人工智能 信任函数 理论 ②

7-12,31

计算机科学1993Vol.20№.4

信任函数理论的研究进展

TP18

杜劲松 黄俊杰 （武汉大学计算机系 武汉430072）
张尧庭 （武汉大学管理学院 武汉430072）

摘 要

This paper surveys the state-of-the-art progress made in the research of belief function theory in the context of its application in Artificial Intelligence. Main contents include:(i) basic concepts and operations of belief function; (ii) graphical belief function model and its local computation; (iii) belief function theory's relation with other numeric approaches to uncertainty in Artificial Intelligence. This paper provides a basic framework for the management of uncertainty

一、引言

通过对人工智能及专家系统的进一步研究，人们发现仅仅采用符号计算难以处理问题求解过程中所遇到的各式各样的不确定信息，而这种不确定性是无处不在的：从知识获取到计算机视觉与模式识别，从状态空间的有效搜索到自然语言理解。导致不确定的原因可能是：(1) 信息不是完全可靠的，例如，在基于规则的系统中，由专家给出的规则可能存在例外，甚至专家自身也会对规则缺乏自信；(2) 信息是不完备的，例如，在状态空间的搜索中，可能不知道全局的信息(代价)，有时候这种全局信息甚至不可能得到；(3) 不同来源的信息矛盾，这些冲突信息各自导致不同的决策；(4) 信息是不精确的，这种不精确性来自于概念本身缺乏严格定义(例如，自然语言中的模糊谓词)，或者来自于观测误差(例如传感器的随机误差)。

迄今，人们提出的处理人工智能中的不确定性的方法可以分为数值方法和非数值方法[1]。数值方法的典型代表有：(1)古典概率论或其变形(如MYCIN的确定性因子，PROSPECTOR的准Bayes方法)，(2)信任函数理论(或称为证据理论，Dempster-Shafer理论)；(3)模糊逻辑及可能性理论。非数值方法的典型代表有：(1)支持理论(Theory of Endorsement)，(2)非单调逻辑。数值方法和非数值方法在处理不确定性的方式上是不同的，它们能有效处理的不确定性的类别也不一样[2]，本文所感兴趣的是处理不确定性的数值方法中的信任函数理论。

信任函数理论是古典概率论的扩充[3,4,5]，它的一个主要优点在于提供给我们在不完备的概率模型上处理不确定性的便利，在人工智能及专家系统的研究领域中，信任函数理论正受到越来越多的研究者的关注。

人工智能及专家系统研究的问题包括：把定性的关系明晰地表示出来；把主观的信任定量化；建立形式化的模型；知识在模型上的有效存储与计算，确定的或不确定的证据在模型上的传播，从例子中学习，结论的解释等等。从目前的实践看，对这些问题进行概率分析是可行的，对于某些问题，概率分析甚至不可避免[6]，用信任函数理论研究上

收到日期：92—11—3。杜劲松博士生、张尧庭教授和黄俊杰教授的研究方向为人工智能中的不确定性。

述问题，是一个十分吸引人的课题。基于这种目的，本文介绍关于信任函数理论计算的研究进展，进一步研究它在人工智能及专家系统中的应用的基本框架。

二、信任函数理论的基本概念及运算

定义1 令Θ为一有限的结果空间，称为识别骨架。令m为从Θ的幂集到[0,1]的函数，满足：(1)$m(\phi)=0$；(2)$\forall A\subseteq\Theta$，$m(A)\geqslant 0$并且$\sum_{A\subseteq\Theta}m(A)=1$。称m为$\Theta$上的基本概率赋值。所有满足$m(A)>0$的$\Theta$的子集A称为焦点元素。

定义2 给定Θ上的基本概率赋值m，其相应的信任函数BEL定义为：

$$\forall A\subseteq\Theta,\ BEL(A)=\sum_{B\subseteq A}m(B)。$$

似真函数PLA定义为：

$$\forall A\subseteq\Theta,\ PLA(A)=1-BEL(\overline{A})。$$

定义3 给定Θ上的二个独立信任函数BEL_1和BEL_2(其基本概率赋值分别为m_1和m_2)，则它们可以用Dempster规则组合，得到Θ上的新的信任函数BEL，记为$BEL=BEL_1\oplus BEL_2$，其基本概率赋值m定义为：

$$(1)\ \forall C\subseteq\Theta,\ C\neq\phi,$$

$$m(C)=K^{-1}\sum_{\substack{A\cap B=C\\ A,\ B\subseteq\Theta}}m_1(A)m_2(B)$$

$$(2)\ m(\phi)=0$$

$$(3)\ K=[1-\sum_{\substack{A\cap B=\phi\\ A,\ B\subseteq\Theta}}m_1(A)m_2(B)]$$

注意：若$K=0$，则认为BEL_1与BEL_2不能组合，BEL无定义。

例1 从罐中取球[7]。设第一个罐中装的全是红球(记为R)，第二个罐中装的是红球和白球(记为R，W)，但不知其比例(可能全是红球，也可能全是白球)。现在要从罐子中取出一个球，设从第一个罐中取球的概率是0.7，从第二罐中取球的概率是0.3，在信任函数理论框架下，$\Theta=\{R,\ W\}$，基本概率赋值m为：$m\{R\}=0.7$，$m\{R,\ W\}=0.3$。于是：$BEL\{R\}=0.7$，$PLA\{R\}=1$，$BEL\{W\}=0$，$PLA\{W\}=0.3$。结果可以解释为取到红球的概率至少是0.7，至多是1。$m\{R,\ W\}=0.3$描述了我们关于第二个罐子的无知。

当结果空间为乘积空间时，我们称其上面的信任函数为多元信任函数[7]。使用多元信任函数不仅可以描述变量之间的概率关系(例如联合分布，条件分布)，也可以表示变量之间的逻辑关系。

例2 蕴含式的表示。设有蕴含$X\rightarrow Y$，这里X与Y均为布尔变量，在信任函数理论框架下，$\Theta=\Theta_X\times\Theta_Y=\{0,\ 1\}\times\{0,\ 1\}$，基本概率赋值为：$m\{(1,1),\ (0,\ 0),\ (0,\ 1)\}=1$，$m\{(1,\ 0)\}=0$。例2给出的信任函数是一类特殊的信任函数。

定义4 给定识别骨架Θ：(1)由基本概率赋值$m(\Theta)=1$，$\forall A\subseteq\Theta$，$m(A)=0$所给出的信任函数称为空信任函数；(2)由基本概率赋值$m(E)=p$，$m(\Theta)=1-p$，$\forall A\subseteq\Theta$，$A\neq E$，$m(A)=0$(其中E为Θ的子集)所给出的信任函数称为简单支持信任函数。

多元信任函数的常见运算如下：

定义5 令BEL_θ为$\Theta=\Theta_X\times\Theta_Y$上的信任函数，其基本概率赋值为$m_\theta$，则$BEL_\theta$到$_\theta X$的边缘化定义为$\Theta_X$上的信任函数$BEL_{\theta\downarrow\theta_X}$，其基本概率赋值m为：

$$\forall A\subseteq\Theta_X,\ m(A)=\sum_{\substack{B\subseteq\Theta\\ P_j\Theta_X(B)=A}}m(B)$$

这里$P_j\Theta_X(B)$表示B在Θ_X上的投影。

定义6 令BEL_{θ_X}为Θ_X上的信任函数，其基本概率赋值为m_{θ_X}，则BEL_{θ_X}到$\Theta=\Theta_X\times\Theta_Y$上的最小扩充定义为$\Theta$上的信任函数$BEL_{\theta_X\uparrow\theta}$，其基本概率赋值m为：

$$\forall A\subseteq\Theta,\ m(A)=\begin{cases}m_{\theta_X}(B), & 若存在B\subseteq\Theta,\ m_{\theta_X}(B)>0且A=B\times\Theta_Y\\ 0, & 否则\end{cases}$$

定义7 令BEL_θ为$\Theta=\Theta_X\times\Theta_Y$上的信任函数，则给定$Y=\theta_y\in\Theta_Y$时，$\Theta_X$上的条件信任函数$BEL_{\theta_X}(\cdot|Y=\theta_y)$定义为：

$$BEL_{\theta_X}(\cdot|Y=_{\theta_y})=(BEL_\theta\oplus BEL_{\theta_y})_{\downarrow\theta_X}$$

其中BEL_{θ_y}为一简单支持信任函数，其唯一

焦点元素为$\Theta_x \times \{\theta_y\}$。

例3 设X,Y为布尔变量，Θ为$\Theta=\Theta_X \times \Theta_Y=\{0, 1\}\times\{0, 1\}$，Θ上信任函数$BEL_\Theta$有基本概率赋值为：$m_\Theta\{(0, 0), (0, 1)\}=0.5$；$m_\Theta\{(1, 0), (1, 1)\}=0.3$；$m_\Theta\{(0, 0), (1, 0)\}=0.2$。则$BEL_{\Theta\downarrow\Theta_X}$的基本概率赋值为：$m\{0\}=0.5$；$m\{1\}=0.3$；$m\{0,1\}=0.2$。而$BEL_{\Theta\downarrow\Theta_X\uparrow\Theta}$的基本概率赋值为：$m\{(0, 0), (0, 1)\}=0.5$；$m\{(1, 0),(1, 1)\}=0.3$；$m\{(0, 0), (0, 1), (1, 0), (1, 1)\}=0.2$。由此可见，在一般情形下，$BEL_{\Theta\downarrow\Theta_X\uparrow\Theta}\neq BEL_\Theta$。

再回到BEL_Θ，设Y=1，则Θ_X上的条件信任函数$BEL_{\Theta_X}(\cdot|Y=1)$的基本概率赋值为：$m\{0, 1\}=1$。若设X=0，则$BEL_{\Theta_Y}(\cdot|X=0)$的基本概率赋值为：$m\{0\}=0.257$；$m\{0, 1\}=0.743$。

定义8 令$\Theta=\Theta_X\times\Theta_Y$，给定$\Theta_X$上的条件信任函数$BEL_{\Theta_X}(\cdot|Y=\theta_y)$，则它的条件嵌套定义为Θ上的信任函数$BEL_{\Theta_X}(\cdot|Y=\theta_y)_{\uparrow\uparrow\Theta}$，其基本概率赋值如下：

$$m(A)=\begin{cases} m(B) & \text{若存在} B\subseteq\Theta_X, m(B)>0, \text{且} A=B\times\theta_y\cup\Theta_X\times(\Theta_Y-\{\theta_y\}) \\ 0 & \text{否则} \end{cases}$$

若$\forall\theta_y\in\Theta_Y$，给定$BEL_{\Theta_X}(\cdot|Y=\theta_y)$，则多重条件嵌套定义为Θ上的信任函数：

$$\bigoplus_{\theta_y\in\Theta_Y} BEL_{\Theta_X}(\cdot|Y=\theta_y)_{\uparrow\uparrow\Theta}。$$

条件嵌套给出了由条件信任函数构造联合信任函数的方法。

例4 在例3中，$BEL_{\Theta_Y}(\cdot|X=0)$的基本概率赋值为：$m\{0\}=0.257$；$m\{0,1\}=0.743$。则它的条件嵌套即Θ上的信任函数$BEL_{\Theta_Y}(\cdot|X=0)_{\uparrow\uparrow\Theta}$的基本概率赋值为：$m\{(1, 0), (1, 1), (0, 0)\}=0.257$；$m\{(1, 1),(1, 0), (0, 1), (0, 0)\}=0.743$。

下面考虑几个计算问题。

(1) 基本概率赋值与信任函数的相互变换。这种变换实际上是Mobius变换[8]，Thoma给出了快速算法。例如从基本概率赋值到信任函数的变换，基本思想是：$\forall X\subseteq\Theta$，对Θ的每一元素v作迭代：

$$f_i(X)=\begin{cases} f_{i-1}(X)+f_{i-1}(X-\{v\}) & \text{若} v\in X \\ f_{i-1}(X) & \text{否则} \end{cases}$$

初始时，$f_0(X)=m(X)$，若$|\Theta|=n$，则f_n为信任函数。

(2) 关于Dempster规则的计算。与Mobius变换一样，由于计算是在Θ的幂集上进行的，因而时间复杂性是指数的。另外，同一识别骨架上的两个信任函数才能用Dempster规则组合；当识别骨架不同时，必须首先最小扩充到共同的识别骨架。为简便计，下文提到Dempster规则组合时，默认已作了最小扩充。注意到，Dempster规则组合运算满足可交换性。

上面二个核心过程的复杂性是导致信任函数理论受到批评的一个主要因素。为了降低复杂性（识别骨架的大小是重要的影响），研究者们提出了信任函数的分解与局部计算[7,8,9]以及近似计算的方法[10]。Dempster规则所要求的独立性后面还要提到。

三、图形信任函数及局部计算

在许多情形下，图是表示知识的自然而有效的途径。在古典概率论的框架下，[11, 12, 13]考虑了Bayes网络的计算（例如边缘化，条件化）。Bayes网络是一个有向图，它分解条件独立性下的联合概率分布。具体地说，对有向图的每一个顶点v，设其父结点集合为Π_v，给定了在Π_v的每一状态组合下的概率分布，即$Pr(v|\Pi_v)$。若$\Pi_v=\phi$，则给出的正好是v的边缘分布，以Bayes网络作为不确定知识的表示工具，Lauritzen和Spiegelhalter[11]研究了概率的局部计算，这里局部计算意味着在计算一组变量的边缘（或给定其它变量的值时的条件分布）时，不必诉诸于联合概率分布，因而使计算复杂性降低。依赖于局部计算，[11]考虑了专家系统中的

常见运算，例如证据的传播与组合。[7，8，9，14]研究了多元信任函数的局部计算。

定义9 一个超图H=〈N，C〉是一个二元组，N是顶点的集合，C是超边的集合，对于每一条超边$c\in C$，$c\subseteq N$。

定义10 图形信任函数模型G=〈N，B，C〉是一个三元组，这里N是变量的集合，〈N，C〉是一个超图，B是成分信任函数的集合，即$\forall c_i\subseteq C$，给定了识别骨架c_i上的信任函数BEL_{c_i}，$BEL_G=\bigoplus\limits_{c_i\in C}BEL_{c_i}$称为G的全信任函数，而G则称为$BEL_G$的一个分解。

例5 一个图形信任函数模型的例子。

G=〈N，B，C〉

$N=\{X_i|1\leqslant i\leqslant 8$，$X_i$为二元变量$\}$

$C=\{(X_1, X_2), (X_2, X_3, X_4), (X_3, X_5), (X_4, X_6, X_7), (X_4, X_8), (X_5, X_7)\}$

$B=\{BEL_{c_i}|c_i\in C\}$

$$BEL_G=\bigoplus_{c_i\in C}BEL_{c_i}$$

如果现在要计算某个变量，例如X_3的边缘信任函数，则$BEL_{X_3}=BEL_{G\downarrow X_3}=$

$$(\bigoplus_{c_i\in C}BEL_{c_i})_{\downarrow X_3}$$

这就是说，为了计算X_3的边缘，我们必须首先计算出全信任函数。这样一来识别骨架的势可达256。

幸好，在具有一定特性的图形模型上，边缘化算子与投影算子是可以交换的（回忆例3），使得局部计算可行。

在通常的图论的树的定义中，把顶点集换成超边的集合，可以得到下面的树模型的定义。

定义11 给定图形信任函数模型G=〈N，B，C〉，它相应的树模型T=〈C，B，E〉是一个三元组，其中C称为结点集，$E\subseteq C\times C$称为边集，并且〈C，E〉为一棵树。

定义12 给定图形信任函数模型G=〈N，B，C〉及相应的树模型T=〈C，B，E〉，对于N的子集A，定义T_A为T的子图，且T_A由C中所有包含A的结点组成。如果$\forall A\subseteq N$，T_A都是树，则称T为Markov树，我们也说T有分离特性或Markov特性。

容易证明[16]，定义12与下面的命题等价。

命题1 一棵树T=〈C，E〉（关于顶点N和成分信任函数B）为Markov树，当且仅当：(1)〈N，C〉是一超图；(2)如果$(c_1, c_2)\subseteq E$，则$c_1\cap c_2\neq\phi$；(3)如果$c_1, c_2\in C, c_1\neq c_2$，且$x\in c_1\cap c_2$，则x含在从$c_1$到$c_2$的路径上的每一个结点上。

例6 一棵Markov树(图3.1)。

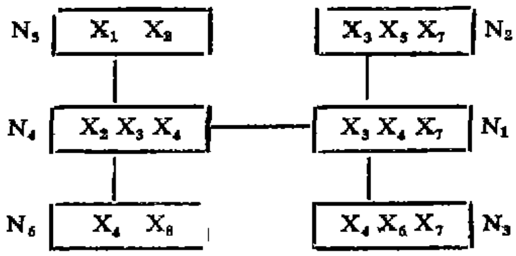

图3.1

容易验证：$N_2\cap N_3\subseteq N_1$

$N_4\cap N_2\subseteq N_1$

$N_6\cap N_1\subseteq N_4$，……等等。

给定图形信任函数模型G=〈N，B，C〉及相应的Markov树模型T=〈C，B，E〉，假定从G到T的变换是等价的，即$BEL_G=BEL_T$。现在设$c^*\in C$，我们要计算c^*的边缘信任函数$BEL_{G\downarrow c^*}$，定义为

$$BEL_{G\downarrow c^*}=BEL_{T\downarrow c^*}=BEL_{c^*}\oplus(\bigoplus_{\substack{c\in C\\ c\neq c^*}}BEL_c)_{\downarrow c^*}$$

现在把c^*看成树T的根结点，则删掉c^*后将得到k个子树组成的森林，设$\{c_1, \cdots, c_k\}$为c^*的邻结点集，以$c_1, \cdots, c_k$为k个子树的根。

利用定理1，它表明何时可以交换边缘化运算与最小扩充运算[7]。

定理1 设C为变量的集合，$C_1, C_2\subseteq C$，BEL_{C_1}和BEL_{C_2}为分别定义在识别骨架C_1及C_2上的信任函数，则：

$(\mathrm{BEL}_{C_1\downarrow(C_1\cap C_2)}\oplus \mathrm{BEL}_{C_2\downarrow(C_1\cap C_2)})\downarrow_{(C_1\cap C_2)}=\mathrm{BEL}_{C_1\downarrow(C_1\cap C_2)}\oplus\mathrm{BEL}_{C_2\downarrow(C_1\cap C_2)}$

考虑到T的Markov特性，可以证明定理[14]。

定理2 $\mathrm{BEL}_{G\downarrow C^*}=\mathrm{BEL}_{C^*}\oplus(\bigoplus_{1\leqslant i\leqslant k}\mathrm{BEL}_{C_i\to C^*})$

其中$\mathrm{BEL}_{C_i\to C^*}$为包含在子树$T_{C_i}$中的信息，它的递归定义如下：

$$\mathrm{BEL}_{C_i\to C^*}=(\mathrm{BEL}_{C_i}\oplus(\bigoplus_{\substack{c_j\neq c^*\\ c_j\in C'}}\mathrm{BEL}_{C_j}))\downarrow_{C^*}$$

这里C′为c_i的邻结点集合。

例7 在例6的Markov树上，设要计算$N_1=(X_3, X_4, X_7)$的边缘信任函数，以N_1为根得到图3.2，从叶结点(即只有一个邻结点的结点）沿箭头方向传播信息，例如，

$$\mathrm{BEL}_{N_5\to N_4}=\mathrm{BEL}_{N_5\downarrow(N_4\cap N_5)\uparrow N_4}$$

$$\mathrm{BEL}_{N_4\to N_1}=\mathrm{BEL}_{N_4}\oplus(\mathrm{BEL}_{N_5\to N_4}\oplus\mathrm{BEL}_{N_6\to N_4})\downarrow(N_4\cap N_1)\uparrow N_1$$

$$\mathrm{BEL}_{G\downarrow N_1}=\mathrm{BEL}_{N_1}\oplus(\bigoplus_{2\leqslant i\leqslant 4}\mathrm{BEL}_{N_i\to N_1})$$

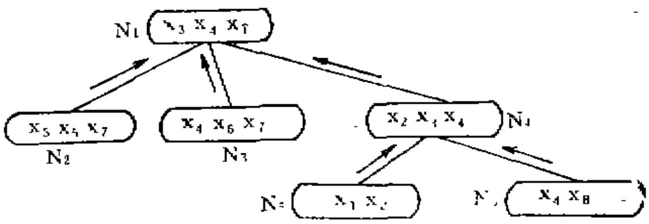

图3.2

定理2保证了传播与组合算法的正确性，局部计算的复杂性显然比直接计算全信任函数要低得多，最大的识别骨架势为3。

把Markov树的每个结点看成独立的处理器，每个结点在发送信息的同时也接受来自邻结点的信息，那么，经过若干周期，可以得到全部结点的边缘信任函数。

以局部计算为核心，就可以完成专家系统的例行过程，例如证据吸收与传播。另外这种局部计算模式在古典概率论的框架下也是可行的[9]。

现在的问题是，一个信任函数BEL如何分解成图形模型信任函数BEL_G，而一图形模型信任函数BEL_G如何转换成Markov树模型BEL_T。

对第一个问题，Dempster和Kong[16]的回答是：信任函数是通过组合独立成分而构造的，这些成分来自于诸如已接受的因果理论，专家的概率断言，或者是经验样本数据，然后在网络上开发的。

对第二个问题，目前正在研究中，一种构造方法是首先获得图形模型的超树覆盖，然后由超树构造Markov树。

定义13 图G=⟨V, E⟩是弦图，如果G的每一长度大于3的圈都有一根弦，即，若$v_1, \cdots, v_k, v_1(k\geqslant 3)$是圈，$\exists v_i, v_j$ $i=j+2$，$(v_i, v_j)\in E$。

定义14 图G=⟨V, E⟩的一个团是G的一个最大完全子图G′=⟨V′, E′⟩，即，$\forall v_i, v_j\in V'$，$(v_i, v_j)\in E'$，并且V′是最大的。

图的团不唯一。

定义15 对超图H=⟨N,C⟩定义图G(H)=⟨N,E⟩，其中$\forall v_i, v_j\in N$，$(v_i, v_j)\in E$ iff $\exists c\in C$，使得$v_i, v_j\in c$。

定义16 超图H=⟨N, C⟩称为conformal，若G(H)的任一个团都包含在一条超边中。H称为超树，若H是conformal的，而且G(H)是弦图。

定义17 超图H′=⟨N, C′⟩是超图H=⟨N, C⟩的一个覆盖，若H的每一条超边都包含在H′的一条超边中。当H′是超树时，称其为H的一个超树覆盖。

在[15]中，我们证明了：

定理3 若H=⟨N, C⟩是超树，则H上的最大势搜索算法[17]给出一个Markov树的构造。

由于Markov树的具有最大势的结点的大小对计算复杂性起着决定作用，我们自然希望找到这样的超树覆盖，即，它的具有最大势的结点的势尽可能小，我们称这样的超树覆盖为最小超树覆盖，我们证明了[18]：

定理4 任给超图H，寻找其最小超树覆盖是NP–完全的。

这启示我们，在超图的规模达到一定程度后，必须放弃最优性，而寻找次优超树覆盖。

四、信任函数理论与其它数值方法的关系

信任函数理论给出了概率的上下界，但是这种上下界不同于古典概率论的框架下区间概率演算的结果[19]。从某种意义上说，这换来的是信任函数理论可以在即使不完备的概率模型上作出推理。根据Bayes理论，如果不引用先验信息则不能对新证据产生任何结果。在[20]中，我们以诊断专家系统为例，证明了在完备的概率模型上，由信任函数理论给出不确定性组合公式与古典概率论给出的完全一致，而当模型不完备时，例如缺少先验信息，则古典概率论无法组合不确定性，但信任函数理论却可以，这一结果也说明观察的统计独立性（即条件独立性）满足Dempster规则所要求的直觉上的独立准则。然而，关于Dempster规则还有待进一步研究。

对Bayes网络，Pearl[13]建立了关于条件概率独立性的公理系统，这种公理系统的作用在于可以检验模型的正确性，同时可以推导出在模型中没有明确表示的独立性断言，而在图形信任函数模型中，超图不明确地定义了变量之间的条件独立性，但是Pearl的工作能否推广到信任函数理论框架下的超图模型上呢？目前我们已获得了初步的结果。

Grosof[21]的研究表明，在工程技术上应用广泛的确定性因子方法，经修正可以作为信任函数的一个特例。[17]中对演绎推理的研究也表明，可能性理论中的必要性测度与可能性测度是信任函数和似真函数的特例。

五、进一步的研究

因为不确定性本身就是一个不严格的定义，所以给形式的处理方法带来困难。尽管如此，从目前对人工智能中的不确定性的研究所取得的结果看，我们获得了关于不确定性的更透彻的认识，由于确定性的事件（或命题）均可视为不确定的一个特例，考虑到人工智能的研究特点（即，在对人的智能缺乏深刻理解的前提下用机器模拟人的智能），是否可以认为不确定性的研究在人工智能中的研究中占据着中心的地位呢？

在这个意义上看，作为比古典概率论等其它数值方法表达更强的形式语言，我们希望不仅能将信任函数理论用于专家系统的有效的不确定推理，而且还能将它用于人工智能的其它研究领域。

让我们引用Barnard在关于Shafer教授的文章的讨论[5]时讲的一句话作为本文的结束："我们花了近50年的时间才对似然性(likelihood)的意义和应用有了合理而清晰的了解；因此关于信任函数我们必须准备花至少类似长的时间"。

参考文献

[1] Mantaras, R. L. D., Approximate Reasoning Models. Ellis Horwood Limited, 1990

[2] Bhatnager, R. K.等, Handling Uncertain Information: A Review of Numeric and Non-numeric Methods, in Kanal, L. N. & Lemmer, J. F. (eds): Uncertainty in Artificial Intelligence, North-Holland, 1986

[3] Dempster, A. P., Upper and Lower Probabilities Induced by a Multivalued Mapping. Ann. Math. Statist., 38

[4] Shafer, G., A Mathematical Theory of Evidence. Princeton University Press, 1976

[5] Shafer, G., Belief Functions and Parametric Models. J. R. Statist. Soc. B, 44 (1982), No.3

[6] Lindley, D. V., Scoring Rules and the Inevitability of Probability. Int.

(转第31页)

的"Advice Taker"思想，Advice Taker是一个表达和求解问题的程序，其中称为情境演算的形式语言执行推导，情境演算扩展一阶语言，如增加了条件表达式(IF P THEN a ELSE b，其中P是一个公式，a，b是项)和一阶λ-表达式作为函词与谓词表达式。(所用条件表达式也是设计和分析LISP必需的。）这样，情境演算能广泛用于表示非数学问题，特别是只含单主体情形。

Advice Taker是第一个系统描述和分析麦卡锡称为语句自动机的尝试，其基本思想是，一个问题及其求解（可运用启发式）的所有基本因素都可以表示在一个形式语言中，形式语言中一些语句作为祈使句被要求执行，另一些语句是描述句，它们是问题各方面情形的描述，程序的基本功能是选择一组描述语句作前提并用一个演绎过程作出推演，这样，系统所知知识—即被告知—被表示为一个公理理论，关键之处在于面对新的问题不需要重新写一个新的程序，而只要告知系统有关新问题的事实就行。对麦卡锡来说，最感兴趣的应用是日常问题，这就需要表示常识概念，即日常生活中我们熟悉的客体及其关系，例如包含信念与知识、能力、和行动的前提与效果等概念。

经过三十多年努力，麦卡锡令人信服地指出一个远大的目标，就是在一种逻辑形式化中，建造一个可运用于广泛应用领域的公共常识库。很早时候，他已看到可靠（真值保护）逻辑演绎不是形式化常识推理的合适模型，常识推理包含非严格或在其它事情不变下的泛化，这包含一定非单调性。一个前提集Γ能是推出P的合适理由，但一个Γ的超集Γ′却未必是推出P的理由，例如Γ′可能包括了在其它事情可变的信息，非单调性很不同于演绎推理的结构(证明)。自1980年以来，麦卡锡的大部分工作集中在分析和形式化这种非单调推理，1977年他提出限制的思想，限制是一条应用于一阶理论的猜测规则，这种猜测规则的想法基于一个一定谓词，或更一般地，一个一定公式，它有与原始理论一致的外延，其直观意义是在其它事情确实相同下，给定原始理论(公理)，我们能用极小化例外或非规范事实。麦卡锡教授表明在一阶逻辑中增加这样一条限制规则（以及灵活使用它)就足以提供一个常识推理模型。许多处理常识推理的其它形式化工作后来被引入，但是，正如LISP和情境演算，麦卡锡的贡献都是最重要的。

我们引用授予麦卡锡教授国家科学勋章的评语来结束这个小传：

他对计算机科学与人工智能做出了巨大贡献，包括开发LISP程序语言，计算的数学理论，分时系统的概念和发展，应用数理逻辑于使用常识及其推理的计算机程序，以及创建人工智能本身。

（接第12页）

Statist. Rev., 50 (1982)

[7] Kong, C. T. A., Multivariate Belief Functions and Graphical Models. Ph. D. dissertation, Dept. of Statist., Harvard University, 1986

[8] Thoma, H. M., Factorization of Belief Functions, Ph. D. dissertation, Dept. of Statist., Harvard University, 1989

[9] Shenoy, P. P.等, Axioms for Probability and Belief Function Propagation. in Shachter, R. D. et al(eds): Uncertainty in Artificial Intelligence, North-Holland, 1990

[10] Gordon, J. 等, A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space. Artificial Intelligence 26 (1985)

[11] Lauritzen, S. L.等, Local Computation with Probabilities on Graphical Structures and their Application to Expert Systems. J. R. Statist. Soc. B, 50(1988), No. 2

[12] Pearl, J., Fusion, Propagation and Structuring in Belief Networks. Artificial Intelligence, 29(1986)

[13] Pearl, J., Probability Reasoning in Intelligence Systems:Networks of Plausible Inference. Morgan Kaufmann, 1988

[14] Almond, R. G., Fusion and Propagation of Graphical Belief Models, 出处同[8]

[15] Du, J. S., Maximal Cardinality Search Algorithm and Construction of Markov Tree Models, Technical Report, Dept, of Computer Science, Wuhan University, 1992

[16] Dempster, A. P.等, Uncertain Evidence and Artificial Analysis, J.Statist. Plan. Inf., 20(1988)

[17] Tarjan, R. E.等, Simple Linear-Time Algorithms To Test Chordality of Graphs, Test Acyclicity of Hypergraphs, and Selectively Reduce Acyclic Hypergraphs. SIAM J. Comput., 13(1984) No. 3

[18] Du, J. S., On Searching Minimal Hypertree Cover. Technical Report, Dept. of Computer Science Wuhan University, 1992

[19] Kyburgh, H. E.Jr., Bayesian and Non-Bayesian Evidential Updating. Artificial Intelligence, 31(1987)

[20] Du, J. S., A Note on Dempster's Rule of Combination. Technical Report, Dept. of Computer Science, Wuhan University, 1992

[21] Grosof, B. N., Evidential Confirmation as Transformed Probability:on the Duality of Prios and Updates. 出处同[2]