

# ECES21 - Assignment 4

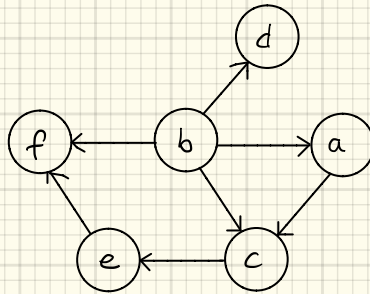
Jianwei Sun (1000009821)

Raymond Ly (999959497)

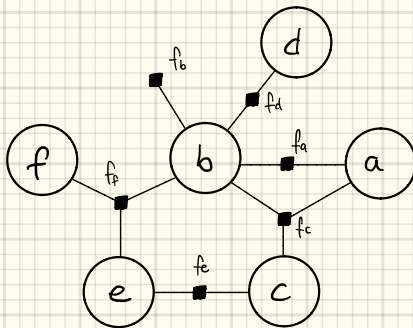
## 1. Graphical Models

### 1.1 Graphical models from factorization

$$1. P(a, b, c, d, e, f) = P(a|b) P(b) P(c|a, b) P(d|b) P(e|c) P(f|b, e)$$

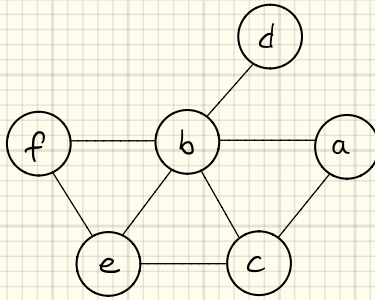


2.

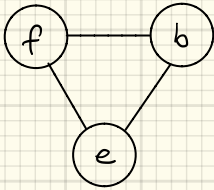


$f_a : p(a|b)$   
 $f_b : p(b)$   
 $f_c : p(c|a, b)$   
 $f_d : p(d|b)$   
 $f_e : p(e|c)$   
 $f_f : p(f|b, e)$

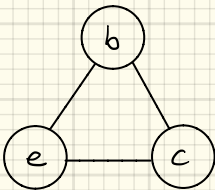
3.



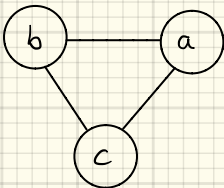
maximum cliques:



$$\psi(b, e, f) = p(f|b, e)$$



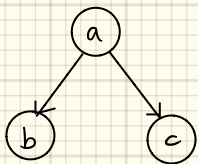
$$\psi(b, e, c) = p(e|c)p(b)$$



$$\psi(a, b, c) = p(a|b)p(c|a, b)$$

## 1.2 Conversion between graphical models

1.1 a.



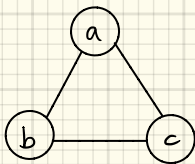
$$p(b|a)p(c|a)p(a)$$

$$f_1(a,b,c) = p(b|a)p(c|a)$$

$$f_2(a) = p(a)$$

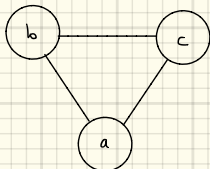
- b. b and c are independent because they each have a "dongle" of a factor node. The factor  $f_5$  must be either  $p(b|c)$  or  $p(c|b)$  because it only connects nodes b and c, which would imply that b and c are not independent. These two statements contradict; hence, an equivalent BN does not exist. In the BN, b and c both have to be parent nodes. There is no placement of the directed edge  $f_5$  that would allow both b and c to be parent nodes.

1.2 a.



$$\psi(a,b,c) = f_1(a,b,c) f_2(a)$$

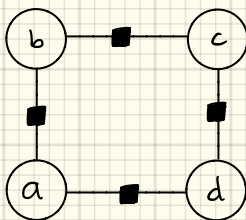
b.



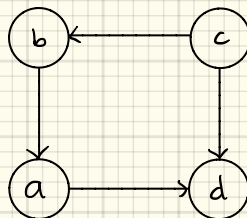
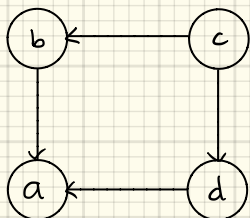
$$\psi(a,b,c) = f_1(b) f_2(c) f_3(a,b) f_4(a,c) f_5(b,c)$$

2.

1.



2. It is not possible to sketch the equivalent BN because the diagonal elements are conditionally independent (using the Markov Blanket), but there is no way to draw the conditional independence with a BN. There would always be at least one dependency. For instance, consider the possible cases (graph must be acyclic):



$c \perp\!\!\!\perp a | b, d$  but  $b \not\perp\!\!\!\perp d | a, c$      $b \perp\!\!\!\perp d | a, c$  but  $a \not\perp\!\!\!\perp c | b, d$

### 1.3 Conditional independence in Bayesian Networks

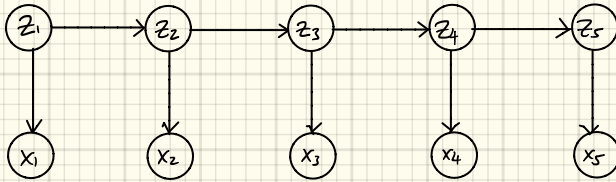
1.  $P(a, b, c, d, e, f) = p(d | b, c) p(f | b) p(c | b, e) p(e) p(b | a) p(a)$

- 2.
- |                            |   |
|----------------------------|---|
| $a \perp\!\!\!\perp c$     | F - cascade, no condition   |
| $a \perp\!\!\!\perp c   b$ | T - cascade, conditional  |
| $e \perp\!\!\!\perp b$     | T - V-structure, no condition   |
| $e \perp\!\!\!\perp b   c$ | F - V-structure, conditional  |
| $a \perp\!\!\!\perp e$     | T - e and b are independent based on V-structure. a is only connected to b, so e and a must be independent                      |
| $a \perp\!\!\!\perp e   c$ | F - e and b are dependent given c, based on the V-structure. b is dependent on a, so e and a are not conditionally independent. |

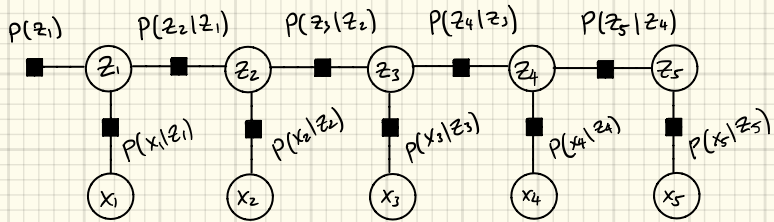
### 3. Hidden Markov Models

#### 3.1 Factor graph representation

1.



2.

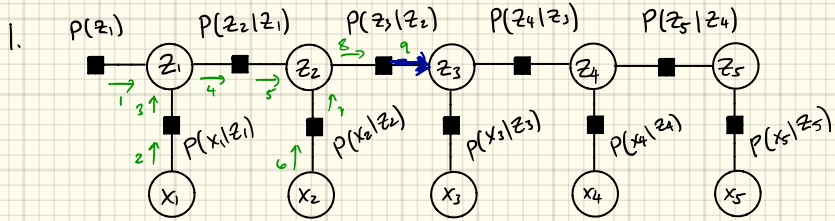


#### 3.2 Inference by passing messages

1. variable to factor

$$\mu_{z_4 \rightarrow f_{z_3 z_4}}(z_4) = \mu_{f_{x_4 z_4} \rightarrow z_4}(z_4) \mu_{f_{z_3 z_4} \rightarrow z_4}(z_4)$$

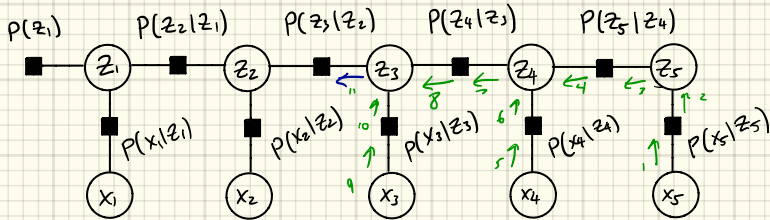
### 3.3 Message-passing as bi-directional RNNs



Let  $\otimes$  denote element-wise multiplication.

$$\mu_{z_2, z_3} \rightarrow z_3(z_3) = T \left( \left[ T \left[ \begin{matrix} T \\ 1 \end{matrix} \otimes \begin{matrix} W^T & x_1 \\ 4 & 2 \end{matrix} \right] \right] \otimes \begin{bmatrix} W^T x_2 \\ 7 \\ 6 \end{bmatrix} \right)$$

2.



$$\mu_{z_3 \rightarrow z_2}(z_3) = T \left( \left[ T \left( \begin{bmatrix} W^T x_5 \\ 8 \end{bmatrix} \otimes \begin{bmatrix} W^T x_4 \\ 7 \\ 6 \\ 5 \end{bmatrix} \right) \otimes \begin{bmatrix} W^T x_3 \\ 11 \\ 10 \\ 9 \end{bmatrix} \right] \right)$$