时态 Dempster-Shafer 理论*)

牟克典 林作铨
(北京大学信息科学系 北京 100871)

Temporal Dempster-Shafer Theory

MU Ke-Dian LIN Zuo-Quan
(Department of Informatics, Peking University, Beijing 100871)

Abstract In this paper, we propose temporal Dempster-Shafer theory to handle the combination of uncertainty and time. In temporal Dempster-Shafer theory, the element of the temporal frame of discernment is defined as an event that associates a hypothesis with corresponding time interval. And the assignment of belief to subset of the temporal frame of discernment is performed by the mass function. It is a representation and reasoning mechanism that combines uncertainty and time by the basic frame of Dempster-Shafer theory.

Keywords Uncertain reasoning, Dempster-Shafer theory, Temporal reasoning

1 引言

不确定性推理是人工智能中一个重要研究方向。在不同的应用领域,基于不同的不确定性度量理论,提出了许多不确定性推理理论和方法,例如确定因子法[1]、概率推理[2]、模糊推理[3]和 Dempster-Shafer 理论[4]等。Dempster-Shafer 理论是 Shafer 在 Dempster 提出的概率区间度量理论[5]的基础上进一步发展的不确定性推理理论。与概率推理等相比,Dempster-Shafer 理论在不确定性的度量上更为灵活,尤其是允许未知的存在更接近人的自然思维习惯。就具体推理机制来说,在 Dempster-Shafer 理论中对于信任更新是通过 Dempster 证据合成规则来实现的,相对于 Bayesian 后验公式,Dempster 规则形式比较直观、简洁,推理过程透明可见,同时适宜机器实现,因此在许多应用领域得到广泛应用[11,12]。

但是,和大多数不确定性推理理论一样,Dempster-Shafer 理论并没有在推理过程中直接考虑时间信息和相应的时态推理。事实上,在很多实际应用中,时间信息在描述所考虑的对象的发展演化等方面具有重要的意义[6,7],在推理过程中具有重要作用。例如,时间信息在诸如"给病人服用药物 D 之后 5 小时做的验血 B"[8]等情况中具有关键作用。事实上,在医学中,无论是预防、诊断还是治疗,对于时间信息的表示和推理都是至关重要的。

实际上,同时考虑对于不确定性和时间的表示和推理要比单纯考虑不确定性推理或时态推理更为复杂。对于时间和不确定性进行联合表示和推理不仅是 Dempster-Shafer 理论,也是其它不确定性推理方法在实际应用中所面临的问题,在不同的应用领域提出了不同的策略和方法[9,10]。在概率推理中,时态节点 Bayesian 网络是一种通过对 Bayesian 网络的节点进行改造从而使用 Bayesian 网络的基本框架来联合表示时间和不确定性的策略[9]。与一般 Bayesian 网络相比,时态节点 Bayesian 网络的节点表示的不是系统中的随机变量,而是由随机变量的状态和与之改变相关的时间区间组成的有序对(也叫事件)。通过这样的拓展,时态节点 Bayesian 网络既可以对时间和不确定性联合表示,同时也可以基于 Bayesian 网络的基本框架进行相关推理。事实上,这种将不确定性和时间信息相结合的策略可以推广到 Dempster-Shafer 理论中,建立比概率推理体系更简洁的推理理论。

在本文中,我们基于 Arroyo-Figueroa 在时态节点 Bayesian 网络中提出的联合处理时间和不确定性的策略[9],对 Dempster-Shafer 理论进行扩展,使之在保持灵活而简洁的度量和推理特性的同时,能够对时间和不确定性进行联合表征和推理。我们简要回顾了 Dempster-Shafer 理论的基本概念;对 Dempster-Shafer 理论进行扩展,提出了时态 Dempster-Shafer 理论;对时态 Dempster-Shafer 理论和 Dempster-Shafer 理论进行了比较。

2 Dempster-Shafer 理论

Dempster-Shafer 理论[4]中最基本的概念是 mass 函数,它体现了 Dempster-Shafer 理论对于不确定性的度量,即信任的分配方式。

设 Θ 是一个由一些两两互斥的假设(元素)构成的完备假设空间,称为辨识框架,$\mathscr{P}(\Theta)$是 Θ 的幂集,mass 函数是从$\mathscr{P}(\Theta)$到[0,1]的映射,满足:

①$mass(\varnothing)=0$,

$$②\sum_{X\subseteq\Theta} mass(X)=1。\tag{1}$$

从 mass 函数的定义可以看出,在信任分配方面,Dempster-Shafer 理论比概率推理更为灵活。基于 mass 函数,对于假设子集 $X\subseteq\Theta$,其信任函数 Bel 定义为:

$$Bel(X)=\sum_{Y\subseteq X} maxx(Y)\tag{2}$$

其合情度 Pls 定义为:

$$Pls(X)=1-Bel(\neg X)=1-\sum_{Y\subseteq\neg X} mass(Y)\tag{3}$$

其中$\neg X$ 为假设 X 的否定。在上面的定义中,信任函数表示证据对于假设的支持程度,而合情度则表示证据不否定假设的程度。在 Dempster-Shafer 理论中,就是采用一个证据区间$[Bel,Pls]$而不是单个的精确值来刻画信任。

*)本文得到国家自然科学基金资助(编号:69925203)。

对于不确定性推理来说，Dempster-Shafer 理论的推理机制是通过证据合成来体现的，即对两个独立的 mass 函数 $mass_1$ 和 $mass_2$，可以通过 Dempster 证据合成规则得到如下的复合 mass 函数：

对于$\varnothing \subset Z \subseteq \Theta$，

$$mass(Z)=mass_1 \oplus mass_2(Z) = \frac{\sum_{X\cap Y=Z} mass_1(X)\cdot mass_2(Y)}{1-\kappa} \tag{4}$$

其中

$$\kappa=\sum_{X\cap Y=\varnothing} mass_1(X) mass_2(Y)$$

为冲突因子，相当于分配给空子集的信任程度。

与概率推理相比，Dempster-Shafer 理论具有如下两个特点：首先，对于任何一个假设 X 来说，X 和 $\neg X$ 的信任函数之和并非一定等于 1，即

$$Bel(X)+Bel(\neg X)\leqslant 1 \tag{5}$$

这意味着允许未知存在，即假设及其否定可能都和证据相容，这是 Dempster-Shafer 理论和概率推理的根本区别。其次，就合成规则来说，对于信任的合成并不涉及原来信任以外的参数，并且合成运算简洁、直观，这也是 Dempster-Shafer 理论在推理机制比概率推理简单的原因之一。

事实上，在一般情况下，如果给定 Θ 和对应于一定证据的 mass 分配，则刻画信任的 Dempster-Shafer 理论框架完全可以用一个描述信任分配的二元组$(\Theta, mass)$来刻画，本文中我们采用$(\Theta, mass)$来表示 Dempster-Shafer 理论。

Dempster 证据合成规则尽管也反映了信任随着证据的增加而动态修正的特性，但这并不反映假设与证据之间的内在时间关系。事实上，在实际应用中，假设和证据之间的时间关系是在推理过程中起着重要作用的信息之一。例如，依据时间生物学的研究成果，在医学领域已经开始注意择时检查[8]，改变既往采用任意时间内取得的各种量值来诊断和治疗疾病的惯例，重视时间信息在诊断等方面的作用。如何将假设和证据之间的时间关系在推理过程中得到体现，这是 Dempster-Shafer 理论在实际应用中所面临的一个挑战。我们将提出时态 Dempster-Shafer 理论来处理这个问题。

3 时态 Dempster-Shafer 理论

根据 Allen 的时间区间推理理论[6]，半开半闭区间作为表示时间的基本单位，由两个时间点 t^- 和 t^+ 来表示，并且一般情况下认为 $t^-<t^+$。如果以 $t(A)^-$ 和 $t(B)^-$ 分别表示时间区间 A 和 B 的左端，以 $t(A)^+$ 和 $t(B)^+$ 分别表示时间区间 A 和 B 的右端，Allen 设计了基本的时间关系集合$\{e_q(=), b(<), b_i, m, m_i, o, o_i, s, s_i, d, d_i, f, f_i\}$，分别定义如下：

A 在 B 之前，以 AbB 或者 Bb_iA 表示，即 $t(B)^->t(A)^+$；

A 等于 B，以 Ae_qB 表示，即 $t(B)^-=t(A)^-$ 且 $t(B)^+=t(A)^+$；

A 遇上 B，以 AmB 或者 Bm_iA 表示，即 $t(B)^-=t(A)^+$；

A 交叉 B，以 AoB 或者 Bo_iA 表示，即 $t(A)^-<t(B)^-<t(A)^+<t(B)^+$；

A 在 B 中，以 AdB 或 Bd_iA 者表示，即 $t(B)^-<t(A)^-<t(A)^+\leqslant t(B)^+$ 或者 $t(B)^-\leqslant t(A)^-<t(A)^+<t(B)^+$；

A 开始 B，以 AsB 或者 Bs_iA 表示，即 $t(A)^-=t(B)^-<t(A)^+<t(B)^+$；

A 结束 B，以 AfB 或者 Bf_iA 表示，即 $t(A)^-<t(B)^-<t(A)^+=t(B)^+$。

本文中我们称这些时间区间及相应的时间关系为 Allen 时态关系。我们采用 Allen 时态关系作为时态表示方式，并且采用时态节点 Bayesian 网络[9]中的策略将时间和不确定性通过事件进行联合表示。

在时态 Dempster-Shafer 理论中，无论是证据还是假设，都可以视为与时间信息相关的事件。我们用 E 或者一个二元组(e, e_t)来表示证据，其中 e 表示证据量值，e_t 表示与它相关的时间区间。例如下午两点到四点的平均云层厚度为 H 可以表示为(平均云层厚度$=H$，(14：00，16：00)。在 Dempster-Shafer 理论中，辨识框架是由两两互斥的假设构成的完备集合，时间信息并没有在这些假设中得到明显表达。我们以 Θ 表示在不包含时间信息的情况下两两互斥的假设构成的完备集合，对于$\forall\ \theta\in\Theta$，以 $\Gamma(\theta)$表示与 θ 相关的时间区间的集合。我们用二元组(θ, θ_t)表示一个与时间相关的假设，其中，$\theta\in\Theta, \theta_t\in\Gamma(\theta)$，并且称这种与时间相关的假设为事件。我们扩展辨识框架为如下的时态辨识框架：$\Theta^T=\{(\theta,\theta_t)\mid\theta\in\Theta, \theta_t\in\Gamma(\theta)\}$并且为了保证事件之间的互斥关系，附加如下约束：若 $\forall\ (\theta^1,\theta_t^1),(\theta^2,\theta_t^2)\in\Theta^T$ 且 $\theta^1=\theta^2$，则 $\theta_t^1\cap\theta_t^2=\varnothing$。

我们首先给出如下的时态关系定义：

定义 1(证据和事件之间的时态关系)　证据(e, e_t)和事件(θ,θ_t)之间的时态关系 $R^T((e,e_t),(\theta,\theta_t))$是指相应的时间区间 e_t 和 θ_t 之间的 Allen 时态关系。

定义 2(证据和证据之间的时态关系)　证据 $E_1=(e_1, e_{1t})$和 $E_2=(e_2, e_{2t})$之间的时态关系 $R^T(E_1, E_2)$是指相应的时间区间 e_{1t}和 e_{2t}之间的 Allen 时态关系。

现在我们定义 mass 函数如下：

定义 3(mass 函数)　$mass$ 函数为 $\mathscr{P}(\Theta^T)$到$[0,1]$的映射，满足：①$mass(\varnothing)=0$；②$\sum_{x\subseteq\Theta^T} mass(X)=1$。

显然，mass 函数的定义实质上没有发生改变，这一点可以确保对于事件的信任分配方式并没有发生改变。对于事件集的信任函数以及合情度我们都可以沿用 Dempster-Shafer 理论中的定义。

基于以上的定义，我们用一个四元组$(\Theta^T, \Gamma, R^T, mass)$来表示时态 Dempster-Shafer 理论，其中：Θ^T 表示时态辨识框架；Γ 表示所有与 Θ 中假设相关的时间区间集合，即 $\Gamma=\bigcup_{\theta\in\Theta}\Gamma(\theta)$；$R^T$ 表示证据和所有事件之间的时态关系集合，即 $R^T=\bigcup_{(\theta,\theta_t)\in\Theta^T}\{R^T((e,e_t),(\theta,\theta_t))\}$；$mass$ 表示定义 3 的 mass 函数。

就不确定性推理的任务来说，还有一个很重要的问题，就是如何随着证据的增加来更新信任分配。对于时态 Dempster-Shafer 理论来说，证据合成要比 Dempster-Shafer 复杂一些，合成规则应该考虑如下两方面的要求：

①对于时态关系集合的更新；

②对于事件信任分配方案的更新。

设$(\Theta^T, \Gamma, R_1^T, mass_1)$和$(\Theta^T, \Gamma, R_2^T, mass_2)$是两个不同的独立信任分配，考虑到上面的两个要求，我们可以通过如下方法进行证据合成，即

$$(\Theta^T, \Gamma, R^T, mass)=(\Theta^T, \Gamma, R_1^T, mass_1)\oplus(\Theta^T, \Gamma, R_2^T, mass_2)=(\Theta^T, \Gamma, R_1^T\otimes R_2^T, mass_1\oplus mass_2) \tag{6}$$

其中

$$R_1^T\otimes R_2^T=R_1^T\cup R_2^T\cup\{R^T(E_1,E_2)\} \tag{7}$$

$mass_1\oplus mass_2$ 表示公式(4)所示的 $mass_1$ 和 $mass_2$ 的正交

和。

显然，这种合成规则一方面在 mass 函数的合成上和 Dempster 合成规则保持了一致，另一方面也充分考虑了时态关系集合的合成或更新。

例　天气预报问题

天气预报是需要对于时间和不确定性联合表征和推理的一个典型例子。为了简洁起见，我们不妨假定关于天气情况的基本假设集合由'晴(S)'，'阴(C)'，'雨(R)'3 种基本假设组成，即 $\Theta=\{S,C,R\}$。在一些比较特殊的日子，如庆典、卫星发射以及大型运动会开幕式等，我们对于可能下雨的时间更为关注。不妨设我们计划在周三下午两点到四点举行某重要活动，现在采用本周一凌晨到周二凌晨的一种气象指标 e_1 的观测资料来预测周三的天气情况，并且我们以周日凌晨作为时间起始点，作为 0:00，则周二凌晨为 24:00，周三凌晨为 48:00。在本例子中，显然

$E_1=(e_1,(0:00,24:00])$

$\Theta^T=\{(S,(48:00,72:00]),(C,(48:00,72:00]),(R,(62:00,64:00])\}$

$\Gamma=\{(48:00,72:00],(62:00,64:00]\}$

$R_1^T=\{(e_1,(0:00,24:00])<(S,(48:00,72:00]),(e_1,(0:00,24:00])<(C,(48:00,72:00]),(e_1,(0:00,24:00])<(R,(62:00,64:00])\}$

不妨设一种预报方法给出如下 mass 的分配：

$mass_1(\{(C,(48:00,72:00]),(R,(62:00,64:00])\})=0.7$

$mass_1(\Theta^T)=0.3$

另一方面，通过对于周一中午十二点到周二下午两点之间的另一种气象指标 e_2 的观测，可以独立给出如下的 mass 函数分配：

$mass_2(\{(R,(62:00,64:00])\})=0.9$

$mass_2(\Theta^T)=0.1$

很显然，

$E_2=(e_2,(12:00,38:00])$

$R_2^T=\{(e_2,(12:00,38:00])<(S,(48:00,72:00]),(e_2,(12:00,38:00])<(C,(48:00,2:00]),(e_2,(12:00,38:00))<(R,(62:00,64:00))\}$,

$R^T(E_1,E_2)=(e_1,(0:00,24:00])o(e_2,(12:00,38:00])$,

综合这两种气象指标，可以得到如下合成结果：

$R_1^T\oplus R_2^T=\{(e_1,(0:00,24:00])<(S,(48:00,72:00]),(e_1,(0:00,24:00])<(C,(48:00,72:00]),(e_1,(0:00,24:00])<(R,(62:00,64:00]),(e_1,(0:00,24:00])o(e_2,(12:00,38:00]),(e_2,(12:00,38:00])<(S,(48:00,72:00]),(e_2,(12:00,38:00])<(C,(48:00,72:00]),(e_2,(12:00,38:00])<(R,(62:00,64:00])\}$

$mass(\{(R,(62:00,64:00])\})=mass_1\oplus mass_2(\{(R,(62:00,64:00])\})=0.9$

$mass(\{(C,(48:00,72:00]),(R,(62:00,64:00])\})=mass_1\oplus mass_2(\{(C,(48:00,72:00]),(R,(62:00,64:00])\})=0.07$

$mass(\Theta^T)=mass_1\oplus mass_2(\Theta^T)=0.03$

则基于上面两个独立的气象观测资料，对于周三下午两点到四点下雨的证据区间为[0.9,1]。

讨论与小结　Dempster-Shafer 理论由于其对不确定性的度量上比较灵活，推理机制比较简洁，因此在专家系统以及不确定决策等领域一直受到一些研究者的关注和提倡。事实上，和其他不确定性推理理论一样，如何对时间和不确定性进行联合表示和推理是它在实际应用中所面临的一个挑战。在本文中，我们引入时间信息将 Dempster-Shafer 理论进行拓展，提出相应的时态 Dempster-Shafer 理论来处理与时间相关的简单的不确定性推理。时态 Dempster-Shafer 理论作为 Dempster-Shafer 理论的扩展，至少在以下两个方面与 Dempster-Shafer 理论有所不同：

①引入了时间信息，定义了事件的概念，通过从结构(Θ, $mass$)到($\Theta^T,\Gamma,R^T,mass$)的扩展实现了对时间和不确定性的联合表示；

②通过证据合成规则($\Theta^T,\Gamma,R_1^T,mass_1)\oplus(\Theta^T,\Gamma,R_2^T,mass_2)=(\Theta^T,\Gamma,R_1^T\oplus R_2^T,mass_1\oplus mass_2$)实现了对时间和不确定性联合进行推理的基本功能。

同时，与遵循概率论的时态节点 Bayesian 网络[9]相比，时态 Dempster-Shafer 理论在不确定性的度量思想和推理方法上基本和 Dempster-Shafer 理论保持了一致，即仍然通过 mass 函数作为信任分配方式，仍然采用 Dempster 合成规则作为信任更新机制，这使得它在信任分配方面比时态节点 Bayesian 网络更灵活，在推理机制方面更加简洁。

另一方面，时态 Dempster-Shafer 理论对于时间和不确定性联合进行表示和推理的机制尽管继承了 Dempster-Shafer 理论在不确定性度量和推理上灵活简洁等特性，但在处理一些更复杂的与时间相关的推理问题的时候具有一定的局限性。例如，在上面的关于天气预报的例子中，我们不能处理像周三下午两点到四点下雨并且早晨七点到八点之间天气晴朗的这类命题。因为在 Dempster-Shafer 理论中，假设之间是互斥的，假设集合$\{\theta_1,\theta_2\}\subseteq\Theta$ 表示假设 θ_1 或者 θ_2 之一成立，而不是假设 θ_1 和 θ_2 同时成立，这是一个值得进一步研究的问题。

参考文献

1 Buchanan B J, Shortcliffe E H. A model of inexact reasoning in medicine. Mathematical Biosciences, 1975, 23: 351～379

2 Pearl J. Probabilistic reasoning intelligent systems: networks of plausible inference. Morgan Kaufmann, San Mateo, CA, 1988

3 Zadeh L A. The role of fuzzy logic in the management of uncertainty in expert systems. Fuzzy Sets and Systems, 1983, 11: 199～227

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

5 Dempster A P. Upper and Lower Probabilities Induced by Multivalued Mappings. Annals of Math. Stat., 1967, 38: 325～329

6 Allen J F. Maintaining knowledge about temporal intervals. Communications of the ACM, 1983, 26(11), 832～843

7 Van Beek P, Manchak W D. The design and experimental analysis of algorithms for temporal reasoning. Journal of Artificial Intelligence Research, 1996, 4: 1～18

8 Santos E Jr. Unifying time and uncertainty for diagnosis. Journal of Experimental and Theoretical Artificial Intelligence, 1996, 8: 75～94

9 Arroyo-Figueroa G, Sucar E, Villavicencio A. Probabilistic temporal reasoning and its application to fossil power plant operation. Expert System with Applications, 1998, 15: 317～324

10 Aliferis C F, Cooper G F. A structurally and temporally extended Bayesian belief network model: definitions, properties and modeling techniques. In: Proc. of the 12th conf. on uncertainty in AI, UAI-96, Portland Bregon: Morgan Kaufmann, 28～39

11 Xu H. Valuation-based systems for decision analysis using belief functions. Decision Support Systems, 1997, 20: 165～184

12 Giarratano J, Riley G. Expert Systems: Principles & Programming, 3rd edition. Pws. Pub. Co, 1998