人工智能

Lecture 9: 不确定性推理 (Uncertain Reasoning)

赵培海

东华大学
计算机科学与技术学院

2023 年 9 月 6 日

内容组织

1. 模糊推理

2. 概率推理

3. 朴素贝叶斯推理

不确定性推理 (Uncertain Reasoning)

现实世界中由于客观上存在的随机性、模糊性，反映到知识以及由观察所得到的证据上

- 不确定性的知识
- 不确定性的证据
- 须对不确定性知识的表示及推理进行研究

不确定性推理方法

- 模糊推理、可信度方法、证据理论等
- 贝叶斯、概率图等

不确定性推理 (Uncertain Reasoning)

推理：一种思维过程

- 从已知事实出发
- 运用相关知识逐步推出结论
- 或者证明某个假设成立或不成立

不确定性推理

- 从不确定性的初始证据出发
- 运用不确定性的知识
- 推出具有一定程度的不确定性但却是合理的结论

不确定性推理中的基本问题

1. 不确定性的表示与量度
2. 不确定性匹配算法及阈值的选择
3. 组合证据不确定性的算法
4. 不确定性的传递算法
5. 结论不确定性的合成

不确定性推理中的基本问题

1. 不确定性的表示与量度

   a. 知识不确定性的表示

   - 在专家系统中知识的不确定性一般是由领域专家给出的，通常是一个数值 – 知识的静态强度

   b. 证据不确定性的表示 – 证据的动态强度

   - 用户在求解问题时提供的初始证据.
   - 在推理中用前面推出的结论作为当前推理的证据.

   c. 不确定性的量度

   - 能充分表达相应知识及证据不确定性的程度.
   - 度量范围的指定便于领域专家及用户对不确定性的估计.
   - 便于对不确定性的传递进行计算，而且对结论算出的不确定性量度不能超出量度规定的范围.
   - 度量的确定应当是直观的，同时应有相应的理论依据.

不确定性推理中的基本问题

2. 不确定性匹配算法及阈值的选择
   - 不确定性匹配算法：用来计算匹配双方相似程度的算法.
   - 阈值：用来指出相似的“限度”.
3. 组合证据不确定性的算法：
   - 最大最小方法、概率方法、有界方法等.
4. 不确定性的传递算法
   - 在每一步推理中，如何把证据及知识的不确定性传递给结论.
   - 在多步推理中，如何把初始证据的不确定性传递给最终结论.
5. 结论不确定性的合成

内容组织

模糊推理

- 1965 年，美国 L. A. Zadeh 发表了 “Fuzzy Set” 的论文，首先提出了模糊理论.
- 从 1965 年到 20 世纪 80 年代，在美国、欧洲、中国和日本，只有少数科学家研究模糊理论.
- 1974 年，英国 Mamdani 首次将模糊理论应用于热电厂的蒸汽机控制.
- 1976 年，Mamdani 又将模糊理论应用于水泥旋转炉的控制.
- 1983 年日本 Fuji Electric 公司实现了饮水处理装置的模糊控制.
- 1987 年日本 Hitachi 公司研制出地铁的模糊控制系统.

模糊集

- 论域：所讨论的全体对象，用 U 等表示．
- 元素：论域中的每个对象，常用 a, b, c, x, y, z 表示．
- 集合：论域中具有某种相同属性的确定的、可以彼此区别的元素的全体，常用 A, B 等表示．
- 元素 a 和集合 A 的关系：a 属于 A 或 a 不属于 A，即只有两个真值“真”和“假”．

模糊集

- 模糊逻辑给集合中每一个元素赋予一个介于 0 和 1 之间的实数，描述其属于一个集合的强度，该实数称为元素属于一个集合的隶属度．集合中所有元素的隶属度全体构成集合的隶属函数．

模糊集表示

- 当论域中元素数目有限时，模糊集合　的数学描述为

$$A = \{(x, \mu_A(x)), x \in X\}$$

其中 $\mu_A(x)$ 是元素 x 属于模糊集 A 的隶属度.

- Zadeh 表示法
  - 论域离散且元素数目有限：

$$A = \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + \cdots + \frac{\mu_A(x_n)}{x_n} = \sum_{i=1}^{n} \frac{\mu_A(x_i)}{x_i}$$

  - 论域连续或元素数目无限：

$$A = \int \frac{\mu_A(x)}{x}$$

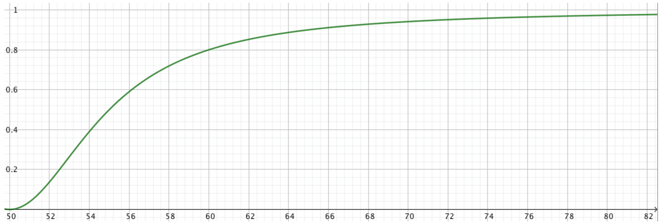
隶属函数

- 常见的隶属函数有正态分布、三角分布、梯形分布等.
- 隶属函数确定方法:
  - 模糊统计法
  - 专家经验法
  - 二元对比排序法
  - $\cdots$

隶属函数

以年龄为论域，取 $U=[0,200]$，Zadeh 给出的年青 Young 与年老 Old 两个模糊集合的隶属函数为

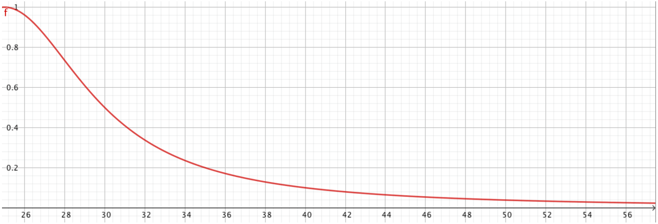
$$\mu_O(\mu)=\begin{cases}0, & 0\le\mu\le 50\\ \left[1+\left(\frac{5}{\mu-50}\right)^2\right]^{-1}, & 50<\mu\le 200\end{cases}$$



隶属函数

以年龄为论域，取 $U=[0,200]$，Zadeh 给出的年青 Young 与年老 Old 两个模糊集合的隶属函数为

$$\mu_Y(\mu)=\begin{cases}1, & 0\le\mu\le 25\\ \left[1+\left(\frac{\mu-25}{5}\right)^2\right]^{-1}, & 25<\mu\le 200\end{cases}$$



模糊集的运算

- 空集判断
  - 设 A 为 U 的模糊子集，当且仅当 $\forall x \in U, \mu_A(x) = 0$ 时，A 为空集，记为 $\emptyset$;
- A 包含于 B
  - A, B 为 U 的任意模糊子集,对 $\forall x \in U, \mu_A(x) \leq \mu_B(x)$,记为 $A \subseteq B$;
- A 等于 B
  - 对 $\forall x \in U, \mu_A(x) = \mu_B(x)$，记为 $A = B$;
- A 的补集
  - $\sim A = \{< x, 1 - \mu_A(x) > | x \in U\}$
  - $\mu_{A^C}(x) = 1 - \mu_A(x)$

模糊集的运算

- A 与 B 的并集
  - $A \cup B = \{< x, \max(\mu_A(x), \mu_B(x)) > | x \in U\}$
  - $\mu_{A\cup B}(x) = \max\{\mu_A(x), \mu_B(x)\} = \mu_A(x) \vee \mu_B(x)$
- A 与 B 的交集
  - $A \cap B = \{< x, \min(\mu_A(x), \mu_B(x)) > | x \in U\}$
  - $\mu_{A\cap B}(x) = \min\{\mu_A(x), \mu_B(x)\} = \mu_A(x) \wedge \mu_B(x)$
- A 与 B 的差集
  - $A - B = \{< x, \min(\mu_A(x), 1 - \mu_B(x)) > | x \in U\}$
  - 显然有 $\sim A = U - A$

模糊集运算的性质

- 交换律：$A \cup B = B \cup A, \quad A \cap B = B \cap A$
- 幂等律：$A \cup A = A \cap A = A$
- 分配律：
  - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- 狄摩根律：
  - $\sim (A \cup B) = \sim A \cap \sim B$
  - $\sim (A \cap B) = \sim A \cup \sim B$
- $A - (B \cup C) = (A - B) \cap (A - C)$
- $A - (B \cap C) = (A - B) \cup (A - C)$
- $A \cap B \subseteq A \subseteq A \cup B$
- $A \subseteq B$ 当且仅当 $A \cap B = A$ 当且仅当 $A \cup B = B$

模糊集的运算

设有论域 $U = \{x_1, x_2, x_3, x_4\}$，$A$ 及 B 是论域上的两个模糊集，已知：

- $A = \frac{0.3}{x_1} + \frac{0.5}{x_2} + \frac{0.7}{x_3} + \frac{0.4}{x_4}$
- $B = \frac{0.5}{x_1} + \frac{1}{x_2} + \frac{0.8}{x_3}$

则 $A^C, B^C, A \cap B, A \cup B$?

模糊集的运算

设有论域 $U=\{x_1,x_2,x_3,x_4\}$，A 及 B 是论域上的两个模糊集，已知:

- $A=\frac{0.3}{x_1}+\frac{0.5}{x_2}+\frac{0.7}{x_3}+\frac{0.4}{x_4}$
- $B=\frac{0.5}{x_1}+\frac{1}{x_2}+\frac{0.8}{x_3}$

则 $A^C, B^C, A\cap B, A\cup B$?

- $A^C=\frac{0.7}{x_1}+\frac{0.5}{x_2}+\frac{0.3}{x_3}+\frac{0.6}{x_4}$
- $B^C=\frac{0.5}{x_1}+\frac{0.2}{x_3}+\frac{1}{x_4}$
- $A\cap B=\frac{0.3\wedge 0.5}{x_1}+\frac{0.5\wedge 1}{x_2}+\frac{0.7\wedge 0.8}{x_3}+\frac{0.4\wedge 0}{x_4}=\frac{0.3}{x_1}+\frac{0.5}{x_2}+\frac{0.7}{x_3}$
- $A\cup B=\frac{0.3\vee 0.5}{x_1}+\frac{0.5\vee 1}{x_2}+\frac{0.7\vee 0.8}{x_3}+\frac{0.4\vee 0}{x_4}=\frac{0.5}{x_1}+\frac{1}{x_2}+\frac{0.7}{x_3}+\frac{0.4}{x_4}$

模糊集的运算

代数积:

- $\mu_{AB}(x) = \mu_A(x)\mu_B(x)$

代数和:

- $\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_{AB}(x)$

有界和:

- $\mu_{A\oplus B}(x) = \min\{1, \mu_A(x) + \mu_B(x)\} = 1 \wedge [\mu_A(x) + \mu_B(x)]$

有界积:

- $\mu_{A\otimes B}(x) = \max\{0, \mu_A(x) + \mu_B(x) - 1\} = 0 \vee [\mu_A(x) + \mu_B(x) - 1]$

模糊集的运算

设有论域 $U=\{x_1,x_2,x_3,x_4,x_5\}$，$A$ 及 B 是论域上的两个模糊集，已知：

- $A=\frac{0.2}{x_1}+\frac{0.4}{x_2}+\frac{0.9}{x_3}+\frac{0.5}{x_5}$
- $B=\frac{0.1}{x_1}+\frac{0.7}{x_3}+\frac{1}{x_4}+\frac{0.3}{x_5}$

求 $A\cdot B, A+B, A\oplus B, A\otimes B$

模糊集的运算

设有论域 $U=\{x_1,x_2,x_3,x_4,x_5\}$，$A$ 及 B 是论域上的两个模糊集，已知：

- $A=\frac{0.2}{x_1}+\frac{0.4}{x_2}+\frac{0.9}{x_3}+\frac{0.5}{x_5}$
- $B=\frac{0.1}{x_1}+\frac{0.7}{x_3}+\frac{1}{x_4}+\frac{0.3}{x_5}$

求 $A\cdot B, A+B, A\oplus B, A\otimes B$

- $A\cdot B=\frac{0.02}{x_1}+\frac{0.63}{x_3}+\frac{0.15}{x_5}$
- $A+B=\frac{0.28}{x_1}+\frac{0.4}{x_2}+\frac{0.97}{x_3}+\frac{1}{x_4}+\frac{0.65}{x_5}$
- $A\oplus B=\frac{0.3}{x_1}+\frac{0.4}{x_2}+\frac{1}{x_3}+\frac{1}{x_4}+\frac{0.8}{x_5}$
- $A\otimes B=\frac{0.6}{x_3}$

模糊关系

模糊关系的定义

- 以集合 $U_1 \sim U_n$ 的笛卡儿乘积 $U_1 \times \cdots \times U_n$ 为基底集的任一模糊子集 R 称为 $U_1 \times \cdots \times U_n$ 间的一个 n 元模糊关系 (Fuzzy Relation)
- 特别地，U_n 的任一模糊子集称为 U 上的一个 n 元模糊关系

模糊关系的表示

- 在传统的有穷二元关系的表示方法基础上加上隶属度数据，作为二元模糊关系的表示
- 矩阵方法

模糊关系

某地区人的身高论域 $X = \{140, 150, 160, 170, 180\}$（单位：cm），体重论域 $Y = \{40, 50, 60, 70, 80\}$.

R X \ Y	40	50	60	70	80
140	1	0.8	0.2	0.1	0
150	0.8	1	0.8	0.2	0.1
160	0.2	0.8	1	0.8	0.2
170	0.1	0.2	0.8	1	0.8
180	0	0.1	0.2	0.8	1

- 从 X 到 Y 的一个模糊关系 R，用模糊矩阵表示：

$$R = \begin{bmatrix} 1 & 0.8 & 0.2 & 0.1 & 0 \\ 0.8 & 1 & 0.8 & 0.2 & 0.1 \\ 0.2 & 0.8 & 1 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.8 & 1 & 0.8 \\ 0 & 0.1 & 0.2 & 0.8 & 1 \end{bmatrix}$$

模糊关系

A, B 模糊集合，模糊关系用叉积 (Cartesian Product) 表示：

$$R : A \times B \to [0, 1]$$

叉积常用最小算子运算：

$$\mu_{A\times B}(a, b) = \min\{\mu_A(a), \mu_B(b)\}$$

A, B 离散模糊集，其隶属函数分别为：

$$\mu_A = [\mu_A(a_1), \mu_A(a_2), \cdots, \mu_A(a_n)], \quad \mu_B = [\mu_B(b_1), \mu_B(b_2), \cdots, \mu_B(b_n)]$$

模糊关系：

$$\mu_{A\times B}(a, b) = \mu_A^T \circ \mu_B \quad \to \quad R = \mu_A^T \circ \mu_B$$

模糊关系

已知输入的模糊集合 A 和输出的模糊集合 B

- $A = \frac{1}{a_1} + \frac{0.8}{a_2} + \frac{0.5}{a_3} + \frac{0.2}{a_4} + \frac{0}{a_5}$
- $B = \frac{0.7}{b_1} + \frac{1}{b_2} + \frac{0.6}{b_3} + \frac{0}{b_4}$

求 A 到 B 的模糊关系 R

$$R = \mu_A^T \circ \mu_B = \begin{bmatrix} 1 \\ 0.8 \\ 0.5 \\ 0.2 \\ 0 \end{bmatrix} \circ \begin{bmatrix} 0.7 & 1 & 0.6 & 0 \end{bmatrix}$$

模糊关系

$$R = \begin{bmatrix} 1.0 \wedge 0.7 & 1.0 \wedge 1.0 & 1.0 \wedge 0.6 & 1.0 \wedge 0.0 \\ 0.8 \wedge 0.7 & 0.8 \wedge 1.0 & 0.8 \wedge 0.6 & 0.8 \wedge 0.0 \\ 0.5 \wedge 0.7 & 0.5 \wedge 1.0 & 0.5 \wedge 0.6 & 0.5 \wedge 0.0 \\ 0.2 \wedge 0.7 & 0.2 \wedge 1.0 & 0.2 \wedge 0.6 & 0.2 \wedge 0.0 \\ 0.0 \wedge 0.7 & 0.0 \wedge 1.0 & 0.0 \wedge 0.6 & 0.0 \wedge 0.0 \end{bmatrix}$$

$$\Downarrow$$

$$R = \begin{bmatrix} 0.7 & 1.0 & 0.6 & 0.0 \\ 0.7 & 0.8 & 0.6 & 0.0 \\ 0.5 & 0.5 & 0.5 & 0.0 \\ 0.2 & 0.2 & 0.2 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

模糊关系合成

设 Q：U 到 V 的模糊关系，R：V 到 W 的模糊关系，则 Q 与 R 的合成 $Q \circ R$ 为 U 到 W 的一个模糊关系，其隶属函数：

$$\mu_{Q \circ R}(u, w) = \bigcup_{v \in V} (\mu_Q(u, b) \wedge \mu_R(v, w))$$

最大 - 最小合成法

最大 - 代数积合成法

$\cdots$

模糊关系合成

设模糊集合

- $X = \{x_1, x_2, x_3, x_4\}$, $Y = \{y_1, y_2, y_3\}$, $Z = \{z_1, z_2\}$
- $Q \in X \times Y$, $R \in Y \times Z$, $S \in X \times Z$，求 S.

$$Q = \begin{bmatrix} 0.5 & 0.6 & 0.3 \\ 0.7 & 0.4 & 1 \\ 0 & 0.8 & 0 \\ 1 & 0.2 & 0.9 \end{bmatrix} \quad R = \begin{bmatrix} 0.2 & 1 \\ 0.8 & 0.4 \\ 0.5 & 0.3 \end{bmatrix}$$

模糊关系合成

解

$$
\begin{aligned}
S &= Q \circ R \\
&= \begin{bmatrix} 0.5 & 0.6 & 0.3 \\ 0.7 & 0.4 & 1 \\ 0 & 0.8 & 0 \\ 1 & 0.2 & 0.9 \end{bmatrix} \circ \begin{bmatrix} 0.2 & 1 \\ 0.8 & 0.4 \\ 0.5 & 0.3 \end{bmatrix} \\
&= \begin{bmatrix} (0.5 \wedge 0.2) \vee (0.6 \wedge 0.8) \vee (0.3 \wedge 0.5) & (0.5 \wedge 1) \vee (0.6 \wedge 0.4) \vee (0.3 \wedge 0.3) \\ (0.7 \wedge 0.2) \vee (0.4 \wedge 0.8) \vee (1 \wedge 0.5) & (0.7 \wedge 1) \vee (0.4 \wedge 0.4) \vee (1 \wedge 0.3) \\ (0 \wedge 0.2) \vee (0.8 \wedge 0.8) \vee (0 \wedge 0.5) & (0 \wedge 1) \vee (0.8 \wedge 0.4) \vee (0 \wedge 0.3) \\ (1 \wedge 0.2) \vee (0.2 \wedge 0.8) \vee (0.9 \wedge 0.5) & (1 \wedge 1) \vee (0.2 \wedge 0.4) \vee (0.9 \wedge 0.3) \end{bmatrix} \\
&= \begin{bmatrix} 0.6 & 0.5 \\ 0.5 & 0.7 \\ 0.8 & 0.4 \\ 0.5 & 1 \end{bmatrix}
\end{aligned}
$$

模糊推理

人类思维判断的基本形式:

- 如果（条件）$\to$ 则（结论）
- 如：如果 压力较高且温度在慢慢上升 $\to$ 阀门略开

模糊规则：从条件论域到结论论域的模糊关系矩阵 R.

通过条件模糊向量与模糊关系 R 的合成进行模糊推理，得到结论的模糊向量.

然后采用“清晰化”方法将模糊结论转换为精确量.

模糊推理

对 $IF\ A\ THEN\ B$ 类型的模糊规则的推理

- 若已知输入为 A，则输出为 B；若现在已知输入为 A'，则输出 B' 用合成规则求取 $B' = A' \circ R$
- 其中模糊关系 R： $\mu_R(x,y) = \min[\mu_A(x), \mu_B(y)]$
- 控制规则库的 N 条规则有 N 个模糊关系： $R_1, R_2, \cdots, R_n$
- 对于整个系统的全部控制规则所对应的模糊关系 R :

$$R = R_1 \cup R_2 \cup \cdots \cup R_n = \bigcup_{i=1}^{n} R_i$$

模糊推理

已知输入的模糊集合 A 和输出的模糊集合 B:

- $A = \frac{1}{a_1} + \frac{0.8}{a_2} + \frac{0.5}{a_3} + \frac{0.2}{a_4} + \frac{0}{a_5}$
- $B = \frac{0.7}{b_1} + \frac{1}{b_2} + \frac{0.6}{b_3} + \frac{0}{b_4}$

已经求得模糊关系为:

$$R = \begin{bmatrix} 0.7 & 1 & 0.6 & 0 \\ 0.7 & 0.8 & 0.6 & 0 \\ 0.5 & 0.5 & 0.5 & 0 \\ 0.2 & 0.2 & 0.2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

模糊推理

当输入 $A' = \frac{0.4}{a_1} + \frac{0.7}{a_2} + \frac{1}{a_3} + \frac{0.6}{a_4} + \frac{0}{a_5}$

$$B' = A' \circ R = \begin{bmatrix} 0.4 \\ 0.7 \\ 1 \\ 0.6 \\ 0 \end{bmatrix}^T \circ \begin{bmatrix} 0.7 & 1 & 0.6 & 0 \\ 0.7 & 0.8 & 0.6 & 0 \\ 0.5 & 0.5 & 0.5 & 0 \\ 0.2 & 0.2 & 0.2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = (0.7, 0.7, 0.6, 0)$$

则 $B' = \frac{0.7}{b_1} + \frac{0.7}{b_2} + \frac{0.6}{b_3} + \frac{0}{b_4}$

模糊决策

模糊决策：由模糊推理得到的结论或者操作是一个模糊向量，转化为确定值的过程；

1. 最大隶属度法：

- 如，得到模糊向量

$$U' = \frac{0.5}{-3} + \frac{0.5}{-2} + \frac{0.5}{-1} + \frac{0}{0} + \frac{0}{1} + \frac{0}{2} + \frac{0}{3}$$

- 取 $U = \dfrac{-3-2-1}{3} = -2$

模糊决策

模糊决策：由模糊推理得到的结论或者操作是一个模糊向量，转化为确定值的过程；

2. 加权平均法：

$$U=\frac{\sum_{i=1}^{n}\mu(u_i)u_i}{\sum_{i=1}^{n}\mu(u_i)}$$

- 如：$U'=\frac{0.1}{2}+\frac{0.6}{3}+\frac{0.5}{4}+\frac{0.4}{5}+\frac{0.2}{6}$
- 则：$U'=\frac{0.1\times 2+0.6\times 3+0.5\times 4+0.4\times 5+0.2\times 6}{0.1+0.6+0.5+0.4+0.2}=4$

模糊推理应用

设有模糊控制规则:

- "如果温度低,则将风门开大".设温度和风门开度的论域为 $\{1,2,3,4,5\}$.
- "温度低" 和 "风门大" 的模糊量:
  - "温度低" $= \frac{1}{1} + \frac{0.6}{2} + \frac{0.3}{3} + \frac{0}{4} + \frac{0}{5}$
  - "风门大" $= \frac{0}{1} + \frac{0}{2} + \frac{0.3}{3} + \frac{0.6}{4} + \frac{1}{5}$
- 已知事实 "温度较低", 可以表示为:
  - 温度较低" $= \frac{0.8}{1} + \frac{1}{2} + \frac{0.6}{3} + \frac{0.3}{4} + \frac{0}{5}$
- 试用模糊推理确定风门开度

模糊推理应用

解：

- 确定模糊关系 R

$$R = \begin{bmatrix} 1 \\ 0.6 \\ 0.3 \\ 0 \\ 0 \end{bmatrix} \circ \begin{bmatrix} 0 & 0 & 0.3 & 0.6 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0.3 & 0.6 & 1 \\ 0 & 0 & 0.3 & 0.6 & 0.6 \\ 0 & 0 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

模糊推理应用

解:

- 模糊推理

$$B' = A' \circ R = \begin{bmatrix} 0.8 \\ 1 \\ 0.6 \\ 0.3 \\ 0 \end{bmatrix}^T \circ \begin{bmatrix} 0 & 0 & 0.3 & 0.6 & 1 \\ 0 & 0 & 0.3 & 0.6 & 0.6 \\ 0 & 0 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = (0, 0, 0.3, 0.6, 0.8)$$

模糊推理应用

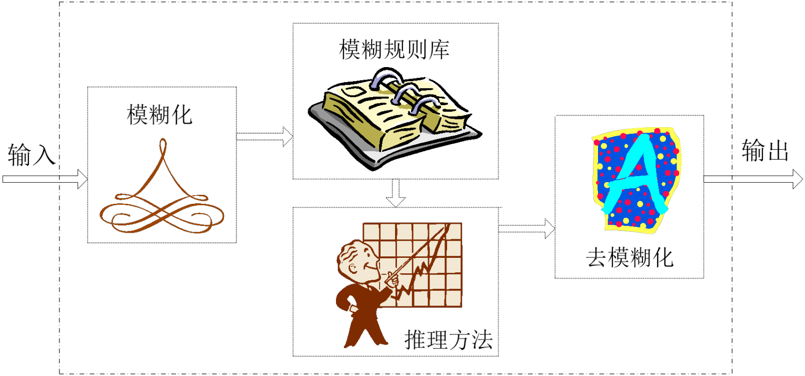
解:

- 模糊决策
- 用最大隶属度法进行决策得风门开度为 5
- 用加权平均判决法进行决策得风门开度为 4

模糊控制应用场景

- 热交换过程的控制
- 暖水工厂的控制
- 污水处理过程控制
- 交通路口控制
- 水泥窑控制
- 飞船飞行控制
- 机器人控制
- 车停靠和转弯控制
- 汽车速度控制
- 水质净化控制
- 电梯控制
- 核反应堆的控制

模糊控制



内容组织

1. 模糊推理
2. 概率推理
3. 朴素贝叶斯推理

概率论

分配赌金问题：

- 甲乙两人赌技相同，各出赌注 500 元
- 约定：谁先胜三局，则谁拿走全部 1000 元
- 现已赌了三局，甲二胜一负而因故要中止赌博
- 问这 1000 元要如何分才算公平?

概率论

分配赌金问题：

- 甲乙两人赌技相同，各出赌注 500 元
- 约定：谁先胜三局，则谁拿走全部 1000 元
- 现已赌了三局，甲二胜一负而因故要中止赌博
- 问这 1000 元要如何分才算公平？

甲：750，乙：250

概率论

古典概率

- 组合数学

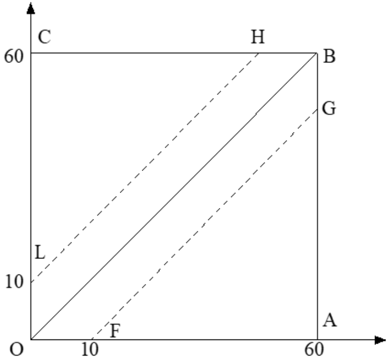
几何概率

- 甲、乙二人约定 1 点到 2 点之间在某处碰头
- 约定先到者等候 10 分钟即离去
- 设想甲、乙二人各自随意地在 $1 \sim 2$ 点之间选一个时刻到达该处
- 问“甲乙二人能碰上”这事件 E 的概率是多少?

概率论

- 甲、乙二人约定 1 点到 2 点之间在某处碰头，约定先到者等候 10 分钟即离去；设想甲、乙二人各自随意地在 $1\sim2$ 点之间选一个时刻到达该处．问“甲乙二人能碰上”这事件 E 的概率是多少?

$$P(E)=\frac{1100}{3600}$$



概率及其公理

随机变量

- 布尔随机变量
- 离散随机变量
- 连续随机变量

原子事件

- 组成世界的所有随机变量的特定赋值
- 构成无法确定的世界状态的完整详细描述
- 如 X 的世界由 $weather =< sunny, rainy, cloudy, snow >$ 和今天是否喝酒 $drink\_today =< T, F >$ 组成，则有 4×2 种不同原子事件

原子事件的性质

1. 原子事件是互斥的：$sunny \wedge drink\_today$ 和 $sunny \wedge \neg drink\_today$ 不能同时成立
2. 由所有原子事件组成的集合是穷尽的
3. 任何特定的原子事件与每个命题 (简单或者复合命题) 的真或假一一对应
4. 任何一个命题逻辑上都等价于所有蕴涵该命题真值的原子事件的析取
    - $sunny$ 等价于 $sunny \wedge drink\_today \vee sunny \wedge \neg drink\_today$

先验概率的表示

- 先验概率：没有任何其它信息存在情况下关于某个命题的信度
- 用向量表示随机变量的先验概率分布
  - $P(weather) =< 0.7, 0.2, 0.08, 0.02 >$
- 对于组成世界的离散随机变量全集，使用诸如：$P(weather, drink\_today)$ 来表示涵盖全集的随机变量集的值的全部组合的概率：全联合概率分布
- 全联合概率分布用概率表表示
  - $P < weather, drink\_today >$ 用 4×2 表格表示

条件概率的表示

条件概率定义

- $P(a|b) = \dfrac{P(a \wedge b)}{P(b)}$

由此有乘法定理

- $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$

如果 a 和 b 相互独立，则

- $P(a|b) = P(a)$
- $P(b|a) = P(b)$
- $P(a \wedge b) = P(a)P(b)$

主观 Bayes

主观 Bayes 概率:

- 现实世界的一些因果关系可以形成一种信念，它并非在所有场合下都正确，可称为部分信念.
- 表示这种信念的最好方法是概率方法
- 对概率的解释有若干种，其中一种称为主观 Bayes 主义
  - 概率是个人的一种合理置信度，每个人的估计 (概率) 虽然各不相同，但应该满足概率的基本规律和其他某些客观规律，因而是合理的

概率公理

Bayes 概率服从如下公理 (Kolmogorov 公理):

1. $0 \leq P(a) \leq 1$
2. $P(T) = 1, P(F) = 0$
3. $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$
   - 当 a, b 互斥有 $P(a \vee b) = P(a) + P(b)$ 此为加法定理

这样的概率公理是不能违反的

全概率公式

- 任何命题 a 等价于所有 a 在其中成立的原子事件的析取，事件集合记为 $e(a)$
- 由所有原子事件是互斥的，得到如下全联合概率分布

$$P(a) = P(e_1 \vee e_2 \vee \cdots \vee e_k) = \sum_{e_i \in e(a)} P(e_i)$$

- 一个命题的概率等于所有它在其中成立的原子事件的概率和

使用全联合概率分布进行推理

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- 全联合概率分布是知识库，从中可得到所有概率的计算
- 命题在其中成立的所有原子事件的概率和
  - $P(cavity \lor toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$
  - $P(catch) = 0.108 + 0.016 + 0.072 + 0.144 = 0.34$

归一化

- 大多数情况下我们对计算某个变量的条件概率感兴趣：

$$P(cavity|toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)} = 0.6$$

$$P(\neg cavity|toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)} = 0.4$$

- $\frac{1}{P(toothache)}$ 保持不变，可把它看成是保证其所包含的概率相加为 1 的常数.
- 引入归一化常数：$\alpha = \frac{1}{[p(a) + p(\neg a)]}$
- 一般公式：$P(X|e) = \alpha P(X, e) = \alpha \sum_y P(X, e, y)$（根据全概率公式）

内容组织

1. 模糊推理
2. 概率推理
3. 朴素贝叶斯推理

Bayes 公式

- Bayes 公式 (也称逆概率公式)
- 从条件概率公式可得

$$P(b|a) = \frac{P(a|b) \times P(b)}{P(a)}$$

- 在某些场合下引入一个证据 e 以后，得更通用的 Bayes 公式

$$P(Y|X, e) = \frac{P(X|Y, e) \times P(Y|e)}{P(X|e)}$$

逆概率公式的例子

- 逆概率公式不仅是条件概率公式的一个简单变形，实际上很有用处
- 若某个条件概率不便计算，则可以先计算其逆概率，而后算出所要的条件概率
- 例: 求 P(肺炎 | 咳嗽) 可能比较困难, 但统计 P(咳嗽 | 肺炎) 可能比较容易 (因为要上医院); 假设 P(肺炎)=1/10000, 而 P(咳嗽)=1/10, 90% 的肺炎患者都咳嗽，则

$$P(\text{肺炎}|\text{咳嗽}) = \frac{P(\text{咳嗽}|\text{肺炎}) \times P(\text{肺炎})}{P(\text{咳嗽})} = \frac{0.9 \times 0.0001}{0.1} = 0.0009$$

修正因子

- 将前面的逆概率公式写成

$$P(H|E) = \left[\frac{P(E|H)}{P(E)}\right] P(H)$$

- 先验概率 $P(H)$ 通过方括号部分作为修正因子，修正为后验概率 $P(H|E)$(证据 E 为真时 H 的后验概率)
- 在上面的例子中，医生认为一个人得肺炎的可能性为万分之一，一旦发现患者咳嗽，就将调整为万分之九

修正因子

- 将 E 看作证据，先验概率 $P(E)$ 越小，且 H 为真时 E 的条件概率 $P(E|H)$ 越大，则修正因子所起作用越大
- 在上例中，如果
  - P(咳嗽)=0.0001， P(咳嗽｜肺炎)=0.9999， P(肺炎) 不变
  - 则 P(肺炎｜咳嗽)=0.9999，远远超过原来的万分之九

后验概率递推公式

- 当有 n 个互相独立的证据，则有公式

$$P(H|E_1\&\cdots\&E_n)=\frac{\prod_{i=1}^{n}P(E_i|H)}{\prod_{i=1}^{n}P(E_i)}\times P(H)$$

- 上式可以写成递推公式形式：

$$P(H|E_1\&\cdots\&E_{m+1})=\frac{P(E_{m+1}|H)}{P(E_{m+1})}\times P(H|E_1\&\cdots\&E_m)$$

- 说明：随着新证据的不断获得，从证据少时的后验概率推出证据多时的后验概率，且每一步都是把上一步的后验概率视为新证据到来时的先验概率.

条件独立性

- 给定第 3 个随机变量 Z 后，X, Y 的条件独立定义为：
  - $P(X, Y|Z) = P(X|Z)P(Y|Z)$ 或
  - $P(X|Y, Z) = P(X|Z)$ 或
  - $P(Y|X, Z) = P(Y|Z)$
- 则牙科领域 3 个随机变量有：
  - $P(Toothache, Catch|Cavity) = P(Toothache|Cavity)P(Catch|Cavity)$
  - $P(Toothache, Cavity, Catch) = P(To, Cat|Cav)P(Cav)$

    $= P(To|Cav)P(Cat|Cav)P(Cav)$

条件独立性的结果

- 条件概率表 (CPT) 的分解
  - 原概率表有 7 个彼此独立的数值 $(23-1)$
  - 新概率表有 5 个独立数值 $(2+2+1)$
  - n 个变量彼此独立后，表示的规模从 $O(2^n)$ 变为 $O(n)$
- 条件独立性允许概率系统进行规模的扩展
- 条件独立性比绝对独立性更容易获得
- 此结论导致了朴素贝叶斯模型
  - $P(Cause, Effect_1, \cdots, Effect_n) = \left(\prod P(E_i|C)\right) P(C)$

独立性条件下的推理

- 使用全联合分布表，可以进行查询 (推理)，但只适用于变量少的情况
  - N 个可能证据变量，则相关条件概率的组合数达到 2^N
- 条件独立性
  - 一旦某个变量的取值确定下来，则其余变量就相互独立
- 对于 $toothache, cavity, catch$ 三者，$cavity$ 决定了其余两者，而它们彼此之间无关系
  - $P(toothache \wedge catch | Cavity) = P(toothache | Cavity) \times P(catch | Cavity)$

Question & Answer