ECE521: Assignment 4

Inference and Learning on Graphical Models

Sunday, April 9, 2017

Rahul Chandan | 999781801

Christine Lee | 999836646

1 Graphical Models [20 pt.]

1.1 Graphical models from factorization [6 pt.]

Consider a joint distribution that factors in the following form:

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

1. *Sketch the corresponding Bayesian network (BN). [2 pt.]*

SOLUTION:

The corresponding Bayesian network (BN) can be found in **Figure 1** below.

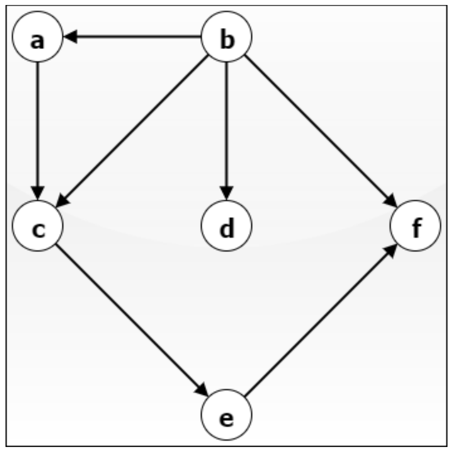


Figure 1: Bayesian Network Representation of $P(a,b,c,d,e,f) = P(a|b)P(b)P(c|a,b)P(d|b)P(e|c)P(f|b,e)$

2. *Sketch the factor graph representation and label the factors with corresponding distributions. [2 pt.]*

SOLUTION:

The corresponding factor graph representation with the labelled factors can be found in **Figure 2** below. The probability with the labels as indicated in the diagram can be expressed as follows:

$$P(a,b,c,d,e,f) = \frac{1}{Z} f_1(a,b) f_2(b) f_3(c,a,b) f_4(d,b) f_5(e,c) f_6(f,b,e)$$

where

$f_1(a,b) = P(a|b)$, $f_1(b) = P(b)$, $f_3(c,a,b) = P(c|a,b)$, $f_4(d,b) = P(d|b)$, $f_5(e,c) = P(e|c)$, and $f_6(f,b,e) = P(f|b,e)$

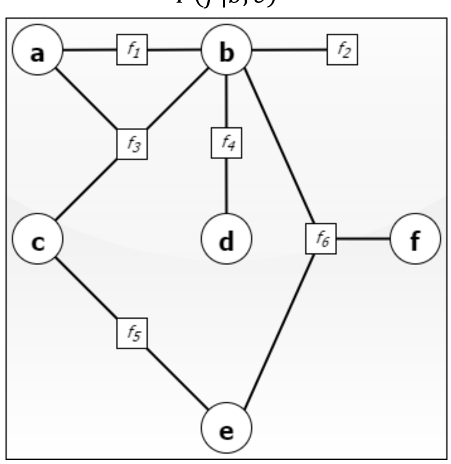


Figure 2: Factor Graph Representation of $P(a,b,c,d,e,f) = P(a|b)P(b)P(c|a,b)P(d|b)P(e|c)P(f|b,e)$

3. *Sketch the Markov random field (MRF) representation and label all the maximum cliques with corresponding distributions. [2 pt.]*

SOLUTION:

The corresponding Markov random field (MRF) representation with the labelled maximum cliques can be found in **Figure 3** below. The maximal cliques are as follows: $\{a, b, c\}$, $\{b, c, e\}$, and $\{b, e, f\}$, and their relevant distributions, respectively, are as follows: $P(c|a, b)$, $P(e|c, b)$, and $P(f|b, e)$.

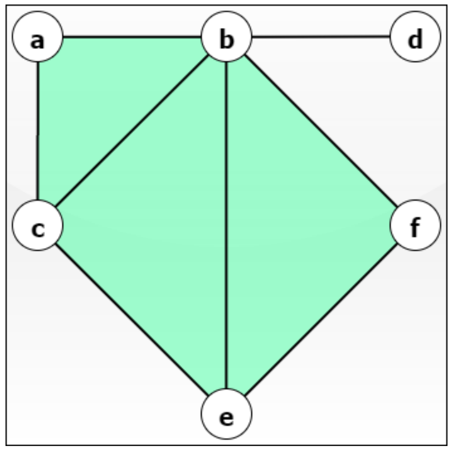
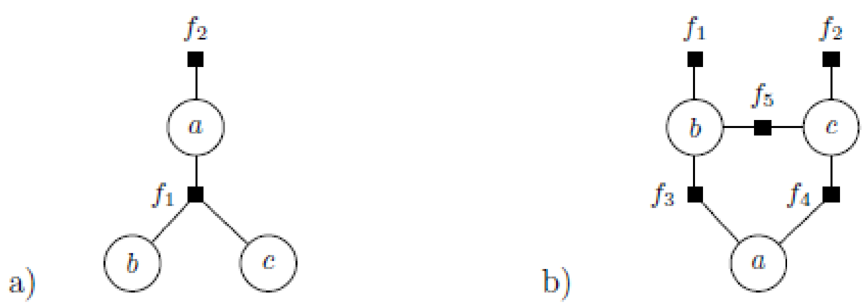


Figure 3: Markov Random Field Representation of $P(a, b, c, d, e, f) = P(a|b)P(b)P(c|a, b)P(d|b)P(e|c)P(f|b, e)$

1.2 Conversion between graphical models [10 pt.]

1.2.1 [4 pt.]



1. *For both factor graphs (a) and (b), if it exists, sketch the equivalent BNs that implies the same conditional independence properties as the factor graphs and write down the conditional probabilities using the factors, $f_1, \dots, f_5$. If it does not exist, explain why. [3 pt.]*

SOLUTION:

a) The equivalent BN for the factor graph shown in a) is shown in **Figure 4**. The joint probability can be expressed as $P(a, b, c) = P(a)P(a|b)P(c|a)$. From this, we can observe that the factors can be expressed as $f_1 = P(a|b)P(c|a)$ and $f_2 = P(a)$.

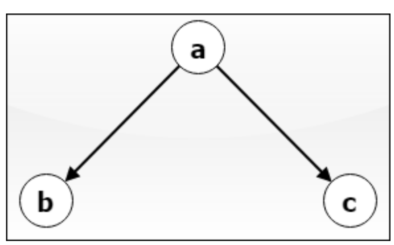


Figure 4: Equivalent BN for Factor Graph a)

b) An equivalent BN for the factor graph shown in b) does not exist. In the factor graph, the factors f_1 and f_2 imply that b and c are independent; however, there also exists a factor f_5 between b and c, which prevent us from representing such in a BN.

2. *For both factor graphs (a) and (b), if it exists, sketch the equivalent MRFs that implies the same conditional independence properties as the factor graphs and write down the maximum clique potentials using the factors, $f_1, \ldots, f_5$. If it does not exist, explain why. [3 pt.]*

SOLUTION:

a) The equivalent MRF for the factor graph shown in a) is shown in **Figure 5**. The maximum clique potential can be expressed as

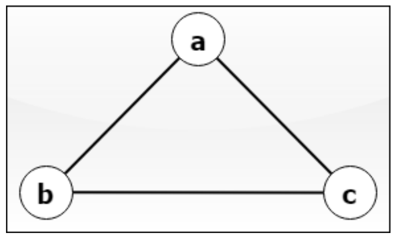
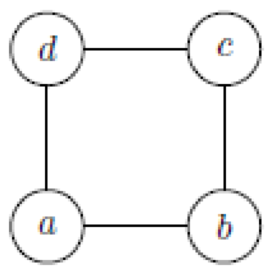


Figure 5: Equivalent MRF for Factor Graph a)

b) Similar to question 1, an equivalent MRF for the factor graph shown in b) does not exist. As explained in the answer to question 1.a) of **1.2.1**, b and c are independent in the factor graph. Therefore, there exists no equivalent MRF.

1.2.2 [4 pt.]



1. *If it exists, sketch the equivalent factor graph representation that implies the same conditional independence properties as the MRF. If it does not exist, explain why. [2 pt.]*

SOLUTION:

The equivalent factor graph for the MRF shown in the question is shown in **Figure 6**.

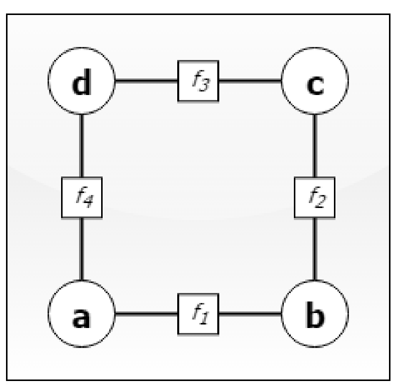


Figure 6: Equivalent Factor Graph for MRF

2. *If it exists, sketch the equivalent BN that implies the same conditional independence properties as the MRF. If it does not exist, explain why. [2 pt.]*

SOLUTION:

An equivalent BN does not exist for the MRF shown in the question. From the given MRF in the question, there exist the following relationships: $a \perp\!\!\!\perp c|(d, b)$ *and* $d \perp\!\!\!\perp b|a$, which cannot be expressed through a BN.

1.3 Conditional Independence in Bayesian Networks [4 pt.]

1. *Express the joint probability $P(a, b, c, d, e, f)$ in a factorized form corresponding to the BN shown. [1 pt.]*

SOLUTION:

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

2. *Determine whether each of the followings statements is TRUE or FALSE and provide your explanations: $a \perp\!\!\!\perp c$? $a \perp\!\!\!\perp c|b$? $e \perp\!\!\!\perp b$? $e \perp\!\!\!\perp b|c$? $a \perp\!\!\!\perp e$? $a \perp\!\!\!\perp e|c$? [3 pt.]*

SOLUTION:

$\boldsymbol{a \perp\!\!\!\perp c}$ → False.

Explanation: As shown in **Figure 7**, there exists active paths from a to c; therefore, a and c are not conditionally independent.

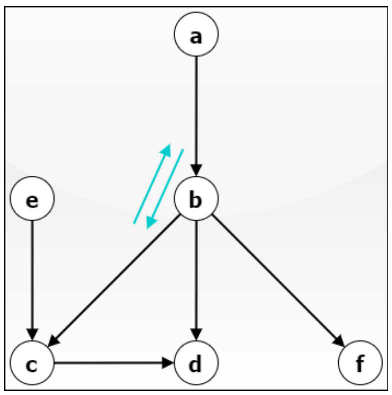


Figure 7: Visual Explanation for $a \perp\!\!\!\perp c$ → False

$\boldsymbol{a \perp\!\!\!\perp c|b}$ → True.

Explanation: As shown in **Figure 8**, when b is shaded, there are no active paths from a to c; therefore, a and c are conditionally independent given b.

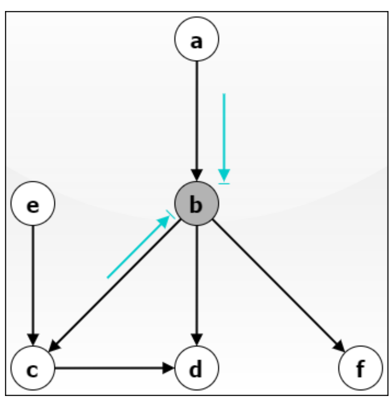


Figure 8: Visual Explanation for $a \perp\!\!\!\perp c|b$ → True

$\boldsymbol{e} \perp\!\!\!\perp \boldsymbol{b}$ → True.

Explanation: As shown in **Figure 9**, there are no active paths from e to b; therefore, a and c are conditionally independent.

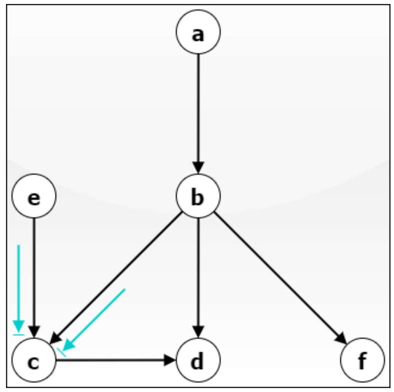


Figure 9: Visual Explanation for $e \perp\!\!\!\perp b$ → True

$\boldsymbol{e} \perp\!\!\!\perp \boldsymbol{b}|\boldsymbol{c}$ → False.

Explanation: As shown in **Figure 10**, when c is shaded, there exists an active path from e to b; therefore, e and b are not conditionally independent given c.

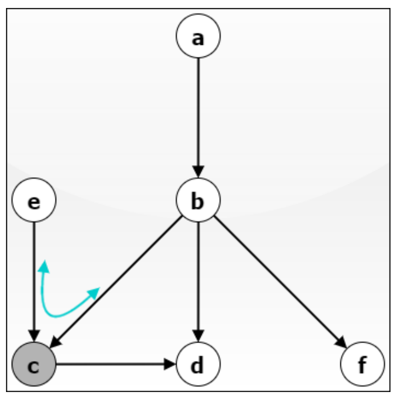


Figure 10: Visual Explanation for $e \perp\!\!\!\perp b|c$ → False

$\boldsymbol{a} \perp\!\!\!\perp \boldsymbol{e}$ → True.

Explanation: As shown in **Figure 11**, there are no active paths from a to e; therefore, a and e are conditionally independent.

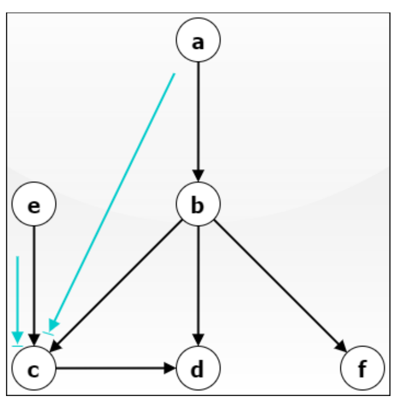


Figure 11: Visual Explanation for $a \perp\!\!\!\perp e$ → True

$\boldsymbol{a} \perp\!\!\!\perp \boldsymbol{e}|\boldsymbol{c}$ → False.

Explanation: As shown in **Figure 12**, when c is shaded, there exists an active path from a to e; therefore, a and e are not conditionally independent given c.

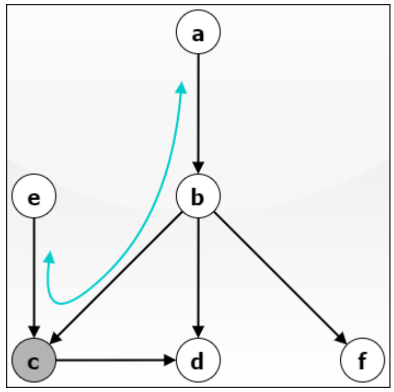


Figure 12: Visual Explanation of $a \perp\!\!\!\perp e|c$ → False

3 Mixtures of Gaussians [14 pt.]

3.1 Factor graph representation [2 pt.]

1. *Sketch a BN representing a HMM of over five observed variables in a sequence $\{x_1, x_2, x_3, x_4, x_5\}$. Label the latent state variables $\{z_1, z_2, z_3, z_4, z_5\}$. [1 pt.]*

SOLUTION:

A BN representing a HMM of over five observed variables in a sequence $\{x_1, x_2, x_3, x_4, x_5\}$ with the labelled latent state variables $\{z_1, z_2, z_3, z_4, z_5\}$ can be found in **Figure 13** below.

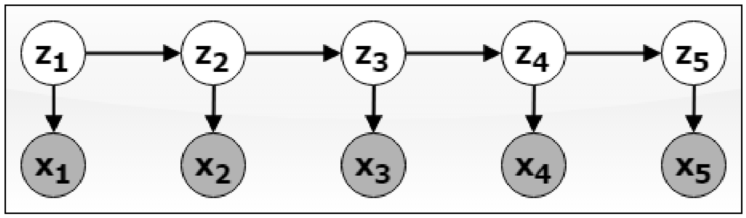


Figure 13: BN for HMM of over 5 observed variables with z_1 independent

2. *Sketch the factor graph representation for the BN above. Annotate each of the factors using either the prior $P(z_1)$, the transition probabilities $P(z_t|z_{t-1})$ and the likelihoods $P(x_t|z_t)$. [1 pt.]*

SOLUTION:

A factor graph representing a HMM of over five observed variables with the labelled latent state variables using the prior, transition probabilities, and the likelihoods can be found in below.

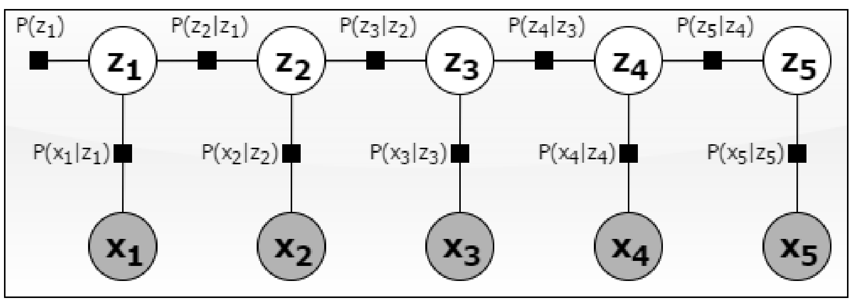


Figure 14: Factor Graph for HMM of over 5 observed variables with z_1 independent

3.2 Inference by passing messages [2 pt.]

1. *Write down the message-passing rule to compute the message from the variable node z_4 to the factor node $f_{z_3 z_4}$, that is $\mu_{z_4 \to f_{z_3 z_4}}(z_4)$, in terms of the other local messages at z_4: $\mu_{z_4 \to f_{z_4 z_5}}(z_4)$, $\mu_{f_{z_4 z_5} \to z_4}(z_4)$, $\mu_{z_4 \to f_{x_4 z_4}}(z_4)$, $\mu_{f_{z_4 z_4} \to z_4}(z_4)$, $\mu_{f_{z_3 z_4} \to z_4}(z_4)$. [2 pt.]*

SOLUTION:

The message from the variable node z_4 to the factor node $f_{z_3 z_4}$ can be expressed as $\mu_{z_4 \to f_{z_3 z_4}}(z_4)$ and is a variable-to-factor message. Therefore, it is governed by the variable-to-factor message equation as follows:

$$\mu_{z_4 \to f_{z_3 z_4}}(z_4) = \prod_{f_n \epsilon Ne(z_4) \backslash f_{z_3 z_4}} \mu_{f_n \to z_4}(z_4)$$

Therefore, the message-passing rule to compute $\mu_{z_4 \to f_{z_3 z_4}}(z_4)$ can be expressed as:

$$\mu_{z_4 \to f_{z_3 z_4}}(z_4) = \left(\mu_{f_{x_4 z_4} \to z_4}(z_4)\right)\left(\mu_{f_{z_4 z_5} \to z_4}(z_4)\right)$$

3.3 Message-passing as bi-direction RNNs [10 pt.]

1. *We continue the example from Section 3.1. Assume the observed variables x_t are one-hot vectors. Write down the expression to compute a vectorized message $\mu_{f_{z_2 z_3 \to z_3}} = \left[\mu_{f_{z_2 z_3 \to z_3}}(z_3 = 1), \ldots, \mu_{f_{z_2 z_3 \to z_3}}(z_3 = K)\right]^T$ in terms of x_1, x_2, W, T, π. [5 pt.]*

SOLUTION:

To computer the vectorized message $\mu_{f_{z_2 z_3 \to z_3}}$ in terms of x_1, x_2, W, T, π, we can begin with the abbreviated version of the factor graph from **3.1**; the abbreviated version can be found in **Figure 15** below.

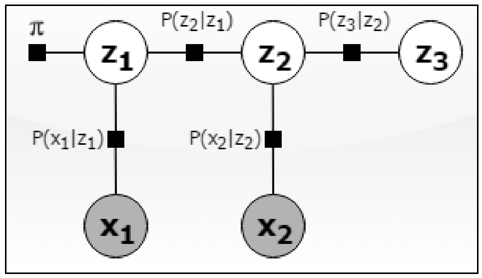


Figure 15: Abbreviated Version of Factor Graph shown in **Figure 14**

By iteratively expanding subsequent messaged using the sum-product algorithm, we can compute the vectorized message $\mu_{f_{z_2 z_3 \to z_3}}$ as follows:

$$
\begin{aligned}
\mu_{f_{z_2 z_3 \to z_3}} &= f_{z_2 z_3} * \mu_{z_2 \to f_{z_2 z_3}} \\
&= f_{z_2 z_3} * \left[\left(\mu_{f_{x_2 z_2} \to z_2}\right) \times \left(\mu_{f_{z_1 z_2} \to z_2}\right)\right] \\
&= f_{z_2 z_3} * \left[\left(f_{x_2 z_2} * \mu_{x_2 \to f_{x_2 z_2}}\right) \times \left(f_{z_1 z_2} * \mu_{z_1 \to f_{z_1 z_2}}\right)\right] \\
&= f_{z_2 z_3} * \left[(f_{x_2 z_2}) \times \left(\mu_{x_2 \to f_{x_2 z_2}}\right)\right] * \left[(f_{z_1 z_2}) * (\mu_{\pi \to z_1}) \times \left(\mu_{f_{x_1 z_1} \to z_1}\right)\right] \\
&= f_{z_2 z_3} * \left[\left\{f_{x_2 z_2} * \left(\mu_{x_2 \to f_{x_2 z_2}}\right)\right\} \times \left\{f_{z_1 z_2} * (\mu_{\pi \to z_1}) \times \left(f_{x_1 z_1} * \mu_{x_1 \to f_{x_1 z_1}}\right)\right\}\right]
\end{aligned}
$$

In this expression, matrix multiplication is indicated by the $*$ symbol, while element-wise multiplication is indicated by the $\times$ symbol. In order to express the computation in a vectorized form, we can make use of the fact that multiplications by factors are the same as matrix multiplication by the appropriate factor matrix (W, T). Assuming that all inputs, (π, x_i) are column vectors, we can compute the vectorized form as follows:

$$\mu_{f_{z_2 z_3 \to z_3}} = T\{(W^t x_1) \times [T * (\pi \times W^T x_2)]\}$$

2. *Write down the expression for $\mu_{z_3 \to f_{z_2 z_3}}$ in terms of x_3, x_4, x_5, W, T, π. [5 pt.]*

SOLUTION:

To compute the vectorized message $\mu_{z_3 \to f_{z_2 z_3}}$ in terms of x_3, x_4, x_5, W, T, π, we can begin with the abbreviated version of the factor graph from **3.1**; the abbreviated version can be found in **Figure 16** below.

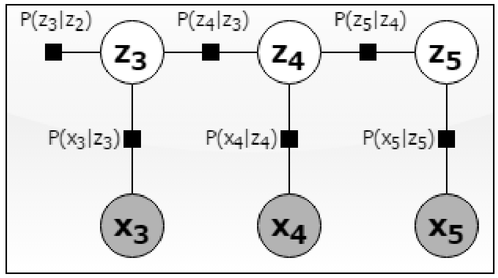


Figure 16: Abbreviated Version of Factor Graph shown in **Figure 14**

Applying the same principles and carrying out the same method as the previous question, we can compute the message as follows:

$$
\begin{aligned}
\mu_{f_{z_3 \to f_{z_2 z_3}}} &= \left(\mu_{f_{z_3 z_4} \to z_3}\right) \times \left(\mu_{f_{z_3 x_3} \to z_3}\right) \\
&= \left\{f_{z_3 z_4} * \left(\mu_{z_4 \to f_{z_3 z_4}}\right)\right\} \times \left\{f_{z_3 x_3} * \left(\mu_{x_3 \to f_{z_3 x_3}}\right)\right\} \\
&= \left\{f_{z_3 z_4} * \left[\left(\mu_{f_{z_4 z_5} \to z_4}\right) \times \left(\mu_{f_{x_4 z_4} \to z_4}\right)\right]\right\} * \left\{f_{z_3 x_3} * \left(\mu_{x_3 \to f_{z_3 x_3}}\right)\right\} \\
&= \left[f_{z_3 z_4} * \left\{f_{z_4 z_5} * \left(\mu_{z_5 \to f_{z_4 z_5}}\right)\right\} \times \left\{f_{x_4 z_4} * \left(\mu_{x_4 \to f_{x_4 z_4}}\right)\right\}\right] \times \left\{f_{z_3 x_3} * \left(\mu_{x_3 \to f_{z_3 x_3}}\right)\right\} \\
&= \left[f_{z_3 z_4} * \left\{f_{z_4 z_5} * \left(f_{x_5 z_5} * \left(\mu_{x_5 \to f_{x_5 z_5}}\right)\right)\right\} \times \left\{f_{x_4 z_4} * \left(\mu_{x_4 \to f_{x_4 z_4}}\right)\right\}\right] \times \left\{f_{z_3 x_3} * \left(\mu_{x_3 \to f_{z_3 x_3}}\right)\right\}
\end{aligned}
$$

In this expression, matrix multiplication is indicated by the $*$ symbol, while element-wise multiplication is indicated by the $\times$ symbol. In order to express the computation in a vectorized form, we can make use of the fact that multiplications by factors are the same as matrix multiplication by the appropriate factor matrix (W, T). Assuming that all inputs, (π, x_i) are column vectors, we can compute the vectorized form as follows:

$$\mu_{z_3 \to f_{z_2 z_3}} = \{T * (TW^T x_5 \times W^T x_4)\} \times W^T x_3$$