

第 4 章 不确定性推理方法

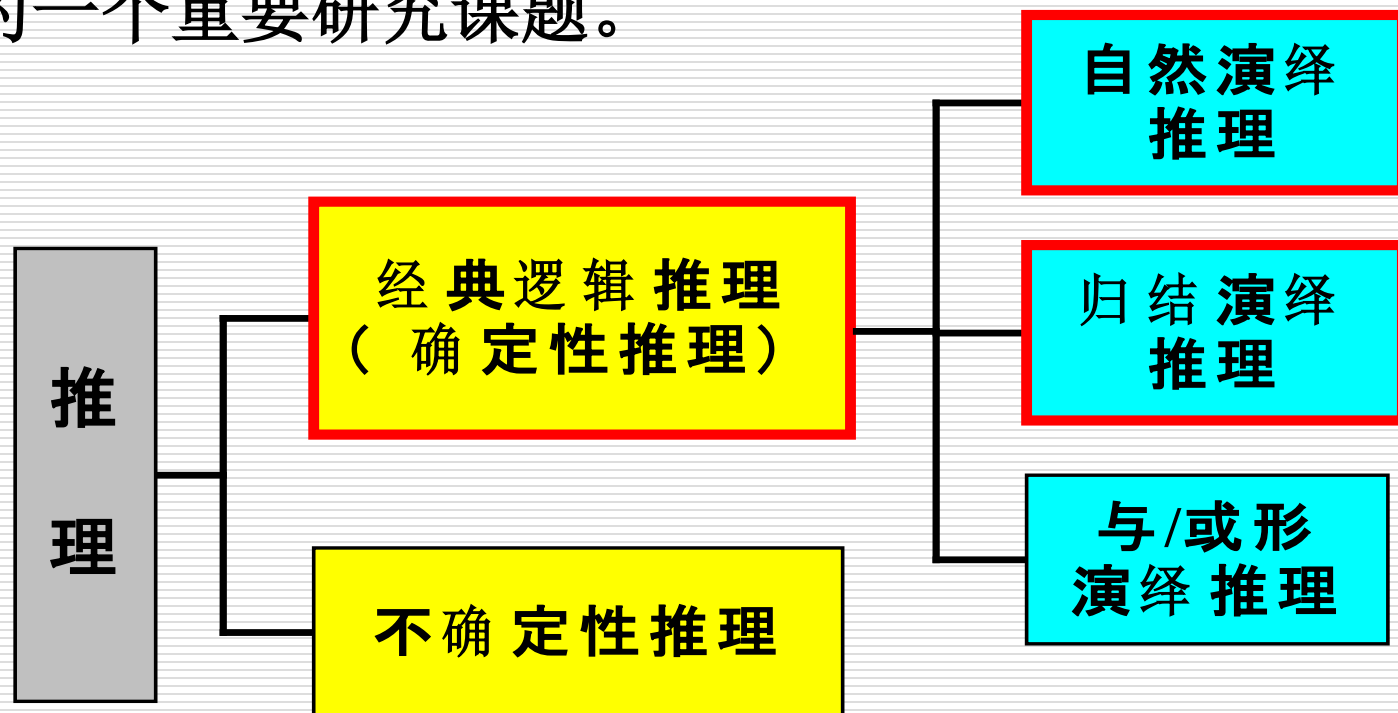


第4章 不确定性推理方法

- 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来，就分别形成了不确定性的知识及不确定性的证据。因而还必须对不确定性知识的表示及推理进行研究。这就是本章将要讨论的不确定性推理。

第4章 不确定性推理方法

前面讨论了把知识用某种模式表示出来存储到计算机中去。但是，为使计算机具有智能，还必须使它具有思维能力。推理是求解问题的一种重要方法。因此，推理方法成为人工智能的一个重要研究课题。



区别	经典逻辑、 经典推理	非经典逻辑、 非经典推理
推理方法	演绎逻辑推理	归纳逻辑推理
辖域取值	二值逻辑	多值逻辑（三值、四值、模糊逻辑等）
运算法则	$\sim (\sim P)=P$ 等	经典逻辑中的许多运算法则在非经典逻辑中不再成立。
逻辑算符	$\sim \vee \wedge \rightarrow \leftrightarrow \forall \exists$ 这些算符组成的谓词逻辑公式只能回答“什么是真/假”的是非判断问题	引用的附加算符（模态算符、算子）可以处理“什么可能真/假”、“什么必然假”之类问题
单调性	单调性，即已知事实（定理）均为充分可信的，不含随着新事实的出现而使原有事实变为假。	非单调性，即当客观情况发生变化或人们对客观情况

第4章 不确定性推理方法

下面首先讨论不确定性推理中的基本问题，然后着重介绍基于概率论的有关理论发展起来的不确定性推理方法，主要介绍

- 1 不确定性推理中的基本问题
- 2 贝叶斯方法
- 3 可信度方法

4.1 不确定性推理中的基本问题

- 推理：从**已知事实（证据）**出发，通过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的思维过程。
- **不确定性推理**：从**不确定性的初始证据**出发，通过运用**不确定性的知识**，最终推出具有一定程度的不确定性但却是合理或者近乎合理的结论的思维过程。

4.1 不确定性推理中的基本问题

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

4.1 不确定性推理中的基本问题

1. 不确定性的表示与量度

(1) 知识不确定性的表示

(2) 证据不确定性的表示

在专家系统中知识的不确定性一般由领域专家给出，通常是一个

(3) 不确定性的量度

用户在求解问题时提供的初始证据。

- ① 能充分表达相应知识及证据不确定性的程度。
- ② 度量范围的指定便于领域专家及用户对不确定性的估计。
- ③ 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。
- ④ 度量的确定应当是直观的，同时应有相应的理论依据。

4.1 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、Hamacher方法、概率方法、有界方法、Einstein方法等。

4.1 不确定性推理中的基本问题

4. 不确定性的传递算法

- (1) 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。
- (2) 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

4.2 贝叶斯决策

问题一：工科院校的男生和女生的比例为3: 1，在不知学生个人信息的情况下，如何预测某个学生的性别？

根据先验概率进行预测

问题二：再给出某个学生的身高和体重，身高180cm，体重70kg，如何来预测学生的性别？

男生的概率记做 $P(w_1 | x)$

女生的概率记做 $P(w_2 | x)$

4.2 贝叶斯决策

- 两个先验知识:

先验概率: $P(w_1)$ 和 $P(w_2)$

类条件概率: $P(x | w_1)$ 和 $P(x | w_2)$

- 后验概率:
$$P(\omega_i | \mathbf{x}) = \frac{P(\mathbf{x} | \omega_i) P(\omega_i)}{P(\mathbf{x})}$$

- 根据后验概率进行决策,这个决策过程就是贝叶斯决策。

- 全概率: $P(x)$

样本 x 在所有类别里出现的概率之和,也称为 x 的全概率

在两类问题情况下: $p(x) = \sum p(x | w_j)P(w_j), j=1,2$

【贝叶斯公式】

(4) 贝叶斯(Bayes)公式

若事件 B_1, B_2, \dots 满足全概率公式条件, 则对于任一事件 $A (P(A) > 0)$, 有

$$P(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{i=1}^{\infty} P(B_i)P(A | B_i)}$$

若 B_i 只有 n 个, 则此公式也成立, 这时右端分母只有 n 项相加。

□ 贝叶斯公式的两个创新点:

- (1) 用概率表示所有形式的不确定性;
- (2) 引入了“先验”与“后验”的概念;

□ 贝叶斯公式由于其权威性、一致性和典雅性而被列入最优美的数学公式之一;

□ 由贝叶斯公式衍生出贝叶斯决策、贝叶斯估计、贝叶斯学习等诸多理论体系, 进而形成一个贝叶斯学派。

【基本概念】

- 先验概率: $P(\omega_i)$

预先已知的或者可以估计的模式识别系统位于某种类型的概率。根据大量统计确定某类事物出现的比例。
如P(男生)

- 类条件概率: $P(\mathbf{x}|\omega_i)$

类别状态为 w_i 时，样本 x 出现的概率密度。

- 全概率: $P(x)$

样本 x 在所有类别里出现的概率之和, 也称为 x 的全概率

在两类问题情况下: $p(x) = \sum p(x|w_j)P(w_j), j=1,2$

- 后验概率: $P(\omega_i|x)$

➤ 一个具体事物属于某种类别的概率.

➤ 后验概率涉及一个具体事物, 而先验概率是泛指一类事物, 因此 $P(\text{男生}|X)$ 和 $P(\text{男生})$ 是两个不同的概念。

贝叶斯推理

□ 利用贝叶斯公式来推理：顾客是一位“买东西的人”的概率

1、进店的人有“买东西的人”和“随便逛逛的人”，售货员根据经验知道：每5位顾客中有1位是“买东西的人”

某种类别的概率（比例）——先验概率

2、假如有“买东西的人”有“前来询问”(0.9)、“不询问”二种行为的概率；“随便逛逛的人”“前来询问”(0.3)、“不询问”二种行为的概率

设置发生“向店员询问”事件的条件概率

3、现有一人前来询问，那么，其买东西的概率有多大呢？

寻求“来买东西的人”的“后验概率”（贝叶斯逆概率）

4.3 可信度方法

- 1975年肖特里菲（E. H. Shortliffe）等人在确定性理论（theory of confirmation）的基础上，结合概率论等提出的一种不确定性推理方法。
- 优点：直观、简单，且效果好。

4.3 可信度方法

- **可信度**：根据经验对一个事物或现象为真的相信程度。
- 可信度带有较大的主观性和经验性，其准确性难以把握。
- **C—F模型**：基于可信度表示的不确定性推理的基本方法。

4.3 可信度方法

1. 知识不确定性的表示

■ 产生式规则表示:

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

$CF(H, E)$: 可信度因子 (certainty factor), 反映前提条件与结论的联系强度。

IF 头痛 AND 流涕 THEN 感冒 (0.7)

- 可信度指标取值范围为 $[-1, 1]$, 其取值原则为:
 - 当 $CF > 0$ 时, 表明证据 E 出现则 H 为真, 值越大则 H 为真的概率越大
 - 当 $CF < 0$ 时, 表明证据 E 出现则 H 为假, 值越小则 H 为假的概率越大
 - $CF = 0$ 时, 证据 E 与 H 中立无关。

4.3 可信度方法

2. 证据不确定性的表示

$CF(E)=0.6$: E 的可信度为0.6

- 证据 E 的可信度取值范围: $[-1, 1]$ 。
- 对于初始证据, 若所有观察 S 能肯定它为真, 则 $CF(E)=1$ 。
若肯定它为假, 则 $CF(E)=-1$ 。
- 若以某种程度为真, 则 $0 < CF(E) < 1$ 。
- 若以某种程度为假, 则 $-1 < CF(E) < 0$ 。
- 若未获得任何相关的观察, 则 $CF(E)=0$ 。

4.3 可信度方法

2. 证据不确定性的表示

- 静态强度 $CF(H, E)$: 知识的强度, 即当 E 所对应的证据为真时对 H 的影响程度。
- 动态强度 $CF(E)$: 证据 E 当前的不确定性程度。

有了可信度指标, 如何进行推理呢?

回顾：确定性推理

- 我们首先回顾确定性推理的过程。假设我们有一条推理规则：

IF X会飞 AND X卵生 THEN X是鸟类

- 如果有事实 Fido会飞, Fido卵生
 - 则得到的结论应该是 Fido是鸟类
- 这个推理，涉及到了两个步骤：
 - 事实匹配：将“Fido会飞”匹配给规则中的“X会飞”，将“Fido卵生”匹配给“X卵生”
 - 推理：当事实证据均成立时，得到结论成立，即Fido是鸟类。

不确定性推理：规则可信度计算

- 当问题扩展到不确定推理时，就有下面几个问题：
- 第一个问题：在推理阶段，对于一个不确定性规则，如

IF E THEN H (0.7)

- 意思是：如果 $CF(E)=1.0$ 即完全成立，则 $CF(H)=0.7$ 。
 - 那么： 如果 $CF(E)=0.5$ ，不确定， $CF(H)$ 应该怎么算？
- 我们用下面公式来计算：

$$CF(H) = CF(H, E) * \max\{0, CF(E)\}$$

- 意思是：结论 H 的可信度，是规则的可信度与证据可信度的乘积。证据可信度最低为0，即不考虑证据不成立的情况。
- 显然，本例中， $CF(H) = 0.5 * 0.7 = 0.35$

4.3 可信度方法

3. 不确定性的传递算法

- C—F模型中的不确定性推理：从不确定的初始证据出发，通过运用相关的不确定性知识，最终得出结论并求出结论的可信度值。结论 H 的可信度由下式计算：

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

当 $CF(E) = 1$ 时，则 $CF(H) = CF(H, E)$

当 $CF(E) < 0$ 时，则 $CF(H) = 0$

4.3 可信度方法

4. 证据组合不确定性的算法

- 组合证据：多个单一证据的合取

$$E = E_1 \text{ AND } E_2 \text{ AND } \cdots \text{ AND } E_n$$

则 $CF(E) = \min\{CF(E_1), CF(E_2), \dots, CF(E_n)\}$

IF E_1 AND E_2 THEN $H(0.7)$

– 假设, $CF(E_1) = 0.5$, $CF(E_2) = 0.9$

– 那么 $CF(E_1 \text{ AND } E_2)$ 是多少?

- 对于合取AND类型的证据组合, 我们认为整个证据应该取每个证据可信度的最小值, 即:

$$CF(E_1 \text{ AND } E_2) = \min\{CF(E_1), CF(E_2)\}$$

4.3 可信度方法

4. 证据组合不确定性的算法

组合证据：多个单一证据的析取

$$E = E_1 \text{ OR } E_2 \text{ OR } \cdots \text{ OR } E_n$$

则

$$CF(E) = \max \{CF(E_1), CF(E_2), \dots, CF(E_n)\}$$

$$CF(E_1 \text{ OR } E_2) = \max \{CF(E_1), CF(E_2)\}$$

回顾：多规则确定性推理

- 确定性的推理规则：

IF X是三角形 AND 底角A=底角B THEN X是等腰三角形

IF X是三角形 AND 边A=边B THEN X是等腰三角形

- 两条规则的结论相同，在推理过程中，只需要满足其一，结论就成立。

多规则的不确定性推理

如： R1: IF 头痛 AND 流涕 THEN 感冒 (0.7)

R2: IF 身体酸痛 AND 咳嗽 THEN 感冒 (0.8)

已知：头痛(0.8),流涕 (0.4) ,身体酸痛 (0.2) ,咳嗽 (0.9)

可得：R1和R2情况下，感冒的可信度

$$R1: \min\{0.8, 0.4\} * 0.7 = 0.28$$

$$R2: \min\{0.2, 0.9\} * 0.8 = 0.16$$

似乎哪一条都无法推出感冒。但患者头痛、咳嗽明显，是明显的感冒症状。

4.3 可信度方法

5. 结论不确定性的合成算法

不确定性推理：多规则合成

- 因此，对于不确定性推理，当多条规则得到同样的结论时，应该共同参与推理，称为规则合成。具体来说，分成三种情况来讨论，得到下面公式：

$$CF_{1,2}(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H)CF_2(H), & \text{if } CF_1(H) \geq 0, CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H)CF_2(H), & \text{if } CF_1(H) < 0, CF_2(H) < 0 \\ CF_1(H) + CF_2(H), & \text{if } CF_1(H)CF_2(H) < 0 \end{cases}$$

- 简单解释：
 - 如果两条规则均支持结论，则两者可信度相加，同时减去两者重复支持的部分
 - 如果两条规则均反对结论，同理。
 - 如果两条规则，一条支持，一条反对，则简单将两者的可信度相加。
- 如果涉及三条以上的规则，则可以两两合成，最终合成为一个结果

已知如下规则：

R1: IF A1 THEN B1 (0.8)

R2: IF A2 THEN B1 (0.5)

R3: IF B1 AND A3 THEN B2 (0.8)

事实: $CF(A1)=1.0$, $CF(A2)=1.0$, $CF(A3)=1.0$

求: $CF(B1)$, $CF(B2)$

- 求解过程：
- 由R1: $CF(B1) = 1 * 0.8 = 0.8$
- 由R2: $CF(B1) = 1 * 0.5 = 0.5$
- 两者合成, 得到 $CF(B1) = 0.8 + 0.5 - 0.8 * 0.5 = 0.9$
- 对R3中的 B1 AND A3 , 有 $CF(B1 \text{ AND } A3) = \min\{0.9, 1.0\} = 0.9$
- 则: $CF(B2) = 0.9 * 0.8 = 0.72$

4.3 可信度方法

在MYCIN系统的基础上形成的专家系统工具EMYCIN中，对多规则的合成可信度 $CF_{1,2}(H)$ 作了修改

求出 E_1 与 E_2 对 H 的综合影响所形成的可信度 $CF_{1,2}(H)$ ：

$$CF_{1,2}(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H)CF_2(H) & \text{若 } CF_1(H) \geq 0, \quad CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H)CF_2(H) & \text{若 } CF_1(H) < 0, \quad CF_2(H) < 0 \\ \frac{CF_1(H) + CF_2(H)}{1 - \min\{|CF_1(H)|, |CF_2(H)|\}} & \text{若 } CF_1(H) \text{ 与 } CF_2(H) \text{ 异号} \end{cases}$$

4.3 可信度方法

□ 例4.1 设有如下一组知识：

$$r_1: \quad IF \quad E_1 \quad THEN \quad H \quad (0.8)$$

$$r_2: \quad IF \quad E_2 \quad THEN \quad H \quad (0.6)$$

$$r_3: \quad IF \quad E_3 \quad THEN \quad H \quad (-0.5)$$

$$r_4: \quad IF \quad E_4 \quad AND \quad (E_5 \quad OR \quad E_6) \quad THEN \quad E_1 \quad (0.7)$$

$$r_5: \quad IF \quad E_7 \quad AND \quad E_8 \quad THEN \quad E_3 \quad (0.9)$$

已知： $CF(E_2)=0.8, CF(E_4)=0.5, CF(E_5)=0.6, CF(E_6)=0.7,$

$CF(E_7)=0.6, CF(E_8)=0.9.$

求： $CF(H)$

4.3 可信度方法

解：

第一步：对每一条规则求出 $CF(H)$ 。

r_4 ：

$$\begin{aligned} CF(E_1) &= 0.7 \times \max\{0, CF[E_4 \text{ AND } (E_5 \text{ OR } E_6)]\} \\ &= 0.7 \times \max\{0, \min\{CF(E_4), CF(E_5 \text{ OR } E_6)\}\} \\ &= 0.7 \times \max\{0, \min\{CF(E_4), \max\{CF(E_5), CF(E_6)\}\}\} \\ &= 0.7 \times \max\{0, \min\{0.5, \max\{0.6, 0.7\}\}\} \\ &= 0.7 \times \max\{0, 0.5\} \\ &= 0.35 \end{aligned}$$

4.3 可信度方法

解：

第一步：对每一条规则求出 $CF(H)$ 。

$$\begin{aligned}r_5: \quad CF(E_3) &= 0.9 \times \max\{0, CF(E_7 \text{ AND } E_8)\} \\&= 0.9 \times \max\{0, \min\{CF(E_7), CF(E_8)\}\} \\&= 0.9 \times \max\{0, \min\{0.6, 0.9\}\} \\&= 0.9 \times \max\{0, 0.6\} \\&= 0.54\end{aligned}$$

$$\begin{aligned}r_1: \quad CF_1(H) &= 0.8 \times \max\{0, CF(E_1)\} \\&= 0.8 \times \max\{0, 0.35\} \\&= 0.28\end{aligned}$$

4.3 可信度方法

解：

第一步：对每一条规则求出 $CF(H)$ 。

$$\begin{aligned}r_2 : \quad CF_2(H) &= 0.6 \times \max\{0, CF(E_2)\} \\ &= 0.6 \times \max\{0, 0.8\} \\ &= 0.48\end{aligned}$$

$$\begin{aligned}r_3 : \quad CF_3(H) &= -0.5 \times \max\{0, CF(E_3)\} \\ &= -0.5 \times \max\{0, 0.54\} \\ &= -0.27\end{aligned}$$

4.3 可信度方法

第二步：根据结论不确定性的合成算法得到：

$$\begin{aligned} CF_{1,2}(H) &= CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H) \\ &= 0.28 + 0.48 - 0.28 \times 0.48 = 0.63 \end{aligned}$$

$$\begin{aligned} CF_{1,2,3}(H) &= \frac{CF_{1,2}(H) + CF_3(H)}{1 - \min\{|CF_{1,2}(H)|, |CF_3(H)|\}} \\ &= \frac{0.63 - 0.27}{1 - \min\{0.63, 0.27\}} = \frac{0.36}{0.73} = 0.49 \end{aligned}$$

综合可信度： $CF(H) = 0.49$

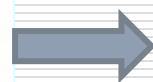
小结：不确定推理

- 不确定性推理的主要问题，与确定性推理的区别
- 基于概率的贝叶斯推理：由先验概率推理出后验概率

两个先验知识：

先验概率: $P(w_1)$ 和 $P(w_2)$

类条件概率: $P(x | w_1)$ 和 $P(x | w_2)$



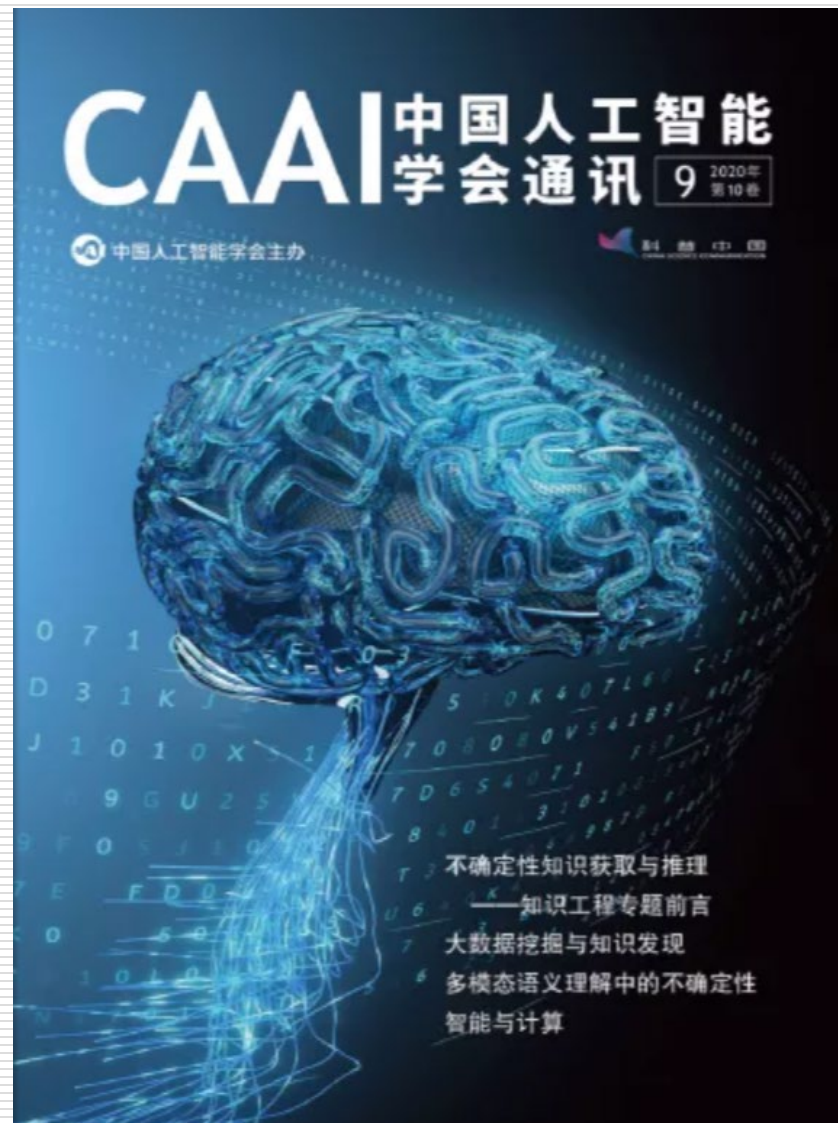
• 后验概率: $P(\omega_i | x) = \frac{P(x | \omega_i) P(\omega_i)}{P(x)}$

$$p(x) = \sum p(x | w_j) P(w_j), j=1,2$$

- 先验概率、类条件概率、全概率、后验概率
- CF推理，又称可信度推理，是不确定性推理方法，是由Shortliffe等人提出；
- 该方法在专家系统MYCIN中得到的成功应用，是不确定性推理的典型代表。

task3: 拓展阅读

- ❑ 中国人工智能学会通讯2020.9月刊（不确定性知识获取与推理）
- ❑ <https://book.yunzhan365.com/poui/moqq/mobile/index.html>





THE END

