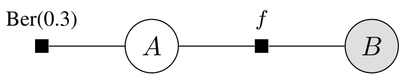
Exercises 5

Undirected graphical models

5.1 Factor-graph basics

Exercise 5.1 Bayes' theorem
Consider the factor graph in Figure 5.1



$f(A,B)$	A	$\bar{A}$
B	0.9	0.2
$\bar{B}$	0.1	0.8

Figure 5.1: Factor graph in Exercise 5.1

Note that $f(A,B)$ can be represented using indicator functions as

$$f(a,b) = 0.9\,\mathbf{1}_A(a)\,\mathbf{1}_B(b) + 0.2\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_B(b) + 0.1\,\mathbf{1}_A(a)\,\mathbf{1}_{\bar{B}}(b) + 0.8\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_{\bar{B}}(b),$$

where $\mathbf{1}(E)$ be the indicator function of event E.

a) Compute the distribution of A using message passing.

b) Let $p(B|A) = f(A,B)$ denote the conditional distribution of B given A, compute $p(A|B)$ using Bayes' theorem and compare with the result in **(a)**.

Exercise 5.2 Convolution of normals
Consider the factor f in the graph in Figure 5.2.

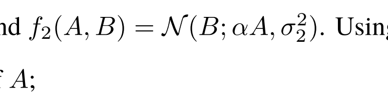


Figure 5.2: Factor graph in Exercise 5.2

Suppose that $f_1(A) = \mathcal{N}(A; \mu_1, \sigma_1^2)$ and $f_2(A,B) = \mathcal{N}(B; \alpha A, \sigma_2^2)$. Using message passing:

a) compute the marginal distribution of A;

b) compute the marginal distribution of B.

Exercise 5.3 Professor Bayes' mid-term exam
Professor Bayes wants to check how well his students are following his course. To this end, he devised a three-question

mid-term exam to verify if the students have learned *Bayesian networks* (BN) and *Factor graphs* (FG). Only BN is needed to answer question 1, and only FG is needed to answer question 2; to answer question 3, both BN and FG are needed (*Hint: represent this condition as a random variable*). In all cases, Professor Bayes estimates that a student with the right knowledge has a 90% probability of answering correctly, and that a student without has 20% probability of answering correctly.

Suppose that a student gives a correct answer to question 1 and fails questions 2 and 3; in addition, suppose that Professor Bayes uses a non-informative prior.

a) Identify the conditional distributions in the model and draw the Bayesian Network.

b) Transform the Bayesian Network into a factor graph.

c) Using message passing, compute the probability that the student is knowledgeable in BN.

d) Using message passing, compute the probability that the student is knowledgeable in FG.

e) Write a program to verify the probabilities using Monte Carlo simulation.

Exercise 5.4 Paperclips inc.

The company *Paperclips inc.* produces clips and pins using two different machines. Each day, the machines produce random amounts of clips and pins with, each with a *Poisson distibution* $\mathrm{Po}(\lambda)$, where the rate λ depends on the quality of the steel the company is using on that specific day. The Poisson distribution is given by:

$$x \sim \mathrm{Po}(\lambda) \quad \Longrightarrow \quad p(x) = \frac{\lambda^x \mathrm{e}^{-\lambda}}{x!}.$$

If the steel is of high quality, $\lambda = 10$, if it is of low quality $\lambda = 7$. The company has a one in four chance of receiving high quality steel on any specific day.

Suppose that at the end of the day the company has produced 10 clips and 8 pins.

a) Identify the conditional distributions in the model and draw the Bayesian Network.

b) Transform the Bayesian Network into a factor graph.

c) Using message passing, compute the probability that the company was using high-quality steel.

d) Write a program to verify the calculated probability using Monte Carlo simulation.

Solutions 5

Undirected graphical models

Solution to Exercise 5.1 **a)**

- The node B is observed, so the first message passed is $\mu_{B\to f}(b) = \mathbf{1}_B(b)$.
- The factor f receives this message and computes

$$\begin{aligned}\mu_{f\to A}(a) &= \sum_b f(a,b)\mu_{B\to f}(b)\\ &= \sum_b \Big(0.9\,\mathbf{1}_A(a)\,\mathbf{1}_B(b) + 0.2\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_B(b) + 0.1\,\mathbf{1}_A(a)\,\mathbf{1}_{\bar{B}}(b) + 0.8\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_{\bar{B}}(b)\Big)\;\mathbf{1}_B(b)\\ &= 0.9\,\mathbf{1}_A(a) + 0.2\,\mathbf{1}_{\bar{A}}(a).\end{aligned}$$

- The factor Ber(0.3) is a leaf factor, so it sends a message to the node A given by

$$\mu_{\text{Ber}(0.3)\to A}(a) = 0.3\,\mathbf{1}_A(a) + 0.7\,\mathbf{1}_{\bar{A}}(a).$$

- The node A collects the messages $\mu_{\text{Ber}(0.3)\to A}$ and $\mu_{f\to A}$ and computes the local marginal

$$\begin{aligned}\mu_A(a) &= \mu_{\text{Ber}(0.3)\to A}(a)\cdot\mu_{f\to A}(a)\\ &= \Big(0.9\,\mathbf{1}_A(a) + 0.2\,\mathbf{1}_{\bar{A}}(a)\Big)\,\Big(0.3\,\mathbf{1}_A(a) + 0.7\,\mathbf{1}_{\bar{A}}(a)\Big)\\ &= 0.27\,\mathbf{1}_A(a) + 0.14\,\mathbf{1}_{\bar{A}}(a)\end{aligned}$$

- We normalize the local marginal with

$$Z_A = \sum_a \mu_A(a) = \mu_A(A) + \mu_A(\bar{A}) = 0.27 + 0.14 = 0.41$$

  and we get that $p(a) = \mu_A(a)/Z_A \approx 0.659\,\mathbf{1}_A(a) + 0.341\,\mathbf{1}_{\bar{A}}(a)$. So the graphical model describes a probability of A of approximately 0.659.

b) From the model, we have that $p(A) = 0.3$, and $p(B|A) = f(A,B)$. Using $f(A,B)$, we can compute the marginal $p(B)$ according to

$$p(B) = p(B|A)p(A) + p(B|\bar{A})p(\bar{A}) = 0.9\cdot 0.3 + 0.2\cdot 0.7 = 0.41.$$

Hence

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} = \frac{0.9\cdot 0.3}{0.41} \approx 0.659,$$

which is the probability found in **(a)**.

Solution to Exercise 5.2 **a)** The leaf node B sends the message $\mu_{B\to f_2} = 1$ to the factor f_2.

This factor computes and passes along the message

$$\begin{aligned}\mu_{f_2\to A}(a) &= \int f_2(a,b)\mu_{B\to f_2}(b)\,\mathrm{d}b \\ &= \int \mathcal{N}(b;\alpha a,\sigma_2^2)\,\mathrm{d}b = 1.\end{aligned}$$

Similarly, the leaf node f_1 sends the message $\mu_{f_1\to A} = \mathcal{N}(\mu_1,\sigma_1^2)$.
Then, the marginal distribution of A can be computed with

$$p_A(a) = \frac{1}{Z_A}\mu_{f_1\to A}(a)\cdot\mu_{f_2\to A}(a) = \mathcal{N}(a;\mu_1,\sigma_1^2);$$

so, in the graphical model, $p(A) = \mathcal{N}(\mu_1,\sigma_1^2)$.

b) The leaf factor f_1 sends the message to the variable node A, which passes it along to the factor f_2 according to

$$\mu_{A\to f_2}(a) = \mathcal{N}(a;\mu_1,\sigma_1^2).$$

Finally, the factor f_2 takes this message and passes it along the graph according to

$$\begin{aligned}\mu_{f_2\to B}(b) &= \int f(a,b)\mu_{A\to f_2}(a)\,\mathrm{d}a \\ &= \int \mathcal{N}(b;\alpha a,\sigma_2^2)\mathcal{N}(a;\mu_1,\sigma_1^2)\,\mathrm{d}a \\ &\propto \mathcal{N}(b;\alpha\mu_1,\alpha^2\sigma_1^2+\sigma_2^2)\end{aligned}$$

where we have used Corollary 2 from Lecture 2. Alternatively, by explicit computation

$$\begin{aligned}\mu_{f_2\to B}(b) &= \int \mathcal{N}(b;\alpha a,\sigma_2^2)\mathcal{N}(a;\mu_1,\sigma_1^2)\,\mathrm{d}a \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\int \exp\left\{-\frac{1}{2}\frac{(b-\alpha a)^2}{\sigma_2^2}\right\}\exp\left\{-\frac{1}{2}\frac{(a-\mu_1)^2}{\sigma_1^2}\right\}\mathrm{d}a \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\int \exp\left\{-\frac{1}{2}\left(\frac{b^2-2\alpha ab+\alpha^2a^2}{\sigma_2^2}+\frac{a^2+\mu_1^2-2a\mu_1}{\sigma_1^2}\right)\right\}\mathrm{d}a \\ &= \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left(\frac{b^2}{\sigma_2^2}+\frac{\mu_1^2}{\sigma_1^2}\right)\right\}\int \exp\left\{-\frac{1}{2}\left(\frac{\alpha^2}{\sigma_2^2}+\frac{1}{\sigma_1^2}\right)a^2+\left(\frac{\alpha b}{\sigma_2^2}+\frac{\mu_1}{\sigma_1^2}\right)a\right\}\mathrm{d}a;\end{aligned}$$

completing the square inside the integral, we get

$$\int \exp\left\{-\frac{1}{2}\left(\frac{\alpha^2}{\sigma_2^2}+\frac{1}{\sigma_1^2}\right)a^2+\left(\frac{\alpha b}{\sigma_2^2}+\frac{\mu_1}{\sigma_1^2}\right)a\right\}\mathrm{d}a \propto \exp\left\{\frac{1}{2}\left(\frac{\sigma_1^2\sigma_2^2}{\alpha^2\sigma_1^2+\sigma_2^2}\right)\left(\frac{\alpha b}{\sigma_2^2}+\frac{\mu_1}{\sigma_1^2}\right)^2\right\}$$

Hence,

$$\begin{aligned}\mu_{f_2\to B}(b) &\propto \exp\left\{-\frac{1}{2}\frac{b^2}{\sigma_2^2}\right\}\exp\left\{\frac{1}{2}\left(\frac{\sigma_1^2\sigma_2^2}{\alpha^2\sigma_1^2+\sigma_2^2}\right)\left(\frac{\alpha b}{\sigma_2^2}+\frac{\mu_1}{\sigma_1^2}\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2}\frac{b^2}{\sigma_2^2}+\frac{1}{2}\left(\frac{\sigma_1^2\sigma_2^2}{\alpha^2\sigma_1^2+\sigma_2^2}\right)\left(\frac{\alpha b\sigma_1^2+\mu_1\sigma_2^2}{\sigma_1^2\sigma_2^2}\right)^2\right\} \\ &= \exp\left\{-\frac{1}{2}\frac{b^2}{\sigma_2^2}+\frac{1}{2}\left(\frac{\alpha^2b^2\sigma_1^4+\mu_1^2\sigma_2^4+2\alpha b\sigma_1^2\mu_1\sigma_2^2}{\sigma_1^2\sigma_2^2(\alpha^2\sigma_1^2+\sigma_2^2)}\right)\right\} \\ &= \exp\left\{-\frac{1}{2}\left(\frac{b^2\sigma_1^2(\alpha^2\sigma_1^2+\sigma_2^2)-\alpha^2b^2\sigma_1^4-\mu_1^2\sigma_2^4-2\alpha b\sigma_1^2\mu_1\sigma_2^2}{\sigma_1^2\sigma_2^2(\alpha^2\sigma_1^2+\sigma_2^2)}\right)\right\} \\ &= \exp\left\{-\frac{1}{2}\left(\frac{b^2\sigma_1^2\sigma_2^2-\mu_1^2\sigma_2^4-2\alpha b\sigma_1^2\mu_1\sigma_2^2}{\sigma_1^2\sigma_2^2(\alpha^2\sigma_1^2+\sigma_2^2)}\right)\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\frac{b^2-2b\mu_1}{\alpha^2\sigma_1^2+\sigma_2^2}\right)\right\} \propto \mathcal{N}(b;\alpha\mu_1,\alpha^2\sigma_1^2+\sigma_2^2)\end{aligned}$$

Solution to Exercise 5.3
Define the following events:

$$\begin{aligned} A&: \text{the student knows Bayesian networks} \\ B&: \text{the student knows factor graphs} \\ C_i&: \text{the student answers question } i \text{ correctly} \\ D = A \wedge B&: \text{the student knows Bayesian networks and factor graphs} \end{aligned}$$

Let $\mathbf{1}_E$ be the indicator function of event E; then we can write

$$p(C_1 = q|A = a) = f_{C_1,A}(q,a) = 0.9\,\mathbf{1}_A(a)\,\mathbf{1}_{C_1}(q) + 0.2\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_{C_1}(q) + 0.1\,\mathbf{1}_A(a)\,\mathbf{1}_{\bar{C}_1}(q) + 0.8\,\mathbf{1}_{\bar{A}}(a)\,\mathbf{1}_{\bar{C}_1}(q).$$

a) The factor graph is represented in Figure 5.1 (left).

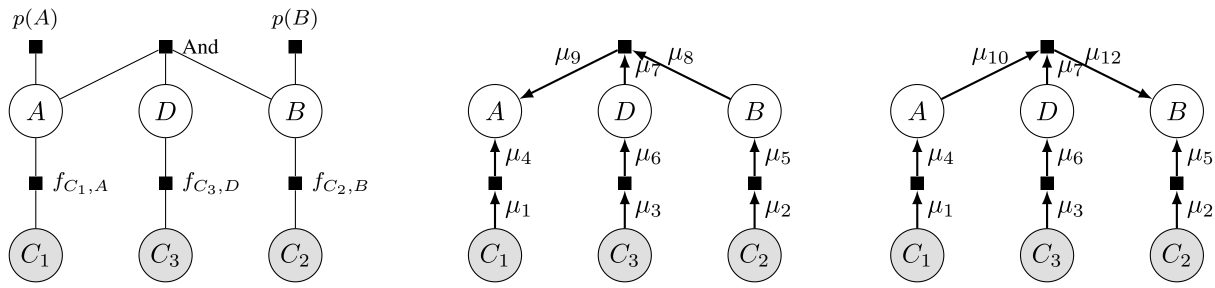


Figure 5.1: Exercise 5.3 – *Left:* Factor graph; *Center:* messages required to compute $p(A)$; *Right:* messages required to compute $p(B)$.

b) We want to compute $p(A|C_1, \bar{C}_2, \bar{C}_3)$. Consider the message-passing schedule in Figure 5.1 (center).

- Messages from observed nodes (note that the messages are functions of one variable node)

$$\mu_1 = \mathbf{1}_{C_1} \qquad \mu_2 = \mathbf{1}_{\bar{C}_2} \qquad \mu_3 = \mathbf{1}_{\bar{C}_3}$$

- Pass μ_1 through $f_{C_1,A}$

$$\begin{aligned} \mu_4 &= \sum_q \mu_1(q) f_{C_1,A}(q,\cdot) \\ &= \sum_q \mathbf{1}_{C_1}(q)\left(0.9\,\mathbf{1}_{C_1}(q)\,\mathbf{1}_A + 0.2\,\mathbf{1}_{C_1}(q)\,\mathbf{1}_{\bar{A}} + 0.1\,\mathbf{1}_{\bar{C}_1}(q)\,\mathbf{1}_A + 0.8\,\mathbf{1}_{\bar{C}_1}(q)\,\mathbf{1}_{\bar{A}}\right) \\ &= 0.9\,\mathbf{1}_A + 0.2\,\mathbf{1}_{\bar{A}}. \end{aligned}$$

- Similarly, pass μ_2 through $f_{C_2,B}$

$$\begin{aligned} \mu_5 &= \sum_q \mu_2(q) f_{C_2,B}(q,\cdot) \\ &= \sum_q \mathbf{1}_{\bar{C}_2}(q)\left(0.9\,\mathbf{1}_{C_2}(q)\,\mathbf{1}_B + 0.2\,\mathbf{1}_{C_2}(q)\,\mathbf{1}_{\bar{B}} + 0.1\,\mathbf{1}_{\bar{C}_2}(q)\,\mathbf{1}_B + 0.8\,\mathbf{1}_{\bar{C}_2}(q)\,\mathbf{1}_{\bar{B}}\right) \\ &= 0.1\,\mathbf{1}_B + 0.8\,\mathbf{1}_{\bar{B}}, \end{aligned}$$

  and μ_3 through $f_{C_3,D}$

$$\mu_6 = 0.1\,\mathbf{1}_D + 0.8\,\mathbf{1}_{\bar{D}}.$$

- At node D, we have that $\mu_7 = \mu_6$.
  At node B, the message from the leaf factor $p(B)$ is non-informative and we can put $\mu_8 = \mu_5$ (note that messages are defined up to a scalar multiple).

- To compute μ_9, define the three-argument And operator as

$$\mathbf{1}_{D=A\wedge B}(a,b,d) = \mathbf{1}_A(a)\,\mathbf{1}_B(b)\,\mathbf{1}_D(d) + \mathbf{1}_A(a)\,\mathbf{1}_{\bar B}(b)\,\mathbf{1}_{\bar D}(d) + \mathbf{1}_{\bar A}(a)\,\mathbf{1}_B(b)\,\mathbf{1}_{\bar D}(d) + \mathbf{1}_{\bar A}(a)\,\mathbf{1}_{\bar B}(b)\,\mathbf{1}_{\bar D}(d)$$

this function is one only when D is equal to $A \wedge B$. Then, we can compute

$$\begin{aligned}
\mu_9 &= \sum_{b,d} \mu_7(d)\mu_8(b)\,\mathbf{1}_{D=A\wedge B}(a,\cdot,d)\\
&= \sum_{b,d} (0.1\,\mathbf{1}_D(d) + 0.8\,\mathbf{1}_{\bar D}(d))(0.1\,\mathbf{1}_B(b) + 0.8\,\mathbf{1}_{\bar B}(b))\,\mathbf{1}_{D=A\wedge B}(\cdot,b,d)\\
&= \sum_{b,d} \big(0.01\,\mathbf{1}_B(b)\,\mathbf{1}_D(d) + 0.08\,\mathbf{1}_{\bar D}(d)\,\mathbf{1}_B(b) + 0.08\,\mathbf{1}_D(d)\,\mathbf{1}_{\bar B}(b) + 0.64\,\mathbf{1}_{\bar D}(d)\,\mathbf{1}_{\bar B}(b)\big)\,\mathbf{1}_{D=A\wedge B}(\cdot,b,d)\\
&= 0.01\,\mathbf{1}_A + 0.08\,\mathbf{1}_{\bar A} + 0.64\,\mathbf{1}_A + 0.64\,\mathbf{1}_{\bar A}\\
&= 0.65\,\mathbf{1}_A + 0.72\,\mathbf{1}_{\bar A}.
\end{aligned}$$

- At node A, we have that

$$p(A=a) \propto \mu_4(a)\mu_9(a) = \big(0.9\,\mathbf{1}_A(a) + 0.2\,\mathbf{1}_{\bar A}(a)\big)\big(0.65\,\mathbf{1}_A(a) + 0.72\,\mathbf{1}_{\bar A}(a)\big) = 0.585\,\mathbf{1}_A(a) + 0.144\,\mathbf{1}_{\bar A}(a)$$

so, according to the model,

$$p(A|C_1,\bar C_2,\bar C_3) = \frac{0.585}{0.585+0.155} \approx 0.8025.$$

c) We want to compute $p(B|C_1,\bar C_2,\bar C_3)$. Consider the message-passing schedule in Figure 5.1 (right); note that the messages μ_1 through μ_7 are the same as the ones computed in point **(b)**.

- At node A, the message from the leaf factor $p(A)$ is non-informative and we can put $\mu_{10} = \mu_4$.
- We use the three-argument And operator to compute

$$\begin{aligned}
\mu_{12} &= \sum_{a,d} \mu_{10}(a)\mu_7(d)\,\mathbf{1}_{D=A\wedge B}(a,\cdot,d)\\
&= \sum_{a,d} \big(0.9\,\mathbf{1}_A(a) + 0.2\,\mathbf{1}_{\bar A}(a)\big)\big(0.1\,\mathbf{1}_D(d) + 0.8\,\mathbf{1}_{\bar D}(d)\big)\,\mathbf{1}_{D=A\wedge B}(a,\cdot,d)\\
&= \sum_{a,d} \big(0.09\,\mathbf{1}_A(a)\,\mathbf{1}_D(d) + 0.72\,\mathbf{1}_A(a)\,\mathbf{1}_{\bar D}(d) + 0.02\,\mathbf{1}_{\bar A}(a)\,\mathbf{1}_D(d) + 0.16\,\mathbf{1}_{\bar A}(a)\,\mathbf{1}_{\bar D}(d)\big)\,\mathbf{1}_{D=A\wedge B}(a,\cdot,d)\\
&= 0.09\,\mathbf{1}_B + 0.72\,\mathbf{1}_{\bar B} + 0.16\,\mathbf{1}_B + 0.16\,\mathbf{1}_{\bar B} = 0.25\,\mathbf{1}_B + 0.88\,\mathbf{1}_{\bar B}
\end{aligned}$$

- At node B, we have that

$$p(B=b) \propto \mu_{12}(b)\mu_5(b) = (0.25\,\mathbf{1}_B + 0.88\,\mathbf{1}_{\bar B})(0.1\,\mathbf{1}_B + 0.8\,\mathbf{1}_{\bar B}) = 0.025\,\mathbf{1}_B + 0.704\,\mathbf{1}_{\bar B}$$

so, according to the model,

$$p(B|C_1,\bar C_2,\bar C_3) = \frac{0.025}{0.025+0.704} \approx 0.034.$$

d) The following listing shows a python program to compute $p(A|C_1,\bar C_2,\bar C_3)$ and $p(B|C_1,\bar C_2,\bar C_3)$.

```python
from random import random as rand
from random import seed

def drawQuestion(Skill):
    """Randomly draws the result of a question based on Skill"""
    return rand() < 0.9 if Skill else rand() < 0.2

def simulateStudent():
    """Randomly draws a student and the results of the questions"""
```

```
    A = rand() < 0.5
    B = rand() < 0.5
    C1 = drawQuestion(A)
    C2 = drawQuestion(B)
    C3 = drawQuestion(A and B)

    return A,B,C1,C2,C3

if __name__ == "__main__":
    seed(123)
    N_total, N_A, N_B = 0, 0, 0

    while N_total < 200000:
        A,B,C1,C2,C3 = simulateStudent()
        if C1 and not C2 and not C3:
            N_total += 1
            if A: N_A += 1
            if B: N_B += 1

    print(f"P(A|C1, ~C2, ~C3) = {N_A/N_total}") # Output: P(A|C1, ~C2, ~C3) = 0.802515
    print(f"P(B|C1, ~C2, ~C3) = {N_B/N_total}") # Output: P(B|C1, ~C2, ~C3) = 0.034115
```

Solution to Exercise 5.4

Define the following random variables:

S: the company was processing high-quality steel

λ: the production rate of the day

C: the number of clips produced

P: the number of pins produced

a) The factor graph is represented in Figure 5.2 (left).

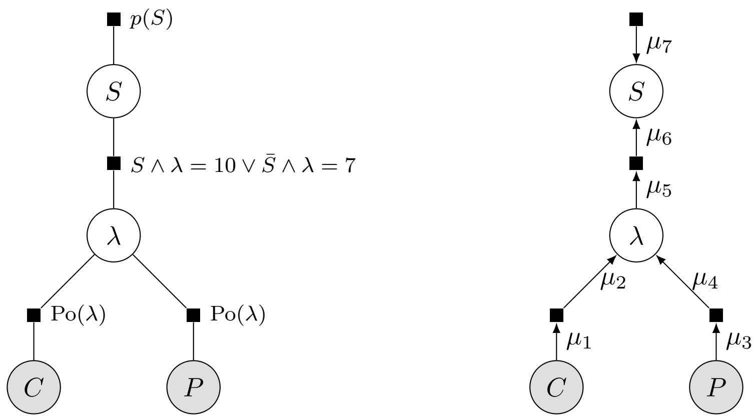


Figure 5.2: Exercise 5.4 – *Left:* Factor graph; *Right:* messages required to compute $p(S)$.

b) We want to compute $p(S|C = 10, P = 8)$. Consider the message-passing schedule in Figure 5.2 (right).

- At node C, we have the observation $C = 10$, so μ_1 is a point mass at 10:

$$\mu_1 = \mathbf{1}_{10}$$

- The message μ_1 passes through the $\mathrm{Po}(\lambda)$ factor to give μ_2 according to

$$\mu_2(\lambda) = \sum_{k=0}^{\infty} \mathrm{Po}(k; \lambda)\mu_1(k) = \sum_{k=0}^{\infty} \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!} \mathbf{1}_{10}(k) = \frac{\lambda^{10} \mathrm{e}^{-\lambda}}{10!}$$

- Using the same argument, we have that $\mu_3 = \mathbf{1}_8$ and $\mu_4(\lambda) = \lambda^8 \mathrm{e}^{-\lambda}/8!$.

- The node λ passes along the message

$$\mu_5(\lambda) = \mu_3(\lambda) \cdot \mu_4(\lambda) = \frac{\lambda^{18} \mathrm{e}^{-2\lambda}}{10!8!}$$

- The message μ_5 arrives at the factor node. This factor describes one point mass at $S \wedge \lambda = 10$ and one point mass at $\bar{S} \wedge \lambda = 7$:

$$f(s, \lambda) = \mathbf{1}_S(s)\delta(\lambda - 10) + \mathbf{1}_{\bar{S}}(s)\delta(\lambda - 7),$$

where $\delta(\cdot)$ is the Dirac measure.
Hence, we have that

$$\begin{aligned}
\mu_6(s) &= \int_0^\infty f(s, \lambda)\mu_5(\lambda)\,\mathrm{d}\lambda \\
&= \int_0^\infty \big(\mathbf{1}_S(s)\delta(\lambda - 10) + \mathbf{1}_{\bar{S}}(s)\delta(\lambda - 7)\big)\,\mu_5(\lambda)\,\mathrm{d}\lambda \\
&= \frac{10^{18}\mathrm{e}^{-20}}{10!8!}\,\mathbf{1}_S(s) + \frac{7^{18}\mathrm{e}^{-14}}{10!8!}\,\mathbf{1}_{\bar{S}}(s)
\end{aligned}$$

- From the prior factor, we have the message

$$\mu_7 = 0.25\,\mathbf{1}_S + 0.75\,\mathbf{1}_{\bar{S}}$$

- At node S, we can compute

$$\begin{aligned}
p(S = s) &\propto \mu_6(s) \cdot \mu_7(s) \\
&= \left(\frac{10^{18}\mathrm{e}^{-20}}{10!8!}\,\mathbf{1}_S(s) + \frac{7^{18}\mathrm{e}^{-14}}{10!8!}\,\mathbf{1}_{\bar{S}}(s)\right)\big(0.25\,\mathbf{1}_S(s) + 0.75\,\mathbf{1}_{\bar{S}}(s)\big) \\
&= \frac{10^{18}\mathrm{e}^{-20}}{4 \cdot 10!8!}\,\mathbf{1}_S(s) + \frac{3 \cdot 7^{18}\mathrm{e}^{-14}}{4 \cdot 10!8!}\,\mathbf{1}_{\bar{S}}(s)
\end{aligned}$$

so, normalizing, we get that

$$p(S|C = 10, P = 8) = \frac{\frac{10^{18}\mathrm{e}^{-20}}{4 \cdot 10!8!}}{\frac{10^{18}\mathrm{e}^{-20}}{4 \cdot 10!8!} + \frac{3 \cdot 7^{18}\mathrm{e}^{-14}}{4 \cdot 10!8!}} = \frac{1}{1 + 3 \cdot \left(\frac{7}{10}\right)^{18}\mathrm{e}^6} \approx 0.3366.$$

- The following listing shows a python program to compute $p(S|C = 10, P = 8)$.

```python
from numpy.random import poisson
from numpy.random import random as rand
from numpy.random import seed

def simulateDay():
    """Simulate one day at the factory"""
    S = rand() < 0.25
    L = 10 if S else 7
    P = poisson(L)
    C = poisson(L)

    return S,L,P,C

if __name__ == "__main__":
    seed(123)
    N_total, N_S = 0, 0

    while N_total < 20000:
        S,L,P,C = simulateDay()
        if P==8 and C==10:
            N_total += 1
            if S: N_S += 1

    print(f"P(S|P=8,C=10) = {N_S/N_total}") # Output: P(S|P=8,C=10) = 0.33685
```

Bibliography

[1] David Barber. *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.

[2] Christopher M Bishop. *Pattern recognition and machine learning*. springer, 2006.

[3] Kevin B Korb and Ann E Nicholson. *Bayesian artificial intelligence*. CRC press, 2010.

[4] Steffen L Lauritzen and David J Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(2):157–194, 1988.

[5] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.