中国科学技术大学
University of Science and Technology of China

人工智能基础习题课 第1/7次作业

助教 冯子洋

第一次作业

- 3.7 给出下列问题的初始状态、目标测试、后继函数和耗散函数。选择精确得足以实现的形式化。
- a. 只用四种颜色对平面地图染色,要求每两个相邻的地区不能染成相同的颜色。

- 初始状态：无染色地区
- 目标测试：所有地区均已着色，且没有两块相邻的地区染有同样的颜色
- 后继函数：给一个未染色地区染色
- 耗散函数：染色次数

第一次作业

- 3.7 给出下列问题的初始状态、目标测试、后继函数和耗散函数。选择精确得足以实现的形式化。
- b. 一间屋子里有一只3英尺高的猴子，屋子的房顶上挂着一串香蕉,离地面8英尺。屋子里有两个可叠放起来、可移动、可攀登的3英尺高的箱子。猴子很想得到香蕉。

- 初始状态：地面上3英尺高的猴子，离地面8英尺高的香蕉，两个可叠放起来、可移动、可攀登的3英尺高的箱子
- 目标测试：猴子可以拿到香蕉
- 后继函数：爬上箱子，爬下箱子，走到一个箱子处，移动一个箱子，叠放箱子，拿香蕉
- 耗散函数：动作次数

第一次作业

- 3.7 给出下列问题的初始状态、目标测试、后继函数和耗散函数。选择精确得足以实现的形式化。
- c. 有一个程序，当送入一个特定文件的输入记录时会输出“不合法的输入记录”。已知每个记录的处理独立于其它记录。要求找出哪个记录不合法。（未要求）

- 初始状态：全部记录待处理
- 目标测试：找到一个特定记录，该记录程序输出“不合法的输入记录”
- 后继函数：将记录的前一半送入程序，将记录的后一半送入程序
- 耗散函数：程序运行次数

第一次作业

- 3.7 给出下列问题的初始状态、目标测试、后继函数和耗散函数。选择精确得足以实现的形式化。
- d. 有三个水壶，容量分别为12加仑、8加仑和3加仑,还有一个水龙头。可以把壶装满或者倒空，从一个壶倒进另一个壶或者倒在地上。要求量出刚好1加仑水。

- 初始状态：三个水壶都为空（[0, 0, 0]）
- 目标测试：某个水壶中存在1加仑水（[1, x, y] or [x, 1, y] or [x, y, 1]）
- 后继函数：装满某个水壶（[x, y, z] -> [12, y, z] or [x, 8, z] or [x, y, 3]）；倒空某个水壶（[x, y, z] -> [0, y, z] or [x, 0, z] or [x, y, 0]）;将某一水壶中的水倒入另一个未满的水壶直到水壶中的水被倒完或另一个水壶装满（eg. [x, y, z] x->y: if x+y < 8 then [0, x+y, z] else [x+y-8, 8, z]）
- 耗散函数：动作次数

第一次作业

- 3.9 传教士和野人问题通常描述如下：三个传教士和三个野人在河的一边，还有一条能载一个人或者两个人的船。找到一个办法让所有的人都渡到河的另一岸，要求在任何地方野人数都不能多于传教士的人数(可以只有野人没有传教士)。这个问题在AI领域中很著名,因为它是第一篇从分析的观点探讨问题形式化的论文的主题（Amarel, 1968）
- a. 精确地形式化该问题，只描述确保该问题有解所必需的特性。画出该问题的完全状态空间图。

- 初始状态：3个传教士和3个野人以及一条船在岸上，另一岸为空
- 目标状态：3个传教士和3个野人都到达另一岸
- 耗散函数：动作次数
- 后继函数：移动1个或者2个人以及一条船到另一岸去。

第一次作业

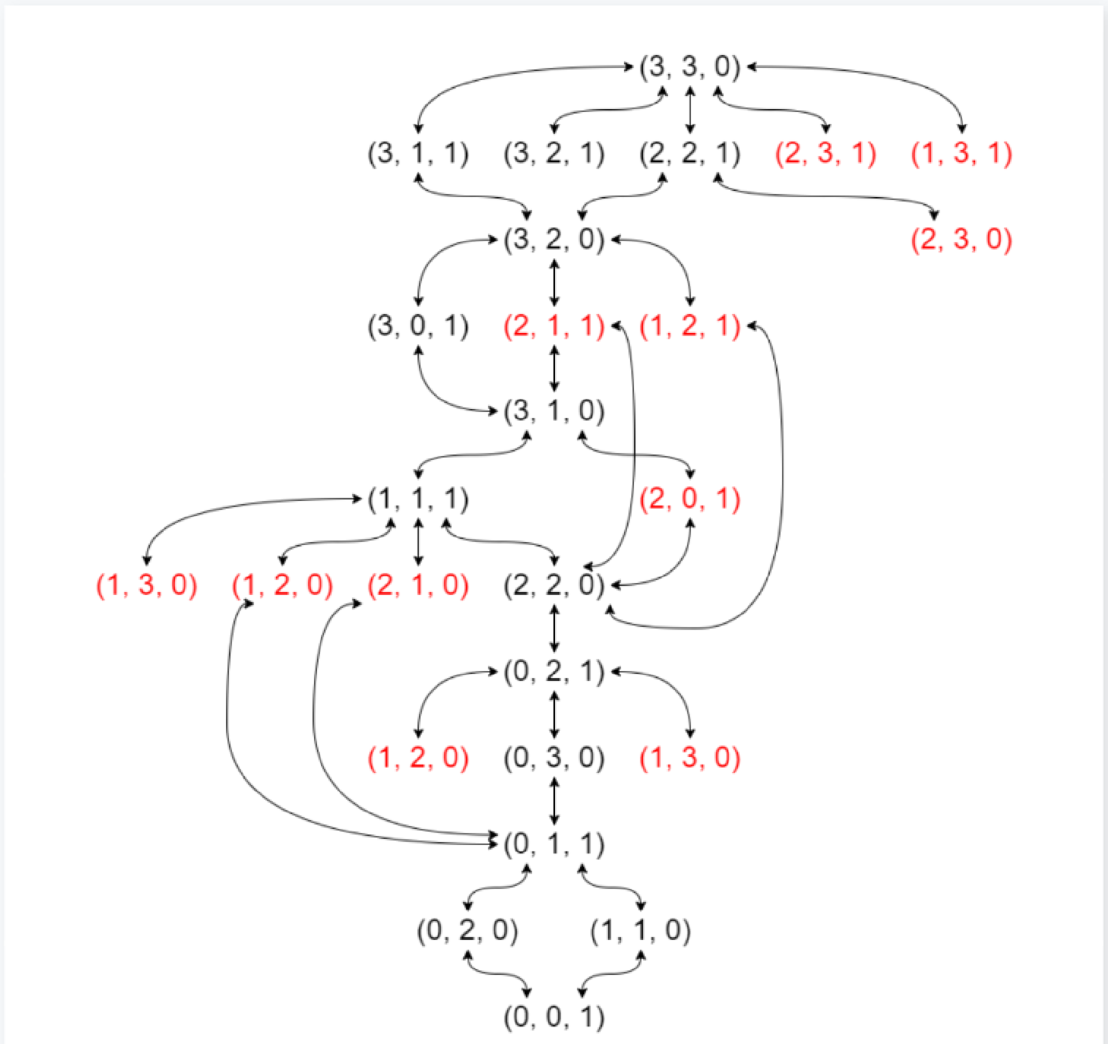
- 初始状态：3个传教士和3个野人以及一条船在岸上，另一岸为空
- 目标状态：3个传教士和3个野人都到达另一岸
- 耗散函数：动作次数
- 后继函数：移动1个或者2个人以及一条船到另一岸去。

- 定义状态表示三元组(A, B, C)，A表示在左岸传教士的数量，B表示在左岸野人的数量，C为二元指示函数，C=0表示船在左岸，C=1表示船在右岸，则可以形式化的表示为：

- 初始状态：[3, 3, 0]
- 目标状态：[0, 0, 1]（注意，最终结束时船不可能在左岸）
- 耗散函数：动作次数
- 后继函数：[a, b, c] -> [a-1, b, ~c] or [a-2, b, ~c] or [a, b-1, ~c] or [a, b-2, ~c] or [a-1, b-1, ~c]

第一次作业

➢ 完全状态空间图：



红色标记的为不合法状态，搜索时，达到不合法状态立即回溯

第一次作业

- b.用一个合适的搜索算法实现和最优地求解该问题。检查重复状态是个好主意吗?
- 下面给出一个BFS的例子

本题状态空间有限，不检查重复状态也是可以搜索出结果的，但就效率而言检查重复状态显然是有利于提高算法搜索速度的

(b)

```
BFS(G)
    // G: the graph of states and actions;
    // G.adj[u]: a set of nodes that can be reached from u by an action;
    // frontier: a FIFO queue;
    // if no solution can be found, return -1; else return cost
    // the optimal path can be infered from u.parent
    cost = 0
    for each node u in G
        u.visited = false
    initial_node.visited = true
    EnQueue(initial_node, frontier)
    while !isEmpty(frontier)
        cost = cost + 1
        u = DeQueue(frontier)
        for each node v in G.f[u] && !v.visited
            if v == goal_node
                return cost
            v.visited = true;
            EnQueue(v, frontier)
    return -1
```

一种可行的方案是：去2个野人，回1个野人，去2个野人，回1个野人，去2个传教士，回1个传教士和1个野人，去2个传教士，回1个野人，去2个野人，回1个野人，去2个野人。

最低路径耗散为11。

第一次作业

- c. 这个问题的状态空间如此简单，你认为为什么人们求解它却很困难?

- 几乎所有的移动要么是非法的，要么就需要返回到上一状态，因此状态空间规模将会变大。这也是为什么检查重复状态有助于提高搜索效率的原因

第七次作业

- 贝叶斯公式，描述两个条件概率之间的关系：
- $P(A∩B) = P(A)*P(B|A)=P(B)*P(A|B)$
- 如上公式也可变形为：
- $P(A|B)=P(B|A)*P(A)/P(B)$

第七次作业

- 13.15 在一年一度的体检之后，医生告诉你一些坏消息和一些好消息。坏消息是你在一种严重疾病的测试中结果呈阳性，而这个测试的准确度为99%(即当你确实患这种病时，测试结果为阳性的概率为0.99；而当你未患这种疾病时测试结果为阴性的概率也是0.99)。好消息是，这是一种罕见的病，在你这个年龄段大约10000人中才有1例。为什么“这种病很罕见”对于你而言是一个好消息？你确实患有这种病的概率是多少？
- 已知P(test|disease)=0.99, P(¬test | ¬disease)=0.99, P(disease)=0.0001. 求P(disease | test)
- 记P(disease) = p，则P(disease|test)= $\frac{P(test|disease)P(disease)}{P(test)}$ =

$$\frac{0.99p}{0.99p+0.01(1-p)} = \frac{0.99}{0.98} - \frac{0.0099}{0.98(0.98p+0.01)}$$(p=0.0001时，P=0.009804)

P(disease|test)随p增大而增大，所以这种病很罕见是一个好消息

第七次作业

- 13.18 假设给你一只袋子，装有n个无偏差的硬币，并且告诉你其中n - 1个硬币是正常的，一面是正面而另一面是反面。不过剩余1枚硬币是伪造的，它的两面都是正面。
- a.假设你把手伸进口袋均匀随机地取出一枚硬币，把它抛出去，硬币落地后正面朝上。那么你取出伪币的（条件）概率是多少?

- 有n个硬币，每个硬币抛掷有两种结果，总共2n种结果，其中n+1种是正面，n−1种是反面。而在正面朝上的n+1种中有两种是伪造的情况，所以拿到伪币得概率是$\frac{2}{n+1}$
- 使用贝叶斯公式计算：P(fake|head) $=\frac{P(heads|fake)P(fake)}{P(heads)}=\frac{1\cdot\frac{1}{n}}{\frac{n+1}{2n}}=\frac{2}{n+1}$

第七次作业

- 13.18 假设给你一只袋子，装有n个无偏差的硬币，并且告诉你其中n - 1个硬币是正常的，一面是正面而另一面是反面。不过剩余1枚硬币是伪造的，它的两面都是正面。
- b.假设你不停地抛这枚硬币，一共抛了k次，而且看到k次正面向上。那么你取出伪币的条件概率是多少?

- 每个币掷k次有2^k种结果，有n个币，即有$n * 2^k$种结果。其中伪币的2^k种结果都是全正面，$n-1$个真币每个币的2^k种结果中，只有一种是k次正面向上。即$n * 2^k$种结果中有$2^k + n - 1$种是k次全正面向上，但是只有2^k种是拿到了伪币。所以P(fake| k-heads)=$\frac{2^k}{2^k+n-1}$

第七次作业

- 13.18 假设给你一只袋子，装有n个无偏差的硬币，并且告诉你其中n - 1个硬币是正常的，一面是正面而另一面是反面。不过剩余1枚硬币是伪造的，它的两面都是正面。
- b.假设你不停地抛这枚硬币，一共抛了k次，而且看到k次正面向上。那么你取出伪币的条件概率是多少?

- 使用贝叶斯公式计算：P(fake|k-heads)

$$=\frac{P(k-heads|fake)P\,(fake)}{P(k-heads)}=\frac{1\cdot\frac{1}{n}}{\frac{1}{n}+\frac{n-1}{n}\cdot\frac{1}{2^k}}=\frac{2^k}{2^k+n-1}$$

第七次作业

- 13.18 假设给你一只袋子，装有n个无偏差的硬币，并且告诉你其中n - 1个硬币是正常的，一面是正面而另一面是反面。不过剩余1枚硬币是伪造的，它的两面都是正面。
- c.假设你希望通过把取出的硬币抛k次的方法来确定它是不是伪造的。如果抛k次后都是正面朝上，那么决策过程返回fake（伪造），否则返回normal（正常）。这个过程发生错误的（无条件）概率是多少?
- “这个过程发生错误”是指拿到了正常的硬币却判定为fake（不会存在拿到了伪币却判定为normal的情况），而拿到伪造的硬币判定为fake并不是错误的。挑选一枚正常硬币的概率是$\frac{n-1}{n}$，而一个正常硬币抛掷k次均为正面的概率是$\frac{1}{2^k}$，故最后的结果应该是$\frac{n-1}{n \cdot 2^k}$

第七次作业

- 13.21 假设你是雅典一次夜间出租车肇事逃逸的交通事的目击者。雅典所有的出租车都是蓝色或者绿色的。而你发誓所看见的肇事出租车是蓝色的。大量测试表明，在昏暗的灯光条件下，区分蓝色和绿色的可靠度为75%。
- a. 有可能据此计算出肇事出租车最可能是什么颜色吗？ （提示:请仔细区分命题“肇事车是蓝色的”和命题“肇事车看起来是蓝色的”。）
- 已知P(appear=blue|blue)=P(appear=green|green)
- a. $P(blue|appear = blue) = \frac{P(appear = blue|blue)P(blue)}{P(appear=blue)} = 0.75 * \frac{P(blue)}{P(appear=blue)}$

$$P(green|appear = blue) = \frac{P(appear = blue|green)P(green)}{P(appear=blue)} = 0.25 * \frac{P(green)}{P(appear=blue)}$$

第七次作业

- 13.21 假设你是雅典一次夜间出租车肇事逃逸的交通事的目击者。雅典所有的出租车都是蓝色或者绿色的。而你发誓所看见的肇事出租车是蓝色的。大量测试表明，在昏暗的灯光条件下，区分蓝色和绿色的可靠度为75%。
- b.假设你知道雅典绿色出租车占总数九成，出租车最有可能的颜色?

- b. $P(blue) = 0.1$代入即可

$$P(blue|appear = blue) = \frac{P(appear = blue|blue)P(blue)}{P(appear=blue)} = 0.75 * \frac{P(blue)}{P(appear=blue)}$$

$$P(green|appear = blue) = \frac{P(appear = blue|green)P(green)}{P(appear=blue)} = 0.25 * \frac{P(green)}{P(appear=blue)}$$

第七次作业

- 13.22 文本分类是基于文本内容将给定的一个文档分类成固定的几个类中的一类。朴素贝叶斯模型经常用于这个问题。在朴素贝叶斯模型中，查询（query）变量是这个文档的类别，而结果（effect）变量是语言中每个单词的存在与否；假设文档中单词的出现是独立的，单词的出现频率由文档类别决定。
- a. 给定一组已经被分类的文档，准确解释如何构造这样的模型。

- a. 模型由先验概率$P(category)$和条件概率$P(word_i|category)$组成。对于每一类来说，利用所有文档中的属于类c的那一部分文档，来近似估计$P(category{=}c)$. 类似的，用属于类 c 的那一部分文档中包含单词 i 的文档来近似估计$P(word_i|category = c)$

第七次作业

- 13.22 文本分类是基于文本内容将给定的一个文档分类成固定的几个类中的一类。朴素贝叶斯模型经常用于这个问题。在朴素贝叶斯模型中，查询（query）变量是这个文档的类别，而结果（effect）变量是语言中每个单词的存在与否；假设文档中单词的出现是独立的，单词的出现频率由文档类别决定。
- b. 准确解释如何分类一个新文档。

- b.当得到一个新文档时，判断文档是否包含某个词$word_i$，最后来计算条件概率$P(category = c| \dots, w_i, \dots, w_j, \dots)$

第七次作业

- 13.22 文本分类是基于文本内容将给定的一个文档分类成固定的几个类中的一类。朴素贝叶斯模型经常用于这个问题。在朴素贝叶斯模型中，查询（query）变量是这个文档的类别，而结果（effect）变量是语言中每个单词的存在与否；假设文档中单词的出现是独立的，单词的出现频率由文档类别决定。
- c. 题目中的条件独立性假设合理吗?请讨论。

- c. 不合理，每个词出现概率不独立，因为实际文档中上下文（context）的单词间存在关联性：显然$P(artificial\ intelligence) \neq P(artificial) * P(intelligence)$

第七次作业

- 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域内恒星的数目 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14.22所示的三种贝叶斯网络结构。

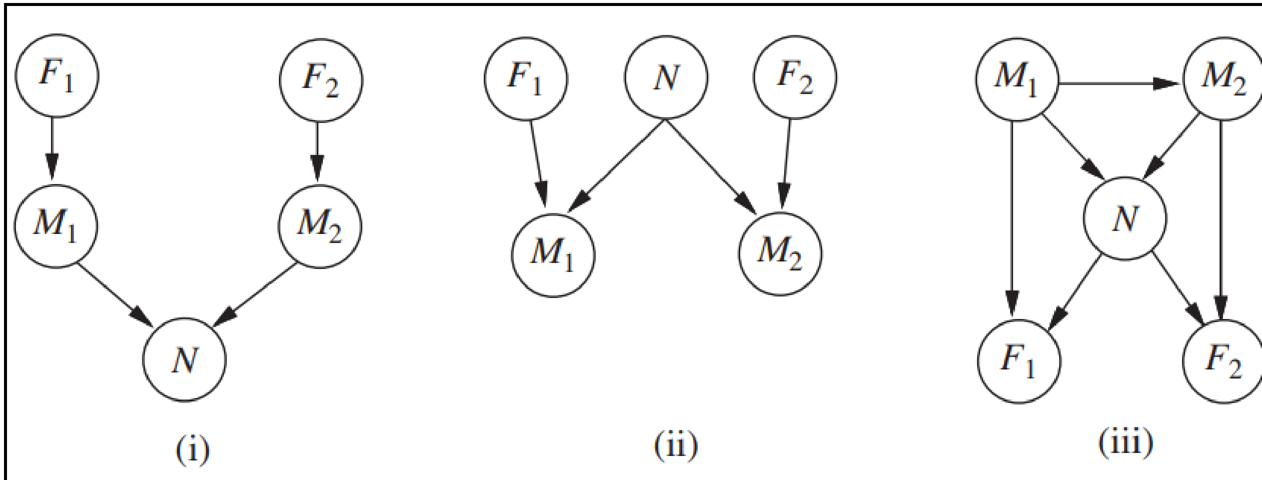


Figure 14.22 Three possible networks for the telescope problem.

第七次作业

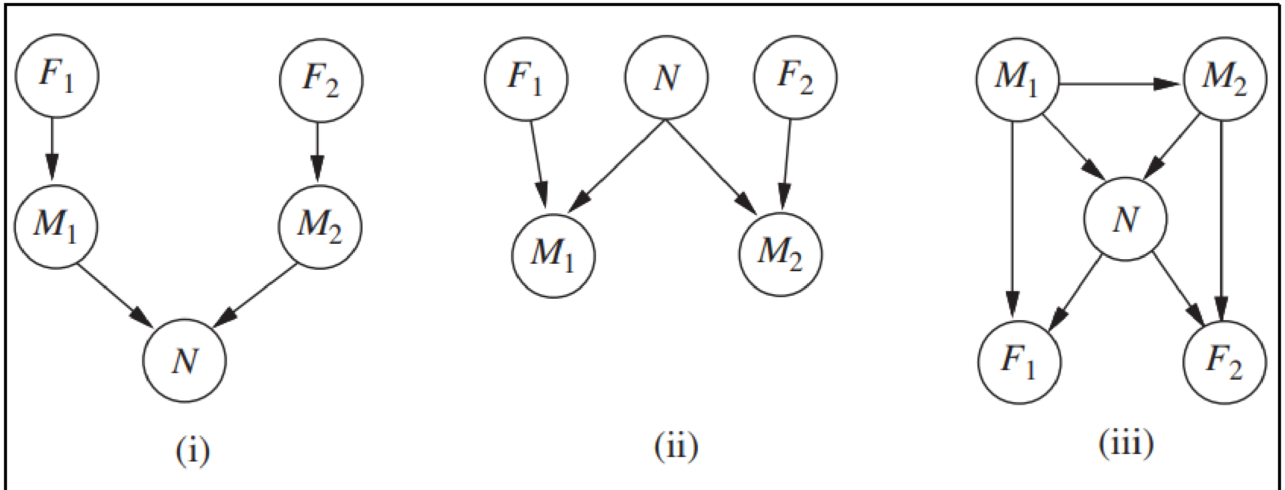


Figure 14.22 Three possible networks for the telescope problem.

- 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?
- (i) 不是。因为它表示在给定M_1和M_2 时，N与F_1, F_2 无关，这是不对的。
- (ii) 是。因为它正确的表达了恒星观测数量M_1, M_2受到实际数量N以及望远镜对焦情况F_1, F_2的影响。
- (iii) 是。因为它同样正确表达了五者之间的相互影响关系，虽然更为复杂。

第七次作业

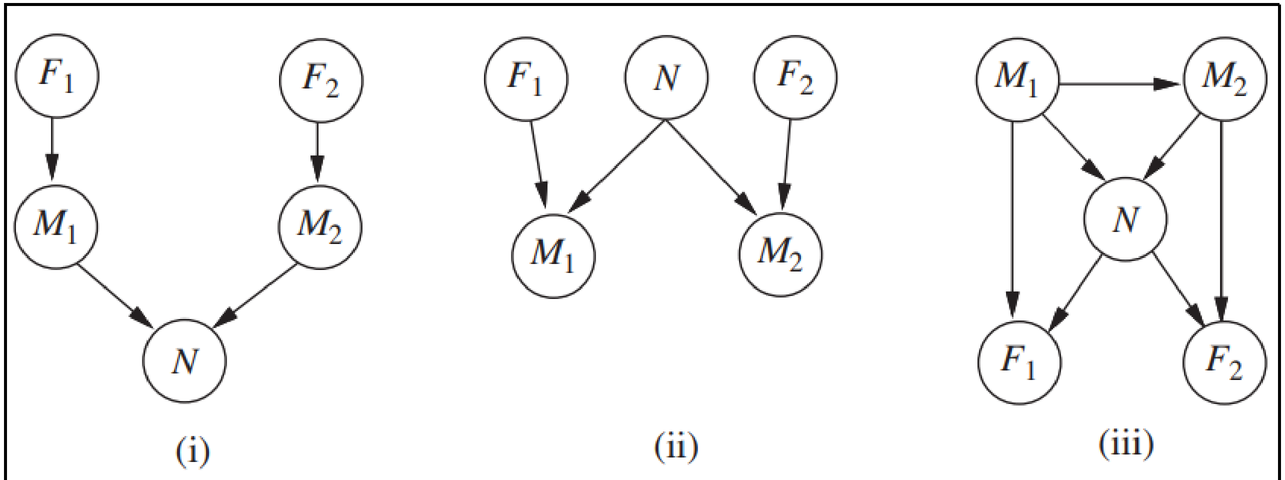


Figure 14.22 Three possible networks for the telescope problem.

- b. 哪一种网络结构是最好的，请解释。
- （ii）最好，因为它需要的参数最少，网络结构更简单。

第七次作业

- c.当$N \in 1,2,3$, $M_1 \in 0,1,2,3,4$ 时，请写出$P(M_1 \mid N)$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。
- 易得 $P(M_1|N) = P(M_1|N,F_1)P(F_1|N) + P(M_1|N,\neg F_1)P(\neg F_1|N)$

$$= P(M_1|N,F_1)P(F_1) + P(M_1|N,\neg F_1)P(\neg F_1)$$

f为望远镜失焦的概率，望远镜失焦时有$M_1 = 0$。若望远镜正常，则存在 e 的概率会少数一颗或多数一颗，$1-2e$的概率计数将准确，打表如下：

	N = 1	N = 2	N = 3
$M_1 = 0$	f + e(1 – f)	f	f
$M_1 = 1$	(1 - 2e)(1 – f)	e(1 – f)	0
$M_1 = 2$	e(1 – f)	(1 – 2e)(1 - f)	e(1 – f)
$M_1 = 3$	0	e(1 – f)	(1 - 2e)(1 – f)
$M_1 = 4$	0	0	e(1 – f)

第七次作业

- d. 假设$M_1 = 1, M_2 = 3$。如果我们假设取值上没有先验概率约束,可能的恒星数目是多少?
- 对于$M_1 = 1$，考虑到测量误差和失焦误差，可能的 $\boldsymbol{N}$ 的取值为$\{2\} \cup \{n|n \geq 4, n \in N\}$
- 对于$M_2 = 3$，考虑到测量误差和失焦误差，可能的 $\boldsymbol{N}$ 的取值为$\{2, 4\} \cup \{n|n \geq 6, n \in N\}$
- 则可能的恒星数量 $\boldsymbol{N}$ 的取值为上述两者的交集，即 $\{2, 4\} \cup \{n|n \geq 6, n \in N\}$

第七次作业

- e. 在这些观测结果下，最可能的恒星数目是多少？解释如何计算这个数目，或者，如果不可能计算，请解释还需要什么附加信息以及它将如何影响结果。
- 由于先验分布 $\boldsymbol{P(N)}$ 未知，因此无法计算最可能的恒星数目。若假设对于先验分布 $\boldsymbol{P(N=2)}, \boldsymbol{P(N=4)}, \boldsymbol{P(N\geq 6)}$ 三者相差不大，则可估计后验概率如下：

$$P(N=2|M_1=1,M_2=3)=\alpha\cdot e^2(1-f)^2\cdot P(N=2)$$

$$P(N=4|M_1=1,M_2=3)=\alpha\cdot ef\cdot P(N=4)$$

$$P(N\geq 6|M_1=1,M_2=3)=\alpha\cdot f^2\cdot P(N\geq 6)$$

其中$\alpha=P(M_1=1,M_2=3)$，考虑到$f\ll e$，则有$P(N=2|M_1=1,M_2=3)$最大，因此在该假设下，最可能的恒星数目为 **2**.

第七次作业

➢ 14.13 考虑图14.22(ii) 的网络，假设两个望远镜完全相同。$N\in 1,2,3, \mathrm{M}_1, M_2 \in 0,1,2,3,4$，CPT表和习题14.12所描述的一样。使用枚举算法计算概率分布$P(N|M_1=2, M_2=2)$

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

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

---

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

Figure 14.9 The enumeration algorithm for answering queries on Bayesian networks.

第七次作业

- 14.13 考虑图14.22(ii) 的网络，假设两个望远镜完全相同。$N\in 1,2,3, \mathrm{M}_1, M_2 \in 0,1,2,3,4$，CPT表和习题14.12所描述的一样。使用枚举算法计算概率分布$P(N|M_1=2,M_2=2)$

- $P(N|M_1=2,M_2=2)=\alpha\sum_{f_1,f_2}P(f_1,f_2,N,M_1=2,M_2=2)$

$$=\alpha\sum_{f_1,f_2}P(f_1)P(f_2)P(N)P(M_1=2|f_1,N)P(M_2=2|f_2,N)$$

展开，化简后得到：

$$P(N|M_1=2,M_2=2)$$

$$=\frac{e^2}{6e^2-4e+1}[P(N=1)+P(N=3)]+\frac{(1-2e)^2}{6e^2-4e+1}P(N=2)$$

谢谢！