

一种集成式不确定推理方法研究

贺怀清 李建伏

(中国民航大学计算机科学与技术学院 天津 300300)

摘要 集成学习是采用某种规则把一系列学习器的结果进行整合以获得比单个学习器更好的学习效果的一种机器学习方法。研究表明集成学习是可行的,能取得比传统学习方法更好的性能。不确定推理是人工智能的重要研究方向之一,目前已经开发出了多种不确定推理方法,这些方法在实际应用中各有优缺点。借鉴集成学习,提出一种集成式不确定推理方法,其基本思想是按照一定的策略集成多种不确定推理方法,以提高推理的准确性。理论分析和实验结果验证了方法的合理性和可行性。

关键词 不确定推理,混合推理,集成学习

中图分类号 TP181 文献标识码 A

Research on an Ensemble Method of Uncertainty Reasoning

HE Huai-qing LI Jian-fu

(School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

Abstract Ensemble learning is a machine learning paradigm where multiple models are strategically generated and combined to obtain better predictive performance than a single learning method. It was proven that ensemble learning is feasible and tends to yield better results. Uncertainty reasoning is one of the important directions in artificial intelligence. Various uncertainty reasoning methods have been developed and all have their own advantages and disadvantages. Motivated by ensemble learning, an ensemble method of uncertainty reasoning was proposed. The main idea of the new method is in accordance with the basic framework of ensemble learning, where multiple uncertainty reasoning methods is used in time and the result of various reasoning methods is integrated by some rules as the final result. Finally, theoretical analysis and experimental tests show that the ensemble uncertainty reasoning method is effective and feasible.

Keywords Uncertainty reasoning, Mixed reasoning, Ensemble learning

1 引言

不确定推理是智能系统开发中的一个重要问题,也是解决传统的专家系统知识脆弱性和推理单调性的一个重要手段^[1]。对它的研究已进行了 30 多年,目前的不确定推理方法主要有主观贝叶斯方法、可信度方法、Dempster-Shafer(简称 D-S)证据理论方法、模糊推理方法、粗糙集方法等^[2]。在实际应用中,这些不确定推理方法各有优缺点。

主观贝叶斯方法是基于贝叶斯规则的计算方法,具有公理基础和易于理解的数学性质。但是,它要求所有假设的概率都是独立的。当这种独立性不被满足时,主观贝叶斯方法就会导致错误的结果。

可信度方法通过计算给定规则下的数值,即可确定结果的确定性因子。该方法简单、直观,易于掌握和使用。但是,如果推理链过长或推理顺序发生改变,该算法得到的推理结果就不准确^[3]。

D-S 证据理论方法有极强的理论基础,可以表示主、客

观信息,区分不确定和不知道,方便地定义各种问题,处理概率、模糊等不确定类型,在 20 世纪 80 年代相当流行。尽管 D-S 证据理论在处理不确定信息时有着独特的优点,但当证据之间是冲突的,即证据分别以较大的概率支持不同的对立命题时;若直接运用 D-S 证据理论的组合公式进行推理,往往会得到与现实相悖的结论,即冲突的证据焦点在推理后往往会变得很小,甚至会变成零,而组合前的概率很小的命题可能会变得很大,或者成为必然事件,很明显和想要得到的结果相悖。Zadeh^[4]曾举例说明 D-S 证据理论处理冲突证据时的不足:

例 1 两个医生检查了同一名病人,认为这个病人可能得的病是:脑膜炎(M)、脑震荡(C)、脑瘤(T)。假设这两个医生都认为该病人得脑瘤的可能性很小,但是脑膜炎还是脑震荡,两医生存在很大的分歧,他们的诊断如下:

$$m_1(M)=0.99, m_1(T)=0.01, m_2(C)=0.99, m_2(T)=0.01$$

将以上两医生的诊断结果利用证据理论组合公式进行组合,得到 $m(M)=0, m(C)=0, m(T)=1$,即合成结果表明这

到稿日期:2010-07-29 返修日期:2010-11-02 本文受国家自然科学基金项目(60879003),中央高校科研业务费(ZXH2009C001),中国民航大学引进人才科研启动费(09qd04S)和天津市应用基础及前沿技术研究计划项目(10JCYBJC00900)资助。

贺怀清(1971-),女,博士,教授,CCF 高级会员,主要研究方向为图形图像与可视化技术, E-mail: hqh@cauc.edu.cn; 李建伏(1979-),女,博士,讲师,主要研究方向为人工智能。

个病人肯定是得了肿瘤(T)——这显然与事实不符。

除了冲突证据问题, D-S 证据理论方法所面临的另一个重要问题是对焦元的基本概率分配敏感, 鲁棒性差。即当某个证据源对焦元的基本概率分配函数发生较小的变化时, 多源证据的组合结果会发生剧烈的变化。例如, 将例 1 的概率分配函数做一个微小的调整得到例 2, 然后利用 D-S 证据理论对例 2 进行推理, 得到 $m'(M)=0, m'(C)=0.99, m'(T)=0.01$ 。可见, 证据 R_1 的微小变化导致合成结果发生了剧烈变化。即对 T 的信任程度由例 1 的几乎完全肯定($m(T)=1$) 改变为几乎完全否定($m'(T)=0.01$); 而对 C 的信任程度由例 1 的完全否定($m(C)=0$) 变为几乎完全肯定($m'(C)=0.99$)。可见, D-S 证据理论对基本概率分配函数是非常敏感的。

例 2 $R_1: m_1'(M)=0.98, m_1'(C)=0.01, m_1'(T)=0.01$
 $R_2: m_2'(M)=0, m_2'(C)=0.99, m_2'(T)=0.01$

针对各种不确定推理算法的缺点, 许多作者进行了大量研究, 如为了解决 D-S 证据理论的冲突证据问题, Yager^[5] 率先发现冲突证据组合时产生的问题, 并提出将冲突信息部分归结为未知以及较小冲突; Dubois^[6] 则进一步提出组合中的冲突应当适当予以保留。此后的学者不断进行改进, 文献[7-11]使用“距离”的概念来衡量证据的相似度以缓解冲突; 文献[12-14]采用统一信念函数的概念建立参数化的合成规则, 而规则根据影响因子的大小确定冲突证据分配给不同识别框架下不同子集的比例。文献[15, 16]则分别对合成规则中的证据损耗和信念函数中的冲突程度进行了分析。虽然以上研究能够在一定程度上避免 D-S 证据理论方法的缺点, 但是以这种“打补丁”的方式来进行修正, 势必会带来新的问题。就目前不确定推理方法的发展趋势, 不同的文献[17-20]均指出, 在实际的应用系统中, 仅靠单一的推理方法难以取得非常理想的效果, 多种方法的联合将成为必然。

目前出现了多种推理方法联合的研究工作, 取得了不同程度的成功。如文献[21]表明 D-S 证据理论方法与模糊推理方法联合能提高系统的精确性和可靠性。文献[18]提出了 3 种联合方案用以目标识别。第 1 种是粗糙集和 D-S 证据理论联合: 首先利用粗糙集理论对源数据进行约简, 然后利用 D-S 证据理论进行合成; 第 2 种是粗糙集理论和人工神经网络(Artificial Neural Networks, 简称 ANN)联合: 首先利用粗糙集理论对源数据进行约简, 然后利用 ANN 进行合成; 第 3 种, 将粗糙集理论、D-S 证据理论和 ANN 3 者联合: 首先利用粗糙集理论对源数据进行约简, 然后利用 ANN 进行初步信息融合以为下一步的 D-S 证据理论计算每个目标的基本可信度分配, 最后利用 D-S 证据理论方法进行融合(如图 1 所示)。仿真实验表明, 以上 3 种联合方法的识别率均比 3 个单一方法, 即粗糙集理论、D-S 证据理论和 ANN 高。

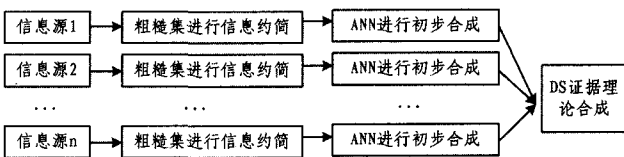


图 1 粗糙集、ANN 和 D-S 证据理论三级联合推理结构

以上推理方法联合的成功示例说明各种不确定方法在性能上具有很好的互补性, 多种不确定推理方法的联合会比单一方法更准确。同时也可以看到, 以上现有的不确定推理方

法的联合方式都是如图 1 所示的串行级联, 即一种不确定推理方法的处理结果作为另一个不确定推理方法的输入, 依此类推, 将最后一个不确定推理方法的输出作为最终结果。采用这种串行级联的联合方式, 其目的一般有两个: 第一, 前一种方法为后一种方法提供了一种信息预处理。如文献[18]中粗糙集理论和 D-S 证据理论方法的联合, 将粗糙集理论约简过的数据作为 D-S 证据理论的输入, 可以减少无关信息对 D-S 证据理论的干扰, 减少 D-S 证据理论的处理时间, 以提高系统的准确性和计算速度。第二, 前一种方法为后一种方法确定一些必要的计算因子。如图 1 中, ANN 就是为 D-S 证据理论计算每个目标的基本可信度分配。虽然在理论上能提高推理的准确性, 但是在实际应用中并没有一种完全精确的推理方法, 这种串行级联方式会使各个不确定推理方法的错误进行层层累积、传播, 进而增加最后推理结果的不准确。如利用粗糙集理论进行约简, 得到的信息不一定能够完全准确地代表原来的知识。以这种约简过的信息作为输入, D-S 证据理论将可能会变得更加不准确。

集成学习是 10 年来机器学习领域最主要的研究方向之一, 其基本思想是在对新实例进行分类时, 通过将多个分类器的分类结果进行某种组合来决定最终的分类并把若干个单个分类器集成起来, 从而达到提高分类器性能的目的。图 2 表示集成学习的基本思想。其中, 集成分类器包括了 n 个单一的分类器, 对于同样的输入, n 个分类器分别给出各自的输出 (O_1, O_2, \dots, O_n), 然后通过某种策略整合这些输出得到集成分类器的最终结果。集成学习始于 Hansen 和 Salamon 的开创性工作^[22]。他们研究发现, 通过训练多个神经网络并将其结果按照一定的规则进行组合, 能显著提高整个学习系统的泛化性能。后来, Schapire^[23] 通过构造性方法提出 Boosting 算法, 证明了这一点。受到这些研究的启发, 人们认识到集成学习所蕴含的巨大潜力和应用前景, 现在已经开发出了多种算法^[24], 如 Boosting, Bagging, Co-training 等。到目前为止, 集成学习已经广泛被应用于模式识别、文本过滤、遥感信息处理等领域, 取得了很好的效果^[25, 26]。

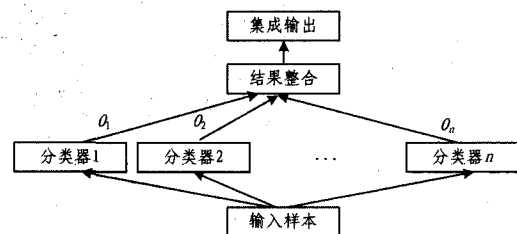


图 2 集成学习结构图

从原理上讲, 集成学习也是多种方法的联合。与目前不确定推理方法所采用的串行级联方式不同, 集成学习的联合方式是并行的, 即首先成员方法各自独立进行“计算”, 然后按照一定的策略将各个计算结果进行“平均”。在并行联合中, 只有当所有成员方法同时都犯错误时, 系统才会给出错误的结果; 只要各种成员方法是有差异的, 各种方法就能够相互纠正。所以集成学习有可能减少单个分类器的误差, 获得对问题空间模型更加准确的表示, 从而提高分类器的分类精度和稳定性。Thomas G. Dieterich^[12, 27] 从统计、计算和表示 3 方面定性分析了集成学习有效的原因, 并指出集成学习有效的一个充分必要条件是个体学习器间存在较大的差异, 各个

成员分类器之间的差异越大,集成学习的优势就越明显。另外,集成学习中各个成员方法相互独立地进行“计算”,从而避免了错误的累积和传播。

集成学习提供了一种解决问题的新范式,该范式被证明是可行的,并取得了比传统的学习方法更好的性能^[28]。受集成学习的启发,研究一种集成式不确定推理方法。其基本思想是采用集成学习的基本框架把多种推理方法联合起来,以提高系统推理的准确性。

2 集成式不确定推理方法

集成式不确定推理方法的基本思想是采用集成学习的基本框架,将多种不确定推理方法联合起来。即分别利用各种不同不确定推理方法各自进行推理,然后按照某种整合策略将各个推理结果进行组合以得到最终推理结果。其示意图如图3所示。

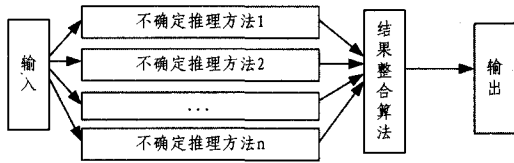


图3 集成式不确定推理方法示意图

集成式不确定推理方法的详细描述如下。

输入:不同的推理规则

输出:多种规则的合成结果

1. 开始
2. For $i=0$ to $n-1$ do
3. 利用不确定推理方法 i 进行推理,得到 O_i
4. End of For
5. 利用某种结果整合算法对 O_i 进行合并,得到输出 O
6. 结束。

值得指出的是,本文所提出的是一种框架,在具体设计一种集成式不确定推理方法时要考虑两个关键点,即选择什么样的不确定方法进行联合以及采用什么样的结果整合方法。根据集成学习理论,各个成员方法之间的差异性和各个方法本身的准确性都会影响到最终集成式不确定推理方法的准确性,因此在联合时尽量选择差异性比较大、比较准确的不确定推理方法进行联合。关于推理结果的整合,可以采用集成学习中一些整合方法,如简单投票、贝叶斯投票等^[29]。

为了在第3节通过算例验证这种集成式推理策略的有效性,这里设计了一个简单的集成式推理方法。该集成式推理方法集成了证据理论、可信度方法以及主观贝叶斯方法3种方法。所采用的结果整合方式是简单的平均,即 $O=(O_1+O_2+O_3)/3$,其中 O 为集成式推理方法的推理结果, O_1, O_2, O_3 分别为3种不确定推理方法的推理结果。

3 数值算例

为了验证集成式推理方法的有效性,针对冲突证据和合成规则的鲁棒性差的问题,本节利用算例1、算例2,将本文第2节设计的集成式不确定推理方法与D-S证据理论方法、目前D-S理论方法的3种较为成功的改进方法即文献[8]、文献[9]和文献[11]的推理结果进行比较。

3.1 冲突性对比实验

根据集成式推理方法的基本原理,利用本文第2节设计

的集成式不确定推理方法对算例1首先利用每个成员不确定推理方法(证据理论、可信度方法以及主观贝叶斯方法)进行推理,然后将3种方法的推理结果进行平均。具体推理过程如下。

根据算例1,有如下两条规则:

$$R_1: m_1(M)=0.99, m_1(C)=0, m_1(T)=0.01$$

$$R_2: m_2(M)=0, m_2(C)=0.99, m_2(T)=0.01$$

根据D-S证据理论方法的合成规则,对 R_1 和 R_2 进行合成,得到 $m(M)=0, m(C)=0, m(T)=1$ 。

为了利用可信度方法计算,重新将算例1描述如下。

已知如下规则:

$$R_1: \text{IF } E_1 \text{ THEN } A \text{ CF}(E_1, M)=0.99$$

$$R_2: \text{IF } E_1 \text{ THEN } B \text{ CF}(E_1, T)=0.01$$

$$R_3: \text{IF } E_1 \text{ THEN } C \text{ CF}(E_1, C)=0$$

$$R_4: \text{IF } E_2 \text{ THEN } A \text{ CF}(E_2, M)=0$$

$$R_5: \text{IF } E_2 \text{ THEN } B \text{ CF}(E_2, T)=0.01$$

$$R_6: \text{IF } E_2 \text{ THEN } C \text{ CF}(E_2, C)=0.99$$

初始证据 E_1, E_2 的可信度均为1。

根据可信度方法的合并规则,合并知识 R_1 和 R_4 ,得到 $\text{CF}(M)=0.99+0-0.99*0=0.99$;合并知识 R_2 和 R_5 ,得到 $\text{CF}(T)=0.01+0.01-0.01*0.01=0.0199$;合并知识 R_3 和 R_6 ,得到 $\text{CF}(C)=0.99+0-0.99*0=0.99$ 。

为了利用主观贝叶斯方法计算,重新将例1描述如下。

已知如下规则:

$$R_1: \text{IF } E_1 \text{ THEN } (9.9, 0.1) M$$

$$R_2: \text{IF } E_1 \text{ THEN } (0.1, 9.9) T$$

$$R_3: \text{IF } E_1 \text{ THEN } (1, 1) C$$

$$R_4: \text{IF } E_2 \text{ THEN } (1, 1) M$$

$$R_5: \text{IF } E_2 \text{ THEN } (0.1, 9.9) T$$

$$R_6: \text{IF } E_2 \text{ THEN } (9.9, 0.1) C$$

M, T, C 的先验概率均为0.5,证据 E_1 和 E_2 必然发生。

根据主观贝叶斯方法的公式,合并知识 R_1 和 R_4 ,得到 $P(M)=0.908$;合并知识 R_2 和 R_5 ,得到 $P(T)=0.0099$;合并知识 R_3 和 R_6 ,得到 $P(C)=0.908$ 。

然后根据结果整合方法将3种成员方法的结果进行平均,得到集成式不确定推理方法的结果,即 $m(M)=0.632, m(T)=0.343, m(C)=0.632$ 。通过归一化处理,得到 $m(M)=0.393, m(T)=0.2134, m(C)=0.393$ 。

本文的集成式推理方法与D-S证据理论方法、文献[8]、文献[9]和文献[11]的D-S理论方法对算例1的推理结果如表1所列。

表1 冲突证据合成实验结果

推理方法	m(A)	m(T)	m(C)
D-S	0	1	0
文献[8]	0.499	0.02	0.499
文献[9]	0.9998	0.002	0
文献[11]	0.495	0.01	0.495
集成式推理方法	0.393	0.213	0.393

通过表1可以看到,与D-S证据理论方法相比,集成式推理方法与文献[8,9,11]中的方法一样,都能得到与事实相符合的、较好的合成结果。但是与文献[8,9,11]中的方法相比,集成式推理方法的合成结果与事实仍有差距。主要原因有以下两点:第一点,根据集成学习理论,成员方法的准确性会影

响最终方法的准确性。而 D-S 证据理论得到的结论完全与事实相悖($m(T)=1$),进而影响了本文的集成式推理方法的结果。第二,本实验中所采用的集成式不确定推理方法只是 3 种不同不确定方法的一个简单平均,其中 D-S 证据本身与事实相差很远,使得平均结果也与事实有一定偏差。在进一步的研究中,通过采用更合理的结果整合方法能进一步提高算法的准确性。在下一节的鲁棒性对比试验中,也会受到以上两点原因的影响。因此,虽然不能完全避免,但本文中的集成性推理方法能够在一定程度上解决 D-S 证据理论方法的冲突证据问题。这说明,这种集成式的推理框架能够在一定程度上解决冲突证据问题。

3.2 鲁棒性对比实验

本文的集成式推理方法对算例 2 的推理过程如下。与上述对算例 1 进行推理的过程一样,首先利用三成员方法对算例 2 进行推理;利用 D-S 证据理论方法,得到 $m(M)=0, m(C)=0.99, m(T)=0.01$;根据可信度方法,得到 $CF(M)=0.98, CF(T)=0.0199, CF(C)=0.9901$;根据主观贝叶斯方法计算,得到 $P(M)=0.907, P(T)=0.0099, P(C)=0.908$ 。然后将以上 3 种成员方法的合成结果进行平均,有 $m(M)=0.392, m(T)=0.0133, m(C)=0.9627$ 。最后通过归一化处理,得到 $m(M)=0.392, m(T)=0.0083, m(C)=0.5998$ 。

用本文中的方法与文献[8]、文献[9]和文献[11]中的方法对算例 2 进行融合,得到的结果如表 2 所列。其中, $\Delta m(M)$, $\Delta m(T)$ 和 $\Delta m(C)$ 分别表示在算例 1 和算例 2 两种概率分配函数下,得到的最终推理结果中对 M, T 和 C 的信任度发生的变化。从数值上, $\Delta m(M)$ 为 $m(M)$ 与 $m'(M)$ 之差的绝对值,同样, $\Delta m(T)$ 为 $m(T)$ 与 $m'(T)$ 之差的绝对值, $\Delta m(C)$ 为 $m(C)$ 与 $m'(C)$ 之差的绝对值。通过表 2 可以看到,与 D-S 证据理论方法相比 ($\Delta m(M)=0, \Delta m(T)=0.99, \Delta m(C)=0.99$),集成式推理方法的合成结果发生的变化不大 ($\Delta m(M)=0.001, \Delta m(T)=0.205, \Delta m(C)=0.2068$),这说明集成式推理方法能够有效抑制 D-S 证据理论的鲁棒性差的问题。但是与文献[8,9,11]中的方法相比,集成式推理方法的合成结果发生的变化较大。其原因与冲突性对比实验中分析的两点原因一样:第一,成员方法 D-S 证据理论方法的鲁棒性差影响了最终结果的鲁棒性;第二,结果整合方式的不合理影响了最终结果。

表 2 鲁棒性对比实验结果

	M			T			C		
	$m(M)$	$m'(M)$	$\Delta m(M)$	$m(T)$	$m'(T)$	$\Delta m(T)$	$m(C)$	$m'(C)$	$\Delta m(C)$
D-S	0	0	0	1	0.01	0.99	0	0.99	0.99
文献 [8]	0.4999	0.4898	0.002	0.0002	0.0002	0	0.4999	0.5100	0.1001
文献 [9]	0.9998	0.9895	0.005	0.0002	0.0002	0	0.0103	0.0103	0
文献 [11]	0.4950	0.4900	0.005	0.01	0.01	0	0.495	0.5	0.005
本文	0.393	0.392	0.001	0.213	0.008	0.205	0.393	0.5998	0.2068

通过以上两个实验,说明了与文献[8,9,11]中提出的推理方法一样,集成式不确定推理方法能在一定程度上避免 D-S 证据理论的冲突证据和鲁棒性差的问题。这里需要指出的是,文献[8,9,11]中的方法只是针对 D-S 证据理论的冲突证据问题和鲁棒性差的问题而提出的,它们能在一定程度上避免冲突证据问题和鲁棒性差的问题。而本文提出的集成式推

理方法不仅能在一定程度上避免冲突证据问题和鲁棒性差的问题,也可以使得各种成员方法中其他缺点在一定程度上得以避免,如 D-S 证据理论的相关证据问题等。

结束语 给出了一种新的不确定推理框架。它采用了集成学习的策略,实现了多种不确定推理方法的联合。与现有的不确定推理方法的串行联合方式不同,本文的不确定推理框架使各个成员方法并行地、独立地进行推理,然后按照一定的策略将各个推理结果进行“平均”。只有当所有成员不确定推理方法同时都犯同一错误时,集成推理方法才会给出错误的结果;但是由于各种不确定推理方法是有差异的,即它们犯同一种错误的可能性很低,这些不确定推理方法能够相互纠正。最后,针对 D-S 证据理论的冲突证据以及鲁棒性差的问题,验证了本文提出的不确定推理方法的合理性和可行性。

在实验部分可以看到,不确定推理方法能够在一定程度上纠正每个成员不确定推理方法的缺点,但是不能完全避免。这主要是因为实验中所采用的结果整合方法过于简单,今后的工作包括研究合适的结果整合方法。

参考文献

- [1] 刘洁,陈小平,蔡庆生,等. 一种新的不确定推理方法[J]. 软件学报,2001,12(11):1675-1679
- [2] 张仰森. 人工智能原理与应用[M]. 北京:高等教育出版社,2004:136-183
- [3] 贾维弟. 不精确推理方法研究[D]. 哈尔滨:哈尔滨工业大学,2007
- [4] Zadeh L A. On the validity of Dempster's rule of combination of evidence[R]. UCB/ERL M79/24. University of California, Berkeley, 1979
- [5] Yager R. Using approximate reasoning to represent default knowledge[J]. Artificial Intelligence, 1987, 31(1):99-112
- [6] Dubois D, Prade H. Default reasoning and possibility theory[J]. Artificial Intelligence, 1988, 35(2):243-257
- [7] Joussleme A L, Grenier D, Bosse E. A new Distance between Two Bodies of Evidence[J]. Information Fusion, 2001, 2(2):91-101
- [8] Deng Yong, Shi Wen-kang, Zhu Zhen-fu, et al. Combining Belief Functions Based on Distance of Evidence[J]. Decision Support Systems, 2004, 38(3):489-493
- [9] 林志贵,徐立中,周金陵. 基于修改模型的冲突证据组合方法[J]. 上海交通大学学报,2006,40(11):1964-1970
- [10] 刘大有,杨鲲,唐海鹰. 凸函数证据理论模型[J]. 计算机研究与发展,2000,37(2):175-181
- [11] 杨善林,罗贺,胡小建. 基于焦元相似度的证据理论合成规则[J]. 模式识别与人工智能,2009,22(2):169-175
- [12] Leferve E, Colot O, Vannoorenberghe P. Belief Functions Combination and Conflict Management [J]. Information Fusion, 2002, 3(2):149-162
- [13] 孙全,叶秀清,顾伟康. 一种新的基于证据理论的合成公式[J]. 电子学报,2000,28(8):117-119
- [14] 何兵,郝爱民,赵沁平. 一种基于不确定信息的决策方法[J]. 计算机科学,2004,27(2):281-285
- [15] 徐凌宇,张博锋,徐炜民,等. D-S 证据理论中证据损耗分析及改进方法[J]. 软件学报,2004,15(1):69-75
- [16] Liu Wei-ru. A nalyising the Degree of conflict among Belief

- Foundations[J]. Artificial Intelligence, 2006, 170 (11): 909-924
- [17] 涂嘉文,徐守时. 贝叶斯方法与 Dempster-Shafer 证据理论的讨论[J]. 红外与激光工程, 2001, 30(2): 61-64
- [18] 王晓丽. 一种混合结构的数据融合算法研究及在目标识别中的应用[D]. 秦皇岛:燕山大学, 2004
- [19] 徐从富,耿卫东,潘云鹤. 面向数据融合的 DS 方法综述[J]. 电子学报, 2001, 29(3): 393-396
- [20] 徐从富,耿卫东,潘云鹤. Dempster-Shafer 证据理论方法与应用的综述[J]. 模式识别与人工智能, 1999, 12(4): 424-430
- [21] 刘同明. 基于证据理论模糊推理的多传感器信息融合海上目标识别[J]. 模式识别与人工智能, 1999, 12(1): 25-31
- [22] Hansen L K, Salamon P. Neural network ensembles[J]. IEEE Trans on Pattern analysis and machine intelligence, 1990, 12 (10): 993-1001
- [23] Schapire R E. The strength of weak learnability[J]. Machine Learning, 1990, 5(2): 197-227
- [24] Sewell M. Ensemble learning [R]. University College Lodon, 2007
- [25] Hampshire J, Waibel A. A novel objective function for improved phoneme recognition using time-delay neural networks[J]. IEEE Trans Neural Networks, 1990, 1(2): 216-228
- [26] Schapire R E, Singer Y, Singhal A. Boosting and Rocchio applied to text filtering[C]//Proc the 21st Annual ACM SIGIR International Conference on Research and Development in Information Retrieval. NY, 1998: 215-223
- [27] Tumer K, Ghosh J. Error correlation and error reduction in ensemble classifiers[J]. Connection Science, 1996, 8(3/4): 385-403
- [28] Diettefich G, learning E. The Handbook of Brain Theory and Neural Networks(Second Edition)[M]. MIT press, 2002: 385-577
- [29] 刘明,袁保宗,等. 从局部分类精度到分类置信的变换[J]. 计算机研究与发展, 2008, 45(9): 1612-1619

(上接第 229 页)

- [76] Lesk M E. Automatic Sense Disambiguation using Machine Readable Dictionaries; How to Tell a Pine Cone from an Ice Cream Cone[C]// Proceedings of the ACM SIGDOC Conference. Toronto, Ontario, 1986: 24-26
- [77] Walker D E. Knowledge Resource Tools for Accessing Large Text Files[C]// Proceedings of Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages. Colgate University, Hamilton, New York, August 1985: 335-347
- [78] Yarowsky D. Word Sense Disambiguation Using Statistical Models of Roget's Categories Trained on Large Corpora[C]// Proceedings, COLING-92. Nantes, France, 1992: 241-246
- [79] Dewey M, et al. Dewey Decimal Classification and Relative Index (22 edition)[M]. OCLC Online Computer Library Center, October 2003
- [80] Gonzalo J, Verdejo F, Peters C, et al. Applying eurowordnet to cross-language text retrieval[J]. Eurowordnet; A Multilingual Database With Lexical Semantic Networks, 1998, 32: 185-207
- [81] Mihalcea R. Using Wikipedia for automatic word sense disambiguation[C]// Human Language Technologies 2007; The Conference of the North American Chapter of the Association for Computational Linguistics. Rochester, New York, April 2007
- [82] Agirre E, Rigau G. A proposal for Word Sense Disambiguation Using Conceptual Distance[C]// Proceedings of the First International Conference on Recent Advanced in NLP. Bulgaria, 1995: 162-171
- [83] Rosso P, Masulli F, Buscaldi D, et al. Automatic noun sense disambiguation[C]// Proceeding of CICLing. 2003: 273-276
- [84] Buscaldi D, Rosso P, Masulli F. Integrating Conceptual Density with Word-Net Domains and CALD Glosses for Noun Sense Disambiguation[J]. Lecture Notes in Artificial Intelligence, 2004, 3230: 183-194
- [85] Litkowski K. Use of machine readable dictionaries for word-sense disambiguation in senseval-2 [C] // Proceedings of the Senseval-2 Workshop. Toulouse, 2001: 107-110
- [86] McCarthy, Diana, Koeling R, et al. Using automatically acquired predominant senses for word sense disambiguation[C]// Proceedings of the ACL Senseval-3 Workshop. Barcelona, Spain, 2004: 151-154
- [87] Morris J, Hirst G. Lexical cohesion computed by thesaural relations as an indicator of the structure of text[J]. Computational Linguistics, 1991, 17(1): 21-48
- [88] Galley M, McKeown K. Improving word sense disambiguation in lexical chaining[C]// Proceeding of the 18th International Joint Conference, on Artificial Intelligence, UCAI 2003. 2003: 1486-1488
- [89] Mihalcea R, Tarau P, Figa E. PageRank on semantic networks with application to word sense disambiguation[C]// Proceeding of the 20th International Conference on Computational Linguistics (COLING-2004)
- [90] Navigli R, Velardi P. Structural semantic interconnections: A knowledge-based approach to word sense disambiguation[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(7): 1075-1086
- [91] Agirre E, Martnez D, de Lacalle O L, et al. Two graph-based algorithms for state-of-the-art WSD[C]// Proceeding of the Conference on Empirical Methods in Natural Language Processing (EMNLP). Barcelona; Association for Computational Linguistics, 2006: 583-593
- [92] Sinha R, Mihalcea R. Unsupervised graph-based word sense disambiguation using measures of word semantic similarity[C]// Proceeding of the IEEE International Conference on Semantic Computing (ICSC). Washington; IEEE Computer Society, 2007: 363-369
- [93] Navigli R. Word sense disambiguation; A survey[J]. ACM Computer. Survey, 2009, 41(2)
- [94] Wu Yun-fang. A survey of Chinese word sense disambiguation: Resources, methods and evaluation[J]. Journal of Contemporary Linguistics, 2009, 11(2): 113-123
- [95] 刘群. 统计机器翻译综述[J]. 中文信息学报, 2003, 17(4): 1-12
- [96] 熊德意, 刘群, 林守勋. 基于句法的统计机器翻译综述[J]. 中文信息学报, 2008, 22(2): 28-39