

人工智能

人工智能概述

主讲: 赵国亮

内蒙古大学电子信息工程学院

May 22, 2020

目录 I

- 1 不确定性推理
 - 统计概率的性质
- 2 证据理论
 - 主观 Bayes 方法
- 3 DS 理论的形式描述
 - 证据理论的推理模型
- 4 模糊推理

不精确性推理方法

现实世界中的大多数问题是不精确、非完备的. 对于这些问题, 若采用前面所讨论的精确性推理方法显然是无法解决的. 为此, 人工智能需要研究不精确性的推理方法, 以满足客观问题的需求.



不确定性推理的定义

什么是不确定性推理

- 不确定性推理泛指除精确推理以外的其它各种推理问题. 包括不完备、不精确知识的推理, 模糊知识的推理, 非单调性推理等.
- 不确定性推理过程实际上是一种从不确定的初始证据出发, 通过运用不确定性知识, 推出具有一定不确定性但却又是合理或基本合理的结论的思维过程.

为什么要采用不确定性推理

- 所需知识不完备.
- 不精确所需知识描述模糊.
- 多种原因导致同一结论.
- 问题的背景知识不足.
- 解题方案不唯一.

不确定性推理的基本问题

(1) 知识的不确定性的表示

- 考虑因素: 问题的描述能力, 推理中不确定性的计算.
- 含义: 知识的确定性程度, 或动态强度.
- 表示: 用概率 $[0,1]$, 0 接近于假, 1 接近于真. 用可信度 $[-1,1]$, 大于 0 接近于真, 小于 0 接近于假.

(2) 证据的非精确性表示

证据来源: 初始证据, 中间结论. 表示方式: 用概率或可信度.

• 含义

不确定的前提条件与不确定的事实匹配.

• 问题

前提是不确定的, 事实也是不确定的

• 方法

设计一个计算相似程度的算法, 给出相似的限度.

解决方法

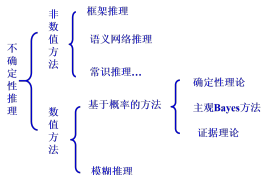
对① 不同推理方法的解决方法不同。

对② 不同推理方法的解决方法基本相同, 即把当前结论及其不确定性作为新的结论放入综合数据库, 依次传递, 直到得出最终结论。

非精确性结论的合成

含义: 多个不同知识推出同一结论, 且不确定性程度不同。

方法: 视不同推理方法而定。



不确定性推理的概率论基础

在概率论中, 把试验中每一个可能出现的结果称为试验的一个样本点, 由全体样本点构成的集合称为样本空间.

表示: 通常用 D 表示样本空间, d 表示样本点.

概念: 由样本点构成的集合称为随机事件.

Example

1. 在掷币试验中, 若用 d_1 表示硬币的正面向上, 用 d_2 表示硬币的反面向上, 则该试验的样本空间为: $D = \{d_1, d_2\}$.
2. 在掷币试验中, 若用 A 表示硬币正面向上这一事件, 则有 $A = \{d_1\}$.

概率运算

常见运算

- 并事件: 事件 A 与事件 B 至少有一个发生, 记为 $A \cup B$.
- 交事件: 事件 A 与事件 B 同时发生, 记为 $A \cap B$.
- 互逆事件: 事件 A 与 B 之间满足 “ $A \cap B = \emptyset, A \cup B = D$ ” .

概率的统计定义

是通过某一事件出现的频率定义的. 频率:

$$f_n(A) = m/n, \quad (1)$$

式中, A 所讨论的事件, n 是试验的总次数, m 是实验中 A 发生的次数. 在同一组条件下所进行大量重复试验时, 如果事件 A 出现的频率总是在区间 $[0, 1]$ 上的一个确定常数 p 附近摆动, 并且稳定于 p , 则称 p 为事件 A 的统计概率. 即

$$P(A) = p. \quad (2)$$

Example

在掷币试验中, 当掷币次数足够多时有

$$f_n(\text{正面向上}) = 0.5. \quad (3)$$

则称正面向上的概率为 0.5, 即

$$P(\text{正面向上}) = 0.5. \quad (4)$$

统计概率的性质

性质

- (1) 对任一事件 A , 有 $0 \leq P(A) \leq 1$.
- (2) 必然事件 D 的概率 $P(D) = 1$, 不可能事件 Φ 的概率 $P(\emptyset) = 0$.
- (3) 对任一事件 A , 有 $P(\neg A) = 1 - P(A)$.
- (4) 设事件 $A_1, A_2, \dots, A_k (k \leq n)$ 是两两互不相容的事件, 即有 $A_i \cap A_j = \emptyset (i \neq j)$, 则 $P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.
- (5) 设 A, B 是两个事件, 则 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

由于事件 B 已经发生, 因此以下事件 {取到红桃 A; 取到红桃 2; 取到红桃 3; ...; 取到红桃 K} 中必有一个出现.

而对事件 A, 在事件 B 发生的前提下, 只有以下事件 {取到红桃 J; 取到红桃 Q; 取到红桃 K 中的一个} 发生时事件 A 才能发生.

因此, 在事件 B 发生的条件下事件 A 发生的概率是 $3/13$.

全概率公式

设事件 A_1, A_2, \dots, A_n 满足:

- (1) 任意两个事件都互不相容, 即当 $i \neq j$ 时, 有 $A_i \cap A_j = \emptyset$, ($i = 1, 2, \dots, n$);
- (2) $P(A_i) > 0$ ($i = 1, 2, \dots, n$);
- (3) $D = \bigcup_{n=1}^n A_n$, 则对任何事件 B 由下式成立:

$$P(B) = \sum_{i=1}^n P(A_i) \times P(B|A_i). \quad (6)$$

该公式称为全概率公式, 它提供了一种计算 $P(B)$ 的方法.

Bayes 定理

设事件 A_1, A_2, \dots, A_n 满足定理 6.1 规定的条件, 则对任何事件 B 有下式成立:

$$P(A_i|B) = \frac{P(A_i) \times P(B/A_i)}{\sum_{j=1}^n P(A_j) \times P(B/A_j)}, i = 1, 2, \dots, n \quad (7)$$

该定理称为 Bayes 定理, 上式称为 Bayes 公式. 其中 $P(A_i)$ 是事件 A_i 的先验概率, $P(B/A_i)$ 是在事件 A_i 发生条件下事件 B 的条件概率; $P(A_i|B)$ 是在事件 B 发生条件下事件 A_i 的条件概率.

如果把全概率公式代入 Bayes 公式, 则有:

$$P(A_i|B) = \frac{P(A) \times P(B|A_i)}{P(B)}, i = 1, 2, \dots, n. \quad (8)$$

即

$$P(A_i|B) \times P(B) = P(B|A_i) \times P(A_i) \quad i = 1, 2, \dots, n. \quad (9)$$

这是 Bayes 公式的另一种形式. Bayes 定理给出了用逆概率 $P(B|A_i)$ 求原概率 $P(A_i|B)$ 的方法.

可信度的概念

可信度

是指人们根据以往经验对某个事物或现象为真的程度的一个判断, 或者说是人们对某个事物或现象为真的相信程度.

例 2.1

沈强昨天没来上课, 理由是头疼. 就此理由, 只有以下两种可能: 一 是真的头疼了, 理由为真; 二是没有头疼, 理由为假. 但就听话人而言, 因不能确切知道, 就只能某种程度上相信, 即可信度.



CF 模型

可信度具有一定的主观性, 较难把握. 但对某一特定领域, 让该领域专家给出可信度还是可行的.

知识不确定性的表示形式

在 CF 模型中, 知识是用产生式规则表示的, 其一般形式为:

IF E THEN H (CF(H, E)),

其中, E 是知识的前提条件; H 是知识的结论; CF(H, E) 是知识的可信度.

可信度的定义与性质

(1) E 可以是单一条件, 也可以是复合条件. 例如:

$$E = (E_1 \text{ OR } E_2) \text{ AND } E_3 \text{ AND } E_4$$

(2) H 可以是单一结论, 也可以是多个结论

(3) CF 是知识的静态强度, $CF(H, E)$ 的取值为 $[-1, 1]$, 表示当 E 为真时, 证据对 H 的支持程度, 其值越大, 支持程度越大.

Example

IF 发烧 AND 流鼻涕 THEN 感冒 (0.8)

表示当某人确实有“发烧”及“流鼻涕”症状时, 则有 80% 的把握是患了感冒.

可信度的定义

在 CF 模型中, 把 $CF(H, E)$ 定义为

$$CF(H, E) = MB(H, E) - MD(H, E)$$

式中 MB 称为信任增长度, $MB(H, E)$ 定义为

$$MB(H, E) = \begin{cases} 1, & \text{若 } P(H) = 1 \\ \frac{\max\{P(H|E), P(H)\} - P(H)}{1 - P(H)}, & \text{否则} \end{cases} \quad (10)$$

MB 和 MD 的关系

MD 称为不信任增长度, $MB(H, E)$ 定义为

$$MD(H, E) = \begin{cases} 1, & \text{若 } P(H) = 0 \\ \frac{\min\{P(H|E), P(E)\} - P(H)}{1 - P(H)}, & \text{否则} \end{cases} \quad (11)$$

- ① 当 $MB(H, E) > 0$ 时, 有 $P(H|E) > P(H)$, 即 E 的出现增加了 H 的概率;
- ② 当 $MD(H, E) > 0$ 时, 有 $P(H|E) < P(H)$, 即 E 的出现降低了 H 的概率.

根据前面对 $CF(H, E)$ 可信度、 $MB(H, E)$ 信任增长度、 $MD(H, E)$ 不信增长度的定义, 可得到 $CF(H, E)$ 的计算公式:

$$CF(H, E) = \begin{cases} MB(H, E) - 0 = \frac{P(H|E) - P(H)}{1 - P(H)} & \text{若 } P(H|E) > P(H) \\ 0 & \text{若 } P(H|E) = P(H) \\ 0 - MD(H, E) = -\frac{P(H) - P(H|E)}{P(H)} & \text{若 } P(H|E) < P(H) \end{cases} \quad (12)$$

分别解释 $CF(H, E) > 0$, $CF(H, E) = 0$, $CF(H, E) < 0$.

可信度的性质

• 互斥性

对同一证据, 它不可能既增加对 H 的信任程度, 又同时增加对 H 的不信任程度, 这说明 MB 与 MD 是互斥的. 即有如下互斥性:

- 当 $MB(H, E) > 0$ 时, $MD(H, E) = 0$.
- 当 $MD(H, E) > 0$ 时, $MB(H, E) = 0$.

当 $CF(H, E) = 1$ 时, 有 $P(H/E) = 1$, 它说明由于 E 所对应证据的出现使 H 为真. 此时, $MB(H, E) = 1, MD(H, E) = 0$.

当 $CF(H, E) = -1$ 时, 有 $P(H/E) = 0$, 说明由于 E 所对应证据的出现使 H 为假. 此时, $MB(H, E) = 0, MD(H, E) = 1$.

当 $CF(H, E) = 0$ 时, 有 $MB(H, E) = 0, MD(H, E) = 0$. 前者说明 E 所对应证据的出现不证实 H ; 后者说明 E 所对应证据的出现不否认 H .

- 典型值对 H 的信任增长度等于对非 H 的不信任增长度.

$$\begin{aligned}
 MD(\neg H, E) &= \frac{P(\neg H|E) - P(\neg H)}{-P(\neg H)} = \frac{(1 - P(H|E)) - (1 - P(H))}{-(1 - P(H))} \\
 &= \frac{-P(H|E) + P(H)}{-(1 - P(H))} = \frac{-(P(H|E) - P(H))}{-(1 - P(H))} \\
 &= \frac{P(H|E) - P(H)}{1 - P(H)} = MB(H, E), \quad \text{信任增长度} \quad (13)
 \end{aligned}$$

再根据 CF 的定义和 MB、MD 的互斥性有

$$\begin{aligned} CF(H, E) + CF(\neg H, E) &= (MB(H, E) - MD(H, E)) + (MB(\neg H, E) - MD(\neg H, E)) \\ &= (MB(H, E) - 0) + (0 - MD(\neg H, E)), \quad (\text{由互斥性}) \\ &= MB(H, E) - MD(\neg H, E) = 0 \end{aligned}$$

它说明

- (1) 对 H 的信任增长度等于对非 H 的不信任增长度
- (2) 对 H 的可信度与非 H 的可信度之和等于 0
- (3) 可信度不是概率, 不满足

$$P(H) + P(\neg H) = 1, 0 \leq P(H), P(\neg H) \leq 1.$$

(5) 对同一前提 E , 若支持若干个不同的结论 H_i ($i = 1, 2, \dots, n$), 则如果发现专家给出的知识有如下情况

$$CF(H_1, E) = 0.7, CF(H_2, E) = 0.4.$$

则因 $0.7 + 0.4 = 1.1 > 1$ 为非法, 应进行调整或规范化.

证据不确定性的表示

不确定性的表示

证据的不确定性也是用可信度来表示的, 其取值范围也为 $[-1, 1]$.

- 若 E 为初始证据, 其值由用户给出.
- 若 E 为中间结论, 其值可通过计算得到.

可信度的含义

- $CF(E) = 1$, 证据 E 肯定它为真.
- $CF(E) = -1$, 证据 E 肯定它为假.
- $CF(E) = 0$, 对证据 E 一无所知.
- $0 < CF(E) < 1$, 证据 E 以 $CF(E)$ 程度为真.
- $-1 < CF(E) < 0$, 证据 E 以 $CF(E)$ 程度为假.

否定证据不确定性的计算

$$CF(\neg E) = -CF(E). \quad (14)$$

组合证据不确定性的计算

对证据的组合形式可分为“合取”与“析取”两种基本情况.

- 合取: 当组合证据是多个单一证据的组合时, 即 $E = E_1 \text{ AND } E_2 \text{ AND } \cdots \text{ AND } E_n$ 时, 若已知 $CF(E_1), CF(E_2), \cdots, CF(E_n)$, 则

$$CF(E) = \min\{CF(E_1), CF(E_2), \cdots, CF(E_n)\}. \quad (15)$$

• 析取

当组合证据是多个单一证据的析取时, 即 $E = E_1 \text{ OR } E_2 \text{ OR } \cdots \text{ OR } E_n$ 时, 若已知 $CF(E_1), CF(E_2), \cdots, CF(E_n)$, 则

$$CF(E) = \max\{CF(E_1), CF(E_2), \cdots, CF(E_n)\}. \quad (16)$$

CF 模型中的不确定性推理实际上是从不确定的初始证据出发, 不断运用相关的不确定性知识, 逐步推出最终结论和该结论可信度的过程.

不确定性的更新

每一次运用不确定性知识, 都需要由证据的不确定性和知识的不确定性去计算结论的不确定性.

不确定性的更新公式

$$CF(H) = CF(H, E) \times \max\{0, CF(E)\}. \quad (17)$$

若 $CF(E) < 0$, 则

$$CF(H) = 0, \quad (18)$$

即该模型没考虑 E 为假对 H 的影响. 若 $CF(E) = 1$, 则

$$CF(H) = CF(H, E), \quad (19)$$

即规则强度 $CF(H, E)$ 实际上是在 E 为真时, H 的可信度.

结论不确定性的合成

当有多条知识支持同一个结论, 且这些知识的前提相互独立, 结论的可信度又不相同时, 可利用不确定性的合成算法求出结论的综合可信度. 设有知识:

$$\text{IF } E_1 \text{ THEN } H(\text{CF}(H, E_1)), \quad (20)$$

$$\text{IF } E_2 \text{ THEN } H(\text{CF}(H, E_2)). \quad (21)$$

则结论 H 的综合可信度可分以下两步计算:

(1) 分别对每条知识求出其 $\text{CF}(H)$. 即

$$\text{CF}_1(H) = \text{CF}(H, E_1) \times \max\{0, \text{CF}(E_1)\}, \quad (22)$$

$$\text{CF}_2(H) = \text{CF}(H, E_2) \times \max\{0, \text{CF}(E_2)\}. \quad (23)$$

(2) 用如下公式求 E_1 与 E_2 对 H 的综合可信度

$$CF(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H), & \text{若 } CF_1(H) \geq 0, \text{ 且 } CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H) \times CF_2(H), & \text{若 } CF_1(H) < 0 \text{ 且 } CF_2(H) < 0 \\ \frac{CF_1(H) + CF_2(H)}{1 - \min\{|CF_1(H)|, |CF_2(H)|\}}, & CF_1(H) < 0 \text{ 且 } CF_2(H) \text{ 异号} \end{cases} \quad (24)$$

Example

设有如下一组知识:

r_1 : IF E1 THEN H(0.9) r_2 : IF E2 THEN H(0.6) r_3 : IF E3 THEN H(-0.5) r_4 : IF E4
AND (E5 OR E6) THEN E1(0.8)

已知: $CF(E_2) = 0.8$, $CF(E_3) = 0.6$, $CF(E_4) = 0.5$, $CF(E_5) = 0.6$, $CF(E_6) = 0.8$, 求
 $CF(H)$.

由 r_4 得到:

$$\begin{aligned}
 CF(E_1) &= 0.8 \times \max\{0, CF(E_4 \text{ AND } (E_5 \text{ OR } E_6))\} \\
 &= 0.8 \times \max\{0, \min\{CF(E_4), CF(E_5 \text{ OR } E_6)\}\} \\
 &= 0.8 \times \max\{0, \min\{CF(E_4), \max\{CF(E_5), CF(E_6)\}\}\} \\
 &= 0.8 \times \max\{0, \min\{CF(E_4), \max\{0.6, 0.8\}\}\} \\
 &= 0.8 \times \max\{0, \min\{0.5, 0.8\}\} \\
 &= 0.8 \times \max\{0, 0.5\} = 0.4.
 \end{aligned}$$

由 r_1

$$\begin{aligned}
 CF_1(H) &= CF(H, E_1) \times \max\{0, CF(E_1)\} \\
 &= 0.9 \times \max\{0, 0.4\} = 0.36.
 \end{aligned} \tag{25}$$

由 r_2 得到:

$$\begin{aligned} CF_2(H) &= CF(H, E_2) \times \max\{0, CF(E_2)\} \\ &= 0.6 \times \max\{0, 0.8\} = 0.48. \end{aligned} \quad (26)$$

由 r_3 得到:

$$\begin{aligned} CF_3(H) &= CF(H, E_3) \times \max\{0, CF(E_3)\} \\ &= -0.5 \times \max\{0, 0.6\} = -0.3. \end{aligned} \quad (27)$$

根据结论不精确性的合成算法, $CF_1(H)$ 和 $CF_2(H)$ 同号, 有:

$$\begin{aligned} CF_{1,2}(H) &= CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H) \\ &= 0.36 + 0.48 - 0.36 \times 0.48 \\ &= 0.84 - 0.17 = 0.67. \end{aligned} \quad (28)$$

$CF_{12}(H)$ 和 $CF_3(H)$ 异号, 有:

$$\begin{aligned} CF_{1,2,3}(H) &= \frac{CF_{1,2}(H) + CF_3(H)}{1 - \min\{|CF_{1,2}(H)|, |CF_3(H)|\}} \\ &= \frac{0.67 - 0.3}{1 - \min\{0.67, 0.3\}} = \frac{0.37}{0.7} \\ &= 0.53. \end{aligned} \tag{29}$$

主观 Bayes 方法

在主观 Bayes 方法中, 知识是用产生式表示的, 其形式为:

$$\text{IF } E \text{ THEN } (LS, LN) H, \quad (30)$$

其中, (LS, LN) 用来表示该知识的知识强度, LS(充分性度量) 和 LN(必要性度量) 的表示形式分别为:

$$LS = \frac{P(E|H)}{P(E|\neg H)}. \quad (31)$$

$$LN = \frac{P(\neg E|H)}{P(\neg E|\neg H)} = \frac{1 - P(E|H)}{1 - P(E|\neg H)}. \quad (32)$$

LS 和 LN 含义的讨论

由本章第 7 节的 Bayes 公式可知:

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)}, \quad (33)$$

$$P(\neg H|E) = \frac{P(E|\neg H) \times P(\neg H)}{P(E)}. \quad (34)$$

两式相除得:

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{P(E|H)}{P(E|\neg H)} \times \frac{P(H)}{P(\neg H)}. \quad (35)$$

为讨论方便, 下面引入概念 几率函数 1

$$O(X) = \frac{P(X)}{1 - P(X)}, P(X) = \frac{O(X)}{1 + O(X)}. \quad (36)$$

可见, X 的几率等于 X 出现的概率与 X 不出现的概率之比, $P(X)$ 与 $O(X)$ 的变化一致, 且有:

$$P(X) = 0, O(X) = 0; \quad (37)$$

$$P(X) = 1, O(X) = +\infty. \quad (38)$$

即把取值为 $[0, 1]$ 的 $P(X)$ 放大为取值为 $[0, +\infty]$ 的 $O(X)$ 。

把 (33) 式代入 (37) 式有:

$$O(H|E) = \frac{P(E|H)}{P(E|\neg H)} \times O(H). \quad (39)$$

再把 LS 代入此式, 可得:

$$O(H|E) = LS \times O(H). \quad (40)$$

同理可得到关于 LN 的公式:

$$\frac{P(H|\neg E)}{P(\neg H|\neg E)} = \frac{P(E|H)}{P(\neg E|\neg H)} \times \frac{P(H)}{P(\neg H)}, \quad (41)$$

$$O(H|\neg E) = LN \times O(H). \quad (42)$$

式 (41) 和 (42) 就是修改的 Bayes 公式.

引言

DS 理论处

DS 理论处理的是集合上的不确定性问题, 为此需要先建立命题与集合之间的一一对应关系, 以把命题的不确定性转化为集合的不确定性。

设 Ω 为样本空间, 且 Ω 中的每个元素都相互独立, 则由 Ω 的所有子集构成的幂集记为 2^{Ω} .

当 Ω 中的元素个数为 N 时, 则其幂集 2^{Ω} 的元素个数为 2^N , 且其中的每一个元素都对应于一个关于 x 取值情况的命题。

Example

设 $\Omega = \{\text{红, 黄, 白}\}$, 求 Ω 的幂集 2^Ω .

Ω 的幂集可包括如下子集:

$$\begin{aligned} A_0 &= \emptyset, A_1 = \{\text{红}\}, A_2 = \{\text{黄}\}, A_3 = \{\text{白}\}, \\ A_4 &= \{\text{红, 黄}\}, A_5 = \{\text{红, 白}\}, A_6 = \{\text{黄, 白}\}, A_7 = \{\text{红, 黄, 白}\} \end{aligned}$$

其中, \emptyset 表示空集, 空集也可表示为 $\{\}$. 上述子集的个数正好是 $2^3 = 8$.

Example

对上例所给出的有限集 Ω 及基本函数 m , 当

- $A = \{\text{红}\}$ 时, 有 $m(A) = 0.3$, 它表示对命题 “ x 是红色” 的精确信任度为 0.3.
- $B = \{\text{红}, \text{黄}\}$ 时, 有 $m(B) = 0.2$, 它表示对命题 “ x 或者是红色, 或者是黄色” 的精确信任度为 0.2, 却不知道该怎么把这 0.2 分给 {红} 还是分给 {黄}.
- $C = \Omega = \{\text{红}, \text{黄}, \text{白}\}$ 时, 有 $m(\Omega) = 0.2$, 表示不知道该对这 0.2 如何分配, 但知道它不属于红, 就一定属于 {黄} 或 {白}.

(2) 概率分配函数不是概率

Example

在例 6.5 中, m 符合概率分配函数的定义, 但

$$m(\{\text{红}\}) + m(\{\text{黄}\}) + m(\{\text{白}\}) = 0.3 + 0 + 0.1 = 0.4 < 1. \quad (45)$$

因此 m 不是概率, 因为概率 P 要求: $P(\{\text{红}\}) + P(\text{黄}) + P(\text{白}) = 1$.

信任函数

信任函数 $\text{Bel}: 2^\Omega \rightarrow [0, 1]$ 为

$$\text{Bel}(A) = \sum_{B \subseteq A} m(B), \quad (46)$$

其中, 2^Ω 是 Ω 的幂集. Bel 又称为下限函数, $\text{Bel}(A)$ 表示对 A 的总的信任度.

Example

对例 6.5 有

$$\text{Bel}(\{\text{红}\}) = 0.3 \quad (47)$$

$$\begin{aligned} \text{Bel}(\{\text{红}, \text{白}\}) &= m(\{\text{红}\}) + m(\{\text{白}\}) + m(\{\text{红}, \text{白}\}) \\ &= 0.3 + 0.1 + 0.2 = 0.6. \end{aligned} \quad (48)$$

根据定义还可以得到:

$$\begin{aligned} \text{Bel}(\Phi) &= m(\Phi) = 0, \\ \text{Bel}(\Omega) &= \sum_{B \subseteq \Omega} m(B) = 1. \end{aligned} \quad (49)$$

Example

对例 7 有

$$\text{Bel}(\emptyset) = m(\emptyset) = 0 \quad (50)$$

$$\begin{aligned} \text{Bel}(\{\text{红, 黄, 白}\}) &= m(\emptyset) + m(\{\text{红}\}) + m(\{\text{黄}\}) + m(\{\text{白}\}) + m(\{\text{红, 黄}\}) \\ &\quad + m(\{\text{红, 白}\}) + m(\{\text{黄, 白}\}) + m(\{\text{红, 黄, 白}\}) \\ &= 0 + 0.3 + 0 + 0.1 + 0.2 + 0.2 + 0 + 0.2 = 1. \end{aligned} \quad (51)$$

似然函数 1 似然函数 $P_1 : 2^\Omega \rightarrow [0, 1]$ 为

$$P_1(A) = 1 - \text{Bel}(\neg A), A \subseteq \Omega, \quad (52)$$

其中, $\neg A = \Omega - A$.

似然函数又称为不可驳斥函数或上限函数. 由于 $\text{Bel}(\neg A)$ 表示对 $\neg A$ 的信任度, 即 A 为假的信任度, 因此, $P_l(A)$ 表示对 A 为非假的信任度.

Example

以例 7 为例:

$$\begin{aligned} P_l(\{\text{红}\}) &= 1 - \text{Bel}(\neg\{\text{红}\}) \\ &= 1 - \text{Bel}(\{\text{黄}, \text{白}\}) \\ &= 1 - (m(\{\text{黄}\}) + m(\{\text{白}\}) + m(\{\text{黄}, \text{白}\})) \\ &= 1 - (0 + 0.1 + 0) = 0.9. \end{aligned} \tag{53}$$

这里的 0.9 是“红”为非假的信任度. 由于“红”为真的精确信任度为 0.3, 而剩下的 $0.9 - 0.3 = 0.6$, 则是知道非假, 但却不能肯定为真的那部分.

Example

再如:

$$P_l(\{\text{黄, 白}\}) = 1 - \text{Bel}(\neg\{\text{黄, 白}\}) = 1 - \text{Bel}(\{\text{红}\}) = 1 - 0.3 = 0.7. \quad (54)$$

这里的 0.7 的含义与上面分析类似.

似然函数的另外一种计算方法:

$$\sum_{\{\text{红}\} \cap B = \emptyset} m(B) = 0.3 + 0.2 + 0.2 + 0.2 = 0.9. \quad (55)$$

由于

$$\sum_{\{\text{黄, 白}\} \cap B = \emptyset} m(B) = 0 + 0.1 + 0 + 0.2 + 0.2 + 0.2 = 0.7. \quad (56)$$

可见, $P_l(\{\text{红}\})$, $P_l(\{\text{黄, 白}\})$ 亦可分别用下式计算:

$$P_l(\{\text{红}\}) = \sum_{\{\text{红}\} \cap B = \emptyset} m(B), \quad P_l(\{\text{黄, 白}\}) = \sum_{\{\text{黄, 白}\} \cap B = \emptyset} m(B). \quad (57)$$

一般可得公式

$$P_l(A) = \sum_{A \cap B \neq \emptyset} m(B). \quad (58)$$

其证明见教材.

信任函数和似然函数之间存在关系:

$$P_l(A) \geq \text{Bel}(A). \quad (59)$$

Proof.

$$\text{Bel}(A) + \text{Be}(\mathcal{T}A) = \sum_{B \in A} m(B) + \sum_{C \in \neg A} m(C) \leq \sum_{E \subseteq \Omega} m(E) = 1.$$

$$\begin{aligned} P_l(A) - \text{Bel}(A) &= 1 - \text{Bel}(\neg A) - \text{Bel}(A) \\ &= 1 - (\text{Bel}(\neg A) + \text{Bel}(A)) \geq 0. \end{aligned} \tag{60}$$

$$P_l(A) \geq \text{Bel}(A),$$

由于 $\text{Bel}(A)$ 和 $P_l(A)$ 分别表示 A 为真的信任度和 A 为非假的信任度, 因此, 可分别称 $\text{Bel}(A)$ 和 $P_l(A)$ 为对 A 信任程度的下限和上限, 记为: $A[\text{Bel}(A), P_l(A)]$. □

Example

在前面的例子中

$$\text{Bel}(\{\text{红}\}) = 0.3, P_l(\{\text{红}\}) = 0.9. \quad (61)$$

即: $\{\text{红}\} = [0.3, 0.9]$. 它表示对 $\{\text{红}\}$ 的精确信任度为 0.3, 不可驳斥部分为 0.9, 肯定不是 $\{\text{红}\}$ 的为 0.1.

同理可以求得

- $\{\text{黄}\}[0, 0.4] \{\text{白}\}[0.1, 0.5];$
- $\{\text{红, 黄}\}[0.5, 0.9] \{\text{红, 白}\}[0.6, 1];$
- $\{\text{黄, 白}\}[0.1, 0.7] \{\text{红, 黄, 白}\}[1, 1];$
- $\{\} [0, 0].$

一些典型值的含义

$A[0, 1]$: 说明对 A 一无所知. 其中, $\text{Bel}(A) = 0$, 说明对 A 无信任; 再由 $P_1(A) = 1 - \text{Bel}(\neg A) = 1$, 可知 $\text{Bel}(\neg A) = 0$, 说明对 $\neg A$ 也没有信任.

- $A[0, 0]$: 说明 A 为假. 即 $\text{Bel}(A) = 0, \text{Bel}(\neg A) = 1$.

- $A[1, 1]$: 说明 A 为真. 即 $\text{Bel}(A) = 1, \text{Bel}(\neg A) = 0$.

• $A[0.6, 1]$: 说明对 A 部分信任. 即 $\text{Bel}(A) = 0.6, \text{Bel}(\neg A) = 0$.

• $A[0, 0.4]$: 说明对 $\neg A$ 部分信任. 即 $\text{Bel}(A) = 0, \text{Bel}(\neg A) = 0.6$.

• $A[0.3, 0.9]$: 说明对 A 和 $\neg A$ 都有部分信任. 其中, $\text{Bel}(A) = 0.3$, 说明对 A 为真有 0.3 的信任度;

$\text{Bel}(\neg A) = 1 - 0.9 = 0.1$, 说明对 A 为假有 0.1 的信任度. 因此, $A[0.3, 0.9]$ 表示对 A 为真的信任度比 A 为假的信任度稍高一些.

当证据来源不同时, 可能会得到不同的概率分配函数. 例如, 对

$$\Omega = \{\text{红}, \text{黄}\}. \quad (62)$$

从不同知识源得到的两个概率分配函数

$$m_1(\{\}, \{\text{红}\}, \{\text{黄}\}, \{\text{红}, \text{黄}\}) = (0, 0.4, 0.5, 0.1). \quad (63)$$

$$m_2(\{\}, \{\text{红}\}, \{\text{黄}\}, \{\text{红}, \text{黄}\}) = (0, 0.6, 0.2, 0.2). \quad (64)$$

求这些概率分配函数的组合.

采用德普斯特提出的求正交和的方法来组合这些函数

概率分配函数的正交和

设 m_1 和 m_2 是两个不同的概率分配函数, 则其正交和 $m = m_1 \oplus m_2$ 满足

$$m(\Phi) = 0, m(A) = K^{-1} \times \sum_{x \cap y = A} m_1(x) \times m_2(y). \quad (65)$$

其中 $K = 1 - \sum_{x \cap y = \Phi} m_1(x) \times m_2(y) = \sum_{x \cap y \neq \Phi} m_1(x) \times m_2(y)$.

如果 $K \neq 0$, 则正交和也是一个概率分配函数; 如果 $K = 0$, 则不存在正交和 m , 称 m_1 与 m_2 矛盾.

Example

设 $\Omega = \{a, b\}$, 且从不同知识源得到的概率分配函数分别为

$$m_1(\{\}, \{a\}, \{b\}, \{a, b\}) = (0, 0.3, 0.5, 0.2), \quad (66)$$

$$m_2(\{\}, \{a\}, \{b\}, \{a, b\}) = (0, 0.6, 0.3, 0.1). \quad (67)$$

求正交和 $m = m_1 \oplus m_2$.

先求 K

$$K = 1 - \sum_{x \cap y = \Phi} m_1(x) \times m_2(y) \quad (68)$$

$$= 1 - (m_1(\{a\}) \times m_2(\{b\}) + m_1(\{b\}) \times m_2(\{a\})) \quad (69)$$

$$= 1 - (0.3 \times 0.3 + 0.5 \times 0.6) = 0.61. \quad (70)$$

再求 $m(\{a\}, \{b\}, \{a, b\})$, 由于

$$\begin{aligned} m(\{a\}) &= \frac{1}{0.61} \times \sum_{x \cap y = \{a\}} m_1(x) \times m_2(y) \\ &= \frac{1}{0.61} \times (m_1(\{a\}) \times m_2(\{a\}) + m_1(\{a\}) \times m_2(\{a, b\}) + m_1(\{a, b\}) \times m_2(\{a\})) \\ &= \frac{1}{0.61} \times (0.3 \times 0.6 + 0.3 \times 0.1 + 0.2 \times 0.6) = 0.54. \end{aligned}$$

(71)

同理可求得

$$m(\{b\}) = 0.43, \quad (72)$$

$$m(\{a, b\}) = 0.03. \quad (73)$$

故有

$$m(\{\}, \{a\}, \{b\}, \{a, b\}) = \{0, 0.54, 0.43, 0.03\}. \quad (74)$$

对于多个概率分配函数的组合, 方法类似.

$\text{Bel}(A)$ 和 $\text{Pl}(A)$ 分别表示命题 A 的信任度的下限和上限, 同时也可用来表示知识强度的下限和上限.

从信任函数和似然函数的定义看, 它们都是建立在概率分配函数之上的, 可见不同的概率分配函数将得到不同的推理模型.

一个特殊的概率分配函数建立的推理模型

设 $\Omega = s_1, s_2, \dots, s_n$, m 为定义在 2^Ω 上的概率分配函数, 且 m 满足

$$\begin{aligned} (1) \quad & m(\{s_i\}) \geq 0 \\ (2) \quad & \sum_{i=1}^n m(\{s_i\}) \leq 1 \\ (3) \quad & m(\Omega) = 1 - \sum_{i=1}^n m(\{s_i\}). \end{aligned} \tag{75}$$

$A \subset \Omega, |A| \gg 1$ 或者 $|A| = 0 \Rightarrow m(A) = 0$, 其中, $|A|$ 表示命题 A 所对应的集合中的元素个数. 该概率分配函数的特殊性:

- ① 只有当子集中的元素个数为 1 时, 其概率分配数才有可能大于 0;
- ② 当子集中有多个或 0 个元素, 且不等于全集时, 其概率分配数均为 0;
- ③ 全集 Ω 的概率分配数按 (3) 计算.

其似然函数

对任何命题 $A \subset \Omega$, 其似然函数为

$$\begin{aligned} P_l(A) &= 1 - \text{Bel}(\neg A) = 1 - \sum_{s_i \in \neg A} m(\{s_i\}) = 1 - \left[\sum_{i=1}^n m(\{s_i\}) - \sum_{s_i \in A} m(\{s_i\}) \right] \\ &= 1 - [1 - m(\Omega) - \text{Bel}(A)] \\ &= m(\Omega) + \text{Bel}(A). \end{aligned} \quad (78)$$

$$P_l(\Omega) = 1 - \text{Bel}(\neg \Omega) = 1 - \text{Bel}(\Phi) = 1. \quad (79)$$

可以看出, 对任意命题 $A \subset \Omega$ 和 $B \subset \Omega$ 均有:

$$P_l(A) - \text{Bel}(A) = P_l(B) - \text{Bel}(B) = m(\Omega). \quad (80)$$

它表示对 A (或 B) 不知道的程度.

Example

设 $\Omega = \{\text{红}, \text{黄}, \text{白}\}$, 概率分配函数

$$m(\{\}, \{\text{红}\}, \{\text{黄}\}, \{\text{红}, \text{黄}\}, \{\text{红}, \text{黄}, \text{白}\}) = (0, 0.6, 0.2, 0.1, 0.1). \quad (81)$$

$A = \{\text{红}, \text{黄}\}$, 求 $m(\Omega)$ 、 $\text{Bel}(A)$ 和 $P_l(A)$ 的值.

$$m(\Omega) = 1 - [m(\{\text{红}\}) + m(\{\text{黄}\}) + m(\{\text{白}\})] = 1 - (0.6 + 0.2 + 0.1) = 0.1. \quad (82)$$

$$\text{Bel}(\{\text{红}, \text{黄}\}) = m(\{\text{红}\}) + m(\{\text{黄}\}) = 0.6 + 0.2 = 0.8. \quad (83)$$

$$P_l(\{\text{红}, \text{黄}\}) = m(\Omega) + \text{Bel}(\{\text{红}, \text{黄}\}) = 0.1 + 0.8 = 0.9. \quad (84)$$

$$\text{或 } P_l(\{\text{红}, \text{黄}\}) = 1 - \text{Bel}(\neg\{\text{红}, \text{黄}\}) = 1 - \text{Bel}(\{\text{白}\}) = 1 - 0.1 = 0.9.$$

基本概率分配函数的正交和

设 m_1 和 m_2 是 2^Ω 上的基本概率分配函数, 它们的正交和定义为

$$\mathbf{m}(\{\mathbf{s}_i\}) = \mathbf{K}^{-1} \times [\mathbf{m}_1(\mathbf{s}_i) \times \mathbf{m}_2(\mathbf{s}_i) + \mathbf{m}_1(\mathbf{s}_i) \times \mathbf{m}_2(\Omega) + \mathbf{m}_1(\Omega) \times \mathbf{m}_2(\mathbf{s}_i)]. \quad (85)$$

Example

设 $\Omega = \{\text{红}, \text{黄}, \text{白}\}$, 概率分配函数

$$m(\{\}, \{\text{红}\}, \{\text{黄}\}, \{\text{白}\}, \{\text{红, 黄, 白}\}) = (0, 0.6, 0.2, 0.1, 0.1). \quad (86)$$

$A = \{\text{红, 黄}\}$, 求 $m(\Omega)$, $\text{Bel}(A)$ 和 $P_l(A)$ 的值.

$$\begin{aligned} m(\Omega) &= 1 - [m(\{\text{红}\}) + m(\{\text{黄}\}) + m(\{\text{白}\})] \\ &= 1 - (0.6 + 0.2 + 0.1) = 0.1. \end{aligned} \quad (87)$$

$$\text{Bel}(\{\text{红}, \text{黄}\}) = m(\{\text{红}\}) + m(\{\text{黄}\}) = 0.6 + 0.2 = 0.8. \quad (88)$$

$$P_l(\{\text{红}, \text{黄}\}) = m(\Omega) + \text{Bel}(\{\text{红}, \text{黄}\}) = 0.1 + 0.8 = 0.9. \quad (89)$$

另一形式

$$P_l(\{\text{红}, \text{黄}\}) = 1 - \text{Bel}(\neg\{\text{红}, \text{黄}\}) = 1 - \text{Bel}(\text{白}) = 1 - 0.1 = 0.9. \quad (90)$$

概率分配函数的正交和

设 m_1 和 m_2 是 2^Ω 上的基本概率分配函数, 分配函数的正交和定义为

$$m(\{s_i\}) = K^{-1} \times [m_1(s_i) \times m_2(s_i) + m_1(s_i) \times m_2(\Omega) + m_1(\Omega) \times m_2(s_i)]. \quad (91)$$

其中,

$$K = m_1(\Omega) \times m_2(\Omega) + \sum_{i=1}^n [m_1(s_i) \times m_2(s_i) + m_1(s_i) \times m_2(\Omega) + m_1(\Omega) \times m_2(s_i)]. \quad (92)$$

类概率函数

设 Ω 为有限域, 对任何命题 $A \subset \Omega$, 命题 A 的类概率函数为

$$f(A) = \text{Bel}(A) + \frac{|A|}{|\Omega|} \times [P_1(A) - \text{Bel}(A)], \quad (93)$$

其中, $|A|$ 和 $|\Omega|$ 分别是 A 及 Ω 中元素的个数.

类概率函数 $f(A)$ 的性质

$$(1) \sum_{i=1}^n f(\{S_i\}) = 1.$$

$$\begin{aligned} \therefore f(\{s_i\}) &= \text{Bel}(\{s_i\}) + \frac{|\{s_i\}|}{|\Omega|} \times [P_l(\{s_i\}) - \text{Bel}(\{S_i\})] \\ &= m(\{s_i\}) + \frac{1}{n} \times m(\Omega), i = 1, 2, \dots, n. \end{aligned}$$

$$\begin{aligned} \therefore \sum_{i=1}^n f(\{s_i\}) &= \sum_{i=1}^n \left[m(s_i) + \frac{1}{n} \times m(\Omega) \right] \\ &= \sum_{i=1}^n m(\{s_i\}) + m(\Omega) = 1. \end{aligned} \tag{94}$$

(2) 对任何 $A \subset \Omega$, 有 $\text{Bel}(A) \leq f(A) \leq P_l(A)$.

$$\begin{aligned} \because P_l(A) - P_l(B) &= m(\Omega), \quad \frac{|A|}{|\Omega|} \geq 0 \\ \therefore \text{Bel}(A) &\leq f(A) \\ \because \frac{|A|}{|\Omega|} &\leq 1, f(A) \leq \text{Bel}(A) + P_l(A) - \text{Bel}(A), \\ \therefore f(A) &\leq P_l(A). \end{aligned} \tag{95}$$

(3) 对任何 $A \subseteq \Omega$, 有 $f(\neg A) = 1 - f(A)$.

$$\therefore f(\neg A) = \text{Bel}(\neg A) + \frac{|A|}{|P_1(\neg A) - \text{Bel}(\neg A)|} \quad (96)$$

$$\begin{aligned} \text{Bel}(\neg A) &= \sum_{\substack{s_i \in \neg A \\ s_i \in A}} m(\{s_i\}) \\ &= 1 - \sum_{\substack{s_i \in A \\ s_i \in d}} m(\{s_i\}) - m(\Omega) = 1 - \text{Bel}(A) - m(\Omega). \end{aligned} \quad (97)$$

$$|\neg A| = |\Omega| - |A|. \quad (98)$$

$$m(\Omega) = P_1(\neg A) - \text{Bel}(\neg A). \quad (99)$$

$$\begin{aligned}
 \therefore f(\neg A) &= 1 - \text{Bel}(A) - m(\Omega) + \frac{|\Omega| - |A|}{|\Omega|} \times m(\Omega) \\
 &= 1 - \text{Bel}(A) - m(\Omega) + m(\Omega) - \frac{|A|}{|\Omega|} \times m(\Omega) \\
 &= 1 - \left[\text{Bel}(A) + \frac{|A|}{|\Omega|} \times m(\Omega) \right] = 1 - f(A). \quad (100)
 \end{aligned}$$

推论

- (1) $f(\emptyset) = 0$;
- (2) $f(\Omega) = 1$;
- (3) 对任何 $A \subseteq \Omega$, 有 $0 \leq f(A) \leq 1$.

Example

设 $\Omega = \{\text{红}, \text{黄}, \text{白}\}$, 概率分配函数

$$m(\{\}, \{\text{红}\}, \{\text{黄}\}, \{\text{白}\}, \{\text{红}, \text{黄}, \text{白}\}) = (0, 0.6, 0.2, 0.1, 0.1). \quad (101)$$

若 $A = \{\text{红}, \text{黄}\}$, 求 $f(A)$ 的值.

$$\begin{aligned}
 f(A) &= \text{Bel}(A) + \frac{|A|}{|\Omega|} \times [P_1(A) - \text{Bel}(A)] \\
 &= m(\{\text{红}\}) + m(\{\text{黄}\}) + \frac{2}{3} \times (1 - \text{Bel}(\neg A) - 0.8) \quad (102) \\
 &= 0.6 + 0.2 + \frac{2}{3} \times 0.1 = 0.87.
 \end{aligned}$$

表示形式

$$\text{IF } E \text{ THEN } H = \{h_1, h_2, \dots, h_n\} \text{ CF} = \{c_1, c_2, \dots, c_n\}. \quad (103)$$

其中: ● E 为前提条件, 它既可以是简单条件, 也可以是用合取或析取词连接起来的复合条件;

- H 是结论, 它用样本空间中的子集表示, h_1, h_2, \dots, h_n 是该子集中的元素;
- CF 是可信度因子, 用集合形式表示. 该集合中的元素 c_1, c_2, \dots, c_n 用来指出 h_1, h_2, \dots, h_n 的可信度, c_i 与 h_i 一一对应. 并且 c_i 应满足如下条件:

$$c_i \geq 0, \sum_{i=1}^n c_i \leq 1, i = 1, 2, \dots, n. \quad (104)$$

Example

设 A 是规则条件部分的命题, E' 是外部输入的证据和已证实的命题, 在证据 E' 的条件下, 命题 A 与证据 E' 的匹配程度为

$$MD(A|E') = \begin{cases} 1, & \text{如果H所要求的证据都已出现} \\ 0, & \text{否则} \end{cases} \quad (105)$$

条件部分命题 A 的确定性

$$\text{CER}(A) = \text{MD}(A|E') \times f(A), \quad (106)$$

其中 $f(A)$ 为类概率函数. 由于 $f(A) \in [0, 1]$, 因此 $\text{CER}(A) \in [0, 1]$.

当组合证据是多个证据的合取时:

$$E = E_1 \text{ AND } E_2 \text{ AND } \cdots \text{ AND } E_n. \quad (107)$$

则

$$\text{CER}(E) = \min\{\text{CER}(E_1), \text{CER}(E_2), \cdots, \text{CER}(E_n)\}. \quad (108)$$

当组合证据是多个证据的析取时:

$$E = E_1 \text{ OR } E_2 \text{ OR } \cdots \text{ OR } E_n. \quad (109)$$

则

$$\text{CER}(E) = \max\{\text{CER}(E_1), \text{CER}(E_2), \cdots, \text{CER}(E_n)\}. \quad (110)$$

设有知识 IF E THEN $H = \{h_1, h_2, \dots, h_n\}$ $CF = \{c_1, c_2, \dots, c_n\}$, 则求结论 H 的确定性 CER(H) 的方法如下: (1) 求 H 的概率分配函数

$$m(\{h_1\}, \{h_2\}, \dots, \{h_3\}) = (CER(E) \times c_1, CER(E) \times c_2, \dots, CER(E) \times c_n) \quad (111)$$

$$m(\Omega) = 1 - \sum_{i=1}^n CER(E) \times c_i. \quad (112)$$

如果有两条或多条知识支持同一结论 H , 例:

$$\text{IF } E \text{ THEN } H = \{h_1, h_2, \dots, h_n\} \quad \text{CF} = \{c_{11}, c_{12}, \dots, c_{1n}\} \quad (113)$$

$$\text{IF } F \text{ THEN } H = \{h_1, h_2, \dots, h_n\} \quad \text{CF} = \{c_{21}, c_{22}, \dots, c_{2n}\}. \quad (114)$$

则按正交和求 $\text{CER}(H)$, 即先求出:

$$m_1 = m(h_1, h_2, \dots, h_n) \quad (115)$$

$$m_2 = m(h_1, h_2, \dots, h_n). \quad (116)$$

然后再用公式 $m = m_1 \oplus m_2$ 求 m_1 和 m_2 的正交和, 最后求得 H 的 m .

(2) 求 $\text{Bel}(H)$ 、 $P_1(H)$ 及 $f(H)$

$$\begin{aligned} \text{Bel}(H) &= \sum_{i=1}^n m(\{h_i\}), \\ P_1(H) &= 1 - \text{Bel}(\neg H), \\ f(H) &= \text{Be}(H) + \frac{|H|}{|\Omega|} \times [P_1(H) - \text{Be}(H)] = \text{Be}(H) + \frac{|H|}{|\Omega|} \times m(\Omega). \end{aligned} \quad (117)$$

(3) 求 H 的确定性 $\text{CER}(H)$

按公式 $\text{CER}(H) = \text{MD}(H|E') \times f(H)$ 计算结论 H 确定性.

Example

设有如下规则:

- r_1 : IF E_1 AND E_2 THEN $A=\{a_1, a_2\}$ $CF=\{0.3, 0.5\}$,
- r_2 : IF E_3 AND (E_4 OR E_5) THEN $B=\{b_1\}$ $CF=\{0.7\}$,
- r_3 : IF A THEN $H=\{h_1, h_2, h_3\}$ $CF=\{0.1, 0.5, 0.3\}$,
- r_4 : IF B THEN $H=\{h_1, h_2, h_3\}$ $CF=\{0.4, 0.2, 0.1\}$.

已知用户对初始证据给出的确定性为:

$$CER(E_1)=0.8, CER(E_2)=0.6, CER(E_3)=0.9, CER(E_4)=0.5, CER(E_5)=0.7.$$

并假定 Ω 中的元素个数 $|\Omega| = 10$, 求 $CER(H)$.

由给定知识形成的推理网络为:

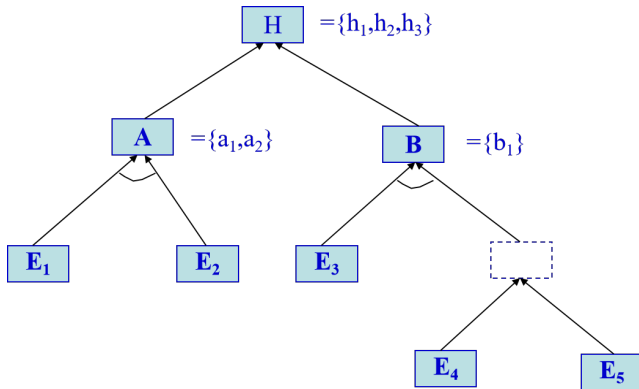


图 1: 给定知识形成的推理网络

(1) 求 CER(A)

$$\therefore \text{CER}(E_1 \text{ AND } E_2) = \min\{\text{CER}(E_1), \text{CER}(E_2)\} = \min\{0.8, 0.6\} = 0.6.$$

$$m(a_1, a_2) = \{0.6 \times 0.3, 0.6 \times 0.5\} = \{0.18, 0.3\}.$$

$$\text{Bel}(A) = m(a_1) + m(a_2) = 0.18 + 0.3 = 0.48.$$

$$P_1(A) = 1 - \text{Bel}(\neg A) = 1 - 0 = 1.$$

$$\begin{aligned} f(A) &= \text{Bel}(A) + |A|/|\Omega| * [P_1(A) - \text{Bel}(A)] \\ &= 0.48 + 2/10 * [1 - 0.48] = 0.584. \end{aligned}$$

$$\therefore \text{CER}(A) = \text{MD}(A|E') \times f(A) = 0.584.$$

(2) 求 CER(B)

$$\begin{aligned}\therefore \text{CER}(E_3 \text{ AND } (E_4 \text{ OR } E_5)) &= \min\{\text{CER}(E_3), \max\{\text{CER}(E_4), \text{CER}(E_5)\}\} \\ &= \min\{0.9, \max\{0.5, 0.7\}\} \\ &= \min\{0.9, 0.7\} = 0.7.\end{aligned}$$

$$m(b_1) = 0.7 \times 0.7 = 0.49.$$

$$\text{Bel}(B) = m(b_1) = 0.49.$$

$$\text{Pl}(B) = 1 - \text{Bel}(\neg B) = 1 - 0 = 1.$$

$$\begin{aligned}F(B) &= \text{Bel}(B) + |B|/|\Omega| * [\text{Pl}(B) - \text{Bel}(B)] \\ &= 0.49 + 1/10 * [1 - 0.49] = 0.541.\end{aligned}$$

$$\therefore \text{CER}(B) = \text{MD}(B|E') \times f(B) = 0.541.$$

(3) 求 $CER(H)$, 由 r_3 可得

$$\begin{aligned} m_1(h_1, h_2, h_3) &= \{CER(A) \times 0.1, CER(A) \times 0.5, CER(A) \times 0.3\} \\ &= \{0.584 \times 0.1, 0.584 \times 0.5, 0.584 \times 0.3\} \\ &= \{0.058, 0.292, 0.175\}. \end{aligned}$$

$$\begin{aligned} m_1(\Omega) &= 1 - [m_1(h_1) + m_1(h_2) + m_1(h_3)] \\ &= 1 - (0.058 + 0.292 + 0.175) = 0.475. \end{aligned}$$

再由 r_4 可得

$$\begin{aligned} m_2(h_1, h_2, h_3) &= CER(B) \times 0.4, CER(B) \times 0.2, CER(B) \times 0.1 \\ &= \{0.541 \times 0.4, 0.541 \times 0.2, 0.541 \times 0.1\} \\ &= \{0.216, 0.108, 0.054\}. \\ m_2(\Omega) &= 1 - [m_2(h_1) + m_2(h_2) + m_2(h_3)] \\ &= 1 - [0.216 + 0.108 + 0.054] = 0.622. \end{aligned}$$

求正交和 $m = m_1 \oplus m_2$

$$\begin{aligned}
 K &= m_1(\Omega) \times m_2(\Omega) + m_1(h_1) \times m_2(h_1) + m_1(h_1) \times m_2(\Omega) + m_1(\Omega) \times m_2(h_1) \\
 &\quad + m_1(h_2) \times m_2(h_2) + m_1(h_2) \times m_2(\Omega) + m_1(\Omega) \times m_2(h_2) \\
 &\quad + m_1(h_3) \times m_2(h_3) + m_1(h_3) \times m_2(\Omega) + m_1(\Omega) \times m_2(h_3) \\
 &= 0.475 \times 0.622 + 0.058 \times 0.216 + 0.058 \times 0.622 + 0.475 \times 0.216 \\
 &\quad + 0.292 \times 0.108 + 0.292 \times 0.622 + 0.475 \times 0.108 \\
 &\quad + 0.175 \times 0.054 + 0.175 \times 0.622 + 0.475 \times 0.054 = 0.855.
 \end{aligned}$$

$$\begin{aligned} m(h_1) &= \frac{1}{K} \times [m_1(\{h_1\}) \times m_2(\{h_1\}) + m_1(\{h_1\}) \times m_2(\Omega) + m_1(\Omega) \times m_2(\{h_1\})] \\ &= \frac{1}{0.855} \times [0.058 \times 0.216 + 0.058 \times 0.622 + 0.475 \times 0.216]. \end{aligned}$$

$$\begin{aligned} m(h_2) &= \frac{1}{K} \times [m_1(\{h_2\}) \times m_2(\{h_2\}) + m_1(\{h_2\}) \times m_2(\Omega) + m_1(\Omega) \times m_2(\{h_2\})] \\ &= \frac{1}{0.855} \times [0.292 \times 0.108 + 0.292 \times 0.622 + 0.475 \times 0.108] \\ &= 0.309. \end{aligned}$$

$$\begin{aligned} m(h_3) &= \frac{1}{K} \times [m_1(\{h_3\}) \times m_2(\{h_3\}) + m_1(\{h_3\}) \times m_2(\Omega) + m_1(\Omega) \times m_2(\{h_3\})] \\ &= \frac{1}{0.855} \times [0.175 \times 0.054 + 0.175 \times 0.622 + 0.475 \times 0.054] \\ &= 0.168. \end{aligned}$$

$$m(\Omega) = 1 - [m(h_1) + m(h_2) + m(h_3)] = 1 - (0.178 + 0.309 + 0.168) = 0.345.$$

$$\text{Bel}(H) = m(h_1) + m(h_2) + m(h_3) = 0.178 + 0.309 + 0.168 = 0.655$$

$$P_l(H) = m(\Omega) + \text{Bel}(H) = 0.345 + 0.655 = 1.$$

$$\begin{aligned} f(H) &= \text{Bel}(H) + \frac{|H|}{|\Omega|} \times [P_l(H) - \text{Bel}(H)] \\ &= 0.655 + \frac{3}{10} \times [1 - 0.655] \\ &= 0.759. \end{aligned}$$

$$\text{CER}(H) = \text{MD}(H|E') \times f(H) = 0.759.$$

DS 理论的优缺点

优点

能处理由“不知道”所引起的非精确性; 并且由于辨别框 (样本空间) 的子集可以是多个元素的集合, 这样更有利于领域专家在不同层次上进行知识表示.

缺点

要求辨别框中的元素满足互斥条件, 这在实际系统中不易实现; 并且, 需要给出的概率分配数太多, 计算比较复杂.

模糊推理表示

用自然语言中的词或句子表示的变量

Example

变量“年龄”在普通集合中为数字变量 $u = [0, 150]$, 而在模糊集和中可使用语言变量, 该语言变量的取值可以是年轻、很年轻、不很年轻、老、很老、不很老等. 这些值可看作是论域 $U = [0, 150]$ 上模糊集的集合名.

模糊谓词

模糊谓词

设 $x \in U$, F 为模糊谓词, 即 U 中的一个模糊关系, 则模糊命题可表示为

$$x \text{ is } F, \quad (118)$$

其中的模糊谓词 F 可以是 大、小、年轻、年老、冷、暖、长、短等.

模糊量词

模糊量词

模糊逻辑中使用的模糊量词, 如极少、很少、几个、少数、多数、大多数、几乎所有等. 这些模糊量词可以很方便地描述类似于下面的命题:

- 大多数成绩好的学生学习都很刻苦.
- 很少有成绩好的学生特别贪玩.

模糊概率、模糊可能性和模糊真值

设 λ 为模糊概率, Π 为模糊可能性, τ 为模糊真值, 则对命题还可以附加概率限定、可能性限定和真值限定:

$$(x \text{ is } F) \text{ is } \lambda(x \text{ is } F) \text{ is } \Pi(x \text{ is } F) \text{ is } \tau, \quad (119)$$

其中, λ 可以是“或许”、“必须”等; π 可以是“非常可能”、“很不可能”等; τ 可以是“非常真”、“有些假”等.

Example

“常青很可能是年轻的” 可表示为

(Age(Chang qing) is young) is likely.

模糊修饰语的表达

模糊修饰语

设 m 是模糊修饰语, x 是变量, F 谓模糊谓词, 则模糊命题可表示为 x is m_F , 模糊修饰语也称为程度词, 常用的程度词有“很”、“非常”、“有些”、“绝对”等.

主要通过以下四种运算实现

① 求补: \neg 表示否定, 如“不”、“非”等, 其隶属函数的表示为

$$\mu_{\neg F}(u) = 1 - \mu_F(u) \quad u \in [0, 1].$$

② 集中表示“很”、“非常”等, 其效果是减少隶属函数的值:

$$\mu_{\text{非常}F}(u) = \mu_F^2(u) \quad u \in [0, 1].$$

③ 扩张表示“有些”、“稍微”等, 其效果是增加隶属函数的值:

$$\mu_{\text{有些}F}(u) = \mu_F^{\frac{1}{2}}(u) \quad u \in [0, 1].$$

④ 加强对比表示“明确”、“确定”等, 其效果是增加 0.5 以上隶属函数的值, 减少 0.5 以下隶属函数的值:

$$\mu_{\text{确实}F}(u) = \begin{cases} 2\mu_F^2(u), & 0 \leq \mu_F(u) \leq 0.5 \\ 1 - 2(1 - \mu_F(u))^2, & 0.5 < \mu_F(u) \leq 1 \end{cases}.$$

在以上 4 种运算中, 集中与扩张用的较多.

语言变量“真实性”取值“真”和“假”的隶属函数

$$\begin{aligned}\mu_{\text{真}}(u) &= u, u \in [0, 1] \\ \mu_{\text{假}}(u) &= 1 - u, u \in [0, 1].\end{aligned}$$

则“非常真”、“有些真”、“非常假”、“有些假”可定义为

$$\mu_{\text{非常真}}(u) = u^2, u \in [0, 1]; \mu_{\text{非常假}}(u) = \mathbb{R} \quad (u) = (1 - u)^2, u \in [0, 1] \quad (120)$$

$$\mu_{\text{有些真}}(u) = u^{\frac{1}{2}}, u \in [0, 1]; \mu_{\text{有些假}}(u) = \pm \lim_{\mathbb{R}} (u) = (1 - u)^{\frac{1}{2}}, u \in [0, 1]. \quad (121)$$

扎德推理模型——产生式规则的表达形式是

IF x is F THEN y is G .

其中: x 和 y 是变量, 表示对象; F 和 G 分别是论域 U 和 V 上的模糊集, 表示概念.

语义距离

用于刻画两个模糊概念之间的差异. 一般使用的距离是海明距离.

离散论域: 设 $U = u_1, u_2, \dots, u_n$ 是一个离散有限论域, F 和 G 分别是论域 U 上的两个模糊概念的模糊集, 则 F 和 G 的海明距离定义为

$$d(F, G) = \frac{1}{n} \times \sum_{i=1}^n |\mu_F(u_i) - \mu_G(u_i)|. \quad (122)$$

连续论域

如果论域 U 是实数域上的某个闭区间 $[a, b]$, 则海明距离为

$$d(F, G) = \frac{1}{b-a} \int_a^b |\mu_F(u) - \mu_G(u)| d(u). \quad (123)$$

Example

设论域 $U = \{-10, 0, 10, 20, 30\}$ 表示温度, 模糊集

$$F = 0.8/-10 + 0.5/0 + 0.1/10, \quad (124)$$

$$G = 0.9/-10 + 0.6/0 + 0.2/10. \quad (125)$$

分别表示“冷”和“比较冷”, 则

$$d(F, G) = 0.2 \times (|0.8 - 0.9| + |0.5 - 0.6| + |0.1 - 0.2|) = 0.2 \times 0.3 = 0.06. \quad (126)$$

即 F 和 G 的海明距离为 0.06.

设 F 和 G 分别是论域 $U = \{u_1, u_2, \dots, u_n\}$ 上的两个模糊概念的模糊集, 则它们的贴近度定义为

$$(F, G) = (1/2)(F \cdot G + (1 - F \odot G)), \quad (127)$$

其中:

$$\begin{aligned} F \cdot G &= \bigvee (\mu_F(u_i) \wedge \mu_G(u_i)), \\ F \odot G &= \bigwedge_U (\mu_F(u_i) \vee \mu_G(u_i)). \end{aligned} \quad (128)$$

称 $F \cdot G$ 为内积, $F \odot G$ 为外积.

Example

设论域 U 及其上的模糊集 F 和 G 如上例所示, 则

$$\begin{aligned} F \cdot G &= 0.8 \wedge 0.9 \vee 0.5 \wedge 0.6 \vee 0.1 \wedge 0.2 \vee 0 \wedge 0 \vee 0 \wedge 0 \\ &= 0.8 \vee 0.5 \vee 0.1 \vee 0 \vee 0 = 0.8 \end{aligned}$$

$$\begin{aligned} F \odot G &= (0.8 \vee 0.9) \wedge (0.5 \vee 0.6) \wedge (0.1 \vee 0.2) \wedge (0 \vee 0) \wedge (0 \vee 0) \\ &= 0.9 \wedge 0.6 \wedge 0.2 \wedge 0 \wedge 0 = 0 \end{aligned}$$

$$(F, G) = 0.5 \times (0.8 + (1 - 0)) = 0.5 \times 1.8 = 0.9.$$

即 F 和 G 的贴度为 0.9.

模糊推理及其构造

模糊推理实际上是按照给定的推理模式, 通过模糊集合与模糊关系的合成来实现的.

模糊关系 R_m

R_m 是由扎德提出的一种构造模糊关系的方法. 设 F 和 G 分别是论域 U 和 V 上的两个模糊集, 则 R_m 定义为

$$R_m = \int_{U \times V} (\mu_F(u) \wedge \mu_G(v)) \vee (1 - \mu_F(u)) / (u, v), \quad (129)$$

其中, \times 号表示模糊集的笛卡尔乘积.

Example

设 $U = V = \{1, 2, 3\}$, F 和 G 分别是 U 和 V 上的两个模糊集, 且
 $F = 1/1 + 0.6/2 + 0.1/3$, $G = 0.1/1 + 0.6/2 + 1/3$, 求 $U \times V$ 上的 R_m .

$$R_m = \int_{U \times V} (\mu_F(u) \wedge \mu_G(v)) \vee (1 - \mu_F(u)) / (u, v).$$

$$R_m = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.4 & 0.6 & 0.6 \\ 0.9 & 0.9 & 0.9 \end{bmatrix}.$$

如: $R_m(2, 3) = (0.6 \wedge 1) \vee (1 - 0.6) = 0.6 \vee 0.4 = 0.6$.

模糊关系 R_c

R_c 是由麦姆德尼 (Mamdani) 提出的一种构造模糊关系的方法.

设 F 和 G 分别是论域 U 和 V 上的两个模糊集, 则 R_c 义为

$$R_c = \int_{U \times V} (\mu_F(u) \wedge \mu_G(v))_Q / (u, v).$$

Example

对例 26 所给出的模糊集

$$\mathbf{F} = 1/1 + 0.6/2 + 0.1/3, \mathbf{G} = 0.1/1 + 0.6/2 + 1/3.$$

其 R_c 为

$$R_c = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 0.6 & 0.6 \\ 0.1 & 0.1 & 0.1 \end{bmatrix},$$

如 $R_c(3, 2)$: $R_c(3, 2) = \mu_F(u_3) \wedge \mu_G(v_2) = 0.1 \wedge 0.6 = 0.1.$

模糊关系 R_g

R_g 是米祖莫托 (Mizumoto) 提出的一种构造模糊关系的方法.

设 F 和 G 分别是论域 U 和 V 上的两个模糊集, 则 R_g 定义为

$$R_g = \int_{U \times V} (\mu_F(u) \rightarrow \mu_G(v)) / (u, v),$$

其中

$$\mu_F(u) \rightarrow \mu_G(v) = \begin{cases} 1, & \mu_F(u) \leq \mu_G(v) \\ \mu_G(v), & \mu_F(u) > \mu_G(v) \end{cases}.$$

Example

对例 26 所给出的模糊集

$$\mathbf{F} = 1/1 + 0.6/2 + 0.1/3, \mathbf{G} = 0.1/1 + 0.6/2 + 1/3.$$

其 R_g 为

$$R_g = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

在这种推理模式下, 模糊知识

IF x is F THEN y is G.

表示在 F 与 G 之间存在着确定的因果关系, 设此因果关系为 R. 则有

$$G' = F' \circ R,$$

其中的模糊关系 R, 可以是 R_m 、 R_c 或 R_g 中的任何一种.

模糊推理的基本方法

Example

对例 26 所给出的 F 、 G , 以及所求出的 R_m , 设有已知事实: $\{x \text{ is 较小}\}$, 并设“较小”的模糊集为: $\text{较小} = 1/1 + 0.7/2 + 0.2/3$, 求在此已知事实下的模糊结论.

例 26 的模糊关系 R_m 已在例 26 中求出, 设已知模糊事实“较小”为 F' , F' 与 R_m 的合成即为所求结论 G' .

$$G' = F' \circ R_m = \{1, 0.7, 0.2\} \circ \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.4 & 0.6 & 0.6 \\ 0.9 & 0.9 & 0.9 \end{bmatrix} = \{0.4, 0.6, 1\}.$$

即所求出的模糊结论 G' 为 $G' = 0.4/1 + 0.6/2 + 1/3$.

模糊拒取式推理

设 F 和 G 分别是 U 和 V 上的两个模糊集, 且有知识

IF x is F THEN y is G .

若有 V 上的一个模糊集 G' , 且 G 可以和 G' 匹配, 则可以推出 x is F' , 且 F' 是 U 上的一个模糊集.

模糊拒取式推理

其表示形式为:

知识 : IF x is F THEN y is G.

证据 : THEN z is H.

— — — — —

结论 : IF x is F'.

在这种推理模式下, 模糊知识

IF x is F THEN y is G.

也表示在 F 与 G 之间存在着确定的因果关系, 设此因果关系为 R, 则有 $F' = R \circ G'$, 其中的模糊关系 R, 可以是 R_m 、 R_c 或 R_g 中的任何一种.

Example

设 F 和 G 如例 26 所示, 已知事实为 {y is 较大}, 且“较大”的模糊集为: 较大 = 0.2/1 + 0.7/2 + 1/3, 若已知事实与 G 匹配, 以模糊关系 R_c 为例, 在此已知事实下可以推出 F' .

本例的模糊关系 R_c 已在前面求出, 设模糊概念“较大”为 G' , 则 R_c 与 G' 的合成即为所求的 F' .

$$F' = R_c \circ G' = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 0.6 & 0.6 \\ 0.1 & 0.1 & 0.1 \end{bmatrix} \circ \begin{bmatrix} 0.2 \\ 0.7 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.6 \\ 0.1 \end{bmatrix}.$$

即所求出的 F' 为 $G' = 1/1 + 0.6/2 + 0.1/3$.

模糊假言三段论推理

设 F, G, H 分别是 U, V, W 上的 3 个模糊集, 且由知识

IF x is F THEN y is G ,
IF y is G THEN z is H .

则可推出:

IF x is F THEN z is H .

模糊假言三段论推理

它可表示为:

IF x is F THEN y is G
IF y is G THEN z is H
—— — — — —
IF x is F THEN z is H.

在模糊假言三段论推理模式下, 模糊知识

$r_1 : \text{IF } x \text{ is } F \text{ THEN } y \text{ is } G,$

表示在 F 与 G 之间存在着确定的因果关系, 设此因果关系为 R_1 .

模糊知识

$r_2 : \text{IF } y \text{ is } G \text{ THEN } z \text{ is } H,$

表示在 G 与 H 之间存在着确定的因果关系, 设此因果关系为 R_2 .
若模糊假言三段论成立, 则模糊结论

$r_3 : \text{IF } x \text{ is } F \text{ THEN } z \text{ is } H$

的模糊关系 R_3 可由 R_1 与 R_2 的合成得到. 即

$$R_3 = R_1 \circ R_2,$$

这里的关系 R_1 、 R_2 和 R_3 都可以是前面所讨论过的 R_m 、 R_c 、 R_g 中的任何一种.

Example

设 $U = W = V = 1, 2, 3$, $E = 1/1 + 0.6/2 + 0.2/3$,
 $F = 0.8/1 + 0.5 + 0.1/3$, $G = 0.2/1 + 0.6 + 1/3$. 按 R_g 求 $E \times F \times G$ 上的关系 R .

先求 $E \times F$ 上的关系 R_1

$$R_1 = \begin{bmatrix} 0.8 & 0.5 & 0.1 \\ 1 & 0.5 & 0.1 \\ 1 & 1 & 0.1 \end{bmatrix}.$$

再求 $E \times G$ 上的关系 R_2 :

$$R_2 = \begin{bmatrix} 0.2 & 0.6 & 1 \\ 0.2 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

最后求 $E \times F \times G$ 上的关系 R ,

$$R = R_1 \circ R_2 = \begin{bmatrix} 0.2 & 0.6 & 0.8 \\ 0.2 & 0.6 & 1 \\ 0.2 & 1 & 1 \end{bmatrix}.$$