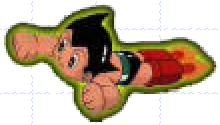

人工智能
Artificial Intelligence

主讲：相明
西安交通大学电信学院计算机系
E_mail：mxiang@mail.xjtu.edu.cn

第四章 不确定性推理

- 概率方法
- 主观Bayes方法
- 可信度方法
- 模糊推理方法

4.1 概述

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（即规则），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

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：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。

2. 不确定性匹配算法

在推理过程中证据和知识的前提的相似程度称为匹配度。确定这个匹配度（相似程度）的算法称为不确定性匹配算法。

3. 组合证据不确定性的计算方法

- 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, 通常取值在0，1之间。

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，即如何计算结论的不确定性。

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。

4.2 基本概率方法

经典概率方法

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。如果我们在实践中经大量统计能得出在E发生条件下H的条件概率（后验概率）P(H/E)，那么就可把它作为在证据E出现时结论H的确定性程度（可信度）。

逆概率方法

假设观测到某事件E，且对应于E有多个可能的结论H1,H2,…,Hn；则可用每个结论的先验概率P(Hi)和条件概率P(E/Hi)来计算在观察到E时结论Hi的后验概率P(Hi/E)

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

上面公式的背景是，存在n条规则 IF E THEN Hi, i=1,...,n。

4.3 主观Bayes方法

4.3.1 不确定性的表示

1、知识不确定性的表示：

在主观Bayes方法中，知识是用产生式规则表示的，具体形式为：

IF　E　THEN　(LS,LN)　H (P(H))

其中，E是知识的前提条件，既可以是简单条件，也可以是复合条件。

- P(H)是结论H的先验概率，由专家根据经验给出。
- LS称为充分性度量，用于指出E对H的支持程度，取值范围为[0,∞)，其定义为：

$$LS=P(E|H)/P(E|\neg H)。$$

- LN称为必要性度量，用于指出¬ E对H的支持程度，取值范围为[0,∞)，其定义为：

$$LN=P(\neg E|H)/P(\neg E|\neg H)=(1-P(E|H))/(1-P(E|\neg H))。$$

- LS和LN的值由领域专家给出，代表知识的静态强度。

2、证据不确定性的表示：

- 在主观Bayes方法中，证据的不确定性用概率表示。对于证据E，由用户根据观察S给出P(E|S)，即动态强度。用P(E|S)描述证据的不确定性（证据E无法直接观测到）。
- 由于主观给定P(E|S)有所困难，所以实际中可以用可信度C(E|S)代替P(E|S)。
- 在PROSPECTOR中C(E|S)取整数：{-5，….5}

C(E|S)=-5表示在观测S下证据E肯定不存在P(E|S)=0

C(E|S)= 5表示在观测S下证据E肯定存在P(E|S)=1

C(E|S)= 0表示S与E无关,即P(E|S)= P(E)

- 给定C(E|S)后，P(E|S)可近似计算如下：

$$P(E|S)=\begin{cases}\dfrac{C(E|S)+P(E)\times(5-C(E|S))}{5} & ,0\le C(E|S)\le 5\\[2ex] \dfrac{P(E)\times(5+C(E|S))}{5} & ,-5\le C(E|S)<0\end{cases}$$

3、组合证据的不确定性：

（1）当组合证据是多个单一证据的合取时，即

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

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

（2）当组合证据是多个单一证据的析取时，即

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

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

（3）对于“¬”运算则：$P(\neg E|S)=1-P(E|S)$

4.3.2 不确定性的传递算法

（1）根据证据E的条件概率P(E|S) 及LS、LN的值，把H的先验概率P(H)更新为后验概率P(H|S) 。

（2） 分以下3种情况讨论：

证据肯定存在：　　$P(E|S)=1$

证据肯定不存在：　$P(E|S)=0$

证据不确定：　　　$0<P(E|S)<1$

（3）引入几率函数Θ(x),它与概率的关系为：

$\Theta(x)=P(x)/(1-P(x))$,　　$P(x)=\Theta(x)/(1+\Theta(x))$

证据肯定存在时

$\Theta(x)=P(x)/(1-P(x))$,　$P(x)=\Theta(x)/(1+\Theta(x))$

计算P(H|E)：

由Bayes公式得：

$$P(H|E)=P(E|H)\times P(H)/P(E) \qquad (1)$$

$$P(\neg H|E)=P(E|\neg H)\times P(\neg H)/P(E) \qquad (2)$$

(1)式除以(2)式得：

$$P(H|E)/P(\neg H|E)=P(E|H)/P(E|\neg H)\times P(H)/P(\neg H)$$

由充分性度量LS和几率函数的定义立即可得：

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

即

$$P(H|E)=LS\times P(H)/[(LS-1)\times P(H)+1]$$

证据肯定不存在时

计算$P(H|\neg E)$：

由Bayes公式得：

$$P(H|\neg E)=P(\neg E|H)\times P(H)/P(\neg E) \quad (1)$$

$$P(\neg H|\neg E)=P(\neg E|\neg H)\times P(\neg H)/P(\neg E) \quad (2)$$

(1)式除以(2)式得：

$$P(H|\neg E)/P(\neg H|\neg E)=P(\neg E|H)/P(\neg E|\neg H)\times P(H)/P(\neg H)$$

根据必要性度量LN和几率函数的定义，可得：

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

即

$$P(H|\neg E)=LN\times P(H)/[(LN-1)\times P(H)+1]$$

证据不确定时

- 当$0<P(E|S)<1$时，可以证明:

  $P(H|S)=P(H|E)\times P(E|S)+P(H|\neg E)\times P(\neg E|S)$
- 当$P(E|S)=1$时，证据肯定存在，此时$P(H|S)=P(H|E)$。
- 当$P(E|S)=0$时，证据肯定不存在，此时$P(H|S)=P(H|\neg E)$。
- 当$P(E|S)=P(E)$时，证据E与观察S无关。由全概率公式得:

  $P(H|S)=P(H|E)\times P(E)+P(H|\neg E)\times P(\neg E)=P(H)$
- 当$P(E|S)$为其它值时，通过分段线性插值计算$P(H|S)$，即

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

- 对于知识：

IF E THEN (LS,LN) H (P(H))

给定证据E的不确定性度量P(E|S)，则结论的可信度可以表示为：

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

该公式称为EH公式。用C(E/S) 代替P(E/S)，可得到等价的CP公式。

LS和LN的性质

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

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

LS>1：表明证据 E是对H有利的证据。
LN>1：表明证据¬E是对H有利的证据。
所以：不能出现LS>1且LN>1的取值。

LS<1：表明证据 E是对H不利的证据。
LN<1：表明证据¬E是对H不利的证据。
所以：不能出现LS<1且LN<1的取值。

一般情况下，取LS>1，LN<1。

4.4.3 结论不确定性的合成算法

- 若有n条规则支持相同的结论H，即

  IF E_i THEN (LS_i,LN_i) H (P(H)),i=1,…,n

且每条知识的前提条件所对应的证据E_i(i=1,2,…,n)都有相应的观察S_i与之对应，则该n条规则均可启用。此时只要先对每条知识分别求出几率函数$\Theta(H|S_i)$，然后就可运用下述公式求出$\Theta(H|S_1S_2\ldots S_n)$:

$$\Theta(H|S_1S_2\cdots S_n)=\frac{\Theta(H|S_1)}{\Theta(H)}\times\frac{\Theta(H|S_2)}{\Theta(H)}\times\cdots\times\frac{\Theta(H|S_n)}{\Theta(H)}\times\Theta(H)$$

主观Bayes方法的特点

优点：

- 主观Bayes方法中的计算公式大多是在概率论的基础上推导出来的，具有较坚实的理论基础。
- 知识的静态强度LS及LN是由领域专家给出，避免了大量的数据统计工作。
- 主观Bayes方法不仅给出了证据肯定存在、肯定不存在时更新后验概率的方法，还给出了证据不确定时的更新方法，实现了不确定性的逐级传递。

缺点：

- 它要求领域专家在给出知识时，同时给出H的先验概率P(H)，这比较困难。

4.4 可信度方法

4.4.1 基本可信度模型

1、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。

2、C-F模型中知识（规则）的不确定性

在C-F模型中，知识（规则）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，CF(H,E)是该知识的可信度，称为可信度因子或规则强度，即静态强度。此处$CF(H,E)\in[-1,1]$。

$CF(H,E)=1$ 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)

CF 为Certainty Factor 的首字母缩写。

3、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。

4、 组合证据不确定性的算法

规则：IF E THEN H (CF(H,E))

若$E=E_1$ AND E_2 AND…AND E_n,则

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

若$E=E_1$ OR E_2 OR…OR E_n,则

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

5、不确定性的传递算法

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

结论H的可信度由下式计算：

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

CF(H)的取值范围：[-1，+1]。

CF(H)>0:表示结论以某种程度为真。

CF(H)<0:表示结论以某种程度为假。

6、 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，则用合成算法求出综合可信度。

设有如下知识：

IF E_1 THEN H $(CF(H,E_1))$

IF E_2 THEN H $(CF(H,E_2))$

则结论H的综合可信度分如下两步算出：

首先分别对每一条知识求出CF(H): 计算$CF_1(H)$、$CF_2(H)$

然后用下述公式求出E_1与E_2对H的综合可信度$CF_{12}(H)$:

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

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 对应证据绝对存在;

CF(E)=-1 对应证据绝对不存在；

CF(E)=0 表示对证据一无所知。

(3) 结论的可信度CF(H)：取值范围[-1，1]

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

该公式隐含了一个知识运用的条件, 即CF(E)>0。（负负得正）

4.4.2 带阈值限度的可信度模型

1. 知识不确定性的表示

知识用下述形式表示：

IF E THEN H (CF(H,E),λ)

其中：

- CF(H,E)为知识的可信度，取值范围为[0,1]。

  CF(H,E)=0 对应于 P(H|E)=0 (证据绝对否定结论)

  CF(H,E)=1 对应于 P(H|E)=1 (证据绝对支持结论)

- λ是阈值，明确规定了知识运用的条件：只有当CF(E)≥λ时，该知识才能够被应用。λ的取值范围为[0,1] 。

IF E THEN H (CF(H,E),λ)

2. 证据不确定性的表示

证据E的可信度仍为CF(E)，但其取值范围为：[0，1]

CF(E)=1 表示证据绝对存在;

CF(E)=0 表示证据绝对不存在。

3. 组合证据的不确定性

4. 不确定性的传递算法

当CF(E)≥λ时，CF(H)=CF(H,E)×CF(E)

注意： CF(H,E)表示证据E为真的条件下结论H为真的可能性。

5. 结论不确定性的合成算法

设有多条规则有相同的结论，即

IF E_1 THEN H $(CF(H,E_1),\lambda_1)$

IF E_2 THEN H $(CF(H,E_2),\lambda_2)$

…

IF E_n THEN H $(CF(H,E_n),\lambda_n)$

如果这n条规则都满足：$CF(E_i)\geq\lambda_i$，$i=1,2,\ldots,n$

且都被启用，则首先分别对每条知识求出它对$CF_i(H)$；

然后求结论H的综合可信度CF(H)。

求综合可信度的几种方法

极大值法：

$CF(H)=\max\{CF_1(H),CF_2(H),\ldots,CF_n(H)\}$

加权求和法：

$$CF(H)=\frac{1}{\sum_{i=1}^{n}CF(H,E_i)}\sum_{i=1}^{n}CF(H,E_i)\times CF(E_i)$$

有限和法：

$$CF(H)=\min\{\sum_{i=1}^{n}CF_i(H),1\}$$

递推法：

$C_1=CF(H,E_1)\times CF(E_1)$

$C_k=C_{k-1}+(1-C_{k-1})\times CF(H,E_k)\times CF(E_k)$

4.4.3 加权的可信度模型

考虑复合前提条件：

E=E1 AND E2 AND…AND En

前面所讨论的模型都认为各个子条件的重要性是完全相等的，各个子条件之间的地位是完全平等的。但是现实中并非都是如此。很可能有一些子条件对结论的影响相对其它更大一些；有些子条件更重要一些。

如果　学生善于思考
并且　动手能力强
并且　经常上自习
并且　坚持锻炼身体
并且　不抽烟
那么　该生成绩较好

1. 知识的不确定性

IF $E_1(\omega_1)$ AND $E_2(\omega_2)$ AND…AND $E_n(\omega_n)$

THEN H (CF(H,E),λ)

其中ω_i(i=1, 2, …, n)是加权因子，λ是阈值，其值均由专家给出。

加权因子的取值范围一般为[0, 1], 且应满足归一条件，即

$$0\le\omega_i\le1,\quad i=1,2,\cdots,n,\quad \sum_{i=1}^{n}\omega_i=1$$

2. 组合证据的不确定性

若有CF(E_1)，CF(E_2)， …， CF(E_n)，则组合证据的可信度为：

$$CF(E)=\frac{1}{\sum_{i=1}^{n}\omega_i}\sum_{i=1}^{n}(\omega_i\times CF(E_i))$$

3. 不确定性的传递算法

当一条知识的CF(E)满足如下条件时，

$$CF(E) \geq \lambda$$

该知识就可被应用。结论H的可信度为：

$$CF(H)=CF(H,E)\times CF(E)$$

- 加权因子的引入不仅可以区分不同证据的重要性，同时还可以解决证据不全时的推理问题。

4.4.4 前件带不确定性的可信度模型

1. 知识不确定性的表示

IF $E_1(cf_1)$ AND $E_2(cf_2)$ AND…AND $E_n(cf_n)$ THEN H (CF(H,E),λ)

其中，cf_i为子条件$E_i(i=1,2,…,n)$的可信度。cf_i在[0,1]上取值，其值由专家给出，反映了专家对子证据可信度的一种要求。

IF $E_1(cf_1,\omega_1)$ AND $E_2(cf_2,\omega_2)$ AND…AND $E_n(cf_n,\omega_n)$ THEN H (CF(H,E),λ)

证据不确定性的表示：子证据E_i的可信度记为cf'_i，其取值范围在[0,1]上，cf'_i即前面的$CF(E_i)$。

2. 不确定性匹配算法

◆ 不带加权因子的不确定性匹配算法：

知识：IF $E_1(cf_1)$ AND $E_2(cf_2)$ AND…AND $E_n(cf_n)$

THEN H (CF(H,E),λ)

证据：$E_1(cf'_1)$，$E_2(cf'_2)$，…，$E_n(cf'_n)$

匹配算法：

$$\max\{0,cf_1-cf'_1\}+\max\{0,cf_2-cf'_2\}+\ldots+\max\{0,cf_n-cf'_n\}\le\lambda$$

◆ 带加权因子的不确定性匹配算法：

知识：IF $E_1(cf_1,\omega_1)$ AND $E_2(cf_2,\omega_2)$ AND…AND $E_n(cf_n,\omega_n)$

THEN H (CF(H,E),λ)

匹配算法：

$$(\omega_1\times\max\{0,cf_1-cf'_1\})+(\omega_2\times\max\{0,cf_2-cf'_2\})+\ldots+(\omega_n\times\max\{0,cf_n-cf'_n\})\le\lambda$$

3. 不确定性的传递算法

- 不带加权因子时：

$CF(H)=[(1-\max\{0,cf_1-cf'_1\})\times(1-\max\{0,cf_2-cf'_2\})\times\ldots\times(1-\max\{0,cf_n-cf'_n\})]\times CF(H,E)$

- 带加权因子时：

$CF(H)=[(\omega_1\times(1-\max\{0,cf_1-cf'_1\}))\times(\omega_2\times(1-\max\{0,cf_2-cf'_2\}))\times\ldots\times(\omega_n\times(1-\max\{0,cf_n-cf'_n\}))]\times CF(H,E)$

$CF(H)=[(1-\omega_1\max\{0,cf_1-cf'_1\})\times(1-\omega_2\max\{0,cf_2-cf'_2\})\times\ldots\times(1-\omega_n\max\{0,cf_n-cf'_n\})]\times CF(H,E)$

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。

完

谢谢