

集对分析理论及其应用研究进展^{*}

蒋云良^{1,2} 徐从富²

(湖州师范学院信息工程学院 湖州 313000)¹ (浙江大学人工智能研究所 杭州 310027)²

摘要 集对分析理论是一种较新的软计算方法,可有效地分析和处理不确定信息。近年来,该理论日益受到学术界的重视,已经在决策、预测、数据融合、不确定性推理、产品设计、网络计划、综合评价等领域得到较为成功的应用。本文简要介绍了集对分析理论的基本概念和理论基础,较详细地论述了该理论的最新研究成果与应用进展情况,最后指出可能的发展趋势和研究方向。

关键词 集对分析, 不确定性, 数据分析, 软计算

Advances in Set Pair Analysis Theory and its Applications

JIANG Yun-Liang^{1,2} XU Cong-Fu²

(School of Information & Engineering, Huzhou Teachers College, Huzhou 313000)¹

(Institute of Artificial Intelligence, Zhejiang University, Hangzhou 310027)²

Abstract Set pair analysis (SPA) theory is a relatively new soft computing tool to deal with vagueness and uncertainty. It has attracted much attention of researchers. SPA has been applied to many areas successfully including decision-making, forecasting, data fusion, uncertainty reasoning, product design, network planning, and comprehensive evaluation, etc. This paper introduces the basic concepts and foundations of SPA, and reviews the recent advances of the development and applications of SPA. Finally, some promising research directions are discussed.

Keywords Set pair analysis, Uncertainty, Data analysis, Soft computing

1 引言

集对分析(Set Pair Analysis, SPA)理论(简称集对论)是一种新型的处理模糊和不确定知识的数学工具,能有效地分析和处理不精确、不一致、不完整等各种不确定信息,并从中发现隐含的知识,揭示潜在的规律^[1,2]。

集对论由中国学者赵克勤在1989年首次提出^[3]。2000年,赵克勤出版了专著^[4],系统、全面地阐述了该理论的基本概念、原理、方法及应用。该书与2001年出版的集对论应用专集^[5]较好地总结了这一时期关于集对论的研究及应用成果,促进了它的进一步发展,现已成为学习和应用集对论的重要文献。

从1995年至今,中国每年都召开以集对论为主题的学术研讨会。2001年起中国人工智能学会把集对论作为中国人工智能学会全国学术大会征文领域之一。许多学术期刊都将集对论列为重要内容之一。所有这些都极大地推动了集对论的拓展和应用。目前集对论已成为国内人工智能(特别是不确定信息处理)领域中的一个较新的学术热点,受到越来越多的科研人员的关注。经过十几年的研究与发展,集对论已经在理论与实际应用上取得了长足的发展。近年来集对论在决策、预测、数据融合、不确定性推理、产品设计、网络计划、综合评价等领域得到了较为成功的应用。

本文首先简要介绍集对论的基本概念和理论基础,论述集对论本身的研究现状,并着重评述集对论在决策、预测、数

据融合、不确定性推理、产品设计中的具体应用方法与实例,最后给出了建议的研究方向。

2 集对论的基本概念和理论基础^[1,2,4,6]

世界是确定性与不确定性的矛盾统一体。各种系统、各种事物,在某种条件下、某种层次上,体现出确定性;而在另一种条件下、另一种层次上,体现出不确定性。集对论是从整体和全局上研究确定性和不确定性的一种新的不确定性理论,其核心思想是:认为任何系统都是由确定性和不确定性信息构成的,在这个系统中,确定性与不确定性相互联系、相互影响、相互制约,甚至在一定条件下可相互转化,并用一个能充分体现上述思想的计算公式来统一地描述模糊不确定、随机不确定、中介不确定,以及由不知道和信息不完全等情况引起的不确定性。

2.1 集对分析的基本思路

所谓集对,就是具有一定联系的两个集合所组成的对子。如系统和环境、系统和工程、信息和控制、军事和国防、电脑和人脑、生产与销售等,在一定条件下都可看成是集对的例子。集对分析的基本思路是:在特定的问题背景下,对所论述的两个集合所具有的特性做相同、相异和相反性分析并加以定量刻画,得到这两个集合的联系度表达式,在此基础上再深入研究系统的有关联系、决策、预测、控制、仿真、评价、演化、突变等问题。

2.2 联系度

^{*}国家自然科学基金资助项目(60573056,60402010)、浙江省自然科学基金资助项目(Y105090, M603169)和湖州市自然科学基金资助项目(2005YZ08)。蒋云良 博士生,副教授,主要研究方向为人工智能、信息融合、GIS等;徐从富 博士,副教授,主要研究方向为人工智能、信息融合、数据挖掘等。

一般情况下,联系度是一个表达式:

$$\mu = \frac{S}{N} + \frac{F}{N}i + \frac{P}{N}j \quad (1)$$

式中的 N 为所讨论集对具有的特性总数, S 为集对中两个集合共同具有的特性个数, P 为集对中两个集合相互对立的特性个数, $F = N - S - P$ 是集对中两个集合既不共同具有、又不互相对立的特性个数, S/N , F/N , P/N 分别称为所讨论集对在指定问题背景下的同一度、差异度、对立(相反)度; j 为对立度的系数, 规定取值 -1 ; i 为差异度的系数, 规定在 $[-1, 1]$ 区间视不同情况取值。在不计 i, j 的值时, i, j 分别做差异度与对度的标记使用。为简便计, 可令 $a = S/N, b = F/N, c = P/N$, 于是(1)式可写成

$$\mu = a + bi + cj \quad (2)$$

由前述定义可知, a, b, c 满足归一化条件: $a + b + c = 1$ 。

据此, 有时又可把(2)式简写成

$$\mu' = a + bi \quad (3)$$

$$\text{或 } \mu' = a + cj \quad (4)$$

$$\text{或 } \mu' = bi + cj \quad (5)$$

上述(2)~(5)式可根据不同情况加以选用。

不难看出, 联系度 $\mu = a + bi + cj$ 是对研究对象所处状态空间进行“一分为三”的一种刻画。这种刻画对于不少问题的研究已有足够的精度。但也存在一些问题, 因为将状态空间简单地“一分为三”, 显得过于粗糙。为此, 可将联系度 $\mu = a + bi + cj$ 根据不同的情况做不同层次的展开。例如, 把 b 做更深层次的细分, 则可将(2)式展开为:

$$\mu = a + b_1i_1 + b_2i_2 + \dots + b_ni_n + cj \quad (6)$$

当 $n=2$ 时, (6)式变为

$$\mu = a + b_1i_1 + b_2i_2 + cj \quad (7)$$

为方便起见, 可把(7)式改写为:

$$\mu = a + bi + cj + dk \quad (8)$$

并规定 $a \in [0, 1], b \in [0, 1], c \in [0, 1], d \in [0, 1]$, 且 $a + b + c + d = 1, i \in [0, 1], j \in [-1, 0], k = -1$ 。在不计 i, j, k 的值时, i, j, k 仅做标记使用, 并称 a 为同一度、 b 为正差异度、 c 为负差异度、 d 为对立(相反)度。称以上规定的(8)式为四元联系度。类似地可以得到 n 元 ($n=5, 6$) 联系度的概念。

联系度源自对决策的研究。由(1)式所示的联系度可以方便地刻画决策中的评价、表决结果。例如, 一个方案, 让 10 个人去表决, 其中 6 人赞同, 3 人反对, 1 人弃权, 则可用

$$\mu = \frac{6}{10} + \frac{1}{10}i + \frac{3}{10}j = 0.6 + 0.1i + 0.3j \text{ 表示表决结果, 并据此做出决策。}$$

联系度的确定可以采用穷举法、逐步分析法、层次分析法、直接比较法、统计法、极限法、比例法等^[1, 2]。而不确定系数 i 的取值可以采用顺势取值法、逆势取值法、计算取值法、比例取值法、随机取值法、特殊值法、确定取值法等^[1, 2]。需要指出的是, 在不少情况下, 我们只能给出 i 的一个取值范围, 即在 $[-1, 1]$ 这个大区间中确定出一个较小的区间, 如何具体地确定出这个小范围, 要根据具体情况来定^[4]。

2.3 联系数

称形如 $a + bi + cj, a + bi, a + cj, bi + cj$ 的数为联系数, 其中 a, b, c 是任意正数, $j = -1, i \in [-1, 1]$ 且不确定取值^[4]。

引进联系数的最初目的是为了应用上的方便, 但其理论意义则在于推广了数的概念。联系数的意义在于一方面它把

可确定数与所在范围、数与值联系起来, 另一方面它把宏观层次上的确定量和微观层次上的不确定量联系起来。

2.4 联系变量与联系函数

称在宏观层次上随时间 t 等因素变动着的联系数为联系变量, 可用 $\mu(t), \mu(x), \mu(y)$ 等加以表示^[4]。如果一个联系变量 $\mu(y)$ 的变化由另一个联系变量 $\mu(x)$ 引起, 则称 $\mu(y)$ 是 $\mu(x)$ 的联系函数。其中 $\mu(x)$ 称为自变不确定量, $\mu(y)$ 称为因变不确定量。

2.5 联系熵

熵是一个重要的概念, 长期以来人们提出了热力学熵、统计熵、信息熵、模糊熵、负熵等多种熵概念。对 $n \geq 2$ 个联系度进行数学处理过程中, 文^[6]提出了一种有别于传统熵概念的“联系熵”。联系熵的一般形式为:

$$S = S_s + S_f + S_p = \sum a_n \ln a_n + i \sum b_n \ln b_n + j \sum c_n \ln c_n \quad (9)$$

研究表明, 热力学熵、统计熵、信息熵、模糊熵、负熵都只是反映了联系熵中的部分熵, 而联系熵却在一定条件下包含了这些各具特征的熵, 即联系熵是一种全局熵。初步的研究表明, 联系熵是研究系统复杂性的一个较为有力的工具^[6]。

2.6 基于集对分析的不确定性理论

基于集对分析的不确定性理论包括以下内容^[1]:

(1) 确定性和不确定性共同构成一个系统。对不确定性的描述应同时从确定和不确定两方面进行, 即一方面有 $b = 1 - a - c$, 另一方面 b 的系数 i 在 $[-1, 1]$ 区间根据不同情况取值, 可参考(2)式。

(2) 在一个系统中, 确定与不确定是相互联系、相互影响和相互渗透的, 确定中有不确定, 不确定中有确定, 在一定条件下可把不确定性里面的确定性分离出来, 可参考(1)式。

(3) 只有在忽略了系统的不确定性时, 系统才呈现为确定, 可参考(4)式。

(4) 不确定性与对立关系密切。可按对立的分类把不确定性分成倒数型不确定(特征方程: $k \times 1/k = 1$)、有无型不确定(特征方程: $k \times 0 = 0$)、正负型不确定(特征方程: $1 \times (-1) = -1$)、虚实型不确定(特征方程: $1 \times \sqrt{-1} = \sqrt{-1}$), 以及互补型不确定(特征方程: $(A+B) = 1$), 这些不确定与我们前面提到的不确定类型的对应关系是: 倒数型不确定-模糊不确定; 有无型不确定-随机不确定; 正负型不确定-中介不确定; 虚实型不确定-由不知道引起的不确定; 互补型不确定-由信息不完全引起的不确定。

(5) 不确定性就是在某个区间变化着的东西, 只是在某个特定条件下, 才取某个定值。

(6) 概率与隶属度由于都是以某个定值出现, 因此只反应系统在某个特定条件下的不确定性, 它们都可以在一定条件下与 μ 中的同一度 a 等价。

(7) 可以根据不确定的程度把不确定性分成一阶不确定(i)、二阶不确定(i^2)、三阶不确定(i^3)、无穷大阶不确定(i^∞), 以及 i 与 a, i 与 c 以不同形式复合的不确定。

(8) 熵可以度量不确定性, 也可以度量确定性, 还可以度量确定与不确定相混杂的情况^[6]。

(9) 所有的不确定性在本质上应该是统一的, 可以统一地加以研究。联系度就是对各种不确定性加以统一研究的一种工具。等等。

3 集对论的研究现状

目前, 集对论的研究主要集中在其数学性质、集对论的拓

广、同异反集合理论、与其它不确定方法的关系和互补等方面。

3.1 集对论的数学性质及其拓广

在集对论数学性质方面,主要研究联系度(或联系数)的四则运算及其性质、联系度的等价关系和优先关系,并且证明了联系度是“半序”的^[7]。

在集对论拓广方面的研究主要涉及广义集对分析模型。有学者认为没有必要限定集对论中两个集合的元素数目相等,并提出了广义集对分析模型,从而拓广了集对论的应用范围。并且,两集合元素间的对比应在元素序偶中展开,非对应元素进行直接比较,对比方式应采用“是否符合某种关系”来进行^[7]。

3.2 同异反集合理论

在同异反集合理论的研究方面,文[8]给出了对象空间(U)、属性空间(F)、表征空间(S)、问题空间($T=(U, F, S)$)、背景空间($T_0=(U, F_0, S_0)$)等概念,并研究了不确定性信息处理的一般性描述。

定义1 以 T_0 为论域,在 T_0 中关于一种属性 $y \in F_0$,定义 T_0 上的一对子集 A, A^c 分别为相对确定性信息和相对不确定性信息,则对于信息 $x \in T_0$,存在一对映射

$$\mu_A : T_0 \rightarrow [0, 1], x \rightarrow \mu_A(x) = a + c$$

$$\mu_{A^c} : T_0 \rightarrow [0, 1], x \rightarrow \mu_{A^c}(x) = b$$

$a+c, b$ 分别称为 x 关于 A, A^c 的确定性隶属度和差异隶属度。 A, A^c 分别称为 T_0 上的确定集和差异集(或不确定集)。记 $M=(A, A^c) = \{(\mu_A(x), x), (\mu_{A^c}(x), x) | x \in T_0\}$,构成 T_0 的一对信息分配,称为关于 T_0 的确定-不确定偶集。

定义2 以相对确定性信息 A 为论域,关于 $y \in F_0, A$ 上的一对子集 B, B^c 分别为同一性信息和对立性信息,则对于信息 $x \in A$,存在一对映射

$$\mu_B : A \rightarrow [0, 1], x \rightarrow \mu_B(x) = a$$

$$\mu_{B^c} : A \rightarrow [0, 1], x \rightarrow \mu_{B^c}(x) = c$$

a, c 分别称为 x 关于 B, B^c 的同一隶属度和对立隶属度。 B, B^c 分别称为 T_0 上的同一集和对立集。记 $N=(B, B^c) = \{(\mu_B(x), x), (\mu_{B^c}(x), x) | x \in A\}$,构成 A 的一对信息分配,称为关于 A 的同一-对立偶集。

将以上刻画归结为一结构函数, $\mu = a + bi + cj$,称为集对在 T_0 上的同异反联系度表达式。其中 $i \in [-1, 1], j = -1$,分别称为差异度、对立度标记系数,用于标识分类信息的方向和确定性程度。

通过上述同异反两次对立刻画,形成了一种 T_0 上的信息分类,它取自 T_0 ,为此提出了同异反集合空间的概念。

定义3 记 $T'=(U, A, A^c)$ 或 $T'=(U, B, B^c, A^c)$,称为关于 T_0 的同异反集合空间,或称为关于 T_0 的同异反生成空间。

文[8]还研究了经典集合意义下的同异反集合的一个实现,并且给出了同异反集合理论在知识发现中的一个应用。

3.3 与其它不确定性方法的关系和互补

在集对论与其它模糊性或含糊性处理方法之间的关系的研究中,目前主要讨论它与模糊集理论^[9]和粗糙集(Rough Sets,也称Rough集)理论^[10]的关系和互补。

文[11]分析了模糊集理论与集对论在研究对象、思维方法、应用范围等方面的不同之处。在研究对象方面,模糊集理论主要研究主体的内在不确定性,而对外在不确定性几乎不涉及,因而它的研究对象只是问题不确定性的一个方面;集对

论则以问题这一整体为研究对象,即视确定性与不确定性为统一的确定-不确定系统,因而它可包括各种不确定性。在思维方法方面,模糊集理论将“肯定-不定-否定”分离开;集对论则采用一种“肯定-不定-否定”相统一的思维方法,是认识论上的矛盾论与思维中的整体观的结合。此外,模糊集理论由于进行的是非对称和非独立刻画,相对于 A 的隶属度 $\mu_A(x)$,其余 A^c 集的隶属度 $1 - \mu_A(x)$ 一般未提供有价值的新信息,难以互相印证,更谈不上考虑差异性信息的影响。而且,其模糊数与模糊集合概念缺乏实质性联系,模糊隶属度的构造又因人而异,处理方法多样,难于统一,分析结果存在着可信用度问题。相反,集对论对事物的刻画较全面,提供信息完整,联系度与联系数形式统一,其中蕴含着较丰富的信息,因而它除具有与模糊数学相同的常规用途外,在应用范围上有很大的扩展。综上所述,集对论从主体与客体联系层次上较好地实现了不确定信息的整体处理,集对论的同异反集合和联系数等概念弥补了模糊集理论的缺陷。文[12,13]给出了模糊集对分析方法,该方法将确定性信息和不确定性信息作为一对模糊集,从确定性和不确定性本身进行刻画,由客体中提取相对确定性信息,同时承认和考虑对应于一种刻画的相对不确定性信息,并视它们均为有用信息。该方法集确定与不确定分析为一体,体现了不确定性对结果的影响。

文[14]在给出了Rough逻辑算子和上、下近似集的值化定义基础上,将联系度概念应用于Rough集中,提出了Rough集联系度的概念。Rough集联系度能同时给出Rough集的数字特征和拓扑特性,这对Rough集理论的逻辑研究提供了较大的方便。文[15]提出了集合型Rough集联系度的概念,以及利用Rough集联系度对决策表进行条件属性简化和属性冗余值简化的计算步骤,并通过实例说明该方法比传统的Rough集理论中的约简方法简单。

4 集对论的应用

对于含有模糊不确定、随机不确定、中介不确定,以及由不知道和信息不完全等情况引起的不确定等各种不确定性,集对论从同、异、反三个方面来研究事物的确定性与不确定性之间的联系,运用联系度(或联系数)的概念来统一处理由上述不确定因素引起的系统的不确定性问题。由于该理论是从不确定因素的全局进行研究,并不需要很明确地区分出哪一部分是属于模糊或随机或其它的不确定因素,所以这种理论具有较强的实用性。从诞生至今虽然只有十几年的时间,集对论已经被较为成功地应用于许多领域。

4.1 集对论在决策中的应用^[16~19]

彭飞等^[16]将集对分析引入模糊、灰色物元空间(FHW)决策支持系统中,建立了基于集对分析和模糊、灰色物元空间决策支持系统的方案评价决策方法。

假设有 $m(m > 0)$ 个专家(e_1, e_2, \dots, e_m)对一个含有 n 个指标(h_1, h_2, \dots, h_n)的方案 H 进行评价(假设专家权重分别为 w_1, w_2, \dots, w_m)。

若第 k 个专家 e_k 对方案 H 的 n 个指标有 s_k 个认为“好”, f_{1k} 个认为“较好”, f_{2k} 个认为“一般”, f_{3k} 个认为“较差”, $p_k = n - s_k - f_{1k} - f_{2k} - f_{3k}$ 个认为“差”,则第 k 个专家对方案 H 的同异反决策联系度为:

$$\mu(k-H) = \frac{s_k}{n} + \frac{f_{1k}}{n}i_1 + \frac{f_{2k}}{n}i_2 + \frac{f_{3k}}{n}i_3 + \frac{p_k}{n}j$$

简记为 $\mu(k-H) = a_k + b_{1k}i_1 + b_{2k}i_2 + b_{3k}i_3 + c_{kj}$ 。

从而 m 个专家对方案 H 的评价就形成了一个方案评价决策矩阵。进一步考虑专家权重,可用加权平均法求解总的同一度 a ,总的差异度 b_1, b_2, b_3 ,总的对立度 c ,从而得到其同异反联系度表达式为:

$$\mu(k-H)_{\Sigma} = a + b_1 i_1 + b_2 i_2 + b_3 i_3 + c_j$$

最后,用“过半决策准则”或其它常用方法进行评价^[16]。

实际应用表明,集对分析可以把专家对方案所做评价的同、异、反各方面综合考虑,并将之量化,使不确定性也能够定量地表达;辅之以权重分析系统,更可以精确表示各方面因素;最终根据方案评价决策矩阵和有关评价准则来确定方案的优劣。

程启月等^[17]给出了集对贴近度(ρ)、集对摆动度(γ)、集对尖锐度(κ)的概念及计算方法以及更有实际意义的联系度公式,给出了同异反决策模型,并应用此决策模型对“射击准备方案”进行定性定量分析,通过给出的理想方案的测度计算各子系统的联系度、系统的联系度,由此排序,达到择优“射击准备方案”的目的。

王坚强^[18]在分析模糊数学和灰色关联分析方法评价多指标、多水平问题的不足基础上,针对多指标、多水平系统决策特点,提出了一种新的基于集对分析的多指标多水平决策模型,同时将该模型成功应用于经济决策问题。此外,黄良骥^[19]则研究了集对分析在交通指挥控制决策支持系统中的应用。

4.2 集对论在预测中的应用^[20]

高洁等^[20]提出了集对分析聚类预测法,该方法融合了集对分析中的同异反模式识别的“择近原则”和聚类分析的基本思路。具体步骤为(设待预测的事物为 N ,相应的待预测系统为 B):(1) 确定事物 N 的分类模式系统;(2) 建立描述事物 N 的分类模式系统与参照系统的同异反联系向量;(3) 建立描述事物 N 的待预测系统 B 与参照系统的同异反联系向量;(4) 计算同异反距离;(5) 确定待预测系统 B 所属的类别。文^[21]将该方法应用于邮电业务总量预测的研究。研究中考虑了邮电业务总量和第一、二、三产业的国内生产总值之间的关系,利用邮电业务总量和三个产业的生产总值的历史数据,建立了邮电业务量水平聚类预测的模型,从而得到邮电业务总量预测结果。利用我国某地区的实际数据进行分析计算,并与其它方法比较,结果表明该预测方法是行之有效的。

4.3 集对论在数据融合中的应用^[23]

雷达与电子支援措施(Electronic Support Measurement, ESM)的数据关联是数据融合领域中的主要研究内容之一,具有重要的军事应用价值。由于雷达可以测量目标的完整位置信息(包括距离和方位),而 ESM 通常采用被动式传感器,只能提供关于目标的属性信息和位置中的方位信息,不能提供距离信息,再由于在现代电子战环境中存在干扰及其他因素的影响,使得在雷达与 ESM 的数据关联中面临着很大的不确定性。文^[21]给出了基于统计理论的雷达-ESM 数据关联算法,该算法的缺点是累积概率的计算时间效率低。文^[22]给出了基于模糊集理论的雷达-ESM 数据关联算法,该算法的缺点是阈值的选取比较困难。文^[23]基于集对论和多假设检验(Multiple Hypothesis Testing, MHT)理论研究了各雷达目标航迹样本容量不相等的情况下利用方位信息进行雷达与 ESM 航迹关联的问题,给出了新的雷达-ESM 数据关联算法。其核心思想是建立雷达航迹和 ESM 航迹之间的同异反联系度,并以此来测量雷达航迹和 ESM 航迹的关联程

度。该算法克服了时间效率低和阈值选取困难的缺点。仿真结果表明该算法特别适用于大数目雷达和 ESM 航迹样本的情况。

4.4 集对论在不确定性推理中的应用^[24~26]

文^[24]提出一种把确定性推理与不确定性推理结合起来的推理方法——同异反推理法。文^[25]把关于联系度 μ 的乘法运算以及关于 i 的取值与同异反推理结合起来,给出同异反推理的一种定量推理模式,同时指出:同异反定量推理中选择合适的推理模式相当重要,它直接影响推理精度。文^[26]则把这一问题归结为“推理路径”的选择,给出了路径选择的一般规则,并用推理实例表明了这一思路的正确性。文^[26]同时指出存在推理的类同性、由样本推理推断总体特征等需要进一步研究的问题。

4.5 集对论在产品中的应用^[27,28]

李志辉等^[27,28]提出了一种基于 CBR 的同异反产品设计方法。该方法的核心在于设计过程中领域专家的经验知识和人类创新思维的信息交合。这种信息交合是以人类专家为主体,在相应的数学基础上建立系统模型而得到的。该方法的基本思路是:从现代设计方法论的基本问题入手,围绕三个统一(技术与艺术的统一、功能与形式的统一、微观与宏观的统一)和以人为核心的设计价值观,利用 SPA 支撑的 CBR 系统产生参考案例;然后对其展开同异反分析,在定性描述的基础上用联系度 μ 定量刻画;进而建立同异反决策矩阵评价方案的优劣,最终形成产品设计方案。

已发表的关于集对论应用的论文已超过 400 篇。集对论的应用领域还包括:网络计划^[29~34]、综合评价^[35,36]、农业^[37]、林业^[38,39]、教育^[40,41]、体育^[42]、经济^[43]、环境^[44]、科研管理^[45]、知识创新^[46]、军事^[17]、交通^[19]、电力系统^[35]、聚类分析^[47]和其他领域^[48~50]。此外,有人提出了针对小数据集的基于 SPA 的知识发现算法^[8],但由于现实中的数据库越来越大,因此有必要研究面向海量数据库的基于 SPA 的知识发现算法。可能的解决方法有采样、并行化等。

结束语 虽然集对论诞生至今只有十几年的发展历史,但其研究和应用已经取得了不少成果,表明它是一种较有前途的软计算方法,为处理不确定性信息提供了较有力的分析手段。同时应该指出,目前其基础理论还较薄弱,尚未形成非常完整的理论和方法体系。因此,如何充分利用相关的哲学思想、数学理论等,并借鉴其它理论方法,深入研究集对论及其应用是当前迫切需要解决的关键问题之一。

其次,目前还没有从语言、语法、语义的角度深入研究集对论中的同异反量化计算问题。利用相关的模型(如 Chomsky 分层框架模型等)深入研究集对论将是十分有前途的。

再次是参数 i 的定义及选取方法。将其限定在 $[-1,1]$ 之间过于简单、粗糙且没有充分揭示客观事物或系统的演化规律,应该考虑它随时间演变的特性,以及在模糊、含糊等情况下的取值等问题。

此外,虽然已有学者对集对论与模糊集理论、粗糙集理论的关系和互补做过一些研究^[11~15],但研究还不够深入,有必要对集对论与模糊集理论、粗糙集理论、Dempster-Shafer(D-S)证据理论^[51~53]等不确定性理论的关系做进一步深入研究。特别是在决策方面的应用上,集对论与 D-S 证据理论的对比值得深入研究。

最后,由于到目前为止该理论及其应用的研究大多是用中文发表的,故该研究目前仍未引起国际学术界的重视。因

此,有必要努力把集对论这一门不确定性理论介绍给各国学者,并使之发扬光大。

参 考 文 献

- 赵克勤. 集对分析对不确定性的描述和处理[J]. 信息与控制, 1995, 24(3): 162~166
- JIANG Yunliang, XU Congfu, LIU Yong, et al. A New Approach for Representing and Processing Uncertainty Knowledge. In: The 2003 IEEE Intl. Conf. on Information Reuse and Integration, Oct. 2003. 466~470
- 赵克勤. 集对与集对分析——一个新的概念和一种新的系统分析方法[A]. 见: 全国系统理论与区域规划研讨会论文集[C], 1989. 87~91
- 赵克勤. 集对分析及其初步应用[M]. 杭州: 浙江科技出版社, 2000
- 赵克勤, 曹鸿兴. 集对分析与界壳论的研究与应用[M]. 北京: 气象出版社, 2001
- 赵克勤. 集对分析与熵的研究[J]. 浙江大学学报(社科版), 1992, 6(2): 65~72
- 张鹏, 王光远. 新集对论[J]. 哈尔滨建筑大学学报, 2000, 33(3): 1~5
- Zadeh L. A. Fuzzy sets [J]. Information and Control, 1965, 8: 338~353
- Pawlak Z, et al. Rough sets [J]. Communications of ACM, 1995, 38(11): 89~95
- 张斌. 不确定性信息处理的集对论思想与方法[J]. 模糊系统与数学, 2001, 15(2): 89~93
- 张斌, 赵秀梅. 处理不确定性问题的模糊集对分析方法[J]. 电子科技大学学报, 1997(增刊): 630~634
- 张斌. 多目标系统决策的模糊集对分析方法[J]. 系统工程理论与实践, 1997, 17(12): 108~114
- 张平, 黄德才. 基于联系度的 Rough 集[J]. 杭州电子工业学院学报, 2001, 21(1): 50~54
- 张平, 黄德才. 基于 Rough 集联系度的决策表简化方法[J]. 浙江工业大学学报, 2002, 30(1): 5~8
- 彭飞, 胡光正. SPA 及其在模糊、灰色物元空间决策支持系统中的应用[J]. 系统工程理论与实践, 1998, 18(11): 103~106
- 程启月, 邱苑华. 射击准备策略的决策分析[J]. 系统工程理论与实践, 2001, 21(11): 134~137
- 王坚强. 一种新的多指标多水平决策方法及应用[J]. 系统工程与电子技术, 2001, 23(11): 16~17, 61
- 黄良骥. 交通指挥控制决策支持系统[J]. 系统工程理论与实践, 2001, 21(1): 120~124
- 高洁, 盛昭瀚. 集对分析聚类预测法及其应用[J]. 系统工程学报, 2002, 17(5): 458~462
- Trunk G V, Wilson J D. Association of DF Bearing Measurements with Radar Tracks [J]. IEEE Trans on Aerospace and Electronic Systems, 1995, 23(4)
- 王国宏. 雷达与 ESM 模糊关联算法[J]. 电子对抗, 1995(4)
- WANG Guohong, HE You, SHEN Ning, et al. A New Radar-ESM Correlation Algorithm in Multisensor Data Fusion. In: CIE 1996 Inter Radar Conf. 1996. 747~750
- 赵克勤. 试论集对分析在人工智能中的应用[A]. 见: CAAI-7 全国人工智能会议论文集[C]. 西安: 西北工业大学出版社, 1992. 379~381
- 蒋云良, 张裔智, 潘云鹤, 等. 基于集对分析的同异反定量推理初探[J]. 计算机工程, 1996, 22(3): 286~290
- 蒋云良. 基于 SPA 联系系数的同异反定量推理路径选择[J]. 微电子学与计算机, 2000, 17(5): 6~9
- 李志辉, 查建中, 张英. 同异反产品设计方法[J]. 工程设计学报, 2003, 10(1): 23~27
- 李志辉, 夏少云, 查建中. 基于案例推理的同异反产品设计及其应用[J]. 计算机辅助设计与图形学学报, 2003, 15(11): 1397~1403
- 黄德才, 赵克勤. 用联系系数描述和处理网络计划中的不确定性[J]. 系统工程学报, 1999, 14(2): 112~117
- 赵克勤, 黄德才, 陆耀忠. 基于 $a+bi+cj$ 型联系系数的网络计划方法初探[J]. 系统工程与电子技术, 2000, 22(2): 29~31
- 赵克勤, 黄德才, 朱艺华, 等. 含有突发性的网络关键路线问题[J]. 管理工程学报, 2000, 14(2): 33~34
- 黄德才, 赵克勤, 陆耀忠. 含有突发事件的网络计划关键路线分类与应用[J]. 系统工程学报, 2001, 16(3): 161~166
- 黄德才, 赵克勤, 陆耀忠. 同异反网络计划的工期预测方法[J]. 系统工程与电子技术, 2001, 23(5): 24~27
- 赵克勤, 黄德才, 陆耀忠. 同异反网络计划的不确定性分类与分析[J]. 系统工程与电子技术, 2000, 22(11): 72~74
- 谢敬东, 王磊, 唐国庆. 集对分析评价法及其在电网规划中的应用[J]. 电网技术, 2002, 2(11): 37~40
- 金英伟, 迟忠先, 李艳红, 等. 基于联系系数系统态势排序综合评判方法及应用[J]. 大连理工大学学报, 2002, 42(6): 759~764
- 周泽, 赵思毅, 简贵儒, 等. 基于 SPA 的数据结构同异反分析在农业上的应用[J]. 农业系统科学与综合研究, 2002, 18(2): 139~145
- 郭建钢, 景芸, 章若鸿, 等. 伐区采育结合作业优化模式的选择[J]. 应用生态学报, 2000, 11(6): 814~818
- 周新年, 沈宝贵, 游明兴, 等. 伐区采集作业综合效益评价[J]. 山地学报, 2002, 20(3): 331~337
- 蒋云良. 基于同异反态势排序的学生成绩分析[J]. 数理统计与管理, 2001, 20(1): 26~29
- 蒋云良. 基于 SPA 的学生成绩态势排序及分析[J]. 数理统计与管理, 2003, 22(6): 5~8
- 徐金尧, 张林凤. 同异反态势排序在体育科学研究中的应用[J]. 中国体育科技, 1999, 35(4): 85~88
- 贺仲雄, 魏小涛. 模糊可拓经济控制[J]. 北方交通大学学报, 1996, 20(6): 657~661
- 李祚冰. 大气环境质量综合评价的集对分析方法[J]. 环境科学研究, 1998, 11(2): 31~33
- 张新华. 基于 SPA 的同异反科研管理理论及应用[J]. 科研管理, 1998, 19(1): 8~11
- 徐忆琳. 用 SPA 同异反系统理论研究知识创新规律[J]. 科学学研究, 2002, 20(3): 327~329
- 汪光先, 陈华豪. 用集对分析研究不同聚类解之间的相似度[J]. 东南大学学报, 1997, 16(6A): 56~59
- 王洪, 魏小涛. 协商支持系统中的冲突协调方法[J]. 系统工程理论与实践, 1998, 18(4): 87~89
- 高淑萍, 刘三阳. 基于联系系数的多资源应急系统调度问题[J]. 系统工程理论与实践, 2003, 23(6): 113~115, 122
- 刘长虹, 陈虬. 集对分析理论在结构强度分析中的应用[J]. 机械设计, 2003, 20(3): 8~10
- Shafer G. A Mathematical Theory of Evidence. Princeton: Princeton University Press, 1976
- 徐从富, 耿卫东, 潘云鹤. Dempster-Shafer 证据推理方法理论与应用的综述[J]. 模式识别与人工智能, 1999, 12(4): 424~430.
- 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