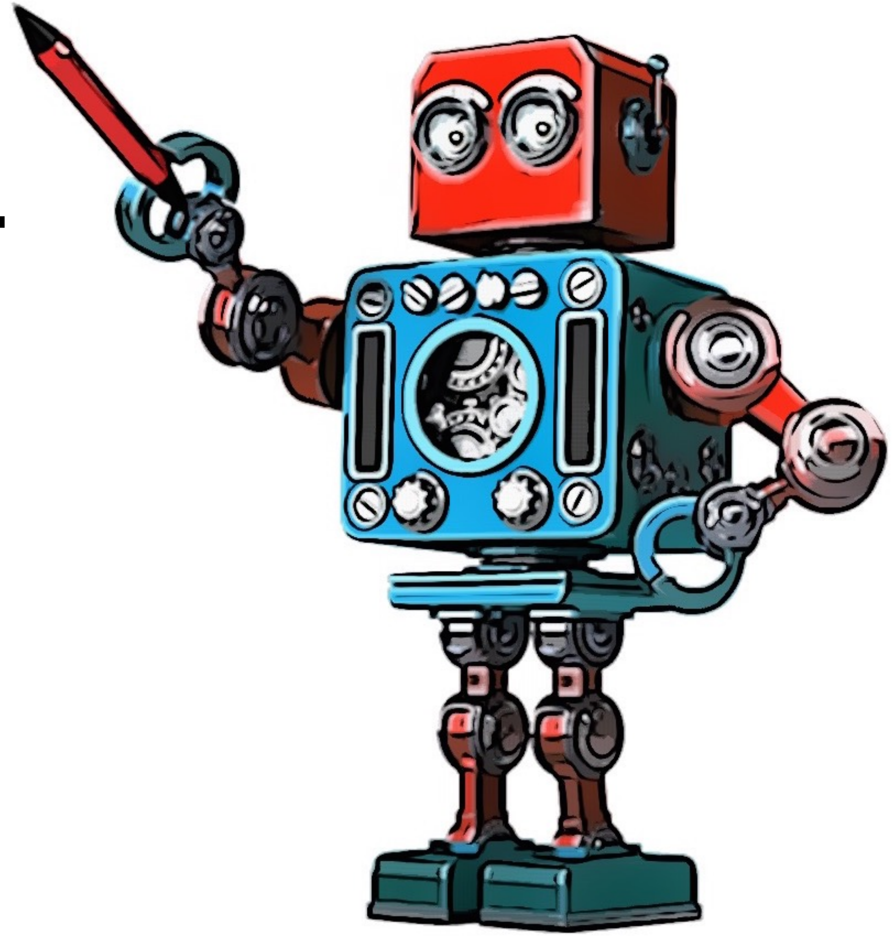
概率图模型

知识点复习



1. 条件独立性

- 独立

$$P(X,Y) = P(X)P(Y)$$

等价

$$P(X) = P(X \mid Y)$$

独立 推不出 条件独立

- 条件独立

$$P(X,Y \mid Z) = P(X \mid Z)P(Y \mid Z)$$

等价

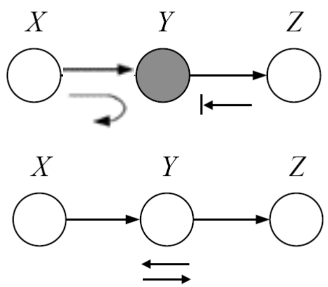
$$P(X \mid Z) = P(X \mid Y,Z)$$

条件独立 推不出 独立

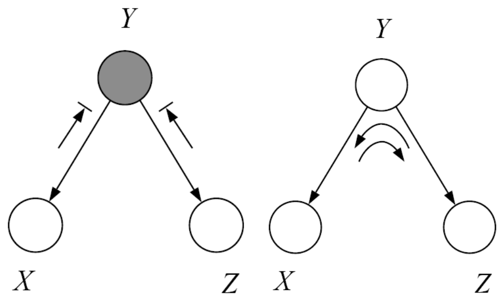
1. 条件独立性 (快速检验)

贝叶斯球规则：假设在贝叶斯网络中，有一个按一定规则运动的球。**已知中间节点（或节点集合）Z，如果球不能由节点X出发到达节点Y（或者由Y到X），则称X和Y关于Z条件独立。**

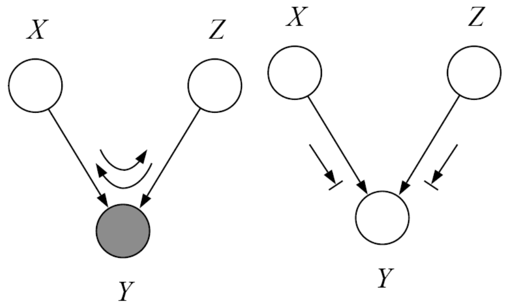
- **未知节点**：总能使贝叶斯球通过，同时还可以反弹从其子节点方向来的球。（父 -> 子）|（子 -> 父/子）
- **已知节点**：反弹从其父节点方向过来的球，截止从其子节点方向过来的球。（父 -> 父）|（子 -> “截止”）



经典的马尔科夫链
“过去”，“现在”，“未来”



共同的起因（Common Cause）
Y “解释” X 和 Z之间所有的依赖



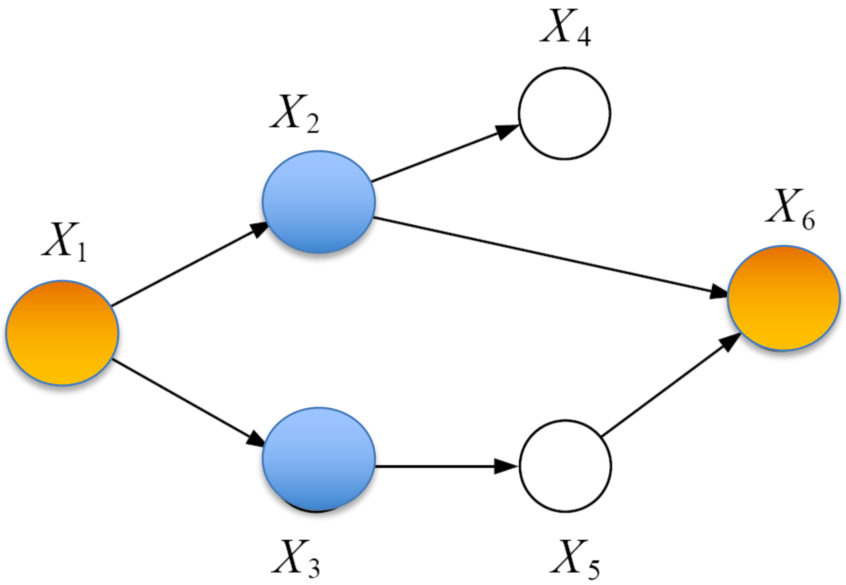
共同效应（Common Effect）
多个相互竞争的解释

1. 条件独立性：例子

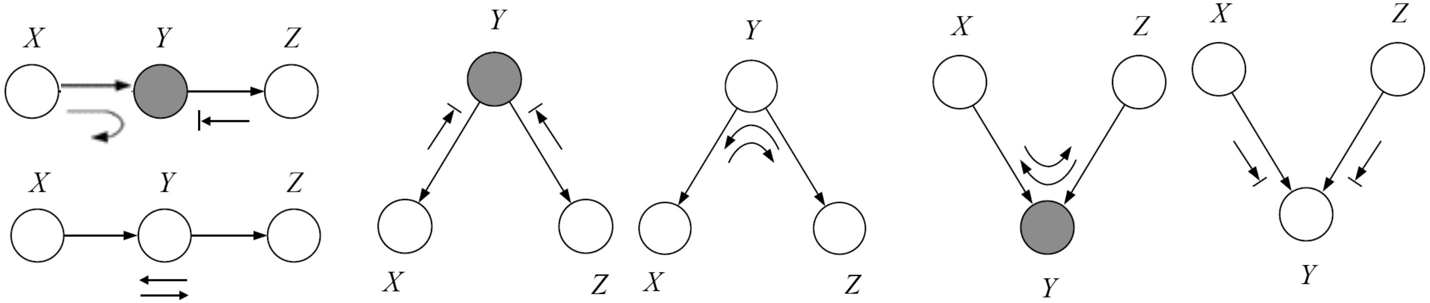
检查通过可达性

$X_2 \perp\!\!\!\perp X_3 \mid \{X_1, X_6\}$ ✗

$X_1 \perp\!\!\!\perp X_6 \mid \{X_2, X_3\}$ ✓

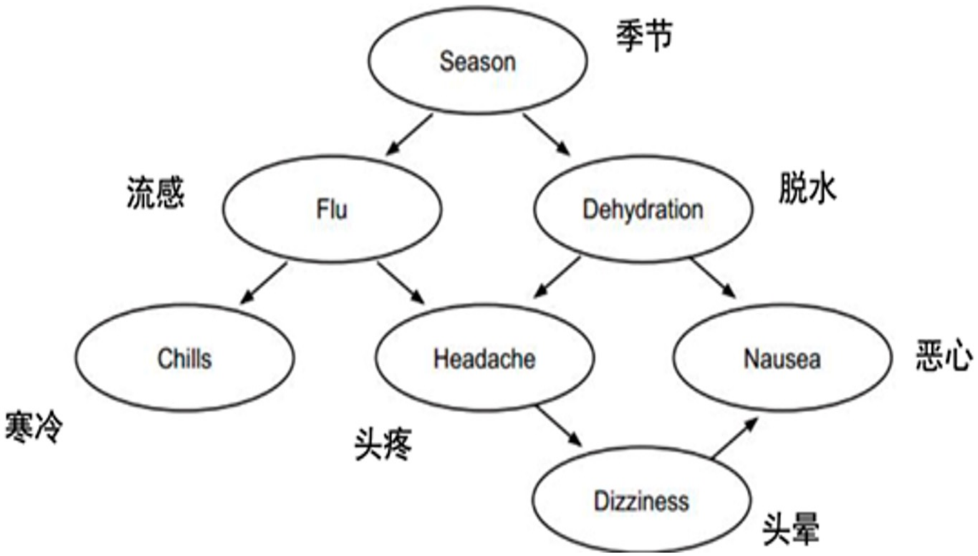


贝叶斯球规则



1. 条件独立性：作业

- **已知以下贝叶斯网络，包含 7 个变量，即 Season （季节），Flu （流感）， Dehydration （脱水）， Chills （发冷），Headache （头疼）， Nausea （恶心）， Dizziness （头晕），则下列条件独立成立的是（ ）**

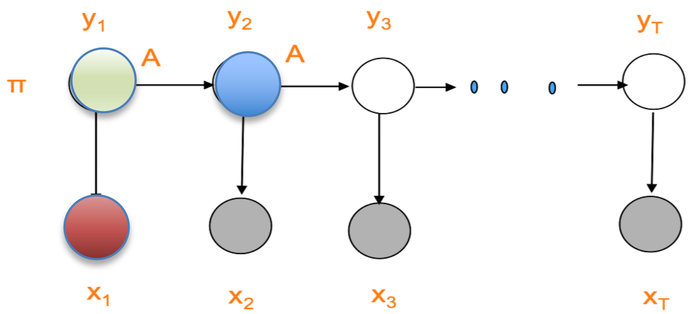


A、 $Season \perp Chills \mid Flu$

B、 $Season \perp Chills$

C、 $Season \perp Headache \mid Flu$

2. HMM表示



- 第一个状态节点y_1对应一个初始状态概率分布

  $$\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N) : \ \pi_i = P\left(y_1^i = 1\right);$$

- 状态转移矩阵$\mathbf{A}$，其中a_{ij}为转移概率：$a_{ij} = P\left(y_{t+1}^j = 1 \middle| y_t^i = 1\right),\quad 1 \le i \le N,\ 1 \le j \le N$；
- 每个输出节点x_t有一个状态节点作为父节点，因此有发射概率矩阵$\mathbf{B}$：$b_{ij} = p\left(x_t^j = 1 \middle| y_t^i = 1\right),\quad 1 \le i \le N,\ 1 \le j \le M$；
- 对于特定的配置，$(\mathbf{x}, \mathbf{y}) = (x_0, x_1, \cdots, x_T, y_0, y_1, \cdots, y_T)$联合概率可以表示为：

T个时刻的发射概率

$$p(\mathbf{x}, \mathbf{y}) = p(y_1) \prod_{t=1}^{T-1} p(y_{t+1} | y_t) \prod_{t=1}^{T} p(x_t | y_t)$$

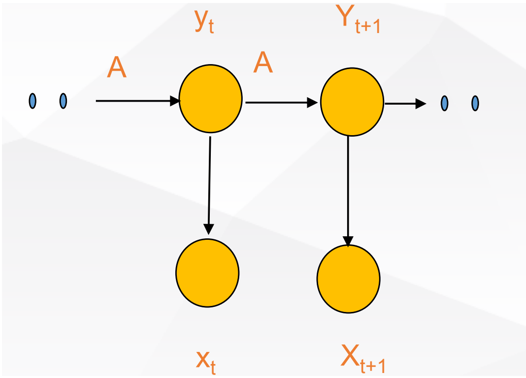
初始状态y1概率　　后续T-1个时刻的状态y的转移概率

3. 推断：前向（递归）算法

$$p(y_t|\mathbf{x}) = \frac{p(\mathbf{x}|y_t)p(y_t)}{p(\mathbf{x})}$$

$$= \frac{p(x_1 \ldots x_t|y_t)\boldsymbol{p}(\boldsymbol{x}_{t+1} \ldots \boldsymbol{x}_T|\boldsymbol{y}_t)p(y_t)}{p(\mathbf{x})}$$

给定yt，x从1到t时刻的观测样本 和x从t+1到T时刻的观测值 条件独立



$$p(y_t|\mathbf{x}) = \frac{\alpha(y_t)\boldsymbol{\beta}(\boldsymbol{y}_t)}{p(\mathbf{x})}$$

$$p(\mathbf{x}) = \sum_{y_t} \alpha(y_t)\boldsymbol{\beta}(\boldsymbol{y}_t)$$

其中$\alpha(y_t)$是产生部分输出序列$x_1, \cdots, x_t$，并结束于y_t的概率

其中$\beta(y_t)$是从y_t状态开始，产生输出序列$x_{t+1}, \cdots, x_T$的概率

α递归计算—前向算法

- 因此可将$p(y_t|\mathbf{x})$转化为计算α，β
- 可以获得$\alpha(y_t)$和 $\alpha(y_{t+1})$的递归关系：

$$\alpha(y_{t+1}) = p(x_1 \ldots x_t, x_{t+1}, y_{t+1})$$

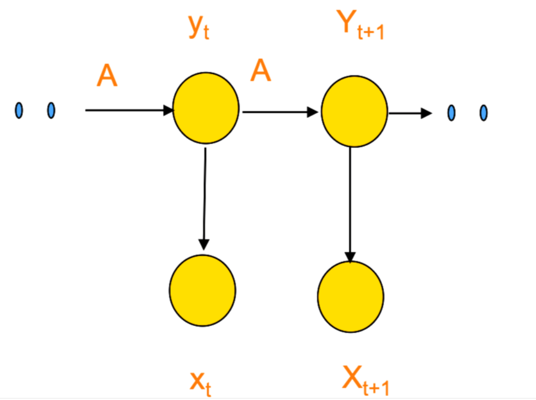
$$= \sum_{y_t} \alpha(y_t)\, a_{y_t,y_{t+1}} b_{y_{t+1},x_{t+1}}$$

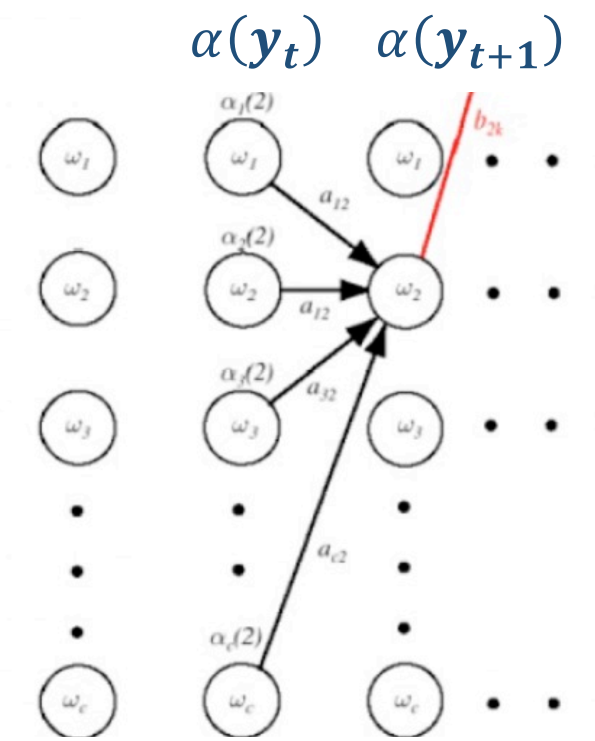
- 初始化:定义α的第一步

$$\alpha(y_1) = p(x_1, y_1)$$
$$= p(x_1|y_1)p(y_1)$$
$$= b_{y_1,x_1}\pi_{y_1}$$

- 终止:

$$p(x_1 \ldots x_T|\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}) = \sum_{y_T} \alpha(y_T)$$





4. Viterbi解码

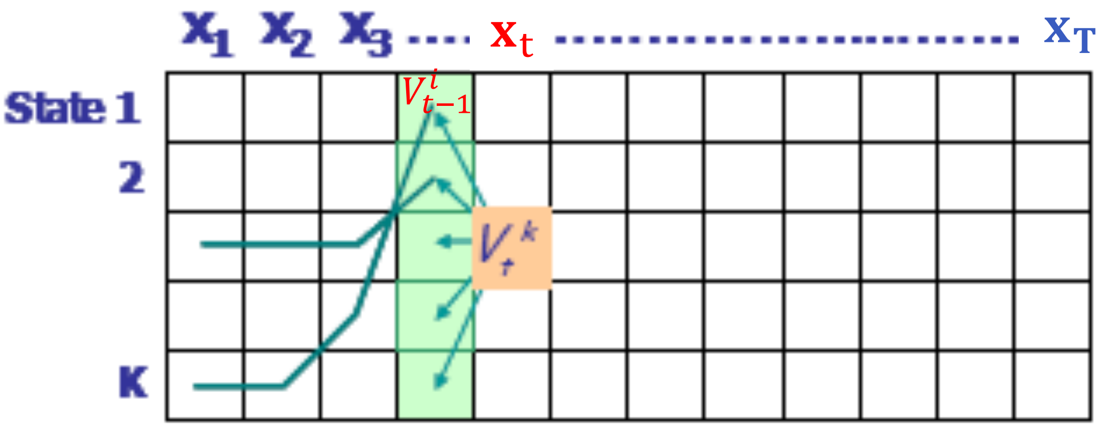
- 给定$x = x_1, \dots, x_T$，我们要找 $y = y_1, \dots, y_T$，使得$p(\mathbf{y}|\mathbf{x})$最大

$$\mathbf{y}^* = \operatorname*{argmax}_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}) = \operatorname*{argmax}_{\mathbf{y}} p(\mathbf{x}, \mathbf{y})$$

- 令$V_t^k = \max_{\{y_1,\dots,y_{t-1}\}} p(x_1, \dots, x_{t-1}, y_1, \dots, y_{t-1}, x_t, y_t = s_k)$

  =结尾状态为$y_t = s_k$时，最可能状态序列的概率

- 递归：$V_t^k = p(x_t|y_t^k = 1)\max_i a_{i,k} V_{t-1}^i$



4. Viterbi解码

- 初始化：$V_1^k = p(x_1, y_1 = s_k) = \pi_k b_{y_1,x_1}$
- 之后就不断迭代：
- $V_t^k = p(x_t | y_t^k = 1) \max_i a_{i,k} V_{t-1}^i$
- 终止：$P^* = \max_{1 \le i \le N} V_T^i$

4. Viterbi 解码

- 假设有 3 个盒子，分别装有不同数量的苹果(A)和桔子(O)：
- 盒子一:2个A，2个O;
  盒子二:3个A，1个O;
  盒子三:1个A，3个O;
- 每次随机选择一个盒子并从中抽取一个水果，观测并记录看到的水果是哪种。不幸的是，忘记去记录所选的盒子号码，只记录了每次看到的水果是 A 还是 O。
- (1)请用 HMM 模型描述上述过程。
- (2)假如观测到水果序列为x={A，A，O，O，O}，请给出最可能的盒子序列。

参考答案：

（1）初始概率$\boldsymbol{\pi}=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$，

盒子间的转移概率矩阵$\mathbf{A}=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{bmatrix}$，

发放概率矩阵（给定盒子时，选择每种水果的概率）$\mathbf{B}=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{3}{4} & \frac{1}{4}\\ \frac{1}{4} & \frac{3}{4}\end{bmatrix}$。

当$t=1$时：已知 $x_1 = apple$，所以

初始化

$$V_1^1 = \pi_1 b_{1,x_1} = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$$

$$V_1^2 = \pi_2 b_{2,x_1} = \frac{1}{3} \times \frac{3}{4} = \frac{1}{4}$$

$$V_1^3 = \pi_3 b_{3,x_1} = \frac{1}{3} \times \frac{1}{4} = \frac{1}{12}$$

	t=1	t=2	t=3	t=4	t=5	
V_t^1	$\frac{1}{6}$					
V_t^2	$\frac{1}{4}$					
V_t^3	$\frac{1}{12}$					

当$t=2$时：已知$x_2=apple$，所以

迭代 $V_t^k = p(x_t|y_t^k=1)\max_i a_{i,k}V_{t-1}^i$

$$V_2^1 = b_{1,1}\max_{\{y_1^i\}} a_{i,1}V_1^i = b_{1,1}\max\left(\frac{1}{3}\times\frac{1}{6}, \frac{1}{3}\times\frac{1}{4}, \frac{1}{3}\times\frac{1}{12}\right) = \frac{1}{2}\times\frac{1}{12} = \frac{1}{24}$$

	t=1	t=2	t=3	t=4	t=5	
V_t^1	$\frac{1}{6}$	$\frac{1}{24}$				
V_t^2	$\frac{1}{4}$					
V_t^3	$\frac{1}{12}$					

当$t=2$时：已知$\mathrm{x}_2 = apple$，所以

迭代 $V_t^k = p(x_t|y_t^k = 1)\max_i a_{i,k} V_{t-1}^i$

$$V_2^1 = b_{1,1} \max_{\{y_1^i\}} a_{i,1} V_1^i = b_{1,1} \max\left(\frac{1}{3}\times\frac{1}{6}, \frac{1}{3}\times\frac{1}{4}, \frac{1}{3}\times\frac{1}{12}\right) = \frac{1}{2}\times\frac{1}{12} = \frac{1}{24}$$

$$V_2^2 = b_{2,1} \max\left(\frac{1}{3}\times\frac{1}{6}, \frac{1}{3}\times\frac{1}{4}, \frac{1}{3}\times\frac{1}{12}\right) = \frac{3}{4}\times\frac{1}{12} = \frac{1}{16}$$

	t=1	t=2	t=3	t=4	t=5	
V_t^1	$\frac{1}{6}$	$\frac{1}{24}$				
V_t^2	$\frac{1}{4}$	$\frac{1}{16}$				
V_t^3	$\frac{1}{12}$					

当$t=2$时：已知$x_2=apple$，所以

迭代 $V_t^k = p(x_t|y_t^k=1)\max_i a_{i,k}V_{t-1}^i$

$$V_2^1 = b_{1,1}\max_{\{y_1^i\}} a_{i,1}V_1^i = b_{1,1}\max\left(\frac{1}{3}\times\frac{1}{6},\frac{1}{3}\times\frac{1}{4},\frac{1}{3}\times\frac{1}{12}\right)=\frac{1}{2}\times\frac{1}{12}=\frac{1}{24}$$

$$V_2^2 = b_{2,1}\max\left(\frac{1}{3}\times\frac{1}{6},\frac{1}{3}\times\frac{1}{4},\frac{1}{3}\times\frac{1}{12}\right)=\frac{3}{4}\times\frac{1}{12}=\frac{1}{16}$$

$$V_2^3 = b_{3,1}\max\left(\frac{1}{3}\times\frac{1}{6},\frac{1}{3}\times\frac{1}{4},\frac{1}{3}\times\frac{1}{12}\right)=\frac{1}{4}\times\frac{1}{12}=\frac{1}{48}$$

	t=1	t=2	t=3	t=4	t=5	
V_t^1	$\frac{1}{6}$	$\frac{1}{24}$				
V_t^2	$\frac{1}{4}$	$\frac{1}{16}$				
V_t^3	$\frac{1}{12}$	$\frac{1}{48}$				

当$t=3$时：已知$x_3=orange$，所以

迭代

$$V_3^1 = b_{1,2}\max_{\{y_2^i\}} a_{i,1}V_2^i = b_{1,2}\max\left(\frac{1}{24}\times\frac{1}{3},\frac{1}{16}\times\frac{1}{3},\frac{1}{48}\times\frac{1}{3}\right) = \frac{1}{2}\times\frac{1}{48} = \frac{1}{96}$$

$$V_3^2 = b_{2,2}\max_{\{y_2^i\}} a_{i,2}V_2^i = b_{2,2}\max\left(\frac{1}{24}\times\frac{1}{3},\frac{1}{16}\times\frac{1}{3},\frac{1}{48}\times\frac{1}{3}\right) = \frac{1}{4}\times\frac{1}{48} = \frac{1}{192}$$

$$V_3^3 = b_{3,2}\max_{\{y_2^i\}} a_{i,3}V_2^i = b_{3,2}\max\left(\frac{1}{24}\times\frac{1}{3},\frac{1}{16}\times\frac{1}{3},\frac{1}{48}\times\frac{1}{3}\right) = \frac{3}{4}\times\frac{1}{48} = \frac{1}{64}$$

当$t=4$时：已知$x_4 = orange$，所以

迭代

$$V_4^1 = b_{1,2} \max_{\{y_3^i\}} a_{i,1} V_3^i = \frac{1}{2} \times \max\left(\frac{1}{96} \times \frac{1}{3}, \frac{1}{192} \times \frac{1}{3}, \frac{1}{64} \times \frac{1}{3}\right) = \frac{1}{384}$$

$$V_4^2 = b_{2,2} \max_{\{y_2^i\}} a_{i,2} V_3^i = \frac{1}{4} \times \max\left(\frac{1}{96} \times \frac{1}{3}, \frac{1}{192} \times \frac{1}{3}, \frac{1}{64} \times \frac{1}{3}\right) = \frac{1}{768}$$

$$V_4^3 = b_{3,2} \max_{\{y_2^i\}} a_{i,3} V_3^i = \frac{3}{4} \times \max\left(\frac{1}{96} \times \frac{1}{3}, \frac{1}{192} \times \frac{1}{3}, \frac{1}{64} \times \frac{1}{3}\right) = \frac{1}{256}$$

当$t=5$时：已知$x_5=orange$，所以

迭代

$$V_5^1 = b_{1,2}\max_{\{y_4^i\}} a_{i,1}V_4^i = \frac{1}{2}\times\max\left(\frac{1}{384}\times\frac{1}{3},\frac{1}{768}\times\frac{1}{3},\frac{1}{256}\times\frac{1}{3}\right)=\frac{1}{6*256}=\frac{1}{1536}$$

$$V_5^2 = b_{2,2}\max_{\{y_4^i\}} a_{i,2}V_4^i = \frac{1}{4}\times\max\left(\frac{1}{384}\times\frac{1}{3},\frac{1}{768}\times\frac{1}{3},\frac{1}{256}\times\frac{1}{3}\right)=\frac{1}{12*256}=\frac{1}{3072}$$

$$V_5^3 = b_{3,2}\max_{\{y_4^i\}} a_{i,3}V_4^i = \frac{3}{4}\times\max\left(\frac{1}{384}\times\frac{1}{3},\frac{1}{768}\times\frac{1}{3},\frac{1}{256}\times\frac{1}{3}\right)=\frac{1}{4*256}=\frac{1}{1024}$$

终止：T=5

$$P^* = \max_{1\le i\le N} V_T^i = V_5^3 = \frac{1}{1024}$$

最优路径回溯： $\boldsymbol{y} = \{2，2，3，3，3\}$。

	t=1	t=2	t=3	t=4	t=5	
V_t^1	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{1}{96}$	$\frac{1}{384}$	$\frac{1}{1536}$	
V_t^2	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{192}$	$\frac{1}{768}$	$\frac{1}{3072}$	
V_t^3	$\frac{1}{12}$	$\frac{1}{48}$	$\frac{1}{64}$	$\frac{1}{256}$	$\frac{1}{1024}$	