

Homework 2

Requirement

- The part marked optional is not compulsory; do it for your interest.
- Compile everything (including your derivations, answers, and source codes) to a pdf file. In the main text, provide your name, student ID, and your email address (marked assignment may be sent to your email). Name your pdf file in the form of “StudentID_name_2nd_homework.pdf”.
- Deadline: 23:59, 31 Oct 2025.
- To be submitted individually.
- Submission method will be announced later.

Exercises

Exercise 1:

(Based on Exercise 3.15 of BRML(Barber, 2020))

A belief network models the relation between the variables *oil*, *inf*, *eh*, *bp*, *rt* which stand for the price of oil, inflation rate, economy health, British Petroleum Stock price, retailer stock price. Each variable takes the states “low”, “high”, except for *bp* which has states “low”, “high”, “normal”. The belief network model for these variables has tables

$p(eh = \text{low}) = 0.2$	
$p(bp = \text{low} \mid oil = \text{low}) = 0.9$	$p(bp = \text{normal} \mid oil = \text{low}) = 0.1$
$p(bp = \text{low} \mid oil = \text{high}) = 0.1$	$p(bp = \text{normal} \mid oil = \text{high}) = 0.4$
$p(oil = \text{low} \mid eh = \text{low}) = 0.9$	$p(oil = \text{low} \mid eh = \text{high}) = 0.05$
$p(rt = \text{low} \mid inf = \text{low}, eh = \text{low}) = 0.9$	$p(rt = \text{low} \mid inf = \text{low}, eh = \text{high}) = 0.1$
$p(rt = \text{low} \mid inf = \text{high}, eh = \text{low}) = 0.1$	$p(rt = \text{low} \mid inf = \text{high}, eh = \text{high}) = 0.01$
$p(inf = \text{low} \mid oil = \text{low}, eh = \text{low}) = 0.9$	$p(inf = \text{low} \mid oil = \text{low}, eh = \text{high}) = 0.1$
$p(inf = \text{low} \mid oil = \text{high}, eh = \text{low}) = 0.1$	$p(inf = \text{low} \mid oil = \text{high}, eh = \text{high}) = 0.01$

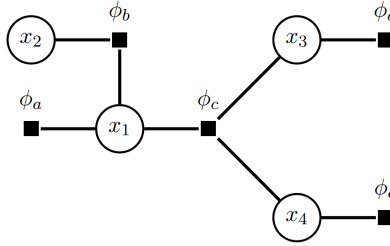
Questions:

1. Write down the joint distribution of all the 5 variables.
2. Draw a belief network for this distribution.
3. Are oil price and retailer stock price independent? How many pathways are there between them? Under what conditions (observations) will they become independent?

4. Without any observation, what is the probability that the oil price is “low”, and the probability that inflation is “low” (aka the marginal probabilities $p(oil = \text{low}), p(inf = \text{low})$)?
5. Convert the belief network into a factor graph, and write down the potential function of each factor constituting the joint distribution.

Exercise 2:

The following factor graph represents a probability distribution $p(\mathbf{x}) = \frac{1}{Z} \prod_{m \in \{a,b,c,d,e\}} \phi_m(\mathbf{x}_m)$ over four binary variables $x_i \in \{0, 1\}$:



The factors $\phi_a, \phi_b, \phi_d, \phi_e$ are defined as follows:

x_1	ϕ_a	x_1	x_2	ϕ_b	x_3	ϕ_d	x_4	ϕ_e
0	2	0	0	5	0	1	0	1
1	1	0	1	2	1	2	1	3
		1	1	6				

and $\phi_c(x_1, x_3, x_4) = 1$ if $x_1 = x_3 = x_4$, and is zero otherwise. Intuitively, factor ϕ_a favors variable x_1 to be in the state $x_1 = 0$ (as $\phi_a(x_1 = 0) > \phi_a(x_1 = 1)$), ϕ_d favors $x_3 = 1$, ϕ_c favors x_1, x_3, x_4 to have the same value, etc.

Though the problem is of small size and exact inference is not difficult, we aim to apply the belief propagation (sum-product message-passing here) algorithm for solving it to improve our understanding.

Questions:

1. Initialize the messages from leaf nodes. How should we set the values of the following messages?
 - $\mu_{x_2 \rightarrow \phi_b}(x_2)$, for $x_2 = 0$ and $x_2 = 1$,
 - $\tilde{\mu}_{\phi_a \rightarrow x_1}(x_1)$, for $x_1 = 0$ and $x_1 = 1$,
 - $\tilde{\mu}_{\phi_d \rightarrow x_3}(x_3)$, for $x_3 = 0$ and $x_3 = 1$,
 - $\tilde{\mu}_{\phi_e \rightarrow x_4}(x_4)$, for $x_4 = 0$ and $x_4 = 1$.
2. Choose x_1 as the root node (this is arbitrary), pass the messages one step inward by computing the values of
 - $\tilde{\mu}_{\phi_b \rightarrow x_1}(x_1)$, for $x_1 = 0$ and $x_1 = 1$,
 - $\mu_{x_3 \rightarrow \phi_c}(x_3)$, for $x_3 = 0$ and $x_3 = 1$,
 - $\mu_{x_4 \rightarrow \phi_c}(x_4)$, for $x_4 = 0$ and $x_4 = 1$.

3. Computing the values of $\tilde{\mu}_{\phi_c \rightarrow x_1}(x_1)$, for $x_1 = 0$ and $x_1 = 1$.
 4. Now you have all the (factor-to-variable) messages arriving at x_1 , what is the marginal probability of $p(x_1 = 0)$? (Do not forget that $p(x_1)$ is normalized.)
 5. To compute the marginal probabilities of x_2, x_3, x_4 , we need to compute the messages outgoing from the root node x_1 . For example, compute $\mu_{x_1 \rightarrow \phi_c}(x_1)$ and then $\tilde{\mu}_{\phi_c \rightarrow x_3}(x_3)$. Based on the messages arriving at x_3 , compute the marginal probability of $p(x_3 = 0)$.
 6. (Optional) Can you use a similar message-passing procedure to compute the conditional probability $p(x_1 | x_2 = 0)$?
- Hint: $p(x_1 | x_2 = 0) = \frac{p(x_1, x_2=0)}{p(x_2=0)} \propto p(x_1, x_2 = 0) = \sum_{x_3, x_4} p(x_1, x_2 = 0, x_3, x_4)$. Apply belief propagation to the graphical model $\tilde{p}(x_1, x_3, x_4) := p(x_1, x_2 = 0, x_3, x_4)$, and normalize the obtained “marginal” $\tilde{p}(x_1)$ (this is the probability obtained by fixing $x_2 = 0$).

Exercise 3:

(By Prof. Jason Eisner)

Imagine that you are a climatologist studying the history of global warming. You cannot find any records of the weather in Baltimore for the summer of 2007, but you do find Jason Eisner’s diary, which lists how many ice creams Jason ate every day that summer.

day #	1	2	3	4	5	6	7	8	9	10	11
ice cream	2	3	3	2	3	2	3	2	2	3	1
day #	12	13	14	15	16	17	18	19	20	21	22
ice cream	3	3	1	1	1	2	1	1	1	3	1
day #	23	24	25	26	27	28	29	30	31	32	33
ice cream	2	1	1	1	2	3	3	2	3	2	2

Your goal is to use these observations to estimate the temperature every day. You will simplify this weather task by assuming there are only two kinds of days: cold (C) and hot (H).

Given a sequence of observations $x_{1:N}$ (here $N = 33$), each observation $x_i \in \mathbb{N}$ corresponds to the number of ice creams eaten on a given day, the task is to infer the correct “hidden” weather states $z_i \in \{C, H\}$ which caused Jason to eat ice creams. Here assume that the weather condition is the only factor influencing Jason to eat ice creams, and that the weather conditions of consecutive days form a Markov chain. Therefore, you have a hidden Markov model (HMM)

$$p(x_{1:N}, z_{1:N}) = p(z_1) \prod_{n=2}^N p(z_n | z_{n-1}) \prod_{n=1}^N p(x_n | z_n).$$

Based on Jason’s dairy and Baltimore’s general summer weather conditions, you have an initial guess of the parameters:

$p(x_n z_n)$	$z_n = C$	$z_n = H$
$x_n = 1$	0.7	0.1
$x_n = 2$	0.2	0.2
$x_n = 3$	0.1	0.7

$p(z_n z_{n-1})$	$z_{n-1} = C$	$z_{n-1} = H$
$z_n = C$	0.8	0.2
$z_n = H$	0.2	0.8

Further assume that $p(z_1 = H) = 0.5$.

Questions:

1. Write a computer program to solve the inference task of smoothing, i.e., to compute the univariate posterior $p(z_n \mid x_{1:N})$, $\forall n$.
(Hint: use the forward-backward algorithm shown in the lecture slide.)
2. Plot x_n and $p(z_n \mid x_{1:N})$ versus time n , and interpret the results.
3. (Optional) Write a computer program to solve the inference task of decoding, i.e., to compute $\arg \max_{z_{1:N}} p(z_{1:N} \mid x_{1:N})$.
(Hint: use the Viterbi algorithm shown in the lecture slide.)