第四章2　不确定性推理

- 基本概念
- 概率方法
- 可信度方法

1 基本概念

什么是不确定性推理

- 不确定性推理是建立在非经典逻辑基础上的一种推理，它是对不确定性知识的运用与处理。
- 具体地说，所谓不确定性推理就是从不确定性的初始证据（即事实）出发，通过运用不确定性的知识，最终推出具有一定程度不确定性的结论。

2 不确定性推理中的基本问题

1. 不确定性的表示与度量

- 不确定性推理中的“不确定性”一般分为两类：一是知识的不确定性，一是证据的不确定性。
- 知识不确定性的表示：目前在专家系统中知识的不确定性一般是由领域专家给出的，通常用一个数值表示，它表示相应知识的不确定性程度，称为知识的静态强度。
- 证据不确定性的表示：证据不确定性的表示方法与知识不确定性的表示方法一致，通常也用一个数值表示，代表相应证据的不确定性程度，称之为动态强度。

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

推理是不断运用知识的过程,为了找到所需的知识,需要在这一过程中用知识的前提与已知证据进行匹配.只有匹配成功的知识才有可能被应用.

- 设计一个不确定性匹配算法；
- 指定一个匹配阈值。

3. 组合证据不确定性的计算方法

即已知证据E_1 和 E_2 的不确定性度量，求证据E_1和 E_2的析取和合取的不确定性，常用的方法有：

- 最大最小法：
  $T(E_1 \text{ AND } E_2)=\min\{T(E_1),T(E_2)\}$
  $T(E_1 \text{ OR } E_2)=\max\{T(E_1),T(E_2)\}$
- 概率法：
  $T(E_1 \text{ AND } E_2)=T(E_1)\times T(E_2)$
  $T(E_1 \text{ OR } E_2)=T(E_1)+T(E_2)-T(E_1)\times T(E_2)$
- 有界法：
  $T(E_1 \text{ AND } E_2)=\max\{0,T(E_1)+T(E_2)-1\}$
  $T(E_1 \text{ OR } E_2)=\min\{1,T(E_1)+T(E_2)\}$

其中，T(E)表示证据E为真的程度（动态强度），如可信度、概率等。

不确定性推理中的基本问题

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

用不同知识进行推理得到了相同结论，但所得结论的不确定性却不同。此时，需要用合适的算法对结论的不确定性进行合成。

不确定性推理方法的分类

- 不确定性推理方法主要可分为模型法与控制法。
- 模型法：在推理一级对确定性推理进行扩展，引入证据的不确定性及知识的不确定性。
- 模型方法又分为数值方法和非数值方法两类。数值方法对不确定性进行定量的描述，按其所依据的理论又可分为基于概率的方法和基于模糊理论的方法。本节主要针对模型方法中相关的典型算法展开.

7

概率推理方法

经典概率方法

（1）设有如下产生式规则：

IF E THEN H

其中，E为前提条件，H为结论。条件概率P(H|E)可以作为在证据E出现时结论H的确定性程度，即规则的静态强度。

（2）对于复合条件

$E=E_1$ AND E_2 AND ... AND E_n

当已知条件概率$P(H|E_1,E_2,\ldots,E_n)$时，就可把它作为在证据$E_1,E_2,\ldots,E_n$出现时结论H的确定性程度。

（3）先验概率：P(H)　　后验概率：P(H|E)

8

概率推理方法

经典概率方法要求给出条件概率P(H/E)，在实际中通常比较困难。例如E代表咳嗽，H代表支气管炎，则P(H/E)表示在咳嗽的人群中患支气管炎的概率，这个比较困难，因为样本空间太大。而逆概率P(E/H)表示在得支气管炎的人群中咳嗽的概率，这个就比较容易获得。

我们可以根据Bayes定理从P(E/H)推出P(H/E)

9

概率推理方法

若$A_1,A_2,\ldots,A_n$是彼此独立的事件，对于事件B，则有以下贝叶斯概率

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

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。

对于一组产生式规则

IF E THEN H_i

同样有后验概率如下（H_i确定性的程度，或规则的静态强度）：

$$P(H_i|E)=\frac{P(H_i)\times P(E|H_i)}{\sum_{j=1}^{n}P(H_j)\times P(E|H_j)},i=1,2,\ldots,n$$

10

对于多个证据

对于有多个证据E1，E2，…，Em和多个结论H1，H2，… Hn，并且每个证据都以一定程度支持结论的情况，上面的式子可进一步扩展为

$$P(H_i|E_1E_2\cdots E_m)=\frac{P(H_i)\times P(E_1|H_i)\times P(E_2|H_i)\times\cdots\times P(E_m|H_i)}{\sum_{j=1}^{n}P(H_j)\times P(E_1|H_j)\times P(E_2|H_j)\times\cdots\times P(E_m|H_j)},i=1,2,\ldots,n$$

11

概率方法举例

例1 设H_1,H_2,H_3分别是三个结论，E是支持这些结论的证据。已知：

$P(H_1)=0.2$,　$P(H_2)=0.5$,　$P(H_3)=0.3$

$P(E|H_1)=0.1$,　$P(E|H_2)=0.3$,　$P(E|H_3)=0.7$

求$P(H_1|E)$,$P(H_2|E)$及$P(H_3|E)$的值各是多少？

解：

$$P(H_1|E)=\frac{P(H_1)\times P(E|H_1)}{P(H_1)\times P(E|H_1)+P(H_2)\times P(E|H_2)+P(H_3)\times P(E|H_3)}$$
$$=\frac{0.02}{0.02+0.15+0.21}$$
$$=0.056$$

同理可得：$P(H_2|E)=0.3947$, $P(H_3|E)=0.5526$

12

对应的产生式规则：

IF E THEN H_1
IF E THEN H_2
IF E THEN H_3

规则的静态强度(H_i为真的程度、或不确定性程度)

$P(H_1|E)=0.056$
$P(H_2|E)=0.3947$
$P(H_3|E)=0.5526$

逆概率法的特点

优点：

- **逆概率法有较强的理论背景和良好的数学特性，当证据彼此独立时计算的复杂度比较低。**

缺点：

- **逆概率法要求给出结论H_i的先验概率$P(H_i)$及条件概率$P(E_j|H_i)$。**

可信度方法

- **可信度方法是在确定性理论的基础上，结合概率论等提出的一种不确定性推理方法，简称C-F模型。该方法首先在医疗系统MYCIN中得到成功的应用。**
- **可信度的概念**
- **根据经验对一个事物和现象为真的相信程度称为可信度。**
- **在可信度方法中，由专家给出规则或知识的可信度，从而可避免对先验概率、或条件概率的要求。**

5.4.2 C-F模型

知识不确定性的表示:

在C-F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:

（1）前提E可以是命题的合取和析取组合

（2）结论H可为单一命题，也可以是复合命题

（3）CF(H, E)为确定性因子(Certainty factor)，简称可信度，用以量度规则的确定性（可信）程度。取值于[-1，1]，表示E为真时，对H的支持程度。CF(H, E)值越大，E就越支持H为真。

可信度因子的定义

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

CF(H,E)定义为:

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

MB反映了证据对结论有利的一面，MD反映了证据对结论不利的一面。MB(Measure Belief) 表示因与E匹配的证据出现，使H为真的信任增长度。MD(Measure Disbelief)指不信任增长度，表示因与E匹配的证据出现，使H为真的不信任增长度。 MB和MD的定义为:

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

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

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

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

- **当P(H|E)>P(H)时：表示证据E支持结论H**
  MB(H,E)>0，MD(H,E)=0 。
- **当P(H|E)<P(H)时，表示E不支持H**
  MD(H,E)>0， MB(H,E) =0。
  当p(H/E)＝p(H)时，表示E对H无影响，则有MB＝MD＝0
- **MB(H,E)与MD(H,E)是互斥的:**
  当MB(H,E)>0时，MD(H,E)＝0
  当MD(H,E)>0时，MB(H,E)＝0

CF(H,E)的计算公式

根据定义CF(H,E)=MB(H,E)-MD(H,E)，及MB(H,E)与MD(H,E)的互斥性，可得:

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

从上式可看出：

CF(H,E)>0对应于P(H|E)>P(H);

CF(H,E)<0对应于P(H|E)<P(H);

CF(H,E)=0对应于P(H|E)=P(H)。

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

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

当且仅当P(H|E)=1时，CF(H,E)=1

当且仅当P(H|E)=0时，CF(H,E)=-1

CF(H,E)定性地反映了P(H|E)的大小,因此可以用CF(H,E)近似表示P(H|E)的大小,从而描述了规则的可信度。

2.证据不确定性的表示

证据的不确定性也用可信度因子表示。如： CF(E)=0.6

CF(E)的取值范围：[-1，+1]。

CF(E)>0:表示证据以某种程度为真。

CF(E)<0:表示证据以某种程度为假。

CF(E)表示证据的强度，即动态强度。

证据不确定性的表示

设证据E所在的环境为 *S*，则可用可信度CF(E, *S*)来表示E在*S*下的确定性程度，并有：

CF(E, S) = MB(E, S) - MD(E, S)

若*S*下E为真，则 CF(E, *S*) = 1；

若E为假，则 CF(E, *S*) =-1；

若*S*对E的真值无影响，则CF(E, *S*)= 0。

类似于规则的不确定性，证据的可信度往往可由领域专家凭经验主观确定。

证据的可信度值来源于两种情况：

（1）初始证据由领域专家或用户给出；

（2）中间结论由不确定性传递算法计算得到。

3、组合证据不确定性的算法

(1)当组合证据是多个单一证据的合取时，即：

E=E1 AND E2 AND…AND En

则CF(E)=min{CF(E1), CF(E2)… CF(En)}

(2)当组合证据是多个单一证据的析取时，即：

E=E1 OR E2 OR…OR En

则CF(E)=max{CF(E1), CF(E2)… CF(En)}

4. 不确定性的传递

不确定性的传递算法定义如下：

CF(H)= CF(H,E) ×max [0，CF(E)]

由上式可以看出:

(1)CF(E)<0时,CF(H)=0,说明该模型没有考虑证据为假时对结论H所产生的影响。

（2）CF(E)=1时,CF(H)=CF(H,E),说明规则可信度CF(H,E)就是证据为真时的结论H的可信度。

5、结论不确定性的合成算法

若由**多条不同知识推出了相同的结论**，但可信度不同，则可用合成算法求出综合的可信度。由于对多条知识的综合可通过两两的合成实现，所以下面只考虑两条知识的情况。

设有如下知识：

IF E1 THEN H （CF(H,E1)）

IF E2 THEN H （CF(H,E2)）

则结论H的综合可信度可分为如下两步算出：

（1）首先分别对每一条知识求出CF(H)

CF1(H)=CF(H,E1) ×max{0,CF(E1)}

CF2(H)=CF(H,E2) ×max{0,CF(E2)}

（2）然后用下述公式求出E_1与E_2对H的综合可信度CF_{12}(H):

$$CF_{12}(H)=\begin{cases} CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H) & ,CF_1(H)\ge 0,CF_2(H)\ge 0 \\ CF_1(H)+CF_2(H)+CF_1(H)\times CF_2(H) & ,CF_1(H)<0,CF_2(H)<0 \\ \dfrac{CF_1(H)+CF_2(H)}{1-\min\{|CF_1(H)|,|CF_2(H)|\}} & ,CF_1(H)\times CF_2(H)<0 \end{cases}$$

C-F模型推理示例(1)

例5.5 设有如下一组知识：

R1: IF E_1 THEN H (0.8)

R2: IF E_2 THEN H (0.6)

R3: IF E_3 THEN H (-0.5)

R4: IF E_4 AND (E_5 OR E_6) THEN E_1 (0.7)

R5: IF E_7 AND E_8 THEN E_3 (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)=？

解：由R4得到：

CF(E_1)=0.7×max{0,CF[E_4 AND (E_5 OR E_6)]}

=0.7×max{0,min{CF(E_4),CF(E_5 OR E_6)}}

=0.35

由R5得到：

CF(E_3)=0.9×max{0,CF[E_7 AND E_8]}

=0.54

R1: IF E_1 THEN H (0.8)

R2: IF E_2 THEN H (0.6)

R3: IF E_3 THEN H (-0.5)

由R1得到：

CF_1(H)=0.8×max{0,CF(E_1)}=0.28

由R2得到：

CF_2(H)=0.6×max{0,CF(E_2)}=0.48

由R3得到：

CF_3(H)=-0.5×max{0,CF(E_3)}=-0.27

根据结论不确定性的合成算法：

CF_{12}(H)=CF_1(H)+CF_2(H)-CF_1(H)×CF_2(H)=0.63

CF_{123}(H)=[CF_{12}(H)+CF_3(H)]/[1-min{|CF_{12}(H)|,|CF_3(H)|}]

=0.49

即最终的综合可信度为CF(H)=0.49。

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

C-F模型的核心问题是三个可信度：

(1) 知识的可信度CF(H,E)：取值范围[-1，1]

CF(H,E)=1 对应于 P(H|E)=1 (证据绝对支持结论)

CF(H,E)=-1 对应于 P(H|E)=0 (证据绝对否定结论)

CF(H,E)=0 对应于 P(H|E)=P(H) (证据与结论无关)

(2) 证据的可信度CF(E)：取值范围[-1，1]

CF(E)=1 对应于 P(E)=1 (证据绝对存在)；

CF(E)=-1 对应于 P(E)=0; (证据绝对不存在)

CF(E)=0 对应于 P(E)=0.5 (对证据一无所知)。

(3) 结论的可信度CF(H)：取值范围[-1，1]

CF(H)=CF(H,E)×max{0,CF(E)}

该公式隐含了一个知识运用的条件，即CF(E)>0。

基于可信度的不确定性推理方法的特点

优点：

- 简单、直观。

缺点：

- 可信度因子依赖于专家主观指定，没有统一、客观的尺度，容易产生片面性。
- 随着推理延伸，可信度越来越不可靠，误差越来越大。当推理深度达到一定深度时，有可能出现推出的结论不再可信的情况。