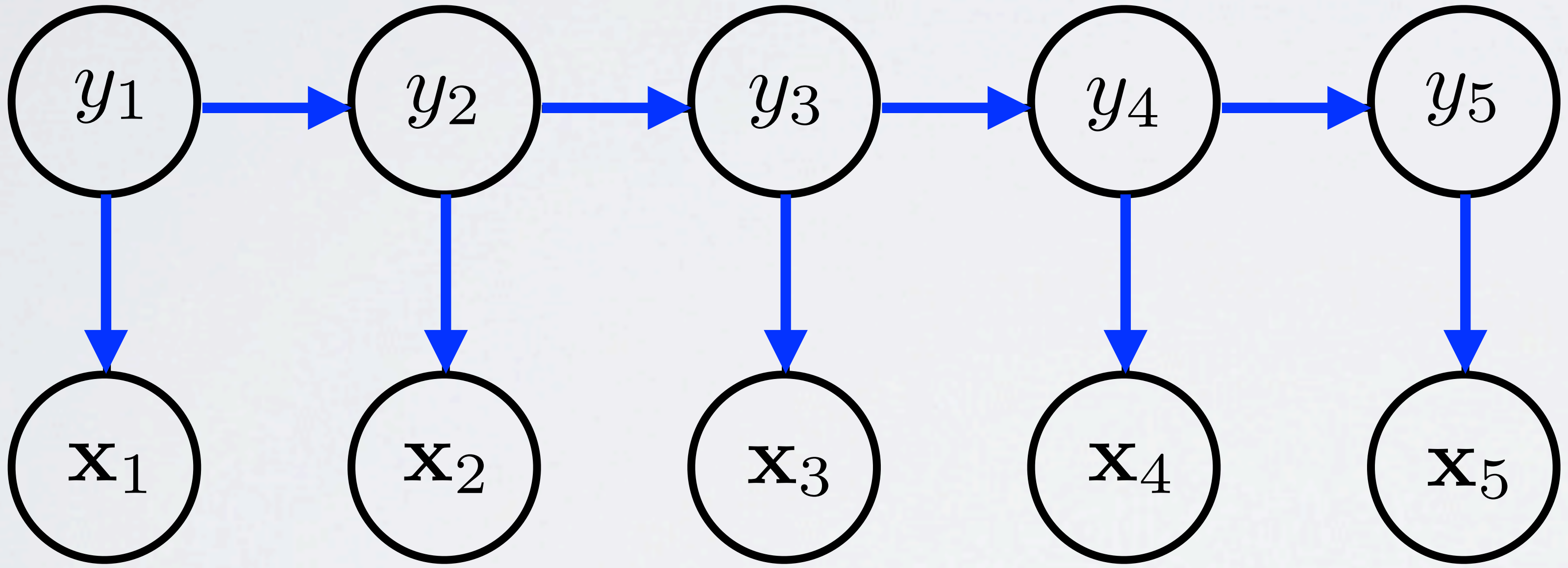
$$\phi_f(y_k, \mathbf{x}_k)$$

$$\phi_f(y_k, y_{k+1})$$

DIRECTED VS. UNDIRECTED

Topics: undirected graphical model, directed graphical model

- HMM is a directed graphical model



$$\phi_f(y_k, \mathbf{x}_k) = p(\mathbf{x}_k | y_k)$$

$$\phi_f(y_k, y_{k+1}) = p(y_{k+1} | y_k)$$