Vague集理论研究及其应用

余建坤 高世健 张文彬 著

科学出版社

Vague 集理论研究及其应用

余建坤　高世健　张文彬　著

科学出版社
北　京

内 容 简 介

本书在介绍Vague集理论的基础上，对云南财经大学信息学院课题组近几年的研究成果进行了汇集，研究 Vague 集向 Fuzzy 集转换的方法、Vague集的相似度、Vague集的模糊熵、基于Vague集的聚类、基于Vague集的关联规则算法和基于Vague集的综合评估方法。全书基于Vague集理论在数据挖掘方面应用的探讨颇具特色，较好地拓展了Vague集的应用领域，为研究Vague集理论和应用的学者提供了有益参考。

本书可以作为高等院校计算机科学与技术、管理科学与工程等相关专业高年级本科生和研究生的教材，也可作为相关研究人员和管理人员的参考书。

图书在版编目(CIP)数据

Vague 集理论研究及其应用/余建坤，高世健，张文彬著. —北京：科学出版社，2017.6

ISBN 978-7-03-052077-7

Ⅰ. ①V… Ⅱ. ①余… ②高… ③张… Ⅲ. ①数据采集 Ⅳ. ①TP274

中国版本图书馆 CIP 数据核字(2017)第 051292 号

责任编辑：于海云/责任校对：桂伟利

责任印制：张 伟/封面设计：迷底书装

科学出版社出版

北京东黄城根北街16号

邮政编码：100717

http://www.sciencep.com

北京东华虎彩印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2017年6月第 一 版 开本：787×1092 1/16

2017年6月第一次印刷 印张：9

字数：214 000

定价：78.00 元

前　言

在现实世界中，存在着大量不确定、不完全和模糊的信息，如何精确描述这些信息是科学研究中很重要的问题。当前，处理模糊信息的方法主要是建立在 Zadeh 提出的 Fuzzy 集的基础之上，然而 Fuzzy 集的理论和方法并不能很好地表示支持、反对和弃权类似的投票模型。

台湾学者 W.L. Gau 和 D.J. Buehrer 提出了 Fuzzy 集的扩展版——Vague 集，它具有比 Fuzzy 集更强的描述不确定信息的能力。Vague 集的提出促进了模糊现象的研究进程，为信息处理和决策分析提供了一种新的、有力的工具。然而，Vague 集从被提出来到现在不过三十多年，其基础理论还不够完善，需要我们进一步研究并开发其广阔的应用领域。

本书在介绍 Vague 集理论的基础上，对云南财经大学信息学院课题组近几年的研究成果进行了汇集，研究了 Vague 集向 Fuzzy 集转换的方法、Vague 集的相似度、Vague 集的模糊熵、基于 Vague 集的聚类、基于 Vague 集的关联规则算法和基于 Vague 集的综合评估方法。全书基于 Vague 集理论在数据挖掘方面应用的探讨颇具特色，较好地拓展了 Vague 集的应用领域，为研究 Vague 集理论和应用的学者提供了有益参考。

在云南省自然科学基金、云南省高校商务智能科技创新团队基金的支持下，云南财经大学信息学院的马冯副教授和几届研究生(陈磊、杨力军、沈小虎、孙宁宁、周孟、张琪)为此付出了辛勤的劳动，其成果是整个研究团队合作的结果。此外，在撰写过程中也参考了大量的国内外文献和研究成果。在此对相关作者和研究人员表示感谢！

由于作者水平有限，书中难免存在纰漏，敬请广大读者批评指正。

作　者

2015 年 10 月于昆明

目　录

第1章　绪　　论

本章介绍 Vague 集产生的背景，简述模糊数据挖掘、模糊决策的概念以及 Vague 集的国内外研究现状。

1.1　Vague 集产生的背景

德国数学家 Cantor 于 1874 年提出了经典集合论的概念，引起了数学界的极大关注。集合概念大大扩充了数学的研究领域，为数学结构提供了一个基础。集合论不仅影响了现代数学，而且也深深影响了现代哲学和逻辑。

20 世纪初的数理逻辑把逻辑的基础归于集合论。经典的逻辑是二值的，即非真即假。传统的二值逻辑在计算机科学中得到了很好的应用，经典集合论为计算机科学发展奠定了基础，如在程序语言、数据结构、编译原理、数据库与知识库、形式语言和人工智能等领域中都得到了广泛的应用。然而，计算机如果只能处理确定的二值逻辑，那么它所表达出来的行为必然是呆板的。例如，让计算机识别一个人的相貌，计算机会记住此人身体上的每一个细节，但下次此人剪了头发再让机器来识别，对不起，机器就会翻脸不认人了。还有在一些控制问题中，由于对系统的运行机理不清而难以定量描述，很难利用经典的方法建立模型。在经典集合论中，对于论域中的任何一个元素，它要么在某个集合中，要么不在某个集合中，即一个元素是否属于某个集合的特征函数只能取 0 或 1，因此它只能处理清楚的、完全的、确定性的信息。然而，当人们在工作中必须对某些事物进行决策时，往往会根据所收集到的信息或数据而作出决定，这些信息或数据大部分都是不确定的或模糊的。生活中也存在着很多模糊不清的概念，如“今天天气很热”“小伙子很帅”“秃子”“年老”“长线段”“高个子”等概念，这些概念无法用精确的数学语言来描述，而在机器模拟人的行为过程中，机器却必须能够处理这些模糊信息。此时，如果再用经典的集合论处理此类问题，就显得无能为力了。

于是，在经典集合论的基础上，Zadeh 教授等人于 1965 年提出了 Fuzzy 集理论[1]。Fuzzy 集理论突破了传统的经典集合论的限制，元素对集合的隶属度从只能取 0 或 1 扩展到可以取区间[0,1]之间的任意一个数，因此一些模糊问题便可以通过 Fuzzy 集理论来处理。Fuzzy 集在模糊控制、模糊专家系统、模糊决策等方面得到了广泛应用，在处理不确定数据或信息方面起到了较好的作用。例如刚才提到的人脸识别，当某人与机器见面时，计算机先建立一个模糊集合，下次见面时计算机又建立一个模糊集合，然后将两个模糊集合进行比较，从而对此人进行识别。这样的识别率显然高于前者，更符合实际。之后，一些学者对投票模型进行了研究，发现 Fuzzy 集理论不能很好地处理类似投票问题，它只考虑了赞成部分，而没有考虑弃权和反对部分。以投票模型为例，假设有 10 人参与投票，其中 7 人同意、1 人反对、2 人弃权，则 $t_V(x)=0.7$，其余为反对，显然不符

合实际情况。为了弥补这一缺陷，Gau 和 Buehrer 于 1993 年提出了 Vague 集理论[2]。即令 U 是一个点空间，其中的任意一个元素用 x 表示，U 中的一个 Vague 集 V 用一个真隶属函数 $t_V(x)$ 和一个假隶属函数 $f_V(x)$ 来表示，$t_V(x)$ 是从支持 x 的证据所导出的 x 肯定隶属度下界，$f_V(x)$ 则是从反对 x 的证据所导出的 x 否定隶属度下界。$t_V(x)$ 和 $f_V(x)$ 将区间 [0,1] 中的一个实数与 U 中的每一个点联系起来，即 $t_V:U\to[0,1], f_V:U\to[0,1]$。其中，$t_V(x)+f_V(x)\leqslant 1$。可见它将 x 的隶属度限制在 [0,1] 的一个子区间 $[t_V,1-f_V]$ 内。以投票模型为例，假设有 10 人参与投票，其中 7 人同意、1 人反对、其余 2 人弃权，则 $t_V(x)=0.7$，$f_V(x)=0.1$，其 Vague 值为[0.7, 0.9]。Vague 集是 Fuzzy 集的推广，当 $t_V(x)$ 等于 $1-f_V(x)$ 时，Vague 集就还原至 Fuzzy 集；当 $t_V(x)$ 与 $1-f_V(x)$ 都同时等于 1 或 0 时，Vague 集就还原至经典集合。由此可见，Vague 集可以更好地刻画现实问题，具有更强的表达数据模糊性和不精确性的能力。这种理论有效地解释了投票模型，它不仅考虑了投票中的赞成票，也考虑了反对票和弃权票这两种情形。这一理论的创立，解决了 Fuzzy 集理论不能解决的问题，为解决模糊概念或问题提供了一种有效的工具。

许多学者对模糊方面的数据或信息进行了较长时间的研究，并取得了一定的研究成果。目前对模糊方面的数据或信息处理的研究主要有两种思路：第一种思路是通过某种方法将 Vague 集转化为 Fuzzy 集，利用 Fuzzy 集的处理方法来解决问题；第二种思路是利用 Vague 集之间的相似度量方法来解决问题。这两种方法是研究中常用的方法，同时也是模糊控制和模糊决策中的重要内容。

1.2　模糊数据挖掘

提到数据挖掘(Data Mining)的相关概念，我们不能不提到另外一个名词——数据库中的知识发现(Knowledge Discovery in Database，KDD)。在第十一届国际人工智能联合会议(1989 年 8 月召开于美国底特律)的专题讨论会上第一次提到了这个术语 KDD。如图 1.1 所示为 U.Fayyad 等给出的 KDD 处理流程，它由数据选择、数据预处理、数据约简、数据挖掘、模式解释和知识评价 5 个处理步骤组成。

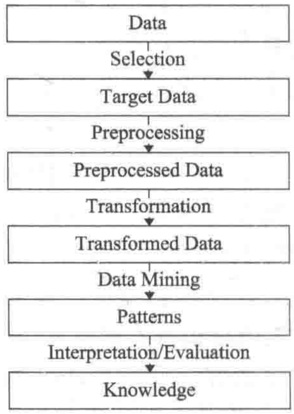

图 1.1　KDD 处理流程图

(1) Selection：在第一个步骤中，我们往往要先知道什么样的数据可以应用于 KDD 工程中。

(2) Preprocessing：当采集到数据后，下一步必须要做的事情是对数据进行预处理，尽量消除数据中存在的错误以及缺失信息。

(3) Transformation：将数据转换为数据挖掘工具所需的格式。这一步可以使得结果更加理想化。

(4) Data mining：应用数据挖掘工具。

(5) Interpretation/Evaluation：了解以及评估数据挖掘结果。

数据挖掘可以认为是知识发现过程的一个关键步骤。数据挖掘进行数据分析常用的方法主要有分类、回归分析、聚类、关联规则、特征、变化和偏差分析及 Web 页挖掘等。它是从大量的、不完全的、有噪声的、模糊的、随机的实际应用数据中，提取隐含在其中的、人们事先不知道的、但又是潜在有用的信息和知识的过程。

我们身边的自然现象可以划分为确定性现象、随机现象和模糊现象。①确定性现象，如水加温到 100℃就沸腾，这种现象的规律性靠经典数学去刻画。②随机现象，如掷骰子，观看哪一面向上，这种现象的规律性靠概率统计去刻画。③模糊现象，如“今天天气很热”“小伙子很帅”等。此话准确吗？有多大的水分？这些只有靠模糊集去刻画。

经典集合论为计算机科学与技术的发展做出了巨大贡献，但是在处理模糊信息时却遇到了困难。近年来，一类被称为软计算的智能化信息处理技术受到人们广泛的关注。软计算和传统的硬计算不同，它适合对包含有不确定性数据、不完全数据、噪声数据以及与模式可理解性有关问题的分析和建模，在不能获得或不追求问题精确解的前提下，使用软计算可以快速地获得一个问题的近似解。软计算主要包括模糊集理论、粗糙集、神经网络和遗传算法等。当数据具有模糊和不确定性时，常采用模糊集理论去处理。模糊集理论是软计算应用于数据挖掘研究中最为成熟和最成功的理论，可用于数据清理、数据选择、数据分析、建模和对数据的分析，在模式识别、智能控制、机器学习、人工智能等诸多领域有着广泛的应用。

模糊数据挖掘的步骤分为模糊数据准备、模糊数据挖掘、结果表达和知识评价[1]，如图 1.2 所示。

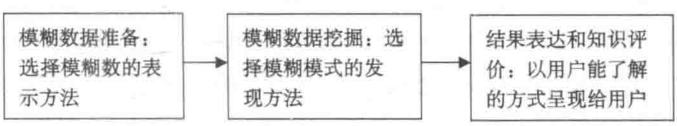

图 1.2　模糊数据挖掘的步骤

(1) 模糊数据准备：选择模糊数的表示方法，确定模糊数据的可信度及模糊模式的可信度的计算方法，检查数据的完整性，对丢失的数据可以利用统计方法进行填补。

(2) 模糊数据挖掘：选择模糊模式的发现方法，说明要发现哪一种知识及有关参数的选择，运用选定的知识发现算法(模糊聚类、模糊关联规则、模糊分类等)，从数据中提出有价值的、用户感兴趣的、带信度的知识。

(3) 结果表达和知识评价：根据用户的目的对发现的知识进行分析，并以用户能了解的方式呈现给用户，这期间也包括对知识的信度进行修正，以确保本次发现的知识的信度与以前发现的知识的信度不相抵触。

实际上，模糊集理论早已应用于数据清理、数据选择、数据分析和建模。对数据的分析至少可采用两种方法[3]：一种是用模糊论方法对传统方法加以扩展；另一种是把数据嵌入到更复杂的数学空间来进行处理，如把模糊数据映射到模糊测度空间来处理。

聚类是模糊集理论研究比较早的一个方向，在机器学习、人工智能、信息检索、统计分析、数据挖掘等领域有着大量的应用。在数据挖掘中，利用模糊聚类对有用模式进行聚类，有助于对有用模式的搜索。所谓聚类，就是把一组对象按其相互间相似程度分成若干簇，使得在同一簇中的对象相似，而在不同簇中对象相异。在传统的聚类中，每个对象被指定到唯一的簇，簇与簇之间有着明显的分界，即簇与簇之间并无重叠。但在现实应用中，大量数据的属性是不确定的或模糊的，因而很难把数据划分到唯一的簇中。在这种情况下，采用模糊逻辑和传统聚类技术结合的方法来处理，结果更符合实际。在模糊聚类中，聚类的对象可以在某种程度上同时属于多个簇，用隶属度这个概念来表示这种程度的大小。分类与聚类不同，分类属有监督的学习，对大规模数据进行分类仍然是一个具有挑战性的问题。在过去几十年里，作为研究智能系统的一个重要工具，模糊逻辑在分类问题上得到广泛的应用。关联挖掘是数据挖掘的重要内容。传统关联规则挖掘主要针对的是定性属性，然而实际生活中，存在着大量的定量数据，许多对象的属性是定量的，因而挖掘定量关联规则有着现实的意义。把模糊集理论用于挖掘具有数量属性的关联规则，人们已提出了不少模糊关联规则挖掘算法。大多数算法是基于 Apriori 算法的，有的算法主要是讨论数量属性如何用语言值表述或对其进行划分，有的是探讨在对属性进行加权的情况下规则发现的算法，此外，模糊集理论还被应用到数据选择、属性泛化、数据总结等数据挖掘任务当中。

随着数据挖掘向结构更复杂、异构、动态、分布数据挖掘方向发展，对模糊集理论提出了新的挑战。例如，为了适应大规模动态数据的挖掘，要求算法具有可扩展性，如何进行降维、如何评价挖掘性能、如何缩小数据和人的认识之间存在的语义鸿沟，以便较好地实现数据的计算机表示与人的认识间的转换等。为应对这些挑战，我们认为可以从以下 3 个方面对模糊理论作进一步的研究[3,4]：①继续探讨模糊集理论与其他传统的和非传统的技术相结合的方法，取长补短，更好地应用到数据挖掘上。在这方面，在机器学习和人工智能上有过成功的例子，如模糊集和神经网络相结合，形成模糊-神经计算模式，它一方面利用了神经网络较强的学习能力和适于信息并行处理能力，另一方面，利用模糊逻辑善于对不精确性数据建模的能力，这使得模糊神经网络模型在模糊知识表达与推理方面得到广泛应用。②随着数据挖掘向复杂结构知识挖掘方向发展，进一步促进了归纳推理与挖掘过程研究的深入，机器学习在当中扮演了十分重要的角色。而模糊集理论在学习能力方面的研究，目前还没有得到更多应有的关注，今后可以沿着这个方向进行更深入的研究。③人们对大规模数据挖掘的结果要求正确、完整、高效的同时，又要求所挖掘出来的模式简单和易于理解，但是全部满足这些特性往往是困难的，如何平衡各种性能指标是一个重要的问题，而且如何去衡量这些指标优劣，目前还没有一个公

认的标准。在模糊数据挖掘中，采用怎样的模糊理论和相关数学表达方式去研究它们之间的关系，以及与之有关的其他问题都有待深入研究。

1.3 模糊决策

所谓决策，就是对某一事物所采取的对策和策略。影响事物的因素不仅仅只有一个，所以一个事物往往需要用多个指标刻画其本质与特征[3]。决策可以认为是从若干方案中做出最佳选择，也可认为是一个提出问题、分析问题、拟定方案、选择方案、实施并修正方案的全过程。决策有不同的分类：按内容的重要程度，可分为战略决策、战术决策和执行决策；按决策时所掌握信息的完备程度，可以分为确定性决策、不确定性决策、风险决策和模糊决策。模糊决策是指在模糊环境下进行决策的数学理论和方法，即将模糊技术应用到决策过程中，使用模糊事实、模糊规则来描述决策过程中存在的不确定性和不准确性，使用模糊推理技术获得决策候选方案，使用模糊综合评判以获得最佳决策方案。在模糊决策中，决策者不能精确定义参数、概念和事件，必须处理成某种适当的模糊集合。

高层次的决策一般以决策者为核心，通过以下5个关键步骤获得最佳方案。

(1) 提出决策问题，将其概念化，并以计算机能够识别的形式表示出来。这个过程是用户同计算机交互的逐步求精的过程。

(2) 收集必要的信息。如何获得决策信息并以统一的方法表示这些信息也是非常重要的一步，因为决策环境信息是否充分、正确在很大程度上决定了决策是否正确。

(3) 为问题求解寻找或建立必要的决策模型。

(4) 通过决策模型，在所掌握情报的基础上获得若干候选方案。

(5) 通过对候选方案进行综合评估，得到最佳解决方案。

学者们提出了多准则决策方法，并结合模糊集理论，产生了模糊多准则决策方法。1994年，Chen 和 Tan[5]首先提出了一些基于 Vague 集理论的处理模糊多准则决策问题的新技术，用 Vague 值来表示可选方案的特征，使用一个分值函数 S 来评估一个方案满足决策要求的度量。随后 Hong 和 Choi[6]提出了一个新的函数，它能提供另外一种有用的方法去有效地帮助决策。后来许多中国学者在这方面也进行了深入研究，提出了不同的基于 Vague 集的模糊决策方法，并且尝试着用于解决实际问题。

1.4 Vague 集的国内外研究现状

自从 Vague 集理论于1993年提出之后，很多学者对其进行了深入研究。例如，De S K、Biswas R 和 Roy A R 在 Vague 集上定义了一些运算规则并在医疗诊断中引入了 Vague 集[7]；Atanassov K 将 Vague 集运用于模糊专家系统、模糊神经网络[8]；Li Dengfeng、Chen S M 和李凡等将 Vague 集运用于模糊决策，并且还提出了 Vague 集间的相似度的概念[9-11]；周晓光等对 Vague 集理论与应用进行了结合[12]，并在这方面取得了很多成果。

与 Fuzzy 集相比，Vague 集能够处理 Fuzzy 集所不能处理的模糊性信息，如投票模

型。随着模糊信息处理技术的进步，Vague 集将更加广泛地应用于各种领域。但从目前的研究成果来看，许多研究人员主要将注意力放在 3 个方面。①基于 Vague 集间的相似度；②基于 Vague 集的模糊决策方法；③Vague 集理论的应用研究。

1.4.1 Vague 集间的相似度量

目前，学者们研究 Vague 集相似度量的方法分为两种：一种是通过某种优势函数来计算 Vague 集间的相似度，另一种是通过距离的方法来衡量 Vague 集的相似度。对于第一种思路，Chen S M 在 1995 年首次提出了通过优势函数来度量 Vague 集间的相似度[13]，即给真、假隶属度函数分配相同的权重来计算 Vague 集的相似度。之后，他对 Vague 集间的相似度又进行了深入研究[14]，这次是给真、假隶属度函数分配不同的权重来计算 Vague 集间的加权相似度。然而这一相似度量方法是有缺陷的，为此 Hong D H 和 Kim C 于 1999 年对此进行了改进[15]。之后，刘华文进一步扩展了 Hong D H 和 Kim C 的相似度量方法[16]，并将其运用于模式识别中。

对于第二种方法，2000 年 Szmidt E 和 Kacprzyk J 依据 Vague 集的几何解释，同时考虑了 Vague 集的支持、反对和弃权 3 部分的隶属函数，提出了新的基于 Vague 集距离的相似度量方法[17]。2002 年，Li Dengfeng 等又提出了新的 Vague 集相似度量方法[18]，并将其运用于模式识别中。目前，基于距离测度的相似度量方法很多，具有代表性的研究如文献[18-22]。

1.4.2 基于 Vague 集的模糊决策方法

近些年来，有很多学者对基于 Vague 集的决策方法进行了研究。Chen S M 和 Tan J M 于 1994 年首次对基于 Vague 集理论的决策问题进行了研究[5]，通过 Vague 集的某种加权排序函数对各种方案进行了排序，以选取最佳的决策方案。之后，Hong D H 和 Choi C H 对其又进行了深入研究[24]，指出了 Chen S M 和 Tan J M 的方法的缺陷，利用新的排序函数对各种方案进行排序，以选择最佳的决策方案。同时，该方案还考虑了决策者对风险的态度问题。2001 年，李凡教授等将其改进的排序函数运用于基于 Vague 集的多目标模糊决策中[11]。2004 年，刘华文研究了基于 Vague 集的模糊条件下的多目标问题[25]，提出了新的目标选择方法。之后，林志贵等提出了新的排序函数和新的基于 Vague 集的多目标模糊决策方法[26]。黄松等于 2005 年提出了基于熵权系数和 Vague 集多目标决策的方法[27]。王珏等通过 Fuzzy 集和 Vague 集的结合[28]，提出了一种新的基于模糊多目标的决策方法。周晓光等对基于 Vague 集的模糊决策方法也进行了深入研究[29]。Szmidt E 等深入探讨了基于 Vague 集的模糊群决策的方法[30]。

1.4.3 Vague 集的应用研究

目前，Vague 集理论已经广泛应用于许多领域，如医疗诊断、模式识别、指标评价和模糊决策等。2001 年，De S K、Biswas R 和 Roy A R 将 Vague 集理论应用于医疗诊断[7]。之后，Li Dengfeng 等在模式识别中引入了 Vague 集理论[18]。2002 年，李凡教授等研究了基于 Vague 集间的相似度量的近似推理[31]。王天江等在李凡等人的基础上进一步研究

了基于 Vague 集的双向近似推理[32]。2005 年，Song 等研究了基于 Vague 集的供应商的选择方法[33]。虽然 Vague 集理论已经得到广泛的应用，但其发展时间比较短，同时又比较新颖，一些理论还不够完善，还需进一步发展。

近些年来，Vague 集理论越来越受到专家们的关注。这主要是因为人们掌握的信息往往是不精确的、不完全的或模糊的，这就要求我们通过一定的方法来解决此类信息，更好地作出判断、推理、预测、决策等。而 Vague 集就是一种很好的处理模糊信息的方法，虽然 Vague 集在许多方面取得了很好的应用效果，如在模式识别、模糊聚类、模糊决策和医疗诊断等领域的应用，但 Vague 集理论还不是很完善，需要进一步地深入研究。

1.5 小　结

为处理模糊信息，Zadeh 于 1965 年创立了 Fuzzy 集理论。其最主要的特征是对于论域 U 中的每一个元素，隶属函数给每个对象分配了一个适中的数作为它的隶属度，这个隶属度是一个单值。在随后的几十年里，Fuzzy 集理论不断发展和完善，并在数据挖掘领域里得到了应用。之后，一些学者对投票模型进行了研究，发现 Fuzzy 集理论不能很好地处理类似投票问题，它只考虑了赞成部分，而没有考虑弃权和反对部分。为了弥补这一缺陷，Gau 和 Buehrer 于 1993 年提出了 Vague 集理论，该理论与 Atanassov 给出的直觉模糊集概念相同。

模糊集理论是软计算中应用到数据挖掘研究中的最为成熟和最成功的理论，可用于数据清理、数据选择、数据分析、建模和对数据的分析，在模式识别、智能控制、机器学习、人工智能等诸多领域有着广泛的应用。

决策就是对某一事物所采取的对策和策略。影响事物的因素不仅仅只有一个，所以一个事物往往需要用多个指标刻画其本质与特征。为此，学者们提出了多准则决策方法，并结合模糊集理论，产生了模糊多准则决策方法。

Vague 集将更加广泛地应用于各种领域。但从目前的研究成果来看，许多研究人员主要将注意力放在 3 个方面：①基于 Vague 集间的相似度；②基于 Vague 集的模糊决策方法；③Vague 集理论的应用研究。

第 2 章　Vague 集理论

本章首先回顾 Fuzzy 集的定义、Fuzzy 集运算法则和性质、Fuzzy 集的截集和分解定理、Fuzzy 集间的贴近度定义及几种常见的贴近度，然后再介绍 Vague 集的运算及其性质、Vague 集的水平集、Vague 集间的距离、Vague 集间的相似度等概念。

2.1　Fuzzy 集

2.1.1　Fuzzy 集

定义 2.1[34]　设 X 是一论域，则 X 上的一个模糊集合 A 用 X 上的一个实值函数

$$\mu_A : X \to [0,1]$$
$$x \mapsto A(x)$$

来表示，对于函数值 $A(x)$ 称为 x 对于 A 的隶属度，而函数 μ_A 称为 A 的隶属函数。

Fuzzy 集 A 是一个抽象概念，集合内的元素是不确定的。换句话说，就是 Fuzzy 集 A 的边界不明确。它与经典集合既有联系，又有区别。在经典集合中，任何一个元素隶属于该集合的取值只能是 0 或 1。也就是说，经典集合具有一定的明确性。而在 Fuzzy 集合中，元素是否一定属于某个集合并不十分清楚，可以出现“亦此亦彼”的情形。

隶属度 $A(x)$ 的数值反映了论域中的元素 x 对于 Fuzzy 集合 A 的隶属程度，如果 $A(x)$ 越来越趋近于 1，则表示元素 x 隶属于 Fuzzy 集 A 的程度就越高；如果 $A(x)$ 越来越趋近于 0，则表示元素 x 隶属于 Fuzzy 集 A 的程度就越低。当 $A(x)=1$ 时，则称元素 x 完全属于 Fuzzy 集 A；当 $A(x)=0$ 时，则称元素 x 完全不属于 Fuzzy 集 A。

Fuzzy 集 A 使某元素以某种程度隶属于该集合，其隶属程度由“0”与“1”之间的某个实数来表示，而隶属度是通过隶属函数来实现的。对于某个问题，隶属函数的确定既要受到主观因素的影响，又要受到客观的制约与限制。

一个 Fuzzy 集可以用向量法、序偶法、Zadeh 表示法、解析式法等方法来表示[35]。

(1) 向量法

当论域 X 为有限个元素，可以用向量来表示 Fuzzy 集，其中各分量表示隶属度。

例 2.1　针对 6 个项目 x_1,x_2,x_3,x_4,x_5 和 x_6 进行成果鉴定，请有关专家按百分制打分，规定一些鉴定的技术指标，考核每个项目对“优秀”的符合程度。经过一些专家打分后，取其平均值，再除以 100，得到结果分别为

x_1：85 分　　$A(x_1)=0.85$；　　x_2：77 分　　$A(x_2)=0.77$；

x_3：90 分　　$A(x_3)=0.90$；　　x_4：73 分　　$A(x_4)=0.73$；

x_5：94 分　　$A(x_5)=0.94$；　　x_6：68 分　　$A(x_6)=0.68$。

这样就得到一个论域 $X=\{x_1,x_2,x_3,x_4,x_5,x_6\}$ 上的 Fuzzy 集 A，用向量法表示就是

$$A=(0.85, 0.77, 0.90, 0.73, 0.94, 0.68)$$

(2) 序偶法

序偶法是将论域中的元素 x 与其隶属度 $A(x)$ 通过有序对的方式来表示 Fuzzy 集 A，则

$$A=\{(x_1, A(x_1)), (x_2, A(x_2)), \ldots, (x_n, A(x_n))\}$$

于是，例 2.1 中的 Fuzzy 集还可以表示为

$$A=\{(x_1, 0.85), (x_2, 0.77), (x_3, 0.90), (x_4, 0.73), (x_5, 0.94), (x_6, 0.68)\}$$

(3) Zadeh 表示法

当论域 X 中为有限个元素时，即 $X=\{x_1, x_2, \ldots, x_n\}$，则论域 X 上的 Fuzzy 集 A 可表示为

$$A=A(x_1)/x_1+A(x_2)/x_2+\ldots+A(x_n)/x_n=\sum_{i=1}^{n} A(x_i)/x_i$$

其中，符号 $A(x_i)/x_i$ 并不表示“分数”，而是表示元素 x_i 隶属于 A 的程度为 $A(x_i)$，符号“+”表示一种联系符号。

当论域 X 是无限集时，则 Fuzzy 集 A 可以表示为

$$A=\int_{x\in X} A(x)/x$$

使用 Zadeh 表示法来表示例 2.1 中的 Fuzzy 集 A，则

$$A=0.85/x_1+0.77/x_2+0.90/x_3+0.73/x_4+0.94/x_5+0.68/x_6$$

(4) 解析式法

当论域是无限集时，也可以用解析式的方法来表示 Fuzzy 集 A。

例 2.2　取论域为 $X=(-\infty, +\infty)$，Fuzzy 集 A 表示“远远大于 0 的实数”，则可取为

$$A(x)=\begin{cases} 0 & x\leqslant 0 \\ \dfrac{1}{1+\dfrac{200}{x^3}} & x>0 \end{cases}$$

2.1.2　Fuzzy 集的运算与相关性质

同经典集合一样，Fuzzy 集也具有自己的运算与性质。

定义 2.2[34]　设 X 为一论域，A 和 B 是 X 上的两个 Fuzzy 集。

(1) 如果 $\forall x\in X$，有 $A(x)\leqslant B(x)$，则称 Fuzzy 集 B 包含 Fuzzy 集 A，或称 Fuzzy 集 A 包含于 Fuzzy 集 B，记为 $B\supseteq A$ 或 $A\subseteq B$，即

$$A\subseteq B \text{ 或 } B\supseteq A \Leftrightarrow A(x)\leqslant B(x) \quad \forall x\in X$$

这时，也称 A 为 B 的子集。

(2) 如果 $A\subseteq B$ 且 $A\supseteq B$，则称 Fuzzy 集 A 与 Fuzzy 集 B 相等，记为 $A=B$，即

$$A=B \Leftrightarrow A(x)=B(x) \quad \forall x\in X$$

(3) A 与 B 的并集记为 $A\cup B$，其隶属函数为

$$(A\cup B)(x)=\max\{A(x), B(x)\}=A(x)\vee B(x) \quad \forall x\in X$$

(4) A 与 B 的交集记为 $A\cap B$，其隶属函数为

$$(A\cap B)(x)=\min\{A(x),B(x)\}=A(x)\wedge B(x)\ \ \forall x\in X$$

(5) A 的补集记为 A^c，其隶属函数为

$$A^c(x)=1-A(x)$$

Fuzzy 集 A,B,C 之间的并、交、补这 3 种运算满足以下规律[36]：

(1) 幂等律

$$A\cup A=A，\ \ A\cap A=A$$

(2) 交换律

$$A\cup B=B\cup A，\ \ A\cap B=B\cap A$$

(3) 结合律

$$A\cup(B\cup C)=(A\cup B)\cup C，\ \ A\cap(B\cap C)=(A\cap B)\cap C$$

(4) 分配律

$$A\cup(B\cap C)=(A\cup B)\cap(A\cup C)，\ \ A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$$

证明：由于 $\forall x\in X$，$A(x)$、$B(x)$、$C(x)$ 之间可能存在如下 6 种关系

① $A(x)\geqslant B(x)\geqslant C(x)$，　② $A(x)\geqslant C(x)\geqslant B(x)$；

③ $B(x)\geqslant A(x)\geqslant C(x)$，　④ $B(x)\geqslant C(x)\geqslant A(x)$；

⑤ $C(x)\geqslant A(x)\geqslant B(x)$，　⑥ $C(x)\geqslant B(x)\geqslant A(x)$。

本节只针对第二种情形来证明 $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$。

对于 $\forall x\in X$

$$(A\cup(B\cap C)(x)=A(x)\vee(B(x)\wedge C(x))=(A(x)\vee B(x))\wedge(A(x)\vee C(x))$$

对于情形二，有

$$A(x)\vee(B(x)\wedge C(x))=A(x)\vee B(x)=A(x)$$

$$(A(x)\vee B(x))\wedge(A(x)\vee C(x))=A(x)\wedge A(x)=A(x)$$

即对第二种情形结论是正确的。类似地，其他情形用同样的方法，也可以证明结论是正确的。

针对第二种情形来证明 $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$。

对于 $\forall x\in X$

$$(A\cap(B\cup C)(x)=A(x)\wedge(B(x)\vee C(x))=(A(x)\wedge B(x))\vee(A(x)\wedge C(x))$$

对于情形二，有

$$A(x)\wedge(B(x)\vee C(x))=A(x)\wedge C(x)=C(x)$$

$$(A(x)\wedge B(x))\vee(A(x)\wedge C(x))=B(x)\vee C(x)=B(x)\vee C(x)$$

即对第二种情形结论是正确的。类似地，其他情形用同样的方法，也可以证明结论是正确的。

(5) 吸收律

$$A\cup(A\cap B)=A，\ \ A\cap(A\cup B)=A$$