二级不确定性的内涵处理方法

刘大有　李岳峰

（吉林大学计算机系　长春　130023）

摘要　本文给出处理二级（或阶）不确定性的一种内涵方法（或称为基于模型的方法）. 该方法用概率簇空间表示证据空间的二级不确定性结构，用L-集合表示规则强度的二级不确定性，用概率簇空间的传递方式刻画二级不确定性的传播过程. 给出了概率簇上限定化关系的传递性质，并用限定化的概念对规则组合的两种不同方式进行了比较. 本文为基于知识的推理提供了新的综合框架.

关键词　二级不确定性，概率簇空间，限定化.

中图法分类号　TP18

规则的可靠性研究，二阶或二级不确定性的描述（表示）与处理，对于提高近似推理的准确性，改善专家系统的性能、评估ES所推出的结论的质量、处理多ES协作系统中的结论冲突都是很重要的.

1985年，Cheeseman 提出了关于用二阶概率描述不确定性的不确定性的建议.[1]1990～1993年，针对复杂的实际应用领域，刘大有教授提出了二级不确定性的概念，并提出了用外延方法分别处理证据信度偏差与规则强度偏差和可信度与可靠度的两个两级不确定性推理模型.[2]1991年，Kruse 等人[3]为了讨论不同信息源的组合问题，提出了二阶不确定性的概念（与二级不确定性概念等价）. 1993年，Zdrahal 使用了二阶不确定性的概念，以分析基于外延推理模型的敏感性问题.[4]

迄今为止，关于二阶或二级不确定性处理方法的研究，大都基于外延方法的. 用外延的方法处理不确定性，在计算方面虽然具有简单性，但计算结果的健壮性却较差. 因此在本章中，我们用内涵方法给出二阶不确定性信息的处理模型.

1995年 Guan 等人讨论了打折扣的 mass 函数，从二阶不确定性的意义上来说该研究是把 mass 函数的二阶不确定性直接作用在 mass 函数上. 与 Guan 的作法不同，我们总是把不确定性结构和它的二阶不确定性分开处理. 根据这一想法，在本文中，我们用内涵方法给出了表示和处理二级不确定性的模型：用概率簇空间表示证据空间上的二阶不确定性结构，

---

* 本文研究得到国家863高科技项目基金、国家自然科学基金、国家教委博士点基金、国家教委符号计算与知识工程开放实验室的资助. 作者刘大有，1942年生，教授，博士导师，主要研究领域为知识工程与专家系统，分布式人工智能，MAS. 李岳峰，1964年生，副教授，主要研究领域为知识工程与专家系统，多agent系统.

本文通讯联系人：刘大有，长春130023，吉林大学计算机系

本文1996-09-11收到修改稿

用L-集合表示规则强度的二阶不确定性,并提出了概率簇空间的传递模型.

1 集合-值映射的扩充

假定 $E=\{e_1,e_2,\ldots,e_n\}$ 是一个证据空间,Θ 是辨别框架.人们通常用集合-值映射 $\delta:E\to 2^{\Theta}$ 来描述知识的不确定性,从二阶不确定性的意义上来看,这种表示刻画了知识的一阶不确定性情形.实际中,集合-值映射不能充分刻画出知识的这种不精确性.

在文献[5]中,Yen 用概率多集合映射给出了集合-值映射的一种扩展方法,该映射把证据 e 对应为假设空间的一组互不相交的子集合,而规则强度表示为条件概率,这种映射用概率分布的方式,给出了不确定性知识的更详细描述.第 2 种有意义的扩展是 Kruse 等人[3]提出的映射 $\delta:E\to F_L(\Theta)$,他们把这种映射解释为模糊观察映射,该映射把知识看成是模糊的.

本节给出了集合值-映射的又一扩展方法,我们把证据 e 对应为假设空间的一组嵌套的子集合,用模糊集合(L-集合[3])表示规则强度.与 Yen 和 Kruse 等人解释方式不同,我们提出的映射描述了知识所具有的二阶不确定性.

设 L 是有限的全序集,它的最小元素和最大元素分别是 $l_{\min}$ 和 $l_{\max}$.对于集合 L,可用标度变换把集合 L 中的元素对应为实数.在如下定义的映射中,我们假定集合 L 中的元素为实数.定义映射 $\Gamma:E\to 2^{2^{\Theta}\times L}$

对任意的 $e\in E$ 令 $\Gamma(e)=\{(A_{e_1},s_e(A_{e_1}|e)),(A_{e_2},s_e(A_{e_2}|e)),\ldots,(A_{e_{n_e}},s_e(A_{e_{n_e}}|e))\}$,

其中 $A_{e_1},A_{e_2},\ldots,A_{e_{n_e}}\in 2^{\Theta}$,$s_e$ 是实数值函数,且满足如下条件:

(L1)$A_{e_i}\neq\varnothing$;$i=1,2,\ldots,n_e$

(L2)$s_e(A_{e_1}|e)=l_{\min}$ 且 $s_e(A_{e_i}|e)\in L,i=1,2,\ldots,n_e$

(L3)$A_{e_n}\subseteq\cdots\subseteq A_{e_2}\subseteq A_{e_1}$

(L4)$s_e(A_{e_1}|e)<s_e(A_{e_2}|e)<\cdots<s_e(A_{e_{n_e}}|e)$,其中"<"是 L 集合中元素的"大小"关系.

我们称 Γ 为 L 集合-值映射,并令 $\Gamma'(e)=\{A_{e_1},A_{e_2},\ldots,A_{e_{n_e}}\}$.对于 L 集合-值映射 Γ,当 $s_e(A|e)=l_{\min}$ 时,我们说集合 A 中有且仅有一种假设为真,而当 $s_e(A|e)>l_{\min}$ 时,此时 $s_e(A|e)$ 所表示的是集合 A 中有且仅有一种假设为真的可能性,$s_e(A|e)$ 越大则可能性越小.这里我们用模糊集合(L-集合)表示规则强度,其直观含义是说:当证据 e 为真时,集合 A_{e_1} 表示结论的不确定性(即 A_{e_1} 中有且仅有一种假设为真,但专家不知道是哪一个),该不确定性的二级不确定性用 L-集合来表示(即用层集合给出为真假设可能出现的更小区域).

根据上述规定,在一个专家系统中,第 i 条规则的一般形式如下:

RULE$_i$:

"IF EVIDENCE e_1 THEN

HYPOTHESIS $A_{e_1 1}$ WITH STRENGTH $s_{e_1}(A_{e_1 1}|e_1)$

HYPOTHESIS $A_{e_1 2}$ WITH STRENGTH $s_{e_1}(A_{e_1 2}|e_1)$

...

HYPOTHESIS $A_{e_1 n_{e_1}}$ WITH STRENGTH $s_{e_1}(A_{e_1 n_{e_1}}|e_1)$

...

ELSE IF EVIDENCE e_n THEN

HYPOTHESIS $A_{e_n 1}$ WITH STRENGTH $s_{e_n}(A_{e_n 1}|e_n)$

HYPOTHESIS A_{e_n2} WITH STRENGTH $s_{e_n}(A_{e_n2}|e_n)$

…

HYPOTHESIS $A_{e_n n_{en}}$ WITH STRENGTH $s_{e_n}(A_{e_n n_{en}}|e_n)$"

对于集合-值映射 $\delta: E\to 2^{\Theta}$,Dempster 等人已经讨论了这种不确定性知识的捕获方式.为了给出捕获多集合-值映射 Γ 所表示知识的合理方式,我们用可能性函数(或相应的 mass 函数)来表示上述规则中的模糊规则强度.

我们令
$$m_e(A_{e_{n_e}})=\frac{l_{\max}-s_e(A_{e_{n_e}}|e)}{l_{\max}-l_{\min}},$$
而当 $i=n_{e-1},\cdots,1$ 时,我们令
$$m_e(A_{e,i})=\frac{s_e(A_{e,i+1}|e)-s_e(A_{e,i}|e)}{l_{\max}-l_{\min}}$$

如果 A 不属于 $\{A_{e_1},A_{e_2},...,A_{e_{n_e}}\}$ 则令 $m_e(A)=0$. 可以验证 m_e 是 Θ 上的一个嵌套的 mass 函数,且 $Bel_{m_e}(A_{e_1})=1$. 称 m_e 是对应于证据 e 的规则强度 mass 函数.

定理 1. 假定 P 是证据空间 E 上的概率函数,对于规则 RULE$_i$,我们定义 Θ 上的函数 $\Gamma[P]:2^{\Theta}\to[0,1]$,满足
$$\Gamma[P](A)=\begin{cases}\Sigma_{j=1}^{n}m_{e_j}(A)\cdot p(\{e_j\}) & \text{如果 } A\subseteq\Theta, A\neq\varnothing;\\ 0 & \text{如果 } A=\varnothing\end{cases}$$
则 $\Gamma[P]$是 Θ 上的一个 mass 函数. 这里 m_{e_j}是对应于 e_j 的规则强度 mass 函数.

证明:显然 $\Gamma[P](\varnothing)=0$;当 $A\subseteq\Theta$ 且 $A\neq\varnothing$ 时,由 $\Gamma[P](A)=\Sigma_{j=1}^{n}m_{e_j}(A)\cdot p(\{e_j\})$ 知
$$\begin{aligned}\Sigma_{A:A\subseteq\Theta}\Gamma[P](A)&=\sum_{A\subseteq\Theta}\Sigma_{j=1}^{n}m_{e_j}(A)\cdot P(\{e_j\})\\&=\Sigma_{j=1}^{n}\Sigma_{A:A\subseteq\Theta}m_{e_j}(A)\cdot P(\{e_j\})\\&=\Sigma_{j=1}^{n}P(\{e_j\})\Sigma_{A:A\subseteq\Theta}m_{e_j}(A)\\&=\Sigma_{j=1}^{n}P(\{e_j\})=1\end{aligned}$$

所以 $\Gamma[P]$是 Θ 上的一个 mass 函数. □

2　概率簇空间的传递模型

实际应用中,专家往往用主观概率来表示证据空间的不确定性结构. 当多个专家共同提供信息时,使用概率簇来描述证据空间的不确定性结构是一种较合理的方法[3];另外,通常信息的不确定性很难用一个单一数值来刻画,因此即使只有一个专家,他也喜欢用一个区间值来度量信息的不确定性.[6]

本节假定有多个专家提供证据空间上的不确定性结构,即有一个概率簇. 因为实际中每个专家的权威是不一样的,所以往往存在该意义下的二阶不确定性信息,在这里我们也用二阶概率来表示这样的二阶不确定性信息.

2.1　不确定性结构的传递方法

假定 $E=\{e_1,e_2,...,e_n\}$是一个证据空间,Δ 是 E 上的一个概率簇(E 上概率函数组成的有限集合),Q 是 Δ 上的二阶概率(second order probability),称(Δ,Q)为 E 上的一个概率簇空间. 在概率簇上引入二阶概率 Q 有两方面的意义,①为了刻画领域专家所提供的不确定性信息的二级不确定性;②为了在结果概率簇上进行有效的决策. 设 $\Gamma: E\to 2^{2^{\Theta}\times L}$是多集

合值映射，则称(E,Δ,Q,Γ)为一个证据源.

设P_r是辨别框架Θ上的先验概率函数(实际中往往存在这样的先验概率)，下面的两个定理给出了不确定性结构从证据空间到假设空间的传递方法，同时还表明了，概率簇空间的传递结果仍然是一个概率簇空间.

定理 2. (Shafer，1976 年)假定P_r是一个概率函数，Bel是一个信任函数. 如果P_r和Bel是可组合的* (combinable)，则$Bel \oplus P_r$是一个概率函数(证明可见文献[7]).

定理 3. 假定(E,Δ,Q,Γ)为一个证据源，对任意的$P\in\Delta$，函数$\Gamma[P]:2^{\Theta}\to[0,1]$的定义与定理 3.1 中的相同. 如果令$Bel_P=Bel_{\Gamma[P]}$，$\Delta r=\{Bel_P\oplus P_r|P\in\Delta\}$，$Q_r:\Delta r\to[0,1]$，满足$Q_r(R)=Q(\{P\in\Delta|R=Bel_P\oplus P_r\})$. 则$(\Delta r,Q_r)$为$\Theta$上的一个概率簇空间.

证明：对任意的$P\in\Delta$，由定理 1 知$\Gamma[P]$是Θ上的一个 mass 函数. 因为Bel_P是$\Gamma[P]$的信任函数，从而由定理 2 知$Bel_P\oplus P_r$是Θ上的一个概率函数，即Δr是一个概率簇；可以验证Q_r是Δr上的二阶概率. □

以上我们用$Bel_P\oplus P_r$来表示假设空间上的后验概率，这种表示与 Shafer，Halpern 和 Fagin 等人的证据捕获(capture)方式是一致的.[8]对任意的$A\subseteq\Theta$，我们规定A的上、下确认度$D^*(A)$和$D_*(A)$定义为A的上、下包络$sup\{R(A)|R\in\Delta r\}$和$inf\{R(A)|R\in\Delta r\}$. 使用上、下包络的性质(见文献[3])，我们很容易证明下述定理.

定理 4. 对任意的$A\subseteq\Theta$，上、下确认度函数D有如下的性质：

(D1) $D^*(\varnothing)=D_*(\varnothing)=0, D^*(\Theta)=D_*(\Theta)=1$；

(D2) $0\leqslant D_*(A)\leqslant D^*(A)\leqslant 1$；

(D3) $D^*(A)=1-D_*(A)$；

(D4) 对任意$A,B\subseteq\Theta$，如果$A\cap B=\varnothing$，则$D_*(A)+D_*(B)\leqslant D_*(A\cup B)\leqslant D_*(A\cup B)\leqslant D_*(A)+D_*(B)$.

由确认度函数的上述性质，我们看到使用确认度区间表示假设的可靠程度是比较合理的. 可以证明基于集合-值映射的传递理论是我们所给模型的特例.[9]

2.2 限定化关系的传递

上节我们给出了概率簇空间的传递模型. 接下来我们证明，在这个传递过程中，原概率簇上的什么性质能够得以保持(本小节的证明见文献[9]).

定理 5. 如果$P_1,P_2\in\Delta$，且$P_1\prec P_2$(P_1是P_2的一个限定化，有关概率函数和信任函数的限定化的概念可见文献[3]或[10]). 则$Bel_{P_1}\oplus Pr\prec Bel_{P_2}\oplus Pr$.

从计算的角度而言，如果概率簇的基数越大，则传递模型的计算复杂性越高. 定理 5 一方面确保了按限定化关系选择用于传递计算的概率簇空间较小，另一方面又使得结果概率簇能够保持原有的限定化关系(即结果概率簇所承载的信息量在相对意义下保持不变).

定理 6. 假定P_0是E上的概率，δ是从E到$2^{\Theta}-\{\varnothing\}$的映射，$\Delta=\{P|P$是$E$上的概率且$P\prec P_0\}$，$\Delta_0=\{Bel_m\oplus Pr|m\leqslant\delta[P_0]$且$Bel_m$是$m$的信任函数$\}$. 则$\Delta r=\Delta_0$.

当Γ是多集合-值映射时，由定理 6 的证明过程可以看出，$\Delta r\subseteq\Delta_0$；并且还能证明(文献[9])，在一般情况下，$\Delta r\subset\Delta_0$.

---

* 这里$\oplus$是 Dempster 组合运算，为了方便计，本章总是假定 Dempster 组合运算$\oplus$是有定义的.

由上述定理，当概率簇是 P_0 的典型集合 $\{P \mid P \prec P_0\}$ 或 $\{P \mid$ 存在 $A \subseteq E$ 使得 $P = P_0(\cdot \mid A)\}$ 时(注意：集合 $\{P \mid$ 存在 $A \subseteq E$ 使得 $P = P_0(\cdot \mid A)\}$ 是集合 $\{P \mid P \prec P_0\}$ 的子集合)，传递模型中与 Pr 作 Dempster 组合运算的 mass 函数皆为 $\Gamma[P_0]$ 的一个严格限定化. 从证据的捕获角度来看，该结论进一步论证了概率簇空间传递模型的合理性.

3　规则的组合策略

上节我们讨论了概率簇空间的传播方式. 下面我们将讨论基于该模型的规则的组合方式. 一般来说，有 2 种组合方式. 第 1 种策略是组合证据源，然后再传递概率簇空间. 第 2 种策略是先传递概率簇空间，然后再组合.

3.1　组合证据源

为了组合证据源，必须给出概率族的组合方式和集合-值映射的组合方式. 设 $(E_1, \Delta_1, Q_1, \delta_1)$ 和 $(E_2, \Delta_2, Q_2, \delta_2)$ 是 2 个证据源，$P_1 \in \Delta_1$，$P_2 \in \Delta_2$. 为了确保不丢失信息，P_1 和 P_2 组合的结果是 $E_1 \times E_2$ 上满足边缘条件(marginal condition)的概率函数的集合. 通常，满足 $P_1(B) = P(\hat{\Pi}^{\{1,2\}}_{\{1\}}(B))$，$P_2(B) = P(\hat{\Pi}^{\{1,2\}}_{\{2\}}(B))$ 的 P 有很多(见文献[3]或[9]). 如果证据源是独立的，则 P_1 和 P_2 组合的结果仅有一个乘积概率 $(P_1 \otimes P_2)$，满足 $(P_1 \otimes P_2)(\{e', e''\}) = P_1(\{e'\}) \cdot P_2(\{e''\})$，其中 $e' \in E_1$，$e'' \in E_2$.

集合值-映射的组合方式有两种方法，一种是析取组合，另一种是合取组合. 我们可仿照这两种组合方式来组合 L 集合-值映射 Γ_1 和 Γ_2. 但是这种组合要复杂的多，这是因为不仅要给出 Γ_1' 和 Γ_2' 中每两个集合的并运算(或交运算)，而且还要组合每个证据所对应的规则强度 mass 函数. 因此当证据空间较大时，就要花费很多的时间来组合两个映射.

3.2　结论组合

与第 1 种策略不同，第 2 种策略是先传递、后组合. 在概率簇的传递模型中，我们采用第 2 种策略. 假定概率簇空间 (Δ_1, Q_1) 和 (Δ_2, Q_2) 产生的概率簇空间分别是 $(\Delta r_1, Qr_1)$ 和 $(\Delta r_2, Qr_2)$；D_1 和 D_2 是相应的确认度函数. 组合确认度函数 D 定义为 $D_*(A) = \min(D_{1*}(A), D_{2*}(A))$ 和 $D^*(A) = \max(D_1^*(A), D_2^*(A))$. 根据这个规定，对于组合确认度函数，定理 4 亦然成立；这是因为对任意的 $A \subseteq \Theta$，组合确认度函数 D 满足 $D^*(A) = sup\{R(A) \mid R \in \Delta r_{12}\}$ 和 $D^*(A) = inf\{R(A) \mid R \in \Delta r_{12}\}$，其中 $\Delta r_{12} = \Delta r_1 \cup \Delta r_2$.

我们采用第 2 种组合策略是基于如下的两点考虑. 一是计算方面的考虑. 如果 Δ_1 和 Δ_2 的元素个数分别为 m 和 n，则在最好情况下(假定证据源独立)，第 1 种策略所得到的组合概率簇的元素个数为 $m \cdot n$ 个，而一般来说映射的组合亦要花费很多的计算时间；而第 2 种策略所传递的概率个数最多为 $m+n$ 个，且不需要组合映射. 因此当概率簇较大时，我们的方法是比较快的. 第 2 种考虑，自然是要回答我们的方法是不是丢失了很多的信息量. 下面我们从捕获证据的角度来讨论这一问题.

定理 7. 如果 $(E_1, \Delta_1, Q_1, \delta_1)$ 和 $(E_2, \Delta_2, Q_2, \delta_2)$ 是两个独立的证据源，δ_1 和 δ_2 是集合-值映射 $P_1 \in \Delta_1$，$P_2 \in \Delta_2$. 则 $(\delta_1 \cup \delta_2)[P_1 \otimes P_2]$ 是 Θ 上的 mass 函数，且 $\delta_1[P_1]$ 和 $\delta_2[P_2]$ 皆为 $(\delta_1 \cup \delta_2)[P_1 \otimes P_2]$ 的限定化(证明见文献[9]).

可以验证当 Γ 是多集合-值映射时，上述定理仍然成立. 由上述定理可以看到，当证据

源独立时，我们的传递方法所捕获的信息量并不比证据源组合策略所捕获的信息量少. 如果 $Q_1, Q_2, \ldots, Q_k$ 分别是 $k(k>1)$ 个证据源上的二阶概率，q_i 是领域专家为第 i 个证据源所赋的一个权值(这里，我们要求 $q_i>0$ 且 $q_1+q_2+\ldots+q_k=1$，Kruse 等人把这些数值称为二阶不确定性)，$Qr_1, Qr_2, \ldots, Qr_k$ 分别是 $k(k>1)$ 个二阶概率的传递后得到的二阶概率. 则对任意的概率函数 $R\in\Delta_{12\ldots n}$，组合二阶概率 Qr 定义为 $Qr(R)=\Sigma_{j, R\in\Delta r_j} q_j \cdot Q_{rj}(R)$，其中 Δr_j 是第 j 个证据源传递后形成的概率簇，$\Delta_{12\ldots n}=\Delta r_1\cup\cdots\cup\Delta r_k$.

参考文献

1 Cheeseman P. In defense of probability. In: Proceedings of IJCAI, 1985. 1002～1009.

2 刘大有，钟绍春，高雅卿. 具有两级不确定性的推理模型. 软件学报，1993，**4**(3)：45～52.

3 Kruse R, Schwecke E, Heinsohn J. Uncertainty and vagueness in knowledge based systems (numerical methods). Springer-Verlag, New York, 1991.

4 Zdrahal Z. Second order measures for uncertainty processing. In: Proceedings of IJCAI, 1993. 626～631.

5 Yen J. GERTIS: a Dempster-Shafer approach to diagnosing hierarchical hypotheses. Communications of ACM, 1989, **32**(5): 573～585.

6 Kyburg H E Jr. Bayesian and non-Bayesian evidential updating. Artificial Intelligence, 1987, **31**: 271～293.

7 Guan J W, Bell D A. Evidence theory and its applications. Studies in Computer Science and Artificial Intelligence 7, Elsevier, North-Holland, 1991.

8 Halpern J Y, Fagin R. Two views of belief: belief as generalized probability and belief as evidence. Artificial Intelligence, 1992, **54**(3): 275～317.

9 李岳峰. 知识系统中处理不确定性信息的内涵方法研究[博士论文]. 吉林大学，1996.

10 刘大有，李岳峰，唐海鹰. 证据函数的限定化关系. 软件学报，1996，**7**(增刊)：216～223.

A MODEL-BASED METHOD FOR PROCESSING THE SECOND-LEVEL UNCERTAINTIES

LIU Dayou LI Yuefeng

(*Department of Computer Science Jilin University Changchun* 130023)

Abstract This paper provides a model-based method for processing the second-level uncertainties. In this paper, the authors use the space of probability class to present the second-level uncertainty structure of evidence space, L-set to present the second-level uncertainties of the rule strengths and use transmitting way of spaces of probability classes to describe the transmitting process of the second-level uncertainties. In this paper, the authors also present some transmitting properties of the specialization relations on classes of probabilities, and compares the two different ways of rule combination by using the concept of specialization. This paper gives a new comprehensive framework for knowledge-based reasoning.

Key words The second-level uncertainties, the space of probability class, specialization.

Class number TP18