

# Homework 7

PB17000297 罗晏宸

April 19 2020

## 1 EXERCISE 14.12

两个来自世界上不同地方的宇航员同时用他们自己的望远镜观测了太空中某个小区内恒星的数目  $N$ 。他们的测量结果分别为  $M_1$  和  $M_2$ 。通常，测量中会有不超过 1 颗恒星的误差，发生错误的概率  $e$  很小。每台望远镜可能出现（出现的概率  $f$  更小一些）对焦不准确的情况（分别记作  $F_1$  和  $F_2$ ），在这种情况下科学家会少数三颗甚至更多的恒星（或者说，当  $N$  小于 3 时，连一颗恒星都观测不到）。考虑图1所示的三种贝叶斯网络结构。

- a. 这三种网络结构哪些是对上述信息的正确（但不一定高效）表示？
- b. 哪一种网络结构是最好的？请解释。
- c. 当  $N \in \{1, 2, 3\}$ ， $M_1 \in \{0, 1, 2, 3, 4\}$  时，请写出  $\mathbf{P}(M_1|N)$  的条件概率表。概率分布表里的每个条目都应该表达为参数  $e$  和/或  $f$  的一个函数。

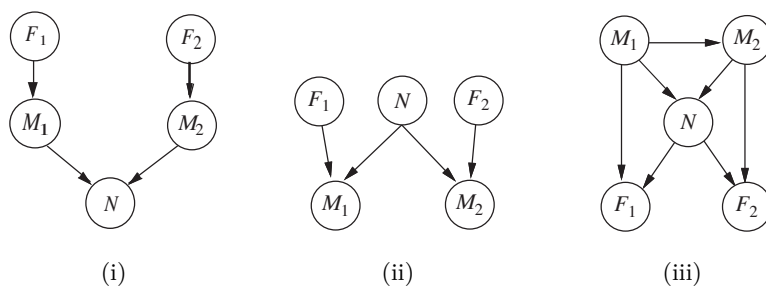


图 1: 望远镜问题的三种可能网络

数。

解

a. 网络 (ii) 和 (iii) 都是正确的表示。虽然 (i) 描述了计算过程，但没有准确描述恒星数目  $N$  和对焦情况的关系，因此不是一个贝叶斯网络结构；(ii) 对因果结构的表示是正确的，两台望远镜的对焦情况彼此独立，测量结果受实际恒星数量和对焦情况的影响；(iii) 也是正确的，其是对结点排序之后的结果。

b. (ii) 比 (iii) 需要更少的参数，因此网络 (ii) 是最好的。

c. 对概率展开有

$$\begin{aligned}\mathbf{P}(M_1|N) &= \mathbf{P}(M_1|N, F_1)\mathbf{P}(F_1|N) + \mathbf{P}(M_1|N, \neg F_1)\mathbf{P}(\neg F_1|N) \\ &= \mathbf{P}(M_1|N, F_1)\mathbf{P}(F_1) + \mathbf{P}(M_1|N, \neg F_1)\mathbf{P}(\neg F_1) \\ &= \mathbf{P}(M_1|N, F_1)f + \mathbf{P}(M_1|N, \neg F_1)(1-f)\end{aligned}$$

假设测量时发生的多数 1 颗恒星和少数 1 颗恒星的概率均为  $\frac{e}{2}$  计算得到条件概率表如下：

$\mathbf{P}(M_1 N) \backslash N$	1	2	3
$M_1$			
0	$f + \frac{e}{2}(1-f)$	$f$	$f$
1	$(1-e)(1-f)$	$\frac{e}{2}(1-f)$	0
2	$\frac{e}{2}(1-f)$	$(1-e)(1-f)$	$\frac{e}{2}(1-f)$
3	0	$\frac{e}{2}(1-f)$	$(1-e)(1-f)$
4	0	0	$\frac{e}{2}(1-f)$

表 1: 当  $N \in \{1, 2, 3\}$ ,  $M_1 \in \{0, 1, 2, 3, 4\}$  时  $\mathbf{P}(M_1|N)$  的条件概率表

```

function ENUMERATION-ASK( $X, \mathbf{e}, bn$ ) returns a distribution over  $X$ 
  inputs:  $X$ , the query variable
            $\mathbf{e}$ , observed values for variables  $\mathbf{E}$ 
            $bn$ , a Bayes net with variables  $\{X\} \cup \mathbf{E} \cup \mathbf{Y}$  /*  $\mathbf{Y} = \text{hidden variables}$  */

   $Q(X) \leftarrow$  a distribution over  $X$ , initially empty
  for each value  $x_i$  of  $X$  do
     $Q(x_i) \leftarrow$  ENUMERATE-ALL( $bn.VARS, \mathbf{e}_{x_i}$ )
    where  $\mathbf{e}_{x_i}$  is  $\mathbf{e}$  extended with  $X = x_i$ 
  return NORMALIZE( $Q(X)$ )

```

---

```

function ENUMERATE-ALL( $vars, \mathbf{e}$ ) returns a real number
  if EMPTY?( $vars$ ) then return 1.0
   $Y \leftarrow$  FIRST( $vars$ )
  if  $Y$  has value  $y$  in  $\mathbf{e}$ 
    then return  $P(y | \text{parents}(Y)) \times$  ENUMERATE-ALL(REST( $vars$ ),  $\mathbf{e}$ )
  else return  $\sum_y P(y | \text{parents}(Y)) \times$  ENUMERATE-ALL(REST( $vars$ ),  $\mathbf{e}_y$ )
  where  $\mathbf{e}_y$  is  $\mathbf{e}$  extended with  $Y = y$ 

```

图 2: 在贝叶斯网络上回答查询的枚举算法

## 2 EXERCISE 14.13

考虑图1(ii) 的网络, 假设两个望远镜完全相同。 $N \in \{1, 2, 3\}, M_1, M_2 \in \{0, 1, 2, 3, 4\}$ , CPT 表和习题 14.12 所描述的一样。使用枚举算法 (图2) 计算概率分布  $\mathbf{P}(N|M_1 = 2, M_2 = 2)$ 。

**解** 使用枚举算法计算的概率分布如下

$$\begin{aligned}
& \mathbf{P}(N|M_1 = 2, M_2 = 2) \\
&= \alpha \sum_{f_1, f_2} \mathbf{P}(f_1, f_2, N, M_1 = 2, M_2 = 2) \\
&= \alpha \sum_{f_1, f_2} \mathbf{P}(f_1) \mathbf{P}(f_2) \mathbf{P}(M_1 = 2|f_1, N) \mathbf{P}(M_2 = 2|f_2, N) \\
&= \alpha (1-f)(1-f) \langle p_1, p_2, p_3 \rangle \left\langle \frac{e}{2}, 1-e, \frac{e}{2} \right\rangle \left\langle \frac{e}{2}, 1-e, \frac{e}{2} \right\rangle \\
&= \alpha' \left\langle p_1 \frac{e^2}{4}, p_2 (1-e)^2, p_3 \frac{e^2}{4} \right\rangle
\end{aligned}$$

```

function ELIMINATION-ASK( $X, \mathbf{e}, bn$ ) returns a distribution over  $X$ 
  inputs:  $X$ , the query variable
            $\mathbf{e}$ , observed values for variables  $\mathbf{E}$ 
            $bn$ , a Bayesian network specifying joint distribution  $\mathbf{P}(X_1, \dots, X_n)$ 

   $factors \leftarrow []$ 
  for each  $var$  in ORDER( $bn.VARS$ ) do
     $factors \leftarrow [MAKE-FACTOR(var, \mathbf{e}) | factors]$ 
    if  $var$  is a hidden variable then  $factors \leftarrow SUM-OUT(var, factors)$ 
  return NORMALIZE(POINTWISE-PRODUCT( $factors$ ))

```

图 3: 用于贝叶斯网络推理的变量消元算法

### 3 EXERCISE 14.15

考虑图3中的变量消元算法。

a 14.4 节对如下查询应用了变量消元算法

$$\mathbf{P}(\text{Burglary} \mid \text{JohnCalls} = \text{true}, \text{MaryCalls} = \text{true})$$

执行必要的计算，并检验计算结果的正确性。

b 统计所执行的算术运算的次数，将其与枚举算法所需的运算次数进行比较。

c 假设贝叶斯网络结构具有链式结构，即由一个布尔随机变量序列  $X_1, \dots, X_n$  构成，其中  $\text{Parents}(X_i) = \{X_{i-1}\}$ ,  $i = 2, \dots, n$ 。请问使用枚举算法计算  $\mathbf{P}(X_1 | X_n = \text{true})$  的复杂度是多少？使用变量消元算法呢？

解

**a** 计算表达式

$$\begin{aligned}
& \mathbf{P}(\text{Burglary} \mid \text{JohnCalls} = \text{true}, \text{MaryCalls} = \text{true}) \\
&= \mathbf{P}(b \mid j, m) \\
&= \alpha \mathbf{P}(b) \sum_e P(e) \sum_a \mathbf{P}(a \mid b, e) P(j \mid a) P(m \mid a) \\
&= \alpha \mathbf{P}(b) \sum_e P(e) \left[ \begin{pmatrix} P(a \mid b, e) & P(a \mid \neg b, e) \\ P(a \mid b, \neg e) & P(a \mid \neg b, \neg e) \end{pmatrix} P(j \mid a) P(m \mid a) \right. \\
&\quad \left. + \begin{pmatrix} P(\neg a \mid b, e) & P(\neg a \mid \neg b, e) \\ P(\neg a \mid b, \neg e) & P(\neg a \mid \neg b, \neg e) \end{pmatrix} P(j \mid \neg a) P(m \mid \neg a) \right] \\
&= \alpha \mathbf{P}(b) \sum_e P(e) \left[ \begin{pmatrix} 0.95 & 0.29 \\ 0.94 & 0.001 \end{pmatrix} \times 0.90 \times 0.70 \right. \\
&\quad \left. + \begin{pmatrix} 0.05 & 0.71 \\ 0.06 & 0.999 \end{pmatrix} \times 0.05 \times 0.01 \right] \\
&= \alpha \mathbf{P}(b) \sum_e P(e) \begin{pmatrix} 0.598525 & 0.183055 \\ 0.59223 & 0.0011295 \end{pmatrix} \\
&= \alpha \mathbf{P}(b) \left[ \begin{pmatrix} 0.598525 \\ 0.183055 \end{pmatrix} \times 0.002 + \begin{pmatrix} 0.59223 \\ 0.0011295 \end{pmatrix} \times 0.998 \right] \\
&= \alpha \mathbf{P}(b) \begin{pmatrix} 0.59224259 \\ 0.001493351 \end{pmatrix} \\
&= \alpha \langle 0.001 \times 0.59224259, 0.999 \times 0.001493351 \rangle \\
&= \alpha \langle 0.00059224259, 0.001491857649 \rangle \\
&\approx \langle 0.284, 0.716 \rangle
\end{aligned}$$

结果是正确的

**b** 在上述的算术运算中，共有 7 次加法，16 次乘法和 2 次除法。枚举算法有 7 次加法，18 次乘法和 2 次除法，相比之下需要 2 次额外的乘法。

**c** 使用枚举算法计算  $\mathbf{P}(X_1|X_n = true)$ , 对于随机变量  $X_1$  的两种取值, 分别需要计算两棵深度均为  $n - 2$  的二叉树, 因此复杂度是  $O(2^n)$  的。

使用变量消元算法计算  $\mathbf{P}(X_1|X_n = true)$ , 可以归纳地得到规模  $n - 1, n - 2, \dots, 1$  的子问题, 且递推是常数时间的, 因此总的复杂度是  $O(n)$  的。