

第4章 不确定性推理

✓ 4.1 概述

4.2 可信度方法

4.3 主观Beyes方法

4.4 模糊推理（见其它ppt）

4.1 概述

4.1.1 知识的不确定性

4.1.2 不确定性推理方法分类

4.1.3 不确定推理要解决的基本问题

4.1 概 述

在确定性推理中，一些结论一定是对的，另一些结论一定是错的。但在不少推理中，有的结论既不是一定对，也不是一定错。它们“可以这样，也可以那样”，具有所谓“**不确定性**”。

不确定性是智能问题的本质特征，无论是人类智能还是人工智能，都离不开不确定性的处理。可以说，**智能主要反映在求解不确定性问题的能力上**。因此，不确定性推理模型是人工智能的一个核心研究课题。

研究**不确定性推理**，首先要研究**知识的不确定性**。

4.1.1 知识的不确定性

- **确定性的量：**经典数学（建立在cantor提出的集合论基础上），如数学分析，代数，几何，和泛函分析

- **不确定性量：**

随机性：清晰事件发生的可能性。这里事件是确定的，预测是否发生是不确定的。

如投硬币，问正面向上的可能性。

概率论

模糊性：客观事物差异（事物与事物的不同）的中间过渡中表现出的亦此亦彼性。事物与事物的之间有界限，但界限可能明确也可能不明确。即事物本身具有不确定性。

如年龄：老，小；成绩：优秀，不优秀；头发：秃，不秃。

模糊数学

■ 知识的不确定性，还来自知识的不完备性、不协调性和非恒常性。

(1) 知识的不完备性：包括知识内容的不完整、知识结构的不完备等。内容的不完整，可能来源于获取知识时观测不充分、设备不精确；知识结构的不完备，可能因为人的认识能力、获取手段的限制等原因，造成对解决某个特定问题的背景和结构认识不全，忽略了一些重要因素。

(2) 知识的不协调性：是指知识内在的矛盾，不协调的程度可以依次为冗余、干扰、冲突等。不协调性是知识不确定性的重要体现，是知识的一种常态。

(3) 知识的非恒常性：主要是指知识随时间的变化而变化的特性。人类对自然、社会乃至自身的认识都是一个由不知到知、由不深刻到深刻，不断更新的过程。人类的知识是无限发展的，永远不可能停留在某个水平上。

4.1.2 不确定推理方法的分类

- 两种不确定性推理方法：
- 在推理一级上扩展不确定性推理的方法
(模型方法)
- 在控制策略级处理不确定性的方法
(控制方法)

把不确定证据和不确定的知识分别与某种量度标准对应起来，并且给出更新结论不确定性算法，从而建立不确定性推理模式。

通过识别领域中引起不确定性的某些特征及相应的控制策略来限制或减少不确定性对系统产生的影响，这类方法没有处理不确定性的统一模型，其效果极大地依赖于控制策略。

■ 模型方法分为：

数值方法

非数值方法

数值方法是对不确定性的一种定量表示和处理方法。又分为**基于概率的方法**和**基于模糊集的方法**

非数值方法是指数值方法之外的其他各种处理不确定性的方法。主要的有**逻辑法**。

- 概率论与模糊数学所研究和处理的是两种不同的不确定性。
- 概率论处理的是由随机性引起的不确定性，模糊数学处理的是由模糊性引起的不确定性。

- **概率论**有着完善的理论和方法，而且具有现成的公式实现不确定性的合成与传递，因此可以用做度量不确定性的重要手段。
- **纯概率方法**虽然有严格的理论依据，但通常要求给出事件的先验概率和条件概率，而这些数据又不易获得，且需要较大量的计算，因此使其应用受到限制。
- 为了解决这个问题，人们在概率论的基础上发展起来了一些新的方法和理论，主要有
 - 可信度方法
 - 主观Bayes方法（又称主观概率论）
 - 证据理论等

尽管这些技术大多数是从实践中总结出来的工程性方法，对不确定性的处理往往不够严格，使用上也有很多局限性，但是它们却能解决一些问题，其结果能够给出令人满意的解释，符合人类认识世界的直觉。

4. 1. 3 不确定推理要解决的基本问题

- **推理**是运用知识求解问题的过程，是**证据**和**规则**相结合得出**结论**的过程。由于知识的不确定性，导致了所产生的结论的不确定性。
- **不确定性推理**反映了知识不确定性的动态积累和传播过程，推理的每一步都需要综合证据和规则的不确定因素，通过某种不确定性测度，寻找尽可能符合客观实际的计算模式，通过不确定测度的传递计算，最终得到结果的不确定测度。

- 要实现对不确定性知识的处理，要解决：
 - 不确定性的表示问题
 - 不确定性的计算问题
 - 计算的语义解释问题
 - 不确定性的量度问题

1. 表示问题

- **表示问题**指的是采用什么方法描述不确定性，这是解决不确定性推理的关键一步。
- **“不确定性”** 一般分为两类：
 - 知识（规则）的不确定性
 - 证据（事实）的不确定性

E: 证据;

H: 结论;

规则 (知识) $E \rightarrow H$;

(1) 知识不确定性 $(E \rightarrow H, f(H, E))$

(2) 证据不确定性 $(E, C(E))$

它表示相应知识的不确定性程度, 称为知识或规则强度。

注意: 一般来说, 证据不确定性的表示方法应与知识不确定性的表示方法保持一致。

它表示证据E的不确定性程度。称之为动态强度。证据有两种来源: 初始证据 (由用户给出); 前面推出的结论作为当前证据 (通过计算得到)。

2. 计算问题

- **计算问题**主要指不确定性的传播与更新，即获得新信息的过程。
- 它是在领域专家给出的**规则强度**和用户给出的**原始证据的不确定性**的基础上，定义一组函数，求出结论的不确定性度量。
- 它主要包括如下三个方面：

(1) 不确定性的传递算法

(A) 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。

(B) 在多步推理中, 如何把初始证据的不确定性传给结论。

也就是说, 已知规则的前提E的不确定性 $C(E)$ 和规则强度 $f(H, E)$, 求假设H的不确定性 $C(H)$, 即定义函数 f_1 , 使得:

$$C(H) = f_1(C(E), f(H, E))$$

(2) 结论不确定性合成

推理中有时会出现这样的一种情况, 用不同的知识进行推理, 得到了相同的结论, 但不确定性的程度却不相同。

即已知由两个独立的证据 E_1 和 E_2 求得的假设 H 的不确定性度量 $C_1(H)$ 和 $C_2(H)$, 求证据 E_1 和 E_2 的组合导致的假设 H 的不确定性 $C(H)$, 即定义函数 f_2 , 使得:

$$C(H) = f_2(C_1(H), C_2(H))$$

(3) 组合证据的不确定性算法

即已知证据 E_1 和 E_2 的不确定性度量 $C(E_1)$ 和 $C(E_2)$ ，求证据 E_1 和 E_2 的析取和合取的不确定性，即定义函数 f_3 和 f_4 使得：

$$C(E_1 \wedge E_2) = f_3(C(E_1), C(E_2))$$

$$C(E_1 \vee E_2) = f_4(C(E_1), C(E_2))$$

目前，关于组合证据的不确定性的计算已经提出了多种算法，用的最多的是如下3种：

(a) 最大最小法

$$C(E1 \wedge E2) = \min(C(E1), C(E2))$$

$$C(E1 \vee E2) = \max(C(E1), C(E2))$$

(b) 概率方法

$$C(E1 \wedge E2) = C(E1) \times C(E2)$$

$$C(E1 \vee E2) = C(E1) + C(E2) - C(E1) \times C(E2)$$

(c) 有界方法

$$C(E1 \wedge E2) = \max\{0, C(E1) + C(E2) - 1\}$$

$$C(E1 \vee E2) = \min\{1, C(E1) + C(E2)\}$$

3. 语义问题

- 语义问题指上述表示和计算的含义是什么？即对它们进行解释。

如 $f(H, E)$ 可理解为当前提 E 为真时，对结论 H 为真的一种影响程度， $C(E)$ 可理解为 E 为真的程度。

■ 4. 量度问题

- 知识（规则）的不确定性度量 $f(H, E)$ ，需要定义在下述3个典型情况下的取值：

若E为真，则H为真，这时 $f(H, E) = ?$

若E为真，则H为假，这时 $f(H, E) = ?$

E对H没有影响，这时 $f(H, E) = ?$

- 对于证据的不确定性度量 $C(E)$ ，需要定义在下述3个典型情况下的取值：

E为真， $C(E) = ?$

E为假， $C(E) = ?$

对E一无所知， $C(E) = ?$

第4章 不确定性推理

4.1 概论

✓ 4.2 可信度方法

4.3 主观 Bayes方法

4.4 模糊推理

4.2 可信度方法

- 可信度方法是由美国斯坦福大学肖特利夫(E.H.Shortliffe)等人在考察了非概率的和非形式化的推理过程后于已于1975年提出的一种不确定性推理模型,并于1976年首次在血液病诊断专家系统MYCIN中得到了成功应用。
- 它是不确定性推理中非常简单且又十分有效的一种推理方法。
- 有许多成功的专家系统都是基于这一方法建立起来的。

1. 可信度的定义

- 在C-F模型中，可信度最初定义为信任与不信任的差，即 $CF(H,E)$ 定义为：

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

由证据E得到假设H的可信度（也称为确定性因子）

MB(Measure Belief, MB)称为信任增长度，它表示因为与前提条件E匹配的证据的出现，使结论H为真的信任的增长程度。

MD(Measure Disbelief, MD)称为不信任增长度，它表示因为与前提条件E匹配的证据的出现，对结论H的不信任的增长程度。

1. 可信度的定义

■ MB(H,E)定义为

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

■ MD(H,E)定义为

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

- 式中：P(H)表示H的先验概率；P(H|E)表示在前提条件E所对应的证据出现的情况下，结论H的条件概率（后验概率）。

由MB与MD的定义可以得出如下结论：

- 当 $MB(H,E)>0$ 时，有 $P(H|E)>P(H)$ ，这说明由于E所对应的证据的出现增加了H的信任程度，但不信任程度没有变化。
- 当 $MD(H,E)>0$ 时，有 $P(H|E)<P(H)$ ，这说明由于E所对应的证据的出现增加了H的不信任程度，而不改变对其信任的程度。
- 根据前面对 $CF(H,E)$ 、 $MB(H,E)$ 、 $MD(H,E)$ 的定义，可得到 $CF(H,E)$ 的计算公式

CF(H,E)的计算公式

若 $CF(H,E) > 0$ ，则 $P(H|E) > P(H)$ 。说明由于前提条件E所对应证据的出现增加了H为真的概率，即增加了H为真的可信度， $CF(H, E)$ 的值越大，增加H为真的可信度就越大。

若 $CF(H,E) < 0$ ，则 $P(H|E) < P(H)$ 。这说明由于前提条件E所对应证据的出现减少了H为真的概率，即增加了H为假的可信度， $CF(H, E)$ 的值越小，增加H为假的可信度就越大。

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

根据CF、MB、MD的定义，可得性质：

■ (1) 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， $MB(H,E) = 0$

■ (2) 值域

$$0 \leq MB(H,E) \leq 1$$

$$0 \leq MD(H,E) \leq 1$$

$$-1 \leq CF(H,E) \leq 1$$

根据CF、MB、MD的定义，可得性质：

■ (3) 典型值

① 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

$$MB(H,E)=1, MD(H,E)=0$$

② 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

$$MB(H,E)=0, MD(H,E)=1$$

③ 当 $CF(H,E)=0$ 时，则 $P(H|E)=P(H)$ ，表示H与E独立即E所对应的证据的出现对H没有影响。

根据CF、MB、MD的定义，可得性质：

(4) 对H的信任增长度等于对非H的不信任增长度

- 根据MB、MD的定义及概率的性质

$$\text{MD}(\neg H, E) = \text{MB}(H, E)$$

同理：

$$\text{MD}(H, E) = \text{MB}(\neg H, E)$$

根据CF、MB、MD的定义，可得性质：

(5) 对H的可信度与对非H的可信度之和等于0

- 再根据CF的定义及MB、MD的互斥性有

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

可信度不是概率。对概率有

$$P(H) + P(\neg H) = 1 \text{ 且}$$

$$0 \leq P(H), P(\neg H) \leq 1$$

而可信度不满足此条件。

注意事项

实际应用中 $P(H)$ 和 $P(H|E)$ 的值是很难获得的，因此 $CF(H, E)$ 的值应由领域专家给出。

- 原则：
- 若相应证据的出现会增加 H 为真的可信度，则 $CF(H, E) > 0$ ，证据的出现对 H 为真的支持程度越高，则 $CF(H, E)$ 的值越大；
- 反之，证据的出现减少 H 为真的可信度，则 $CF(H, E) < 0$ ，证据的出现对 H 为假的支持程度越高，就使 $CF(H, E)$ 的值越小；
- 若相应证据的出现与 H 无关，则使 $CF(H, E) = 0$ 。

2. 可信度的计算

■ (1) 规则不确定性的表示

在C-F模型中，规则用产生式规则表示：

If E Then H (CF(H, E))

- E是规则的前提条件；
- H是规则的结论；
- 注意：

CF(H, E)是规则的可信度，也称为规则强度或知识强度，它描述的是知识的静态强度。这里前提和结论都可以是由复合命题组成。

2. 可信度的计算

(2) 证据不确定性的表示

单个证据：

在CF模型中，证据E的不确定性也是用可信度因子 $CF(E)$ 来表示的，其取值范围同样是 $[-1, 1]$ ，其典型值为：

当证据E肯定为真时： $CF(E)=1$ ；

当证据E肯定为假时： $CF(E)=-1$ ；

当证据E一无所知时： $CF(E)=0$ 。

2. 可信度的计算

注意事项:

- (1) 证据可信度的来源有以下两种情况：如果是初始证据，其可信度是由提供证据的用户给出的；如果是先前推出的中间结论又作为当前推理的证据，则其可信度是原来在推出该结论时由不确定性的更新算法计算得到的。
- (2) $CF(E)$ 所描述的是证据的强度。尽管它和知识的强度在表示方法上类似，但二者的含义却完全不同。知识的强度 $CF(H, E)$ 表示的是规则的强度，即当 E 所对应的证据为真时对 H 的影响程度，而强度 $CF(E)$ 表示的是证据 E 当前的不确定性程度。

2. 可信度的计算

组合证据:

对证据的组合形式可分为“合取”与“析取”两种基本情况。
当组合证据是多个单一证据的合取时，即

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

时，若已知 $CF(E_1)$, $CF(E_2)$, ..., $CF(E_n)$, 则

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

当组合证据是多个单一证据的析取时，即

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

时，若已知 $CF(E_1)$, $CF(E_2)$, ..., $CF(E_n)$, 则

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

2. 可信度的计算

(3) 不确定性的推理算法

- C-F模型中的不确定性推理实际上是从不确定性的初始证据出发，不断运用相关的不确定性知识(规则)，逐步推出最终结论和该结论的可信度的过程。而每一次运用不确定性知识，都需要由证据的不确定性和规则的不确定性去计算结论的不确定性。

2. 可信度的计算

(3) 不确定性的推理算法

- 1) 只有单条知识支持结论时，结论可信度的计算方法：

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

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

由上式可以看出，若 $CF(E) \leq 0$ ，即相应证据以某种程度为假，则 $CF(H)=0$

这说明在该模型中没有考虑证据偏向为假时或对证据真假一无所知时对结论H所产生的影响。

若 $CF(E) > 0$ 则 $CF(H)=CF(H,E) \times CF(E)$

其中若 $CF(E)=1$,则 $CF(H)=CF(H,E)$

2. 可信度的计算

(3) 不确定性的推理算法

■ 2) 多条知识支持同一结论时，结论不确定性的合成计算方法

即如果有两条规则推出一个相同结论，并且这两条规则的前提相互独立，结论的可信度又不相同，则可用不确定性的合成算法求出该结论的综合可信度。

2. 可信度的计算

- 设有如下规则：

If E1 Then H (CF(H, E1))

If E2 Then H (CF(H, E2))

- 则结论H的综合可信度可分以下两步计算：

第一步：分别对每条规则求出其CF(H)。即

$$CF1(H) = CF(H, E1) \times \max(0, CF(E1))$$

$$CF2(H) = CF(H, E2) \times \max(0, CF(E2))$$

2. 可信度的计算

第二步:用如下公式求E1与E2对H的综合可信度:

$$CF(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H) \times CF_2(H) & \text{若 } CF_1(H) \geq 0 \text{ 且 } CF_2(H) \geq 0 \\ CF_1(H) + CF_2(H) + CF_1(H) \times CF_2(H) & \text{若 } CF_1(H) < 0 \text{ 且 } CF_2(H) < 0 \\ CF_1(H) + CF_2(H) & \text{若 } CF_1(H) \text{ 与 } CF_2(H) \text{ 异号} \end{cases}$$

在后来基于MYCIN基础上形成的EMYCIN中, 对上式做了如下的修改:

如果 $CF_1(H)$ 和 $CF_2(H)$ 异号, 则:

$$CF(H) = \frac{CF_1(H) + CF_2(H)}{1 - \min\{|CF_1(H)|, |CF_2(H)|\}}$$

其他情况不变。

如果可由多条规则推出同一个结论, 并且这些规则的前提相互独立, 结论的可信度又不相同, 则可以将上述合成过程推广应用到多条规则支持同一条结论, 且规则前提可以包含多个证据的情况。这时合成过程是先把第一条与第二条合成, 然后再用该合成后的结论与第三条合成, 依次进行下去, 直到全部合成完为止。

ex

■ 例 设有如下一组规则:

R1: IF E1 THEN H (0.9)

R2: IF E2 THEN H (0.6)

R3: IF E3 THEN H (-0.5)

R4: IF E4 AND (E5 OR E6) THEN E1 (0.8)

已知: $CF(E2)=0.8$, $CF(E3)=0.6$, $CF(E4)=0.5$,

$CF(E5)=0.6$, $CF(E6)=0.8$

求H的综合可信度 $CF(H)$ 。

ex

- 解:由R4得到:

$$\begin{aligned} CF(E1) &= 0.8 \times \max\{0, CF(E4 \text{ AND } (E5 \text{ OR } E6))\} \\ &= 0.8 \times \max\{0, \min\{CF(E4), CF(E5 \text{ OR } E6)\}\} \\ &= 0.8 \times \max\{0, \min\{CF(E4), \max\{CF(E5), CF(E6)\}\}\} \\ &= 0.8 \times \max\{0, \min\{0.5, 0.8\}\} \\ &= 0.8 \times \max\{0, 0.5\} \\ &= 0.4 \end{aligned}$$

- 由R1得到:

$$\begin{aligned} CF1(H) &= CF(H, E1) \times \max\{0, CF(E1)\} \\ &= 0.9 \times \max\{0, 0.4\} = 0.36 \end{aligned}$$

ex

- 由R2得到:

$$\begin{aligned} CF2(H) &= CF(H, E2) \times \max\{0, CF(E2)\} \\ &= 0.6 \times \max\{0, 0.8\} = 0.48 \end{aligned}$$

- 由R3得到:

$$\begin{aligned} CF3(H) &= CF(H, E3) \times \max\{0, CF(E3)\} \\ &= -0.5 \times \max\{0, 0.6\} = -0.3 \end{aligned}$$

- 根据结论不确定性的合成算法得到:

$$\begin{aligned} CF_{1,2}(H) &= CF1(H) + CF2(H) - CF1(H) \times CF2(H) \\ &= 0.36 + 0.48 - 0.36 \times 0.48 \\ &= 0.84 - 0.17 = 0.67 \end{aligned}$$

ex

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

- 这就是所求出的综合可信度，即 $CF(H) = 0.53$ 。

结论可信度更新计算方法

- 已知 $CF(E)$, 规则 IF E THEN H($CF(H,E)$) 及 $CF(H)$, 求 $CF(H/E)$
分3种情况:

1. $CF(E)=1$ 时

$$CF(H/E) = \begin{cases} CF(H) + CF(H,E) - CF(H,E) \times CF(H) & \text{若 } CF(H) \geq 0, CF(H,E) \geq 0 \\ CF(H) + CF(H,E) + CF(H,E) \times CF(H) & \text{若 } CF(H) < 0, CF(H,E) < 0 \\ \frac{CF(H,E) + CF(H)}{1 - \min(|CF(H)|, |CF(H,E)|)} & \text{若 } CF(H) \text{ 与 } CF(H,E) \text{ 异号} \end{cases}$$

2. $0 < CF(E) < 1$ 时

$$CF(H/E) = \begin{cases} CF(H) + CF(H,E) \times CF(E) - CF(H) \times CF(H,E) \times CF(E) & \text{当 } CF(H) \geq 0, CF(H,E) \geq 0 \\ CF(H) + CF(H,E) \times CF(E) + CF(H) \times CF(H,E) \times CF(E) & \text{当 } CF(H) < 0, CF(H,E) < 0 \\ \frac{CF(H,E) \times CF(E) + CF(H)}{1 - \min(|CF(H)|, |CF(H,E) \times CF(E)|)} & \text{当 } CF(H) \text{ 和 } CF(H,E) \text{ 异号时} \end{cases}$$

3. $CF(E) \leq 0$ 时规则 IF E THEN H不可用, 对结论H的可信度无影响

第4章 不确定性推理

4.1 概论

4.2 可信度方法

✓ 4.3 主观 Bayes方法

4.4 模糊推理

4.3 主观 Bayes方法

- 主观Bayes方法是由杜达 (R. O. Duda) 等人在1976年在概率论的基础上，通过对Bayes公式的修正而形成的一种不确定性推理模型，并成功地应用在他们自己开发的地矿勘探专家系统PROSPECTOR中。

4.3.1 Bayes公式

- 在许多情况下, 同类事件发生的频率不高, 甚至很低, 无法做概率统计, 这时一般是根据观测的数据, 凭领域专家的经验给出一些主观上的判断, 称为**主观概率**。
- 主观概率一般可以解释为对证据和规则的**主观信任度**。
- 概率推理中起关键作用的就是所谓的**Bayes公式**, 它也是主观Bayes方法的基础。

1) . Bayes公式

■ 定义4.1（全概率公式）

设有事件 A_1, A_2, \dots, A_n 满足：

(1) $P(A_i) > 0 (i=1, 2, \dots, n)$;

(2) A_1, A_2, \dots, A_n 为样本空间 S 的一个划分，则对任何事件 B 来说，有下式成立：

$$P(B) = P(A_1) \cdot P(B|A_1) + P(A_2) \cdot P(B|A_2) + \dots + P(A_n) \cdot P(B|A_n)$$

全概率公式提供了计算 $P(B)$ 的方法。

■ 定义4.2 (Bayes公式) (由条件概率和全概率公式得到)

设有事件 A_1, A_2, \dots, A_n 满足:

(1) $P(A_i) > 0 (i=1, 2, \dots, n)$

(2) A_1, A_2, \dots, A_n 为样本空间 S 的一个划分, 则对任何事件 B 来说, 则有以下式成立:

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

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, 称为**后验概率**。

2) . 利用Bayes公式进行推理

- 在专家系统中，假设有：

If **E** Then **H**

其中E为前提条件，H为结论。

那么条件概率 $P(H|E)$ 就表示在E发生时，H的概率，可以用它作为证据E出现时结论H的确定性程度。

用 $P(H|E)$ 表示规则 $E \rightarrow H$ 的不确定程度

有多个条件 $E_i, i=1, 2, \dots, n$

- 同样对于复合条件 $E=E_1 \wedge E_2 \wedge \dots \wedge E_n$, 也可以用条件概率 $P(H|E_1 \dots E_n)$ 作为证据 E_1, \dots, E_n 出现时, 结论 H 的确定性程度。

即用 $P(H|E_1, E_2, \dots, E_n)$ 表示规则 $E_1 \wedge E_2 \wedge \dots \wedge E_n \rightarrow H$ 的不确定性程度。

有多个结论 $H_i, i=1, 2, \dots, n$

- 对于产生式规则 If E Then H_i ，用条件概率 $P(H_i | E)$ 作为证据 E 出现时，结论 H_i 的确定性程度。根据 Bayes 公式，可以得到

$$P(H_i | E) = \frac{P(H_i) \times P(E | H_i)}{\sum_{j=1}^n P(H_j) \times P(E | H_j)} \quad i=1, \dots, n$$

这就是说，当已知结论 H_i 的先验概率 $P(H_i)$ ，根据前提条件 E 所对应的证据出现的条件概率 $P(E | H_i)$ 就可以用上式求出相应证据出现时结论 H_i 的条件概率 $P(H_i | E)$ 。

- 当有多个证据 E_1, \dots, E_m 和多个结论 H_1, \dots, H_n , 并且每个证据都以一定程度支持每个结论时, 若 E_1, \dots, E_m 间相互独立, 根据独立事件的概率公式, Bayes公式可变为

$$P(H_i | E_1 \cdots E_m) = \frac{P(E_1 | H_i) \times P(E_2 | H_i) \times \cdots \times P(E_m | H_i) \times P(H_i)}{\sum_{j=1}^n P(E_1 | H_j) \times P(E_2 | H_j) \times \cdots \times P(E_m | H_j) \times P(H_j)} \quad (i=1, 2, \dots, n)$$

- 此时, 只要已知 H_i 的先验概率 $P(H_i)$ 以及 H_i 成立时证据 E_1, \dots, E_m 出现的条件概率 $P(E_1 | H_i), \dots, P(E_m | H_i)$, 就可利用上式计算出在 E_1, \dots, E_m 出现情况下 H_i 的条件概率 $P(H_i | E_1, \dots, E_m)$ 。

Example: 如果把 H_i ($i=1, 2, \dots, n$)当作一组可能发生的疾病, 把 E_j ($j=1, \dots, m$)当作相应的症状, $P(H_i)$ 是从大量实践中经统计得到的疾病 H_i 发生的先验概率, $P(E_j | H_i)$ 是疾病 H_i 发生时观察到症状 E_j 的条件概率, 则当对其病人观察到有症状 E_1, \dots, E_m 时, 应用上述Bayes公式就可计算出 $P(H_i | E_1, \dots, E_m)$, 从而得知病人患疾病 H_i 的可能性。

- Bayes推理的优点是它有较强的理论背景和良好的数学特性，当证据和结论都彼此独立时，计算的复杂度比较低。
- 但是它也有其局限性：

(1) 因为需要

$$\sum_{j=1}^n P(H_j) = 1$$

如果又增加一个新的假设，则对所有的 $1 \leq j \leq n+1$ ， $P(H_j)$ 都需要重新定义。

(2) Bayes公式(3)的应用条件是很严格的，它要求各事件互相独立，如证据间存在依赖关系，就不能直接使用此方法。

(3) 在概率论中，一个事件或命题的概率是在大量统计数据的基础上计算出来的，因此尽管有时 $P(E_j | H_i)$ 比 $P(H_i | E_j)$ 相对容易得到，但总的来说，要想得到这些数据仍然是一件相当困难的工作。

4.3.2 知识不确定性的表示

主观Bayes方法是在对Bayes公式修正的基础上形成的一种不确定性推理模型。主观Bayes方法通过使用专家的主观概率，避免了所需的大量统计计算工作。

主观Bayes方法

- 主观Bayes方法的推理计算的任务就是根据证据E的概率 $P(E)$ 及影响结论的知识的规则强度 (LS, LN) 把H的先验概率 $P(H)$ 更新为后验概率 $P(H|E)$ 或
- $P(H|\neg E)$

(1) 信任几率

■ 几率 $O(X)$:

$$O(X) = P(X) / (1 - P(X))$$

$$P(X) = O(X) / (1 + O(X))$$

- 用 $P(X)$ 表示 X 出现的可能性， $O(X)$ 表示 X 的几率。显然随着 $P(X)$ 的增大， $O(X)$ 也在增大，并且

$$P(X) = 0 \quad \text{时有} \quad O(X) = 0$$

$$P(X) = 1 \quad \text{时有} \quad O(X) = \infty$$

$$0 < P(X) < 1 \quad \text{时有} \quad O(X) = (0, \infty)$$

这样，就可以把取值为 $[0, 1]$ 的 $P(X)$ 放大到取值为 $[0, +\infty)$ 的 $O(X)$ 。

(2) 信任几率和概率的关系

- 将两式相除，得

$$\frac{P(H|E)}{P(\neg H|E)} = \frac{P(E|H)}{P(E|\neg H)} \times \frac{P(H)}{P(\neg H)} \quad (5-8)$$

- 根据几率定义：

$$O(X) = \frac{P(X)}{1 - P(X)} \quad \text{或} \quad O(X) = \frac{P(X)}{P(\neg X)} \quad (5-9)$$

- 将(5-9)式代入(5-8)式有

$$O(H|E) = \frac{P(E|H)}{P(E|\neg H)} \times O(H) \quad (5-10)$$

- 其中：O(H)和O(H | E)分别表示H的先验几率和后验几率。

■ 定义似然率 (Likelihood Ratio)LS

$$LS = \frac{P(E | H)}{P(E | \neg H)} \quad (5-11) \leftarrow$$

代入 (5-10) 式, 可得

$$O(H|E) = LS \times O(H) \quad (5-12)$$

即

$$LS = O(H|E) / O(H) \quad (5-13)$$

■ 定义:

$$LN = P(\neg E | H) / P(\neg E | \neg H) \quad (5-14) \leftarrow$$

$$O(H | \neg E) = LN \times O(H) \quad (5-15) \leftarrow$$

■ 即 $LN = O(H | \neg E) / O(H) \quad (5-16)$

从公式 (5-12) 与公式 (5-15) 这两个公式可以看出：

(5-12)：当E为真时，可用LS将H的先验几率
 $O(H)$ 更新为其后验几率 $O(H|E)$ ；

(5-15)：当E为假时，可用LN将H的先验几率
 $O(H)$ 更新为其后验几率 $O(H|\neg E)$ ；

LS和LN与几率的关系

■ LS的性质

当 $LS > 1$ 时, $O(H|E) > O(H)$, 说明E支持H; LS越大, $O(H|E)$ 比 $O(H)$ 大得越多, 即LS越大,E对H的支持越充分。

当 $LS \rightarrow \infty$ 时, $O(H|E) \rightarrow \infty$, 即 $P(H|E) \rightarrow 1$, 表示由于E的存在, 将导致H为真。

当 $LS = 1$ 时, $O(H|E) = O(H)$, 说明E对H没有影响;

当 $LS < 1$ 时, $O(H|E) < O(H)$, 说明E不支持H;

当 $LS = 0$ 时, $O(H|E) = 0$, 说明E的存在使H为假。

由上述分析可以看出, LS反映的是E的出现对H为真的影响程度。因此称**LS为知识的充分性度量**。

LS和LN与几率的关系

■ LN的性质

当 $LN > 1$ 时, $O(H|\neg E) > O(H)$, 说明 $\neg E$ 支持 H , 即由于 E 的不出现, 增大了 H 为真的概率。并且 LN 越大, $P(H|\neg E)$ 就越大, 即 $\neg E$ 对 H 为真的支持就越强。

当 $LN \rightarrow \infty$ 时, 即 $O(H|\neg E) = 0$, 表示当 $\neg E$ 为真, H 必为假。

当 $LN = 1$ 时, $O(H|\neg E) = O(H)$, 说明 $\neg E$ 对 H 没有影响;

当 $LN < 1$ 时, $O(H|\neg E) < O(H)$, 说明 $\neg E$ 不支持 H ;

当 $LN = 0$ 时, $O(H|\neg E) = 0$, 说明 $\neg E$ 的存在将导致 H 为假。

由上述分析可以看出, LN 反映的是 E 不存在时对 H 为真的影响。因此称 LN 为知识的必要性度量。

LS和LN的关系

由于E和 $\neg E$ 不会同时支持或同时排斥H，因此只有下述三种情况存在：

LS>1且LN<1

LS<1且LN>1

LS=LN=1

事实上，如果LS>1，即

$$\begin{aligned}LS > 1 &\equiv P(E|H) / P(E|\neg H) > 1 \equiv P(E|H) > P(E|\neg H) \\ &\equiv 1 - P(E|H) < 1 - P(E|\neg H) \equiv P(\neg E|H) < P(\neg E|\neg H) \\ &\equiv P(\neg E|H) / P(\neg E|\neg H) \equiv LN < 1\end{aligned}$$

同理可证明“LS<1 \equiv LN>1”和“LS=1 \equiv LN=1”。

4.3.3 知识（规则）不确定性的表示

- 在主观Bayes方法中，规则是用产生式表示的，其形式为

IF E THEN (LS, LN) H

其中，(LS, LN)用来表示该规则的强度。

- 在实际系统中，LS和LN的值均是由领域专家根据经验给出的，而不是计算出来的；当证据E愈是支持H为真时，则LS的值应该愈大；当证据E对H愈是必要时，则相应的LN的值应该愈小。因此，LS和LN除了在推理过程中使用以外，还可以作为领域专家为LS和LN赋值的依据。

4.3.4 证据（规则）不确定性的表示

- 证据通常可以分为**全证据(Complete Evidence)**和**部分证据(Partial Evidence)**。
 - **全证据**就是所有的证据，即所有可能的证据和假设，它们组成证据E。
 - **部分证据**S就是我们所知道的E的证据，这一部分证据也可以称为**观察**。
- 全证据的可信度依赖于部分证据，表示为 $P(E|S)$ 。如果知道所有的证据，则 $E=S$ 。
 - 其中 $P(E)$ 就是证据E的**先验似然性**， $P(E|S)$ 是已知全证据E中部分知识S后对E的信任度，为E的**后验似然性**。

- 在主观Bayes方法中，证据E的不确定性可以用证据的似然性或几率来表示。
- 似然性与几率之间的关系为

$$O(E) = \frac{P(E)}{1 - P(E)} = \begin{cases} 0 & \text{当E为假时} \\ \infty & \text{当E为真时} \\ (0, +\infty) & \text{当E非真也非假时} \end{cases}$$

- 一般地，原始证据的不确定性通常由用户给定，作为中间结果的证据可以由下面的不确定性传递算法确定。

组合证据不确定性的计算

- 当组合证据是多个单一证据的合取时，即

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

如果已知在当前观察 S 下，每个单一证据 E_i 有概率 $P(E_1|S), P(E_2|S), \dots, P(E_n|S)$ ，则

$$P(E|S) = \min\{P(E_1|S), P(E_2|S), \dots, P(E_n|S)\}$$

- 当组合证据是多个单一证据的析取时

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

如果已知在当前观察 S 下，每个单一证据 E_i 有概率 $P(E_1|S), P(E_2|S), \dots, P(E_n|S)$ ，则

$$P(E|S) = \max\{P(E_1|S), P(E_2|S), \dots, P(E_n|S)\}$$

- 对于“非”运算，用下式计算：

（此处课本有问题，应该是 E 和 S 独立时下式才成立）

$$P(\neg E|S) = 1 - P(E|S)$$

4.3.5. 不确定性的传递算法

- 由于一条规则所对应的证据可能肯定为真，也可能肯定为假，还可能既非真又非假，因此，在把H的先验概率或先验几率更新为后验概率或后验几率时，需要根据证据的不同情况去计算其后验概率或后验几率。
- 下面就来分别讨论这些不同情况。

■ (1) 证据肯定为真

当证据E肯定为真，即全证据一定出现时，此时
 $P(E) = P(E|S) = 1$ 。将H的先验几率更新为后验几率的公式为：

$$O(H|E) = LS \times O(H)$$

如果是把H的先验概率更新为其后验概率，则根据几率和概率的对应关系有：

这是把先验概率 $P(H)$ 更新为其后验概率 $P(H|E)$ 的计算公式。

$$P(H | E) = \frac{LS \times P(H)}{(LS - 1) \times P(H) + 1}$$

■ (2) 证据肯定为假

当证据E肯定为假，即证据不出现时， $P(E)=P(E|S)=0$ ， $P(\neg E)=1$ 。将H的先验几率更新为后验几率的公式为：

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

如果是把H的先验概率更新为其后验概率，则有：

这是把先验概率 $P(H)$ 更新为其后验概率 $P(H|\neg E)$ 的计算公式。

$$P(H|\neg E) = \frac{LN \times P(H)}{(LN - 1) \times P(H) + 1}$$

■ (3) 证据既非为真又非为假

当证据既非为真又非为假时，不能再用上面的方法计算H的后验概率。这时因为H依赖于证据E，而E基于部分证据S，则 $P(H|S)$ 是H依赖于S的似然性。

这时可以使用下面的公式来计算 $P(H|S)$ 的值：

$$P(H|S) = P(H|E) \times P(E|S) + P(H|\neg E) \times P(\neg E|S)$$

分4种情况讨论：

■ ① $P(E|S)=1$

当 $P(E|S)=1$ 时， $P(\neg E|S)=0$ 。则有

$$P(H|S) = P(H|E) = \frac{LS \times P(H)}{(LS - 1) \times P(H) + 1}$$

这实际上就是证据肯定存在的情况。

分4种情况讨论：

■ ② $P(E|S)=0$

当 $P(E|S)=0$ 时， $P(\neg E|S)=1$ 。则有

$$P(H|S) = P(H|\neg E) = \frac{LN \times P(H)}{(LN - 1) \times P(H) + 1}$$

这实际上是证据肯定不存在的情况。

分4种情况讨论：

- ③ $P(E|S)=P(E)$

当 $P(E|S)=P(E)$ 时，表示 E 与 S 无关。

$$\begin{aligned} P(H|S) &= P(H|E) \times P(E|S) + P(H|\neg E) \times P(\neg E|S) \\ &= P(H|E) \times P(E) + P(H|\neg E) \times P(\neg E) \\ &= P(H) \end{aligned}$$

- 通过上述分析，已经得到了 $P(E|S)$ 上的 3 个特殊值 0, $P(E)$ 及 1, $P(H|S)$ 分别取得了对应值 $P(H|\neg E)$, $P(H)$ 及 $P(H|E)$ 。这样就构成了 3 个特殊点：

$$(0, P(H|\neg E)), (P(E), P(H)), (1, P(H|E))$$

分4种情况讨论：

■ ④ $P(E|S)$ 为其他值

当 $P(E|S)$ 为其他值时， $P(H|S)$ 的值可通过上述3个特殊点的分段线性插值函数求得。

■ 函数的解析表达式为

$$P(H|S) = \begin{cases} P(H|\neg E) + \frac{P(H) - P(H|\neg E)}{P(E)} \times P(E|S), & \text{若 } 0 \leq P(E|S) < P(E) \\ P(H) + \frac{P(H|E) - P(H)}{1 - P(E)} \times [P(E|S) - P(E)], & \text{若 } P(E) \leq P(E|S) \leq 1 \end{cases}$$

该分段线性插值函数 $P(H|S)$ 如图5-6所示：

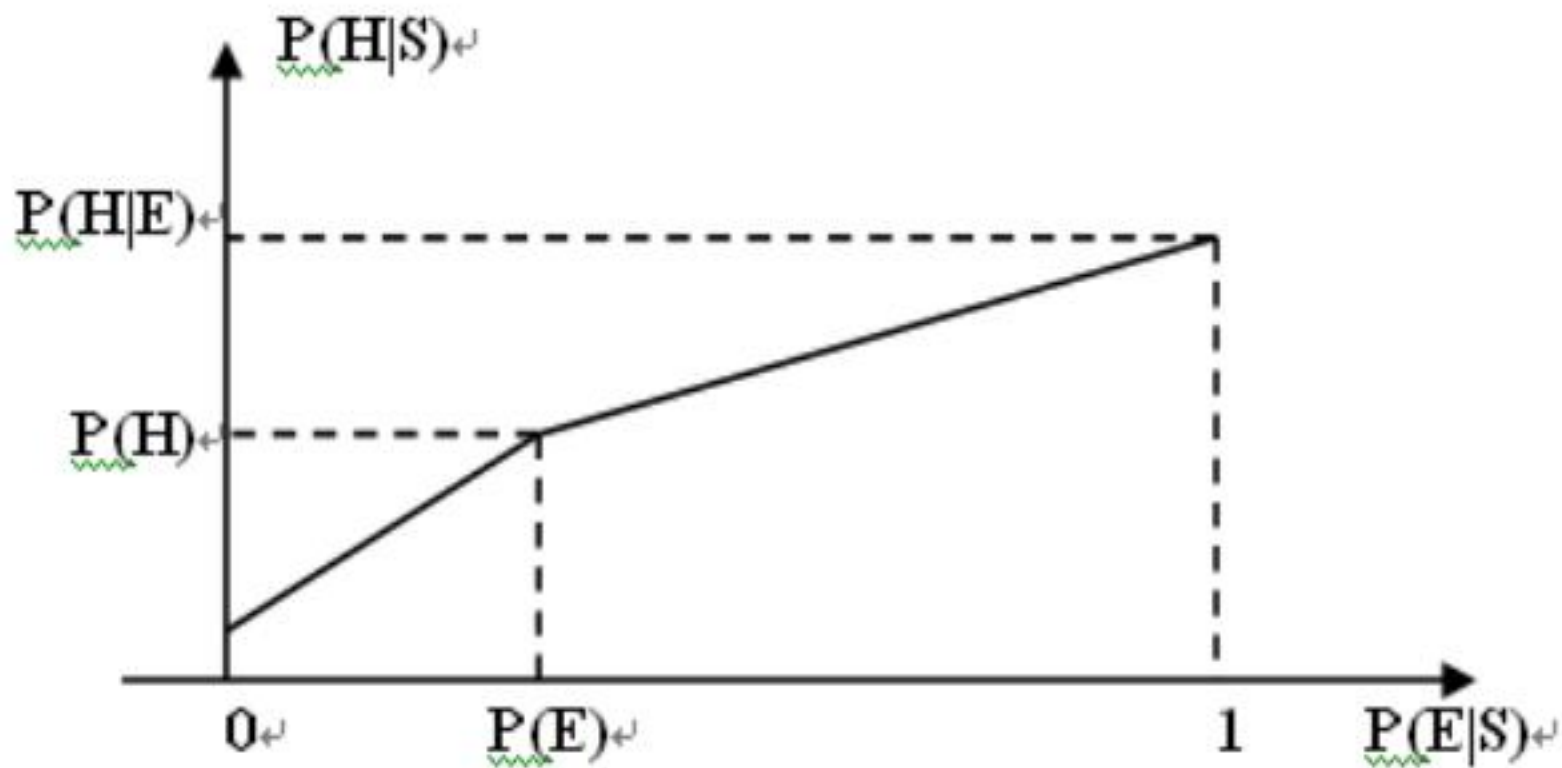


图 5-6 分段线性插值函数

4.3.6 结论不确定性的合成

假设有n条知识都支持同一结论H，且这些知识的前提条件分别是n个相互独立的证据E1，E2...，En，而每个证据所对应的观察又分别是S1，S2...，Sn。

在这些观察下，求H的后验概率的方法：

(1)首先对每条知识分别求出H的后验几率是 $O(H|S_i)$;

(2)利用这些后验几率并按下述公式求出所有观察下H的后验几率：

$$O(H | S_1, S_2, \dots, S_n) = \frac{O(H | S_1)}{O(H)} \times \frac{O(H | S_2)}{O(H)} \times \dots \times \frac{O(H | S_n)}{O(H)} \times O(H)$$

$$P(H | S_1, S_2, \dots, S_n) = \frac{O(H | S_1, S_2, \dots, S_n)}{1 + O(H | S_1, S_2, \dots, S_n)}$$

例 设有规则

R1: If E1 Then (20,l) H

R2: If E2 Then (300,l) H

已知证据E1和E2必然发生，并且 $P(H)=0.03$ ，求H的后验概率。

解: 因为 $P(H)=0.03$ ，则

$$O(H)=0.03/(1-0.03)=0.030927$$

根据R1有:

$$O(H|E1)=LS1 \times O(H)=20 \times 0.030927=0.6185$$

根据R2有:

$$O(H|E2)=LS2 \times O(H)=300 \times 0.030927=9.2781$$

■ 那么

$$O(H | E_1 E_2) = \frac{O(H | E_1)}{O(H)} \times \frac{O(H | E_2)}{O(H)} \times O(H)$$

$$= 0.6185 \times 9.2781 / 0.030927 = 185.55$$

所以H的后验概率为

$$P(H | E_1 E_2) = \frac{O(H | E_1 E_2)}{1 + O(H | E_1 E_2)}$$

$$= 185.55 / (1 + 185.55) = 0.99464$$

- **主观Bayes方法有优点：**

- 1) 该方法基于概率理论，具有坚实的理论基础，是目前不确定推理中最成熟的方法之一；
- 2) 计算量适中。

- **主观Bayes方法不足：**

- 1) 要求有大量的概率数据来构造知识库，并且难于对这些数据进行解释；
- 2) 在原始证据具有相互独立性，并能提供精确且一致的主观概率数据的情况下，该方法可以令人满意地处理不确定推理。但在实际当中，这些概率值很难保证一致性。

第4章 不确定推理

4.1 概论

4.2 主观 Bayes方法

4.3 可信度方法

✓ 4.4 模糊推理