第 5 章

不确定性推理

5.1 概述

在现实生活中遇到的问题通常都具有不确定性，能够进行精确描述的问题只占较少的一部分。对于这些不确定性的问题，若采用前述确定性推理方法显然是无法解决的。因此，为满足现实世界问题求解的需求，人工智能需要研究不确定性推理方法。可以将知识分成确定性知识和不确定性知识，不确定性是智能问题的一个本质特征，是建立在不确定性知识和证据基础上的推理。而知识的不确定性主要体现在两个方面：随机性和模糊性。

随机性主要由概率知识体现，因此研究随机性即通过概率和马尔可夫链等知识进行。模糊性与随机性一样，在生活中几乎无处不在，我们常说的“高、矮、长、短、大、小”都体现了一种模糊性现象，研究模糊性则通过模糊集理论，该理论以及模糊推理将在 5.6 节阐述。

5.1.1 为什么要采用不确定性推理

采用不确定性推理是客观问题的需求，其原因主要包括以下几个方面。

(1) 所需知识不完备、不精确。在很多情况下，解决问题所需要的知识往往是不完备、不精确的。所谓知识的不完备是指在解决某一问题时，不具备解决该问题所需要的全部知识。例如，医生在看病时，一般是从病人的部分症状开始诊断的。所谓知识的不精确是指既不能完全确定知识为真，又不能确定知识为假。例如，专家系统中的知识多为专家经验，而专家经验又多为不确定性知识。

(2) 所需知识描述模糊。知识描述模糊是指知识的边界不明确。例如，平常人们所说的“很好”“好”“比较好”“不很好”“不好”“很不好”等概念，其边界都是比较模糊的。那么，当用这类概念描述知识时，所描述的知识当然也是模糊的。例如，“如果王刚这个人比较好，那么我就把他当作好朋友”描述的就是比较模糊的知识。

(3) 多种原因导致同一个结论。在现实世界中,可由多种不同原因导出同一结论的情况有很多。例如,引起人体低烧的原因至少有几十种,医生在看病时,只能根据病人的症状、低烧的持续时间以及病人的体质、病史等做出猜测性的推断。

(4) 解决方案不唯一。现实生活中存在的问题一般都存在多种不同的解决方案,并且这些方案之间又很难绝对地判断其优劣。对于这种情况,人们往往是优先选择主观上认为相对较优的方案,这也是一种不确定性推理。

总之,在人类的知识和思维行为中,确定性只是相对的,而不确定性才是绝对的。人工智能要解决这些不确定性问题,必须采用不确定性的知识表示和推理方法。

5.1.2 不确定性推理要解决的问题

可将不确定性推理要解决的问题概括为以下5类。

1. 如何进行知识和证据表示的问题

不确定性表示要解决的问题包括知识的不确定性表示和证据的不确定性表示。

不确定性推理中,知识是否能够很好地被表示将直接影响推理的运行效率。一般地,用数值刻画知识的不确定性,该数值称为知识的静态强度或者知识的可信度,描述了知识的不确定性程度。静态强度可以用概率、可信度或者隶属度表示,如果用概率或者隶属度表示,其取值范围是[0,1]。取值越接近1,说明确定性程度越高;反之,取值越接近0,说明确定性程度越低。通常情况下,静态强度是由领域专家给出的或是由实验统计方法得到的。

证据通常包括两部分,其一是求解问题时已有的初始证据;其二是将求解问题中得到的中间结果放入综合数据库,作为后续推理的证据。为了使推理能对不确定性进行处理,证据的不确定性表示通常与知识的不确定性表示一致,即证据的不确定性表示也是用数值表示,代表相应事实的不确定性程度,称为动态强度。

2. 如何进行匹配的问题

在确定性推理中,当从规则库中取出的某规则的前提与综合数据库中的已知事实相一致时,便认为是匹配成功。只有当匹配成功时,相关的规则才会被激活使用。在不确定性推理过程中,同样要解决匹配问题,但由于知识和证据都是不确定的,所以确定性匹配方式不能采用。目前,常用的方法是事先设定一个阈值,用来衡量知识和证据的相似程度,如果知识和证据的不确定性程度在阈值限度内,则认为可以进行匹配,否则不能匹配。这里比较关键的是如何计算出相似程度,往往要视实际具体情况的不同而采用不同的算法。

3. 如何进行证据组合的问题

在不确定性推理过程中,证据可以是复杂的组合条件。此时则需要有合适的算法计算证据组合的不确定性。即已知证据 E_1 和 E_2 的不确定性值 $C(E_1)$ 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_1 和 f_2,使得

$$C(E_1 \wedge E_2) = f_1(C(E_1), C(E_2))$$

$$C(E_1 \vee E_2) = f_2(C(E_1), C(E_2))$$

具体常用来计算证据组合不确定性的方法有以下3种:

(1) 最大最小方法:

$$C(E_1 \wedge E_2) = \min(C(E_1), C(E_2))$$
$$C(E_1 \vee E_2) = \max(C(E_1), C(E_2))$$

(2) 概率方法:

$$C(E_1 \wedge E_2) = C(E_1) \times C(E_2)$$
$$C(E_1 \vee E_2) = C(E_1) + C(E_2) - C(E_1) \times C(E_2)$$

(3) 有界方法。

4. 不确定性的遗传问题

在不确定性推理中,存在两个主要的问题:一是如何用证据的不确定性得到结论的不确定性;二是如何在推理中把初始证据的不确定性传递给最终结论。

一方面,按照某种算法由证据和知识的不确定性计算出结论的不确定性,至于不确定性推理方法的处理方式各有不同;另一方面,不同不确定性推理方法的处理方式基本相同,都是把当前推出的结论及其不确定性作为新的证据放入综合数据库,方便以后推理使用,这样就实现了将不确定性传递给结论。重复这样的过程,可以将不确定性传递给最终结论。

5. 如何合成结论的问题

在不确定性推理过程中,可能出现由多个不同知识推出相同的结论,且不确定性程度不同的情况,需要采用相应算法对这些不同的不确定性进行合成,求出该结论的综合不确定性。

5.1.3 不确定性推理类型

不确定推理方法的研究主要沿着两条不同的路线发展:模型方法和控制方法。

模型方法是对确定性推理框架的一种扩展。模型方法把不确定性证据和不确定性知识分别与某种度量标准对应起来,并且给出了更新结论不确定性的算法,从而构成相应不确定性推理的模型。一般来说,模型方法与控制策略无关,即无论使用何种控制策略,推理的结果都是唯一的。

模型方法又分为数值方法和非数值方法两大类。数值方法是对不确定性的一种定量表示和处理方法。对于该类情况下,根据其所依据的理论可以将其分为两种不同的类型。一类是基于概率的方法,如确定性推理、主观贝叶斯方法、证据理论等;另一种是基于扎德提出的模糊集理论及其在此基础上发展的可能性理论。非数值方法是指除数值方法外的处理不确定性的方法,如框架推理、语义网络推理、常识推理等。

控制方法主要在控制策略一级处理不确定性。其特点是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限制或者减少不确定性对系统的影响。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。目前,常用的控制方法有启发式搜索和相关性制导回溯等。

本章只对模型方法展开讨论,有兴趣的读者可自行查阅文献了解控制方法。

5.2　概率基础

概率论的对象是随机现象。在概率中,把随机现象的某些样本点组成的集合称为随机事件,简称事件。而概率描述的就是随机事件发生的可能性。下面给出概率的公理化定义。

定义 5.1　设 Ω 为一个样本空间,F 为 Ω 的某子集组成的一个事件域,$P(A)$是 F 上的一个实值函数,对于任意一事件 $A\in F$,称 $P(A)$为事件 A 的概率,如果满足以下 3 条性质:

(1) 非负性。对于任意 $A\in F$,都有 $P(A)\geqslant 0$。

(2) 正则性。$P(\Omega)=1$。

(3) 可列可加性。若 $A_1,A_2,\cdots,A_n,\cdots$互不相容,有

$$P\bigcup_{i=1}^{+\infty}A_i=\sum_{i=1}^{+\infty}P(A_i)$$

在概率的学习中,通常还会给出概率的具体定义形式,包括统计概率、古典概率、集合概率、条件概率等。

定义 5.2(统计概率)　若在大量重复实验中,事件 A 发生的频率稳定地接近于一个固定的常数 p,它表明事件 A 出现的可能性,则称此常数 p 为事件 A 发生的概率,记为 $P(A)$,即

$$p=P(A)$$

定义 5.3(古典概率)　设一种实验有且仅有有限的 N 个等可能结果,即 N 个基本事件,而 A 事件包含着其中的 L 个可能结果,则称$\frac{L}{N}$为事件 A 的概率,记为 $P(A)$,即

$$P(A)=\frac{L}{N}$$

定义 5.4(集合概率)　假设 Ω 是集合型随机实验的基本事件空间,F 是 Ω 中一切可测集的集合,则对于 F 中的任意事件 A 的概率 $P(A)$为 A 与 Ω 的体积之比,即

$$P(A)=\frac{V(A)}{V(\Omega)}$$

定义 5.5(条件概率)　把事件 B 已经出现的条件下事件 A 发生的概率记作 $P(A|B)$,并称为在 B 出现的条件下 A 出现的条件概率。

定理 5.1(加法定理)　两个不相容(互斥)事件之和的概率等于两个事件概率之和,即

$$P(A+B)=P(A)+P(B)$$

两个互逆事件 A 和 A^{-1} 的概率之和为 1。即当 $A+A^{-1}=\Omega$,且 A 与 A^{-1} 互斥,则 $P(A)+P(A^{-1})=1$,或 $P(A)=1-P(A^{-1})$。

若 A、B 为两个任意事件,则 $P(A+B)=P(A)+P(B)-P(AB)$成立。此定理可推广到 3 个以上事件的情形:

$$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(BC)-P(CA)+P(ABC)$$

定理 5.2(乘法定理)　设 A、B 为两个不相容的非零事件,则其乘积的概率等于 A 和 B 概率的乘积,即

$$P(AB)=P(A)P(B)\quad 或\quad P(AB)=P(B)P(A)$$

设 A、B 为两个任意的非零事件,则其乘积的概率等于 A(或 B)的概率与在 A(或 B)

出现的条件下 B(或 A)出现的条件概率的乘积。

$$P(AB)=P(A)P(B\mid A)\quad 或\quad P(AB)=P(B)P(A\mid B)$$

此定理可以推广到3个以上事件的乘积情形,即当 z 个事件的乘积 $P(A_1A_2\cdots A_{z-1})>0$ 时,则乘积的概率为

$$P(A_1A_2\cdots A_z)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1A_2)\cdots P(A_z\mid A_1A_2\cdots A_{z-1})$$

当事件独立时,则有

$$P(A_1A_2\cdots A_z)=P(A_1)P(A_2)P(A_3)\cdots P(A_z)$$

下面给出还需要了解的其他概率知识。

(1) 先验概率。指根据历史资料或主观判断所确定的各事件发生的概率,该类概率没能经过实验证实,属于检验前的概率,所以称为先验概率。先验概率一般分为两类,一是客观先验概率,是指利用历史资料计算得到的概率;二是主观先验概率,是指在无历史资料或历史资料不全的时候,只能凭借人们的主观经验判断取得的概率。

(2) 后验概率。一般是指利用贝叶斯公式,结合调查等方式获取了新的附加信息,对先验概率进行修正后得到的更符合实际的概率。

(3) 全概率公式。如果影响事件 A 的所有因素 $B_1,B_2,\cdots,B_n$ 满足:$B_iB_j=\varnothing$,$(i\neq j,i=1,2,\cdots,n,j=1,2,\cdots,n)$,且 $\sum_{i=1}^{t}P(B_i)=1$,则必有

$$P(A)=\sum_{i=1}^{t}P(B_i)P(A\mid B_i)$$

(4) 贝叶斯公式。也称后验概率公式或逆概率公式。设先验概率为 $P(B_i)$,调查所获的新附加信息为 $P(A_j|B_i)$,其中 $i=1,2,\cdots,z,j=1,2,\cdots,z$。则贝叶斯公式计算的后验概率为

$$P(B_i\mid A_j)=\frac{P(B_i)P(A_j\mid B_i)}{\sum_{t=1}^{z}P(B_t)P(A_j\mid B_t)}$$

5.3 主观贝叶斯方法

5.3.1 不确定性的表示

下面介绍在主观贝叶斯方法中如何进行知识和证据的不确定性表示。

1. 知识的不确定性表示

在主观贝叶斯方法中,采用产生式表示知识,其规则为

$$\text{IF } E \text{ THEN } (\mathrm{LS},\mathrm{LN})\,H$$

其中,E 是知识的前提条件,可以是简单条件,也可以是复合条件;H 是结论;(LS,LN)表示知识的静态强度;LS表示充分性度量,即表示 E 对 H 的支持度量;LN表示必要性度量,即表示 $\neg E$ 对 H 的支持度量。它们的表现形式分别为

$$\mathrm{LS}=\frac{P(E\mid H)}{P(E\mid \neg H)}$$

$$\mathrm{LN}=\frac{P(\neg E \mid H)}{P(\neg E \mid \neg H)}=\frac{1-P(E \mid H)}{1-P(E \mid \neg H)}$$

LS 和 LN 的取值范围均为$[0,+\infty)$。

2. 证据的不确定性表示

在主观贝叶斯方法中，证据的不确定性是用概率或几率表示的，二者之间的关系为

$$O(E)=\frac{P(E)}{1-P(E)}=\begin{cases}0, & E\text{ 为假}\\ +\infty, & E\text{ 为真}\\ (0,+\infty), & \text{其他}\end{cases}$$

5.3.2 组合证据不确定性的计算

组合证据包括合取和析取两种基本情形。

当组合证据是多个单一证据的合取时，即

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

如果已知

$$P(E_1 \mid S),P(E_2 \mid S),\cdots,P(E_n \mid S)$$

则

$$P(E \mid S)=\min\{P(E_1 \mid S),P(E_2 \mid S),\cdots,P(E_n \mid S)\}$$

当组合证据是多个单一证据的析取时，即

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

如果已知

$$P(E_1 \mid S),P(E_2 \mid S),\cdots,P(E_n \mid S)$$

则

$$P(E \mid S)=\max\{P(E_1 \mid S),P(E_2 \mid S),\cdots,P(E_n \mid S)\}$$

5.3.3 不确定性的传递算法

主观贝叶斯方法推理的任务就是根据 E 的概率 $P(E)$及 LS 和 LN 的值，把 H 的先验概率 $P(H)$或先验几率 $O(H)$更新为后验概率或后验几率。由于一条知识所对应的证据可能为真，也可能为假，还可能既非为真又非为假，因此，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。

1. 证据肯定为真

当证据 E 肯定为真时，有 $P(E)=P(E|S)=1$。将 H 的先验几率更新为后验几率的公式为

$$O(H \mid E)=\mathrm{LS}\cdot O(H)$$

如果把 H 的先验概率 $P(H)$更新为后验概率 $P(H|E)$，则可以得到几率和概率的对应关系：

$$P(H \mid E)=\frac{\mathrm{LS}\cdot P(H)}{(\mathrm{LS}-1)\cdot P(H)+1}$$

2. 证据肯定为假

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

$$O(H \mid \neg E)=\mathrm{LN}\times O(H)$$

如果把 H 的先验概率 $P(H)$ 更新为后验概率 $P(H|E)$，则可以得到几率和概率的对应关系：

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

3. 证据既非为真又非为假

当证据既非为真又非为假时，这时因为 H 依赖于证据 E，而 E 基于部分证据 S，则 H 依赖于 S 的似然性。根据下面公式来计算证据不确定性的传递问题：

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

这里不再讨论 $P(E|S)=1$ 和 $P(E|S)=0$ 两种情况，只讨论剩下的两种情况：

(1) $P(E|S)=P(E)$。此时 E 与 S 无关，根据全概率公式可得

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

(2) $P(E|S)$ 为其余情况。在此种情况下，依据 3 个特殊值 $(0, P(E), 1)$ 的分段线性插值函数求得。该分段线性插值函数 $P(H|S)$ 如图 5.1 所示，函数的解析表达式为

$$P(H \mid S)=\begin{cases}P(H \mid \neg E)+\dfrac{P(H)-P(H \mid \neg E)}{P(E)}\times P(E \mid S), & 0\leqslant P(E \mid S)<P(E)\\[2ex] P(H)+\dfrac{P(H \mid E)-P(H)}{1-P(E)}\times[P(E \mid S)-P(E)], & P(E)\leqslant P(E \mid S)\leqslant 1\end{cases}$$

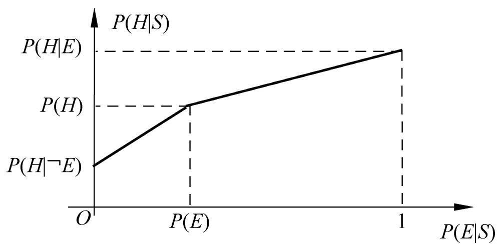

图 5.1 分段性插值函数

5.3.4 结论不确定性的合成

现有 n 条知识都支持同一结论 H，且每条知识的前提条件分别是 n 个相互独立的证据 $E_1, E_2, \cdots. E_n$，而这些证据所对应的观察分别为 $S_1, S_2, \cdots, S_n$。此时，可以对每条知识分别求出 H 的后验几率 $O(H|S_i)$，然后利用这些后验几率并按照下列公式可以求出所有观察下 H 的后验几率。

$$O(H \mid S_1, S_2, \cdots, S_n)=\frac{O(H \mid S_1)}{O(H)}\times\frac{O(H \mid S_2)}{O(H)}\times\cdots\times\frac{O(H \mid S_n)}{O(H)}\times O(H)$$

下面通过实例进一步说明主观贝叶斯方法的推理过程。

例 5.1　设有下列规则：

r_1：IF E_1 THEN (2,0.0001) H_1

r_2：IF E_1 AND E_2 THEN (100,0.0001) H_1

r_3：IF H_1 THEN (200,0.001) H_2

已知 $P(E_1)=P(E_2)=0.5$，$P(H_1)=0.092$，$P(H_2)=0.01$，用户提供的证据如下：$P(E_1|S_1)=0.76$，$P(E_2|S_2)=0.68$。求 $P(H_2|S_1,S_2)$。

解：由已知知识得到的推理网络如图 5.2 所示。

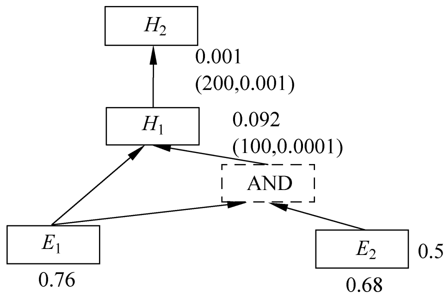

图 5.2　推理网络

(1) 计算 $O(H_1|S_1)$。

先把 H_1 的先验概率 $P(H_1)$ 更新为在 E_1 下的后验概率 $P(H_1|E_1)$：

$$P(H_1 \mid E_1)=\frac{\mathrm{LS}_1 \times P(H_1)}{(\mathrm{LS}_1-1)\times P(H_1)+1}$$
$$=\frac{2\times 0.092}{(2-1)\times 0.092+1}$$
$$=0.1685$$

由于 $P(E_1|S_1)=0.76>P(E_1)$，所以在当前观察 S_1 下 H_1 的后验概率 $P(H_1|S_1)$ 为

$$P(H_1 \mid S_1)=P(H_1)+\frac{P(H_1 \mid E_1)-P(H_1)}{1-P(E_1)}\times(P(E_1 \mid S_1)-P(E_1))$$
$$=0.092+\frac{0.1685-0.092}{1-0.5}\times(0.76-0.5)$$
$$=0.1318$$

$$O(H_1 \mid S_1)=\frac{P(H_1 \mid S_1)}{1-P(H_1 \mid S_1)}=\frac{0.1318}{1-0.1318}=0.1518$$

(2) 计算 $O(H_1|(S_1 \text{ AND } S_2))$。

由于 r_2 的前件是 E_1、E_2 合取关系，且已知 $P(E_1|S_1)=0.76$，$P(E_2|S_2)=0.68$，即 $P(E_2|S_2)<P(E_1|S_1)$。按合取取最小的原则，这里仅考虑 E_2 对 H_1 的影响，即把计算 $P(H_1|(S_1 \text{ AND } S_2))$ 的问题转化为计算 $O(H_1|S_2)$ 的问题。

把 H_1 的先验概率 $P(H_1)$ 更新为在 E_2 下的后验概率 $P(H_1|E_2)$：

$$P(H_1 \mid E_2)=\frac{\mathrm{LS}_2\times P(H_1)}{(\mathrm{LS}_2-1)\times P(H_1)+1}=\frac{100\times 0.092}{(100-1)\times 0.092+1}=0.9102$$

又由于 $P(E_2|S_2)>P(E_2)$，得到在当前观察 S_2 下 H_1 的后验概率 $P(H_1|S_2)$：

$$P(H_1 \mid S_2)=P(H_1)+\frac{P(H_1 \mid E_2)-P(H_1)}{1-P(E_2)}\times(P(E_2 \mid S_2)-P(E_2))$$
$$=0.092+\frac{0.9102-0.092}{1-0.5}\times(0.68-0.5)=0.3866$$

$$O(H_1 \mid S_2)=\frac{P(H_1 \mid S_2)}{1-P(H_1 \mid S_2)}=\frac{0.3866}{1-0.3866}=0.6306$$

(3) 计算 $O(H_1|S_1,S_2)$。

先将 H_1 的先验概率转换为先验几率：

$$O(H_1)=\frac{P(H_1)}{1-P(H_1)}=\frac{0.092}{1-0.092}=0.1013$$

然后根据合成公式计算 H_1 的后验几率：

$$O(H_1\mid S_1,S_2)=\frac{O(H_1\mid S_1)}{O(H_1)}\times\frac{O(H_1\mid S_2)}{O(H_1)}\times O(H_1)$$

$$=\frac{0.1518}{0.1013}\times\frac{0.6303}{0.1013}\times 0.1013=0.9445$$

再将后验几率转换为后验概率：

$$P(H_1\mid S_1,S_2)=\frac{O(H_1\mid S_1,S_2)}{1+O(H_1\mid S_1,S_2)}=\frac{0.9445}{1+0.9445}=0.4857$$

(4) 计算 $P(H_2|S_1,S_2)$。

对 r_3，H_1 相当于已知事实，H_2 为结论。将 H_2 的先验概率 $P(H_2)$ 更新为在 H_1 下的后验概率 $P(H_2|H_1)$：

$$P(H_2\mid H_1)=\frac{\mathrm{LS}_3\times P(H_2)}{(\mathrm{LS}_3-1)\times P(H_2)+1}=\frac{200\times 0.001}{(200-1)\times 0.001+1}=0.1668$$

由于 $P(H_1|S_1,S_2)=0.4857>P(H_1)$，得到在当前观察 S_1、S_2 下 H_2 的后验概率 $P(H_2|S_1,S_2)$：

$$P(H_2\mid S_1,S_2)=P(H_2)+\frac{P(H_2\mid H_1)-P(H_2)}{1-P(H_1)}\times(P(H_1\mid S_1,S_2)-P(H_1))$$

$$=0.01+\frac{0.1668-0.01}{1-0.092}\times(0.4857-0.092)=0.0780$$

从上例可以看出，H_2 先验概率是 0.01，通过运用知识 r_1、r_2、r_3 及初始证据的概率进行推理，最后推出的 H_2 的后验概率为 0.0780，相当于概率增加到 7 倍多。

主观贝叶斯方法的主要优点是理论模型精确，灵敏度高，不仅考虑了证据间的关系，而且考虑了证据存在与否对假设的影响，因此是一种较好的方法。其主要缺点是所需要的主观概率太多，专家不易给出。

5.4 可信度方法

可信度方法是肖特里菲(Shortliffe)等人于 1975 年在确定性理论上结合概率论等理论提出的一种不确定性推理模型。在专家系统等领域有广泛的应用。

本节主要介绍 CF 模型，它是基于可信度概念和产生式规则构建的不确定性推理模型。

5.4.1 不确定性的表示

下面介绍在 CF 模型中如何进行知识和证据的不确定性表示。

1. 知识的不确定性表示

在 CF 模型中，不确定性推理规则的一般形式为

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

其中,E 表示前提条件;H 表示知识的结论;$CF(H,E)$表示该规则的可信度,也称可信度因子或规则强度。可信度是指人们根据以往经验对某个事物或现象为真的程度的一个判断,其值为$[-1,1]$,$CF(H,E)>0$ 表示该证据对结论为真的支持度,CF 值越趋近 1,则该证据对结论为真的支持度就越大,$CF(H,E)=1$ 则表示该证据使结论成立为真;$CF(H,E)<0$ 表示该证据对结论为假的支持度,CF 值越趋近-1,则该证据对结论为假的支持度就越大,$CF(H,E)=-1$ 则表示该证据使结论成立为假;$CF(H,E)=0$ 则表示证据和结论没有关系。

2. 证据的不确定性表示

证据的不确定性也是用可信度 $CF(E)$表示的。$CF(E)$所描述的是证据的动态强度。区别于知识的静态强度 $CF(H,E)$表示的是规则的强度。证据的可信度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。

可信度取值范围是$[-1,1]$。其典型取值如下:

- 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,$CF(E)=0$。

5.4.2 组合证据不确定性的计算

组合证据包括合取和析取两种情形。

(1) 当组合证据是多个单一证据的合取时,即

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

若已知 $CF(E_1),CF(E_2),\cdots,CF(E_n)$,则

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

(2) 当组合证据是多个单一证据的析取时,即

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

若已知 $CF(E_1),CF(E_2),\cdots,CF(E_n)$,则

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

5.4.3 不确定性的传递算法

CF 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,运用相关的不确定性知识,逐步推出最终结论和该结论的可信度的过程。而每一次的不确定性推理都需要由证据的不确定性和知识的不确定性计算结论的不确定性,其计算公式为

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

从上式可以看出,当 $CF(E)=1$ 时,有 $CF(H)=CF(H,E)$,这表明,规则强度 $CF(E)$实际上就是在前提条件对应的证据为真时结论 H 的可信度。而当 $CF(E)<0$ 时,则相应的证据以某种程度为假,若 $CF(H)=0$,则表明在该模型中没有考虑证据为假时对结论 H 所产生的影响。

5.4.4 结论不确定性的合成

设有如下两条规则:

```
IF E1 THEN H (CF(H,E1))
IF E2 THEN H (CF(H,E2))
```

则结论 H 的综合可信度可按照如下步骤求得：

(1) 分别求出 $CF_1(H)$、$CF_2(H)$

$$CF_1(H)=CF(H,E_1)\times\max\{0,CF(E_1)\}$$
$$CF_2(H)=CF(H,E_2)\times\max\{0,CF(E_2)\}$$

(2) 求出 E_1 和 E_2 的综合可信度 CF(H)：

$$CF(H)=\begin{cases}CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H), & CF_1(H)\geqslant 0,CF_2(H)\geqslant 0\\ CF_1(H)+CF_2(H)+CF_1(H)\times CF_2(H), & CF_1(H)<0,CF_2(H)<0\\ 1-\min\{|CF_1(H)|,|CF_2(H)|\} & CF_1(H)\times CF_2(H)<0\end{cases}$$

例 5.2 设有如下一组知识：

r_1：IF E_1 THEN H (0.95)

r_2：IF E_2 THEN H (0.65)

r_3：IF E_3 THEN H (−0.51)

r_4：IF E_4 AND (E_5 OR E_6) THEN E_1(0.81)

已知：$CF(E_2)=0.81$，$CF(E_3)=0.65$，$CF(E_4)=0.51$，$CF(E_5)=0.65$，$CF(E_6)=0.81$，求 $CF(H)$。

解：由 r_4 得

$$\begin{aligned}CF(E_1)&=0.81\times\max\{0,CF(E_4\ \text{AND}\ (E_5\ \text{OR}\ E_6))\}\\&=0.81\times\max\{0,\min\{CF(E_4),CF(E_5\ \text{OR}\ E_6)\}\}\\&=0.81\times\max\{0,\min\{CF(E_4),\max\{CF(E_5),CF(E_6)\}\}\}\\&=0.81\times\max\{0,\min\{CF(E_4),\max\{0,65,0.81\}\}\}\\&=0.81\times\max\{0,\min\{0.51,0.81\}\}\\&=0.81\times\max\{0,0.51\}\\&=0.4131\end{aligned}$$

由 r_1 得

$$\begin{aligned}CF_1(H)&=CF(H,E_1)\times\max\{0,CF(E_1)\}\\&=0.95\times\max\{0,0.4131\}=0.392\,445\end{aligned}$$

由 r_2 得

$$\begin{aligned}CF_2(H)&=CF(H,E_2)\times\max\{0,CF(E_2)\}\\&=0.65\times\max\{0,0.81\}=0.5265\end{aligned}$$

由 r_3 得

$$\begin{aligned}CF_3(H)&=CF(H,E_3)\times\max\{0,CF(E_3)\}\\&=-0.51\times\max\{0,0.65\}=-0.3315\end{aligned}$$

根据结论不确定性的合成算法，得

$$\begin{aligned}CF_{1,2}(H)&=CF_1(H)+CF_2(H)-CF_1(H)\times CF_2(H)\\&=0.392\,445+0.5265-0.392\,445\times 0.5265=0.712\,323\end{aligned}$$

$$\mathrm{CF}_{1,2,3}(H)=\frac{\mathrm{CF}_{1,2}(H)+\mathrm{CF}_3(H)}{1-\min\{|\mathrm{CF}_{1,2}(H)|,|\mathrm{CF}_3(H)|\}}$$
$$=\frac{0.712\,323-0.3315}{1-\min\{0.712\,323,0.3315\}}=0.569\,668$$

这就是所求的综合可信度,即 $\mathrm{CF}(H)=0.569\,668$。

5.5 证据理论

证据理论是由 G. Shafer 拓展了 A. P. Dempster 的工作而来的,因此也称为 DS 理论,该理论实则是对简单概率的一种推广,称为广义概率。广义概率能够处理由"不知道"所引起的不确定性,并且由于辨别框的子集可以是多个元素的集合,因而知识的结论部分不必限制在由单个元素表示的最明显的层次上,而可以是一个更一般的不明确的假设,这样更有利于领域专家在不同细节、不同层次上进行知识表示。

5.5.1 理论基础

在证据理论中,常用的概念有概率分配函数、信任函数、似然函数以及类概率函数等。下面一一介绍这些概念。

1. 概率分配函数

定义 5.6 设函数 $m: 2^{\Omega}\to[0,1]$,且满足

$$m(\varnothing)=0$$
$$\sum_{A\subseteq\Omega}m(A)=1$$

则称 m 是 2^{Ω} 上的概率分配函数,$m(A)$称为 A 的基本概率数。其中,2^{Ω} 表示由 Ω 的所有子集构成的集合(下同)。

对概率分配函数有以下两点说明:

其一,概率分配函数的作用是把 Ω 的任意一个子集 A 都映射为$[0,1]$上的一个数 $m(A)$。当 A 由单元素组成时,$m(A)$表示对 A 的精确信任度;当 A 由多元素组成,且不为全集时,$m(A)$也表示对 A 的精确信任度,但不知道这部分信任度该分配给 A 中的哪些元素;当 A 是全集时,则 $m(A)$是对全集的各个子集进行信任分配后剩余的部分,表示不知道该如何对它进行分配。

其二,概率分配函数不是概率。

2. 信任函数

定义 5.7 信任函数(Belief Function,Bel 函数)如下:

$$\mathrm{Bel}: 2^{\Omega}\to[0,1]$$
$$\text{对 }\forall A\subseteq\Omega,\mathrm{Bel}(A)=\sum_{B\subseteq A}m(B)$$

Bel 函数又称为下限函数,表示当前环境下对假设集 A 的信任度,其值为 A 的所有子集的基本概率之和,表示对 A 的总信任度。

3. 似然函数

定义 5.8 似然函数(Plausibility Function,Pl 函数)如下:

$$\mathrm{Pl}: 2^{\Omega} \rightarrow [0,1]$$

$$\text{对 } \forall A \subseteq \Omega, \mathrm{Pl}(A) = 1 - \mathrm{Bel}(\neg A)$$

其中,$\neg A = \Omega - A$。

似然函数又称为上限函数或不可驳斥函数。由于 Bel(A)表示对 A 为真的信任度,Bel($\neg A$)表示对 $\neg A$ 的信任度,即对 A 为假的信任度,因此,Pl(A)表示对 A 为非假的信任度。

信任函数与似然函数有如下性质:

(1) $\mathrm{Bel}(\varnothing)=0$,$\mathrm{Bel}(\Omega)=1$,$\mathrm{Pl}(\varnothing)=0$,$\mathrm{Pl}(\Omega)=1$。

(2) IF $A \subseteq B$,THEN $\mathrm{Bel}(A) \leqslant \mathrm{Bel}(B)$,$\mathrm{Pl}(A) \leqslant \mathrm{Pl}(B)$。

(3) $\forall A \subseteq \Omega$,$\mathrm{Pl}(A) \geqslant \mathrm{Bel}(A)$。

(4) $\forall A \subseteq \Omega$,$\mathrm{Bel}(A)+\mathrm{Bel}(\neg A) \leqslant 1$,$\mathrm{Pl}(A)+\mathrm{Pl}(\neg A) \geqslant 1$。

另外,$\mathrm{Pl}(A)-\mathrm{Bel}(A)$表示既不信任 A 也不信任 $\neg A$ 的程度,即是真是假不知道的程度。

5.5.2 不确定性表示

下面介绍在证据理论下如何进行知识和证据的不确定性表示。

1. 知识的不确定性表示

在证据理论中,不确定性知识的表示形式为

$$\text{IF } E \text{ THEN } H\{h_1, h_2, \cdots, h_n\} \quad (\mathrm{CF}=\{c_1, c_2, \cdots, c_n\})$$

其中,E 为前提条件,它既可以是简单条件,也可以是用合取或析取词联结起来的复合条件;H 是结论,它用样本空间中的子集表示,$h_1, h_2, \cdots, h_n$ 是该子集中的元素;CF 是可信度因子,用集合形式表示,该集合中的元素 $c_1, c_2, \cdots, c_n$ 用来表示 $h_1, h_2, \cdots, h_n$ 的可信度,c_i 与 h_i 一一对应,并且 c_i 应满足如下条件:

$$\begin{cases} c_i \geqslant 0, \\ \sum_{i=1}^{n} c_i \leqslant 1, \end{cases} \quad i=1,2,\cdots,n$$

2. 证据的不确定性表示

证据 A 的不确定性用类概率函数 $f(A)$表示,类概率函数的定义可以由信任函数和似然函数得到,原始证据的 $f(A)$是由用户给出的,如果是推理过程中得到的中间结论,则其确定性是由推理得到的。

下面给出类概率函数的概念。

定义 5.9 设 Ω 是有限域,对 $\forall A \subseteq \Omega$,$A$ 的类概率函数为

$$f(A)=\mathrm{Bel}(A)+\frac{|A|}{|\Omega|} \cdot (\mathrm{Pl}(A)-\mathrm{Bel}(A))$$

其中，$|A|$和$|\Omega|$分别是 A 和 Ω 中元素的个数。

类概率函数 $f(A)$具有以下性质：

(1) $\sum_{i=1}^{n} f(\{s_i\})=1$。

(2) 对 $\forall A\subseteq\Omega$，有 $\mathrm{Bel}(A)\leqslant f(A)\leqslant \mathrm{Pl}(A)$。

(3) 对 $\forall A\subseteq\Omega$，有 $f(\neg A)=1-f(A)$。

5.5.3 组合证据不确定性的计算

组合证据包括合取和析取两种情形。当组合证据是多个单一证据的合取时，即

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

若已知 $f(E_1),f(E_2),\cdots,f(E_n)$，则

$$f(E)=\min\{f(E_1),f(E_2),\cdots,f(E_n)\}$$

当组合证据是多个单一证据的析取时，即

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

若已知 $f(E_1),f(E_2),\cdots,f(E_n)$，则

$$f(E)=\max\{f(E_1),f(E_2),\cdots,f(E_n)\}$$

当有多条规则支持同一结论时，如果 $A=\{a_1,a_2,\cdots,a_n\}$，则

$$\text{IF } E_i \text{ THEN } H \text{ CF}_i(\text{CF}_i\{c_{i1},c_{i2},\cdots,c_{in}\}),\quad i=1,2,\cdots,m$$

如果这些规则相互独立地支持结论 H 的成立，可以先计算

$$m_i(\{a_1\},\{a_2\},\cdots,\{a_n\})=(f(E_i)\cdot c_{i1},f(E_i)\cdot c_{i2},\cdots,f(E_i)\cdot c_{im}),\quad i=1,2,\cdots,m$$

然后可以根据求解正交和的方法，对 m_i 求正交和，以组合所有规则对结论 H 的支持。一旦组合的正交和 $m(H)$计算出来，就可以计算 $\mathrm{Bel}(H)$、$\mathrm{Pl}(H)$、$f(H)$。

下面介绍求正交和的方法。

定义 5.10　设 m_1 和 m_2 是两个不同的概率分配函数，则其正交和 $m=m_1\oplus m_2$ 满足

$$m(\varnothing)=0$$

$$m(A)=K^{-1}\cdot\sum_{x\cap y=\varnothing} m_1(x)m_2(y)$$

其中，$K=1-\sum_{x\cap y=\varnothing} m_1(x)m_2(y)=\sum_{x\cap y\neq\varnothing} m_1(x)m_2(y)$。

如果 $K\neq 0$，则正交和 m 也是一个概率分配函数；如果 $K=0$，则不存在正交和 m，称 m_1 和 m_2 矛盾。

5.5.4 不确定性的更新

设有知识 IF E THEN $H\{h_1,h_2,\cdots,h_n\}$(CF$=\{c_1,c_2,\cdots,c_n\}$)定义

$$m(\{a_i\})=f(E)\cdot c_i \quad i=1,2,\cdots,m$$

或表示为

$$m(\{a_1\},\{a_2\},\cdots,\{a_m\})=(f(E)\cdot c_1,f(E)\cdot c_2,\cdots,f(E)\cdot c_m)$$

规定

$$m(\Omega)=1-\sum_{i=1}^{m} m(\{a_i\})$$

而对于 Ω 的所有其他子集 H,均有 $m(H)=0$。

当 H 为 Ω 的真子集时,有

$$\mathrm{Bel}(H)=\sum_{B\subseteq H} m(B)=\sum_{i=1}^{m} m(\{a_i\})$$

进一步可以计算 $\mathrm{Pl}(H)$ 和 $f(H)$。

例 5.3 有如下规则:

r_1: IF E_1 AND E_2 THEN $A=\{a_1,a_2\}$,CF=$\{0.3,0.5\}$

r_2: IF E_3 AND (E_4 OR E_5)THEN $B=\{b_1\}$,CF=$\{0.7\}$

r_3: IF A THEN $H=\{h_1,h_2,h_3\}$,CF=$\{0.1,0.5,0.3\}$

r_4: IF B THEN $H=\{h_1,h_2,h_3\}$,CF=$\{0.4,0.2,0.1\}$

已知用户对初始证据给出的确定性为:$\mathrm{CER}(E_1)=0.8$,$\mathrm{CER}(E_2)=0.6$,$\mathrm{CER}(E_3)=0.9$,$\mathrm{CER}(E_4)=0.5$,$\mathrm{CER}(E_5)=0.7$,并假设 Ω 中的元素个数 $|\Omega|=10$,求 $\mathrm{CER}(H)$。

解:由给定知识形成的推理网络如图 5.3 所示。

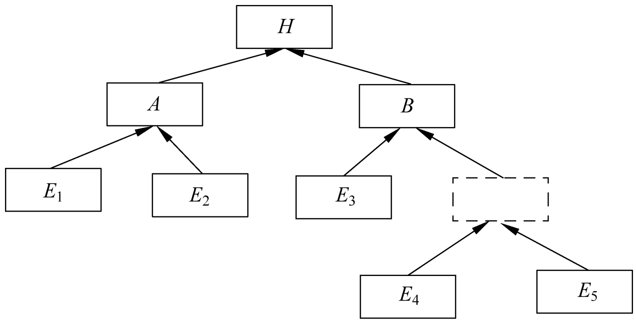

图 5.3 推理网络

(1) 求 $\mathrm{CER}(A)$。

因为

$$\mathrm{CER}(E_1 \text{ AND } E_2)=\min\{\mathrm{CER}(E_1),\ \mathrm{CER}(E_2)\}=\min\{0.8,0.6\}=0.6$$

$$m(\{a_1\},\{a_2\})=\{0.6\times 0.3,0.6\times 0.5\}=\{0.18,0.3\}$$

$$\mathrm{Bel}(A)=m(\{a_1\})+m(\{a_2\})=0.18+0.3=0.48$$

$$\mathrm{Pl}(A)=1-\mathrm{Bel}(\neg A)=1-0=1$$

$$\begin{aligned} f(A)&=\mathrm{Bel}(A)+|A|/|\Omega|\times(\mathrm{Pl}(A)-\mathrm{Bel}(A))\\ &=0.48+2/10\times(1-0.48)=0.584 \end{aligned}$$

所以,$\mathrm{CER}(A)=\mathrm{MD}(A/E')\times f(A)=0.584$。

(2) 求 $\mathrm{CER}(B)$。

因为

$$\begin{aligned} &\mathrm{CER}(E_3 \text{ AND } (E_4 \text{ OR } E_5))\\ &=\min\{\mathrm{CER}(E_3),\max\{\mathrm{CER}(E_4),\mathrm{CER}(E_5)\}\}\\ &=\min\{0.9,\max\{0.5,0.7\}\}=\min\{0.9,0.7\}=0.7\\ &m(\{b_1\})=0.7\times 0.7=0.49\\ &\mathrm{Bel}(B)=m(\{b_1\})=0.49 \end{aligned}$$

$$\mathrm{Pl}(B)=1-\mathrm{Bel}(\neg B)=1-0=1$$

$$f(B)=\mathrm{Bel}(B)+|B|/|\Omega|\times(\mathrm{Pl}(B)-\mathrm{Bel}(B))$$
$$=0.49+1/10\times(1-0.49)=0.541$$

所以，$\mathrm{CER}(B)=\mathrm{MD}(B/E')\times f(B)=0.541$。

(3) 求 $\mathrm{CER}(H)$。

由 r_3 可得：

$$m_1(\{h_1\},\{h_2\},\{h_3\})=\{\mathrm{CER}(A)\times 0.1,\mathrm{CER}(A)\times 0.5,\mathrm{CER}(A)\times 0.3\}$$
$$=\{0.584\times 0.1,0.584\times 0.5,0.584\times 0.3\}$$
$$=\{0.058,0.292,0.175\}$$

$$m_1(\Omega)=1-[m_1(\{h_1\})+m_1(\{h_2\})+m_1(\{h_3\})]$$
$$=1-[0.058+0.292+0.175]=0.475$$

再由 r_4 可得：

$$m_2(\{h_1\},\{h_2\},\{h_3\})=\{\mathrm{CER}(B)\times 0.4,\mathrm{CER}(B)\times 0.2,\mathrm{CER}(B)\times 0.1\}$$
$$=\{0.541\times 0.4,0.541\times 0.2,0.541\times 0.1\}$$
$$=\{0.216,0.108,0.054\}$$

$$m_2(\Omega)=1-[m_2(\{h_1\})+m_2(\{h_2\})+m_2(\{h_3\})]$$
$$=1-[0.216+0.108+0.054]=0.622$$

求正交和 $m=m_1\oplus m_2$。

$$K=m_1(\Omega)\times m_2(\Omega)$$
$$+m_1(\{h_1\})\times m_2(\{h_1\})+m_1(\{h_1\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_1\})$$
$$+m_1(\{h_2\})\times m_2(\{h_2\})+m_1(\{h_2\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_2\})$$
$$+m_1(\{h_3\})\times m_2(\{h_3\})+m_1(\{h_3\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_3\})$$
$$=0.475\times 0.622$$
$$+0.058\times 0.216+0.058\times 0.622+0.475\times 0.216$$
$$+0.292\times 0.108+0.292\times 0.622+0.475\times 0.108$$
$$+0.175\times 0.054+0.175\times 0.622+0.475\times 0.054$$
$$=0.855$$

$$m(h_1)=\frac{1}{K}\times(m_1(\{h_1\})\times m_2(\{h_1\})+m_1(\{h_1\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_1\}))$$
$$=\frac{1}{0.855}\times(0.058\times 0.216+0.058\times 0.622+0.475\times 0.216)$$
$$=0.178$$

$$m(h_2)=\frac{1}{K}\times(m_1(\{h_2\})\times m_2(\{h_2\})+m_1(\{h_2\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_2\}))$$
$$=\frac{1}{0.855}\times(0.292\times 0.108+0.292\times 0.622+0.475\times 0.108)$$
$$=0.309$$

$$m(h_3)=\frac{1}{K}\times(m_1(\{h_3\})\times m_2(\{h_3\})+m_1(\{h_3\})\times m_2(\Omega)+m_1(\Omega)\times m_2(\{h_3\}))$$
$$=\frac{1}{0.855}\times(0.175\times0.054+0.175\times0.622+0.475\times0.054)$$
$$=0.168$$
$$m(\Omega)=1-(m(\{h_1\})+m(\{h_2\})+m\{(h_3)\})$$
$$=1-(0.178+0.309+0.168)$$
$$=0.345$$

再根据 m 可得：

$$\mathrm{Bel}(H)=m(\{h_1\})+m(\{h_2\})+m(\{h_3\})=0.178+0.309+0.168=0.655$$
$$\mathrm{Pl}(H)=m(\Omega)+\mathrm{Bel}(H)=0.345+0.655=1$$
$$\mathrm{CER}(H)=\mathrm{MD}(H/E')\times f(H)=0.759$$

尽管该理论运用较为广泛，但由于要求 Ω 中的元素满足互斥条件，这在系统中不容易实现，并且概率分配数构成的维数空间很大，计算比较复杂。

5.6 模糊知识与模糊推理

5.6.1 模糊知识的表示

1. 语言变量

模糊逻辑使用的变量可以是语言变量，所谓语言变量是指用自然语言中的词表示并可以取语言值的变量。例如，变量“年龄”在普通集合中一般取 0～150 的数值，而在模糊集合中可以使用语言值“年轻、很年轻、不很年轻、老、很老、不很老”等。这些语言值可看成是论域 $U\in[0,150]$上模糊集的集合名。

通常，语言变量的值可以由一个或多个原始值再加上一组修饰词和连词组成。例如，上面给出的语言变量“年龄”，其原始词为“年轻”“老”，若加上修饰词“不很”可得到“不很年轻”“不很老”，若再加上连词“且”，则可得到“不很年轻且不很老”。

2. 模糊命题的描述

模糊逻辑是通过模糊谓词、模糊量词、模糊概率、模糊可能性、模糊真值、模糊修饰语等对命题的模糊性描述的。

1）模糊谓词

设 x 为在 U 中取值的变量，F 为模糊谓词，即 U 中的一个模糊关系，则命题可表示为

$$x \text{ is } F$$

其中的模糊谓词可以是“大、小、年轻、老、冷、暖、长、短”等。

2）模糊量词

模糊逻辑中使用了大量的模糊量词，如“极少、很少、几个、少数、很少、多数、大多数、几乎所有”等。这些模糊量词可以很方便地描述类似下面的命题：

大多数成绩好的学生学习都很刻苦。

很少有成绩好的学生特别贪玩。

3）模糊概率、模糊可能性和模糊真值

设 λ 为模糊概率，π 为模糊可能性，τ 为模糊真值，则对命题还可以附加概率限定、可能性限定和真值限定：

$$(x \text{ is } F) \text{ is } \lambda$$
$$(x \text{ is } F) \text{ is } \pi$$
$$(x \text{ is } F) \text{ is } \tau$$

式中，λ 可以是“或许、必须”等，π 可以是“非常可能、很不可能”等，τ 可以是“非常真、有些假”等。

4）模糊修饰语

模糊修饰语有以下 4 种：

(1) 求补。表示否定，如“不、非”等。

(2) 集中。表示“很、非常”等，其效果是减少隶属函数的值。

(3) 扩张。表示“有些、稍微”等，其效果是增加 0.5 以上隶属函数的值。

(4) 加强对比。表示“明确、确定”等，其效果是增加 0.5 以上隶属函数的值，减少 0.5 以下隶属函数的值。

3. 模糊知识的表示方法

在扎德的推理模型中，产生式规则的表示形式是

$$\text{IF } x \text{ is } F \text{ THEN } y \text{ is } G$$

其中，x 和 y 是变量，表示对象；F 和 G 分别是论域 U 及 V 上的模糊集，表示概念。条件部分(IF)可以是多个 x_i is F_i 的逻辑组合，此时，诸隶属函数间的运算按照模糊集的运算进行。

模糊推理中所用的证据都是用模糊命题表示的，其一般形式为

$$X \text{ is } F'$$

其中，F'是论域 U 上的模糊集。

5.6.2 模糊概念的匹配

模糊概念的匹配是指对两个模糊概念相似程度的比较与判断，而两个模糊概念的相似程度又称为匹配度。本节主要讨论语义距离和贴近度这两种计算匹配度的方法。

1. 语义距离

语义距离刻画的实际上是两个模糊概念之间的差异，常用的计算语义距离的方法有多种，这里主要介绍汉明(Hamming)距离。

设 $U=\{u_1,u_2,\cdots,u_n\}$ 是一个离散有限论域，F 和 G 分别是论域 U 上的两个模糊概念的模糊集，则 F 和 G 的汉明距离定义为

$$d(F,G)=\frac{1}{n}\sum_{i-1}^{n} |\mu_F(u_i)-\mu_G(u_i)|$$

例 5.4 设论域 $U=\{-10,0,10,20,30\}$ 表示温度，模糊集

$$F=0.8/-10+0.5/0+0.1/10$$

$$G = 0.9/-10 + 0.6/0 + 0.2/10$$

分别表示“冷”和“比较冷”,则

$$\begin{aligned} d(F,G) &= 0.2 \times (|0.8-0.9| + |0.5-0.6| + |0.1-0.2|) \\ &= 0.2 \times 0.3 = 0.06 \end{aligned}$$

即 F 和 G 的汉明距离为 0.06。

对求出的汉明距离,可通过下式

$$1 - d(F,G)$$

将其转换为匹配度。当匹配度大于某个事先给定的阈值时,认为两个模糊概念是相匹配的。当然,也可以直接用语义距离判断两个模糊概念是否匹配。这时,需要检查语义距离是否小于某个给定的阈值,距离越小,两者越相似。

2. 贴近度

贴近度是指两个概念的贴近程度,可直接用来作为匹配度。

设 F 和 G 分别是论域

$$U = \{u_1, u_2, \cdots, u_n\}$$

上的两个模糊概念的模糊集,则它们的贴近度定义为

$$(F,G) = \frac{1}{2}[F \cdot G + (1 - F \odot G)]$$

式中,$F \cdot G$ 为 F 与 G 的内积;$F \odot G$ 为 F 与 G 的外积。

5.6.3 模糊推理

模糊推理是按照给定的推理模式通过模糊集的合成来实现的。而模糊集的合成实际上又是通过模糊集与模糊关系的合成实现的。可见,模糊关系在模糊推理中占有重要的位置。为此,在讨论模糊推理方法之前,先对模糊关系的构造问题进行简单的介绍。

1. 模糊关系的构造

前面曾经介绍过模糊关系的概念,这里主要讨论由模糊集构造模糊关系的方法。目前已有多种构造模糊集关系的方法,下面仅介绍其中最常用的几种。

1) 模糊关系 R_m

模糊关系 R_m 是由扎德提出的一种构造模糊关系的方法。设 F 和 G 分别是论域 U 和 V 上的两个模糊集,则 R_m 定义为

$$R_m = \int_{U \times V} (\mu_F(u) \wedge \mu_G(v)) \vee (1 - \mu_F(u)) / (u, v)$$

其中,× 号表示模糊集的笛卡儿乘积。

例 5.5 设 $U = V = \{1,2,3\}$,F 和 G 分别是 U 和 V 上的两个模糊集,并设

$$F = 1/1 + 0.6/2 + 0.1/3$$

$$G = 0.1/1 + 0.6/2 + 1/3$$

则 R_m 为

$$R_m = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.4 & 0.6 & 0.6 \\ 0.9 & 0.9 & 0.9 \end{bmatrix}$$

下面以 $R_m(2,3)$为例说明 R_m 中元素的求法。

$$\begin{aligned} R_m(2,3) &= (\mu_F(u_2) \wedge \mu_G(v_3)) \vee (1-\mu_F(u_2)) \\ &= (0.6 \wedge 1) \vee (1-0.6) = 0.6 \vee 0.4 = 0.6 \end{aligned}$$

2）模糊关系 R_c

模糊关系 R_c 是由麦姆德尼(Mamdani)提出的一种构造模糊关系的方法。设 F 和 G 分别为论域 U 和 V 上的两个模糊集，则 R_c 定义为

$$R_c = \int_{U\times V} (\mu_F(u) \wedge \mu_G(v))/(u,v)$$

对例 5.5 所给出的模糊集，其 R_c 为

$$R_c = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 0.6 & 0.6 \\ 0.1 & 0.1 & 0.1 \end{bmatrix}$$

下面以 $R_c(3,2)$为例说明 R_c 中元素的求法：

$$R_c(3,2) = \mu_F(u_3) \wedge \mu_G(v_2) = 0.1 \wedge 0.6 = 0.1$$

3）模糊关系 R_g

模糊关系 R_g 是米祖莫托(Mizumoto)提出的一种构造模糊关系的方法。设 F 和 G 分别是论域 U 和 V 上的两个模糊集，定义为

$$R_g = \int_{U\times V} (\mu_F(u) \rightarrow \mu_G(v))/(u,v)$$

对例 5.5 给出的模糊集，其 R_g 为

$$R_g = \begin{bmatrix} 0.1 & 0.6 & 1 \\ 0.1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

2. 模糊推理的基本方法

与自然演绎推理相对应，模糊推理也有相应的 3 种基本模式，即模糊假言推理、拒取式推理及模糊三段论推理。

1）模糊假言推理

设 F 和 G 分别是 U 和 V 上的两个模糊集，且有知识

IF x is F THEN y is G

若有 U 上的一个模糊集 F'，且 F 可以和 F' 匹配，则可以推出 y is G'，且 G' 是 V 上的一个模糊集。这种推理模式称为模糊假言推理，其表示形式为

知识：IF x is F THEN y is G

证据：x is F'

结论：y is G'

在这种推理模式下，模糊知识

$$\text{IF } x \text{ is } F \text{ THEN } y \text{ is } G$$

表示在 F 与 G 之间存在确定的模糊关系，设此模糊关系为 R。那么，当已知的模糊事实 F' 可以与 F 匹配时，则可通过 F'与 R 的合成得到 G'。

例 5.6 对例 5.5 所给出的 F、G 以及所求出的 R_m，设有已知事实

$$x \text{ is 较小}$$

并设"较小"的模糊集为

$$\text{较小}=1/1+0.7/2+0.2/3$$

求在此已知事实下的模糊结论。

解：本例的模糊关系在前文中已求出，设已知模糊事实"较小"为 F'，F'与 R_m 的合成即为所求 G'。

$$G'=F'\circ R_m=\{1,0.7,0.2\}\circ\begin{bmatrix}0.1 & 0.6 & 1\\0.4 & 0.6 & 0.6\\0.9 & 0.9 & 0.9\end{bmatrix}=\{0.4,0.6,1\}$$

即所求出的模糊结论 G'为

$$G'=0.4/1+0.6/2+1/3$$

2）模糊拒取式推理

设 F 和 G 分别是 U 和 V 上的两个模糊集，且有知识

$$\text{IF } x \text{ is } F \text{ THEN } y \text{ is } G$$

若有 V 上的一个模糊集 G'，且 G'可以与 G 的补集匹配，则可以推出 x *is* F'，且 F'是 U 上的一个模糊集。这种推理模式称为模糊拒取式推理。它可以表示为

知识：IF x is F THEN y is G

证据：y is G'

结论：x is F'

在这种推理模式下，模糊知识

$$\text{IF } x \text{ is } F \text{ THEN } y \text{ is } G$$

也表示在 F 与 G 之间存在确定的模糊关系，设此模糊关系为 R。那么，当已知的模糊事实 G'可以与 $\neg G$ 匹配时，则可以通过 R 与 G'的合成得到 F'，即

$$F'=R\circ G'$$

式中的模糊关系 R 可以是 R_m、R_c 或 R_g 中的任何一种。

例 5.7 设 F、G 如例 5.5 所示，已知事实为

$$y \text{ is 较大}$$

且模糊概念"较大"的模糊集 G'为

$$G'=0.2/1+0.7/2+1/3$$

若 G'与 $\neg G$ 匹配，以模糊关系 R_c 为例，推出 F'。

解：本例中的模糊关系 R_c 已在前面求出，通过 R_c 与 G'的合成即可得到所求的 F'。

$$F'=R_c\circ G'=\begin{bmatrix}0.1 & 0.6 & 1\\0.1 & 0.6 & 0.6\\0.1 & 0.1 & 0.1\end{bmatrix}\circ\begin{bmatrix}0.2\\0.7\\1\end{bmatrix}=\begin{bmatrix}1\\0.6\\0.1\end{bmatrix}$$

即所求出的 F' 为

$$F'=1/1+0.6/2+0.1/3$$

模糊拒取式推理和模糊假言推理类似，也可以把计算 R_m、R_c 的公式代入求 F' 的公式中，得到求 F' 的一般公式。对 R_m 有

$$F'=R_m G'\int_{u\in U}\vee\{[(\mu_F(u)\wedge\mu_G(v))\vee(1-\mu_F(u))]\wedge\mu_G(v)\}/u$$

同理，对模糊关系 R_c 也可推出求 F' 的一般公式

$$F'=R_c G=\int_{u\in U}\vee\{[\mu_F(u)\wedge\mu_G(u)]\wedge\mu_G(v)\}/v$$

在实际应用中，也可以直接利用这些公式由 F、G 和 G' 求出 F'。

3）模糊三段论推理

设 F、G、H 分别是 U、V、W 上的 3 个模糊集，且由知识

IF x is F THEN y is G

IF y is G THEN z is H

则可推出

IF x is F THEN z is H

这种推理模式称为模糊假言三段论，可以表示为

知识：IF x is F THEN y is G

证据：IF y is G THEN z is H

结论：IF x is F THEN z is H

在这种推理模式下，模糊知识

r_1：IF x is F THEN y is G

表示在 F 和 G 之间存在确定的模糊关系，设此模糊关系为 R_1。模糊知识

r_2：IF y is G THEN z is H

表示在 G 和 H 之间存在确定的模糊关系，设此模糊关系为 R_2。

若模糊假言三段论成立，则 r_3 的模糊关系 R_3 可由 R_1 与 R_2 的合成得到，即

$$R_3=R_1\circ R_2$$

这里的关系 R_1、R_2、R_3 可以是前面所讨论过的 R_m、R_c 与 R_g 中的任何一种。为说明这种方法，下面讨论一个例子。

例 5.8　设 $U=W=V=\{1,2,3\}$。

$$E=1/1+0.6/2+0.2/3$$

$$F=0.8/1+0.5/2+0.1/3$$

$$G=0.2/1+0.6/2+1/3$$

按 R_g 求 $E\times F\times G$ 上的关系 R。

解：先求 $E\times F$ 上的关系 R_{g1}：

$$R_{g1}=\begin{bmatrix}0.8 & 0.5 & 0.1\\ 1 & 0.5 & 0.1\\ 1 & 1 & 0.1\end{bmatrix}$$

再求 $F\times G$ 上的关系 R_{g2}：

$$R_{g2}=\begin{bmatrix}0.2 & 0.6 & 1\\0.2 & 1 & 1\\1 & 1 & 1\end{bmatrix}$$

最后求 $E\times F\times G$ 上的关系 R：

$$R=R_{g1}\circ R_{g2}=\begin{bmatrix}0.2 & 0.6 & 0.8\\0.2 & 0.6 & 1\\0.2 & 1 & 1\end{bmatrix}$$

在线视频

5.7 实践：基于 T-S 模型的模糊推理

模糊推理是一种基于行为的仿生推理方法，主要用来解决带有模糊现象的复杂推理问题。由于模糊现象的普遍存在，模糊推理系统被广泛的应用。模糊推理系统主要由模糊化、模糊规则库、模糊推理方法以及去模糊化组成，其基本流程如图 5.4 所示。

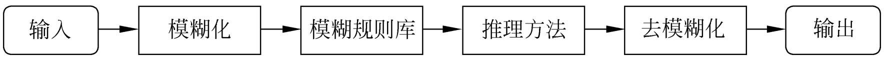

图 5.4　模糊推理流程图

传统的模糊推理是一种基于规则的控制，它通过语言表达的模糊性控制规则实现对难以精确描述系统的控制，在设计中不需要建立被控对象的精确数学模型。T-S 模糊推理模型是将正常的模糊推理规则及其推理转换成一种数学表达形式。T-S 模型本质上是将全局非线性系统通过模糊划分建立多个简单的线性关系，对多个模型的输出再进行模糊推理和判决，可以表示复杂的非线性关系。

5.7.1 T-S 模型的模糊推理过程

T-S 模糊模型基本思想是用线性状态空间模型作为后件表达每条语句对应所表证的局部动态特征，则全局的模糊模型就由这些线性模型通过隶属度函数综合而成，全局模型是一个非线性模型，利用模糊逻辑系统的非线性映射能力，就可以逼近一个复杂的非线性系统，而且能够对定义在一个致密集上的非线性系统做到任意精度上的一致逼近。

(1) 多输入多规则模糊推理系统的工作原理：

① 通过模糊化模块将输入的精确量进行模糊化处理，转换成给定论域上的模糊集合。

② 激活规则库中对应的模糊规则。

③ 选用合适的模糊推理方法，根据模糊事实推理出结果。

④ 对模糊结果进行去模糊化处理。

(2) 模糊化的原则及方法：

原则 1：在精确值处模糊集合的隶属度最大。

原则 2：当输入有干扰时，模糊化的结果具有一定的抗干扰能力。

原则 3：模糊化运算应尽可能简单。

各约束的隶属函数为(本实验采用)

```
Fun(a,b,x):  if x < a return 0;
         if a < x <= b return ((x - a)/(b - a))^2;
         if x > b return 1;
```

当隶属度函数为 f(x)时,模糊化隶属度函数准则(本实验采用):

高:f(x)^(0.5)

中:min(f(x),1−f(x))

低:(1−f(x))^0.5

(3) 模糊规则库(本实验采用):

IF 科研经费低 AND 人数高 AND 作品数低 AND 获奖数低 THEN 评价差

IF 科研经费高 AND 人数低 AND 作品数低 AND 获奖数高 THEN 评价高

IF 科研经费中 AND 人数中 AND 作品数中 AND 获奖数中 THEN 评价中

IF 科研经费高 AND 人数高 AND 作品数低 AND 获奖数低 THEN 评价差

对于评价的等级高中差分别用 3、2、1 表示。

约束等级划分标准:

约束与等级	高	中	低
科研经费	>20	5～20	<5
人数	>10	5～10	<5
作品数	>30	10～30	<10
获奖数	>15	5～15	<5

(4) 去模糊化的原则与方法:

原则 1:所得到的精确值,能够直观地表达该模糊集合。

原则 2:去模糊化运算要足够简单,保证模糊推理系统实时使用。

原则 3:模糊集合的微小变化不会使精确值发生大幅变化。

采用最小法和乘积法进行去模糊化处理(本实验采用)。

5.7.2　T-S 模型的模糊推理实验

具体实现及主要代码如下(详细代码参见附录),实验结果如图 5.5 所示。

1. 科研经费隶属度函数

```
def fun1(m):                          #科研经费隶属度函数
    if m <= 5:
        return 0
    if m > 5 and m <= 20:
        return ((m - 5)/15) * ((m - 5)/15)
    if m > 20:
        return 1
```

2. 规则 1

```
def rule1(self):
W[0] = math.sqrt(1 - T_S.fun1(jf))
W[1] = math.sqrt(T_S.fun2(rs))
W[2] = math.sqrt(1 - T_S.fun3(zp))
W[3] = math.sqrt(1 - T_S.fun4(hj))
pj[0] = 1
for i in range(4):
    if(W[i]< 0.0000000001):
        W[i] = 0
minTemp = 999 # 取小法
fori in range(4):
    if(W[i]! = 999):
        minTemp = min(minTemp,W[i])
MIN[0] = minTemp
mulTemp = 1 # 乘积法
fori in range(4):
    if(W[i]! = 999):
        mulTemp = mulTemp * W[i]
MUL[0] = mulTemp
```

3. 实现结果

经费：22
人数：3
作品：5
获奖：20
取小法评价：2.9999999969999998
乘积法评价：2.9999999969999998
评价高

经费：15
人数：8
作品：20
获奖：10
取小法评价：1.6410256401840893
乘积法评价：1.538460717703204
评价中

经费：22
人数：20
作品：5
获奖：3
取小法评价：0.9999999989999999
乘积法评价：0.9999999989999999
评价差

图 5.5 实验结果

实践示例程序参考附录 D。

5.7.4 思考与练习

现给出农业生产中评价经济指标的模糊推理规则：

IF 亩产低 AND 费用高 AND 用工高 AND 收入低 AND 肥力低 THEN 评价差
IF 亩产中 AND 费用低 AND 用工低 AND 收入中 AND 肥力高 THEN 评价中
IF 亩产高 AND 费用低 AND 用工低 AND 收入高 AND 肥力高 THEN 评价高
IF 亩产中 AND 费用中 AND 用工中 AND 收入中 AND 肥力高 THEN 评价中
IF 亩产高 AND 费用高 AND 用工高 AND 收入低 AND 肥力低 THEN 评价差

请自行设计评价等级区间数据以及不同隶属度函数，例如：采用上梯形隶属函数。去

模糊化时也请自行选择去模糊化方法，例如：中心平均法、最大隶属度法等。

5.8 习题

1. 什么是不确定性推理？有哪几类不确定性推理方法？不确定性推理中需要解决的基本问题有哪些？

2. 什么是可信度？

3. 设有如下一组推理规则：

r_1：IF E_1 THEN E_2(0.6)

r_2：IF E_2 AND E_3 THEN E_4(0.7)

r_3：IF E_4 THEN H (0.8)

r_4：IF E_5 THEN H (0.9)

且已知 $CF(E_1)=0.5$，$CF(E_3)=0.6$，$CF(E_5)=0.7$。求 $CF(H)$。

4. 设有如下推理规则：

r_1：IF E_1 THEN (2,0.00001) H_1

r_2：IF E_2 THEN (100,0.0001) H_1

r_3：IF E_3 THEN (200,0.001) H_2

r_4：IF H_1 THEN (50,0.1) H_2

且已知 $P(E_1)=P(E_2)=P(H_3)=0.6$，$P(H_1)=0.091$，$P(H_2)=0.01$，又由提供证据的用户告知 $P(E_1|S_1)=0.84$，$P(E_2|S_2)=0.68$，$P(E_3|S_3)=0.36$。

用主观贝叶斯方法求 $P(H_2|S_1,S_2,S_3)$。

5. 设有如下一组推理规则：

r_1：IF E_1 AND E_2 THEN $A=\{a\}$ (CF=\{0.9\})

r_2：IF E_2 AND (E_3 OR E_4) THEN $B=\{b_1,b_2\}$ (CF=\{0.5,0.4\})

r_3：IF A THEN $H=\{h_1,h_2,h_3\}$ (CF=\{0.2,0.3,0.4\})

r_4：IF B THEN $H=\{h_1,h_2,h_3\}$ (CF=\{0.3,0.2,0.1\})

且已知初始证据的确定性分别为：$CER(E_1)=0.6$，$CER(E_2)=0.7$，$CER(E_3)=0.8$，$CER(E_4)=0.9$。

假设 $|\Omega|=10$，求 $CER(H)$。

6. 简述模糊概念、模糊集和隶属函数三者之间的关系。

7. 设 $U=V=\{1,2,3,4\}$，且有如下推理规则：

$$\text{IF } x \text{ is 少 THEN } y \text{ is 多}$$

其中，“少”与“多”分别是 U 与 V 上的模糊集，设

$$少=0.9/1+0.7/2+0.4/3$$

$$多=0.3/2+0.7/3+0.9/4$$

已知事实为

$$x \text{ is 较少}$$

“较少”的模糊集为

$$较少=0.8/1+0.5/2+0.2/3$$

试用模糊关系 R_m 求出模糊结论。

8. 设有论域 $U=\{u_1,u_2,u_3,u_4,u_5\}$，并设 F、G 是 U 上的两个模糊集，且有

$$F=0.9/u_1+0.7/u_2+0.5/u_3+0.3/u_4$$

$$G=0.6/u_3+0.8/u_4+1/u_5$$

请分别计算 $F\cap G$、$F\cup G$、$\neg F$。

9. 设有如下两个模糊关系：

$$\boldsymbol{R}_1=\begin{bmatrix}0.3 & 0.7 & 0.2\\ 1 & 0 & 0.4\\ 0 & 0.5 & 1\end{bmatrix}\quad \boldsymbol{R}_2=\begin{bmatrix}0.2 & 0.8\\ 0.6 & 0.4\\ 0.9 & 0.1\end{bmatrix}$$

请写出 $\boldsymbol{R}_1$ 与 $\boldsymbol{R}_2$ 的合成 $\boldsymbol{R}_1\circ\boldsymbol{R}_2$。

10. 设 $\boldsymbol{F}$ 是论域 U 上的模糊集，$\boldsymbol{R}$ 是 $\boldsymbol{U}\times\boldsymbol{V}$ 上的模糊关系，$\boldsymbol{F}$ 和 $\boldsymbol{R}$ 分别为

$$\boldsymbol{F}=\{0.4,0.6,0.8\}$$

$$\boldsymbol{R}=\begin{bmatrix}0.1 & 0.3 & 0.5\\ 0.4 & 0.6 & 0.8\\ 0.6 & 0.3 & 0\end{bmatrix}$$

求模糊变换 $\boldsymbol{F}\circ\boldsymbol{R}$。