西安电子科技大学 XIDIAN UNIVERSITY

人工智能概论

刘若辰
西安电子科技大学 人工智能学院

第七章 不确定推理

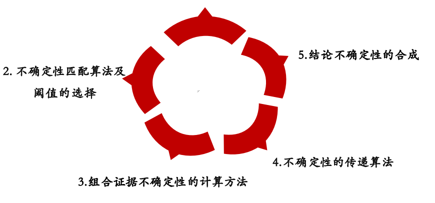
1. 基本概念
2. 概率方法-逆概率推理
3. 可信度方法

基本概念

什么是不确定性推理

- 推理：从已知的事实或证据出发，通过运用相关知识逐步推出结论或者某个假设成立或不成立的过程。
- 不确定性推理：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。

不确定性推理中的基本问题



不确定性推理中的“不确定性”一般分为两类：

知识的不确定性　证据的不确定性。

知识不确定性的表示：

- 在知识表示时，1）要根据领域问题的特征把不确定准确的表示出来，满足问题求解的需要，2）另一方面要考虑推理中对不确定性的推算。
- 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。静态强度可以是相应知识在应用中成功的概率、或是该条知识在运用中的可信度或其他。

证据不确定性的表示：

- 证据：用户在求解问题时提供的初始证据或在推理过程中得到的结论作为当前的证据
- 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。

不确定性的量度：

- 不确定性的程度是不同的，需要用数值表示其确切程度，一般在推理中会规定其取值范围，比如可信度的范围是【-1 1】

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：计算证据与知识的匹配程度
- 指定一个匹配阈值：指定一个相似程度范围。

3. 组合证据不确定性的计算方法

知识的前提条件不一定都是简单条件，可能是有很多个简单条件组合起来的复合条件，一个简单条件对应于一个单一的证据，一个复合条件对应于一组证据，称这一组证据为组合证据。

有三种计算组合证据不确定性的方法：

- 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

- 不确定性程度由用户在证据给出，通过推理过程中运用知识（也有不确定性）传递到结论，而结论又可作为新的证据供以后的推理用，这样不确定性依次传递，到最后的结论。

5. 结论不确定性的合成

- 当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。基于概率的方法包括可信度方法，逆概率方法以及证据理论。

经典概率推理方法

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

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

当已知条件概率$P(H|E_1,E_2,\cdots,E_n)$时，就可把它作为在证据$E_1,E_2,\cdots,E_n$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)

逆概率推理方法

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。

根据Bayes定理从P(E/H)推出P(H/E)

逆概率推理方法

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

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

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）：

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

逆概率推理方法

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

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

逆概率方法举例

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：

$P(H_1)=0.3$，$P(H_2)=0.4$，$P(H_3)=0.5$，$P(E|H_1)=0.5$，$P(E|H_2)=0.3$，

$P(E|H_3)=0.4$

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

解：
$$P(H_1|E)=\frac{P(H_1)\times P(E|H_1)}{P(H_1)\times P(E|H_1)+P(H_2)\times P(E|H_2)+P(H_3)\times P(E|H_3)}$$
$$=\frac{0.15}{0.15+0.12+0.2}=0.32$$

同理可得：$P(H_2|E)=0.26$，$P(H_3|E)=0.43$

逆概率方法举例

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、或确定性程度)

$P(H_1|E)=0.32$

$P(H_2|E)=0.26$

$P(H_3|E)=0.43$

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$P(E_j|H_i)$
- 要求时间独立，实际中经常不满足。

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：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。

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，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。

IF 发热38° 以上AND 四肢关节疼痛无力 AND 胸闷咳嗽 THEN 患SARS (0.7)

表示病人如有上述症状则有七成把握认为他患SARS

可信度因子的定义

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

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

- 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：

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

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

可信度因子的定义

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

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

- 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。
- p(H/E)＝p(H)时，表示E对H无影响，则有MB＝MD＝0
- 同一个证据不可能既增加对H的信任程度，又增加对H的不信任程度，所以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:

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

从上式可看出：

CF(H,E)>0对应于P(H|E)>P(H)：表示证据E支持结论H

CF(H,E)<0对应于P(H|E)<P(H)：表示E不支持H

CF(H,E)=0对应于P(H|E)=P(H)：表示E对H无影响

CF(H,E)的计算公式

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

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

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

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

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

来源：作为初始证据，其值由用户给出；由先前推出的结论作为当前推理的证据，其值通过传递算法计算得到。

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)=0：无关证据

CF(E)表示证据的强度，即动态强度。

CF(H,E)和CF(E)的区别：
静态强度CF(H,E)表示知识的强度，即E为真时对H的影响程度。
动态强度CF(E)表示证据的不确定性程度。

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)}

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)］

其中H可作为E的证据来供下次推理使用，因而可信度由初始证据开始通过传递算法传递到最终结论。

由上式可以看出：

（1）CF(E)<0时, E为假。 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。

$$CF(H)=\begin{cases} CF(H,E)\times max[0,\ CF(E)] & ,CF(E)>0 \\ CF(H,E) & ,CF(E)=1 \\ 0 & ,CF(E)<0 \end{cases}$$

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))

结论不确定性的合成算法

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$CF_{12}(H)$:

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

西安电子科技大学 XIDIAN UNIVERSITY

C-F模型推理示例

CF(H)= CF(H,E) ×max［0，CF(E)］

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)=?

人工智能学院

西安电子科技大学 XIDIAN UNIVERSITY

C-F模型推理示例(

CF(H)= CF(H,E) ×max［0，CF(E)］

例5.5 设有如下一组知识：

R1: IFE_1 THEN H (0.8)

R2: IFE_2 THEN H (0.6)

R3: IFE_3 THEN H (-0.5)

R4: IFE_4 AND (E_5 OR E_6) THEN E_1 (0.7)

R5: IFE_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：

$CF(E_1)=0.7\times\max\{0,CF[E_4 \text{ AND } (E_5 \text{ OR } E_6)]\}$

$=0.7\times\max\{0,\min\{CF(E_4),CF(E_5 \text{ OR } E_6)\}\}=0.35$

由R5得到：

$CF(E_3)=0.9\times\max\{0,CF[E_7 \text{ AND } E_8]\}=0.54$

人工智能学院

西安电子科技大学 XIDIAN UNIVERSITY

C-F模型推理示例

R1: IF E_1 THEN H (0.8)

R2: IF E_2 THEN H (0.6)

R3: IF E_3 THEN H (-0.5)

$CF(E_1)=0.35$，$CF(E_3)=0.54$

由R1得到：$CF_1(H)=0.8\times\max\{0,CF(E_1)\}=0.28$

CF1(H)=CF(H,E1) ×max{0,CF(E1)}

由R2得到：$CF_2(H)=0.6\times\max\{0,CF(E_2)\}=0.48$

由R3得到：$CF_3(H)=-0.5\times\max\{0,CF(E_3)\}=-0.27$

根据结论不确定性的合成算法：

$CF_{12}(H)=CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H)=0.63$

$CF_{123}(H)=[CF_{12}(H)+CF_3(H)]/[1-\min\{|CF_{12}(H)|,|CF_3(H)|\}]=0.49$

即最终的综合可信度为CF(H)=0.49。

人工智能学院

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。