

Granular Computing Methods and Applications
Based on Inclusion Degree

基于包含度的 粒计算方法与应用

徐伟华 米据生 吴伟志◎编 著



科学出版社

基于包含度的粒计算方法与应用

徐伟华 米据生 吴伟志 编著



科学出版社

北京

内 容 简 介

作为知识获取和数据挖掘的重要工具,粒计算是在解决大规模复杂问题时模拟人类思维问题自然模式的一种全新理论、技术和方法。全书以包含度在各种信息系统的实践应用为主线,系统深入地介绍了基于包含度的粒计算方法与应用的最新研究成果。全书共分10章,着重阐述了包含度在粗糙近似集、直觉模糊集、变精度序信息系统、区间值模糊粗糙集、结构粗糙集、概念格等方面的理论及应用。

本书可作为高等院校计算机、应用数学、信息科学专业的高年级本科生及研究生教材,也可供相关专业,特别是从事人工智能、粒计算研究的科研工作者参考阅读。

图书在版编目(CIP)数据

基于包含度的粒计算方法与应用/徐伟华,米据生,吴伟志编著. —北京:科学出版社, 2015.7

ISBN 978-7-03-045297-9

I. ①基… II. ①徐… ②米… ③吴… III. ①人工智能-计算方法 IV. ①TP18

中国版本图书馆 CIP 数据核字(2015) 第 177422 号

责任编辑: 杨 岭 王 玮 / 责任校对: 王 玮

责任印制: 余少力 / 封面设计: 墨创文化

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>

四川煤田地质制图印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2015年7月第 一 版 开本: B5(720×1000)

2015年7月第一次印刷 印张: 15 1/2

字数: 300 000

定价: 85.00 元

(如有印装质量问题, 我社负责调换)

编委会成员

顾 问：张文修

主 编：徐伟华 米据生 吴伟志

编 委：（以姓氏笔画为序）

万 青	马 丽	马江洪	马建敏
王 宁	王 虹	王 霞	仇国芳
龙卫江	安秋生	米据生	许 伟
许 超	祁建军	杨宏志	杨淑云
苏亚娟	李同军	李鸿儒	吴伟志
宋笑雪	张文修	张先韬	张红英
张晓燕	陈德刚	邵明文	周 磊
姚燕青	袁修久	钱 婷	徐 萍
徐伟华	梅长林	解 滨	潘笑晨
魏 玲	魏立力		

谨以此书献给敬爱的张文修教授！

序

不确定性是客观存在的大量现象和事物的特征。随着研究范围的扩大,研究内容的深入,不可能回避对不确定现象与实物的研究。Zadeh L A 于 1965 年提出的模糊集是对经典集合的扩充,从而刻画了“对象”的不确定性。包含度理论是对“包含关系”的“扩充”,从而包容了“关系”的不确定性。模糊集理论与包含度理论相辅相成,成为研究不确定性的重要工具。

1. 不确定性是一种客观存在

在客观实际中,不确定性处处存在。

首先,客观事物是发展变化的,它本身就是不确定的。例如,“年轻人”和“老年人”本身就是一个过渡,不可能把某些人简单地称为“年轻人”,也不可能简单地称某些人为“老年人”;同样,不可能简单地把某个女孩归为“漂亮的女孩”或“不漂亮的女孩”;也不能把体操运动员的动作简单地分为“优美”和“不优美”,把某项决策行动简单地归结为“好的决策”或“不好的决策”等。这是因为客观事物既然是发展变化的,它必然存在着过渡状态,而不能简单地划分为两部分,在这种过渡阶段就存在着不确定性。即在客观世界中存在着大量的不确定的研究对象。

另外一种不确定性即为不确定的关系。例如,“西红柿红了,那么西红柿熟了”,并未表明红到什么程度西红柿才算熟了,也未表明有点红的时候西红柿熟到什么程度;又如“前面有车,汽车要开慢一点”,并未表明汽车要慢到什么程度,也未表明前面的车速和车距多大时要求汽车慢到什么程度;“天上下雨地下湿”,既未表明下多大雨地下才湿,也未表明下雨的大小与时间长短和地下湿的明确关系;“股市大涨时要迅速购买股票”,也未明确涨幅与购买股票量的关系,等等。现实生活的经验给人们以知识,这种知识反映了两种不确定事物的关系,而这种关系也是不确定的,不确定的关系导致不确定的结果。

从人的认识过程来讲,客观的对象是无限的,而人存在的空间是有限的,在具体历史条件和具体环境条件下,人对客观的认识不可能不受历史与环境的限制。人们不可能得到客观世界的全部信息,更不可能得到今后发展变化的全部信息,所以人的知识与推断也有不确定的一面。特别在社会科学、管理科学、思维科学的某些领域,表现出更多的不确定性。

2. 经典逻辑对于不确定性研究的局限性

经典逻辑是一种“非此即彼”的二值逻辑, 要么“是”, 要么“不是”, 有着分明的界限, 从包含关系来讲, 要么“包含”, 要么“不包含”。这样, 既抹杀了事物本身发展变化的过渡状态, 也抹杀了事物之间关系的多样性与复杂性。我们精密地研究客观世界的对象与关系, 首先必须把变化的事物静止下来, 把有联系的事物割裂开来, 把无限的事物变为有限的事物, 把复杂的关系变为简单的关系, 以给出客观事物的一个分明界限。人们规定什么对象属于这个事物, 什么对象不属于这个事物, 这就是经典集合。经典集合反映了一个对象作为个体的确定性, 以及人为强制的规定性。当以确定性为主时, 人们用这样的方法加以研究, 并得出比较精确的符合实际的结果。但在不确定性表现突出的情况下, 用这样的方法进行研究的結果虽然精密, 但常常不符合实际。正如 Zadeh L A 所指出的: 系统越复杂, 人们对其精密而有意义的描述能力将相应降低, 以致“精密性”与“有意义”两者几乎达到相互排斥的地步。因此, 要想确切地描述复杂的现象和系统的任何现实的物理状态, 实际上是办不到的。为了使整个问题的描述有意义, 我们必须在精密与有意义之间寻求某种平衡, 在逻辑与实际之间寻求某种折中, 既不能不要逻辑地回到经验生活, 又不能不顾实际地去追求逻辑的严密性。回顾数学发展历史, 正是经历了这样一个过程。随着研究领域的不断扩大, 数学方法有了大幅度进展, 微分的引入使数学从常量数学发展到变量数学, 概率的引入使数学从确定性数学发展到随机性数学, 模糊集与包含度理论使数学从确定性数学发展到不确定性数学, 使数学可以描述一类更广泛的不确定现象。

3. 包含度理论及其科学意义

模糊集理论最重要的是模糊集的概念, 包含度理论最重要的是包含度的概念。所谓包含度, 是指对于偏序集 L 上的任意两个元素 a 和 b , 称为“ a 包含于 b ”的程度, 如果它满足以下公理:

公理 1 $0 \leq D(b/a) \leq 1$,

公理 2 当 $a \leq b$ 时, $D(b/a) = 1$,

公理 3 当 $a \leq b \leq c$ 时, $D(a/c) \leq D(a/b)$ 。

无论经典集的包含关系、模糊集的包含关系, 还是实数集 $\mathbf{R}$ 上的“ $\leq$ ”关系, 都是一种偏序关系。因此包含度理论可以应用到任何偏序集上。

包含度理论包容了不确定性推理的所有结果。在不确定性推理中, 最早出现的概率推理, 利用条件概率进行推理, 而条件概率恰好是一种包含度。由于在条件变量较多的情况下, 条件概率的数量大大增加, 这在实际中是难于获得的, 因此出现了概率推理的两种变形: 一种是主观 Bayes 方法; 另一种是 MYCIN 不确定因子

方法。主观 Bayes 方法给出条件成立时假设成立的强度 LS, 和条件不成立时的强度 LN, 用两个规则强度 LS 和 LN 计算条件概率, 如果用 LS 和 LN 计算其他包含度就拓广了主观 Bayes 方法。MYCIN 不确定因子是利用给定条件信任增加的程度 MB 和怀疑增加的程度 MD 给出不确定因子, 而不确定因子规范化以后也是包含度。Dempster 和 Shafer 提出证据理论以后, 广泛地应用到不确定性推理中。证据理论主要使用 mass 函数, 它是由人的经验给出的一种评价函数, 形成弱于概率测度的似然测度和信任测度。在关系数据库上利用包含度理论可以生成各种形式的信任测度和似然测度。模糊推理中的蕴涵算子是包含度, 从而由模糊规则生成的模糊关系是包含度。在信息推理中使用的条件信息量是包含度。在王国俊教授的计算逻辑学中的真值即为一种包含度。

由此可见, 包含度理论概括了已有的各种不确定性推理方法。只要有一种包含度, 就有一种不确定性推理。同时, 可以从包含度出发, 研究包含度合成、传播与修正方法。这样, 包含度理论为不确定性推理提供了统一的理论框架与一般研究方法。

专家系统中有两类问题非常重要: 一是知识库中知识的获取, 二是对于矛盾知识的排除。专家要获得知识首先要有案例, 案例由两部分组成: 一部分是条件属性的不同状态, 另一部分是假设的不同结果。案例是可以重复、不一致, 甚至是矛盾的。它是多个专家的经验, 一般量是比较大的。知识获取是在大量的案例中提炼出几条规则, 这些规则构成了知识库中的知识。用包含度理论从案例中提取知识是一种简单而成功的方法。包含度理论的另外一个成功应用是排除矛盾规则。知识库中的知识既可以用上述方法得到, 也可以是专家的经验, 无论哪种情况, 都难以保证知识库中的知识是协调的。知识库中的知识必须有一定的协调性, 否则会影响知识库的使用效果。利用包含度理论可以解决知识库的规则协调问题。

因此, 包含度理论对人工智能、专家系统、模式识别、系统分析、管理决策、经济规划都有着重要意义。

4. 包含度理论与我们的研究团队

客观实物的不确定性必然导致思维的不确定性, 因此研究思维科学与人工智能一直是我的一个梦想。为此我研究过模糊数学、集值测度与随机集、集值随机过程、遗传算法, 它们都反映了客观或主观上的某种不确定性。1993 年, 我在香港中文大学与梁怡、梁广锡教授合作时, 他们论文中对于专家系统的一些公式引起我的兴趣, 我求教于他们这些公式的来源和意义, 他们告诉我是直觉思维与仿真实验的结果。我在试图解释他们的计算公式时提出了包含度公理, 并经过反复修改, 使它能用于所有人工智能计算公式。1996 年我与梁怡教授合作出版了《不确定性推理

原理》，没想到该书在国内外产生了很好的影响。

2000 年左右，在我的学生推动下，我开始了粗糙集与概念格的研究。当时我即将从研究生院院长位置上退下来，有较多时间阅读文献和研究。于是以后的几年时间里招收了比较多的博士生，这样就形成了后来 20 多人的团队，当然还包括少量博士后和硕士生。这个团队的最大优点是对科学研究意义的充分认识，以及对于科学研究群体互动作用的充分理解。这是一个团结奋进的团队，一个讨论探索的团队，一个生动活泼的团队，一个到处闪烁着火花的团队。我置身于他们之中，再一次激活了研究的激情。我是他们的导师，但我更像他们中的一员。我们一起学习，一起讨论，一起探索，一起为一点点的成果喜悦着。在新的世纪里形成了我们这样一个独特的研究团队，他们分布于近 20 所高等院校，已经成为各个学校教学与科研的骨干。更让我感到欣慰的是，我虽然已经退休多年，但这个团队一直在活跃着，他们这么多年一直通过会议、网络等方式不断地共同探讨着新的课题，他们中有不少人取得了更加优秀的研究成果。这个团队任何人的研究成果不仅是他个人的辛勤劳动，同时凝聚着这个团队的集体智慧。

最近他们将这个团队利用包含度理论深入研究的一些最新科研成果总结成本书，我非常高兴地为本书作序，并希望他们沉下心来取得更多更优秀的成果。

张文修

2015 年 5 月于西安

前 言

随着信息技术的日新月异,一些具有海量、高维、分布式、动态等特征的大规模复杂数据不断涌现,人们迫切要去分析处理这些复杂数据,从中找到有价值的信息。然而,直接对这些数据进行处理面临着严重的计算问题和维数灾难问题。千古流传的瞎子摸象的故事就反映出在缺少视觉信息的条件下对不完全信息进行融合会产生认知的偏差。因此,如何有效、快速地处理这些复杂数据,并提取出隐含其中的、潜在有用的知识,一直是智能信息处理领域的一个研究热点。

从多视角、多层次观察、分析与解决问题,是人类智能的特点之一。作为知识获取和数据挖掘的重要工具,粒计算 (granular computing, GrC) 是在解决大规模复杂问题时模拟人类思考问题自然模式的一个新的理论、技术和方法,其基本思想是在问题求解过程中使用信息粒,从不同角度、不同层次对现实问题进行描述、推理与求解。粒计算改变了传统的计算观念,使信息处理更科学、合理和易操作,在海量数据挖掘的研究中有着独特的优势,在智能系统的设计和实现中有着重要的作用。目前粒计算的理论与方法已成为人工智能领域一个非常活跃的研究方向。

粒计算方法是以前 (granule) 为基本计算单位,以处理大规模复杂数据集和信息等建立有效的计算模型为目标。“粒计算”作为术语正式提出于 1997 年,进入新世纪以来,粒计算研究在国际上逐渐活跃起来,特别是自 2005 年在北京召开第一届粒计算国际学术会议“2005 IEEE International Conference on Granular Computing(IEEE GrC 2005)”以来,有关粒计算研究的著作和文献大量增加。最近几年,其他一些国际会议,如自 2002 年开始每年召开的“IEEE International Conference on Machine Learning and Cybernetics(ICMLC)”和自 2006 年起每年召开的“International Conference on Rough Set and Knowledge Technology (RSKT)”都将粒计算作为会议主题进行报道和交流。在国内,除了期刊论文以外,有关粒计算的研究成果都会在每年举办的“中国 Rough 集与软计算学术会议 (CRSSC)”上报道和交流,并且自 2007 年开始在我国每年召开一次全国性的学术会议——“中国粒计算学术会议 (CGrC)”。另外,近几年,每年还举办有关粒计算的专门论坛和类似于暑期学校供交流和讨论。国家科技部已将粒计算列入“863”研究的指南,粒计算也成为近几年国家自然科学基金的研究热点。目前,粒计算已经成为了人工智能领域研究中一个非常活跃的新方向。

在众多的粒计算模型中, Rough 集方法、模糊集方法、形式概念分析方法、证

据理论方法等是公认的经典的粒计算分类方法,对于粒计算研究的发展和推动起着非常重要的作用。需要特别说明的是,这些方法都是不确定推理的定量方法,是通过给出命题的数值计算,给出因果关系的数值趋势。不管是粗糙集、模糊集还是形式概念分析、证据理论,都首先需要对不确定信息进行表示和度量。不同的信息表示方法与度量方法即构成不同的不确定推理。多年来,西安交通大学的张文修教授及其团队在不确定推理的研究中发现,这几种经典的不确定推理方法的共同点是用一种测度来度量假设,这种测度可以是各种特殊的模糊测度,而不确定推理的实质是一种广义的包含关系。基于这种认识,1990年左右,张文修老师提出了包含度概念。1993年,张老师在香港中文大学与计算机科学系梁广锡博士和环境研究中心梁怡博士合作期间,进一步明确了包含度的概念,出版了《不确定性推理原理》一书,建立了包含度理论体系。该书的出版从逻辑和实践上都证明了包含度理论是对已有的各种不确定推理的概括和抽象。包含度理论不仅简洁、概括、广泛,而且便于进行信息的合成、传播和修正。特别是在各种关系数据库中有着直接的应用。包含度理论为不确定推理的理论研究提供了一般的方法,对产生新的不确定推理有着明显的指导作用。同时,像本书中某些研究结果所显示的那样,包含度理论对研究专家系统中其他重要问题,如知识的获取、规则的协调性及矛盾规则的排除,都是一种非常有效的方法。包含度理论不仅是研究不确定推理的一般理论,而且是研究不确定性现象的方法论。

基于上述出发点,全书以包含度在各种信息系统的实践应用为主线,系统深入地介绍了基于包含度的粒计算方法与应用的最新研究成果。本书是由国内从事包含度及其粒计算研究领域的20多位一线科研工作者和学者合作撰写的。全书共分10章,着重阐述了包含度在粗糙近似集、直觉模糊集、变精度序信息系统、区间值模糊粗糙集、结构粗糙集、概念格等方面的理论及应用。本书研究内容为解决复杂系统的知识发现、不确定决策分析、规则获取等问题提供了有利工具。书中汇集了当前包含度、粒计算研究团队的最新研究成果,目的是让更多的学生、老师、学者了解包含度及其粒计算方法与应用,共同促进该领域的发展,并希望能为包含度及粒计算的研究发展做出更大的贡献。

此外,感谢西安交通大学张文修教授及各位编委委员等对本书提出的宝贵意见。同时也感谢研究生余建航、胡猛、李蒙蒙、郭艳婷等对本书提出的宝贵意见,以及在整理本书中付出的辛勤劳动。本书的出版得到了国家自然科学基金项目(61105041、61472463、61402064、61170107)的资助,在此一并表示感谢。

本书注重系统性、严谨性、理论性和可读性,可以作为高等院校数学、计算机、信息科学以及系统工程等专业高年级本科生及研究生的教学用书,也可作为相关专业的科研工作者的参考辅导用书。

由于作者水平有限, 加上时间仓促, 书中难免存在不足之处, 敬请读者批评指正。

编 者

2015 年 5 月

目 录

第 1 章 包含度与粒计算的数学基础	1
1.1 模糊集合与模糊测度	1
1.2 包含度及其生成方法	9
1.3 粗糙集与信息系统	21
1.4 形式概念分析方法	28
参考文献	34
第 2 章 粗糙近似与包含度	36
2.1 经典粗糙近似与包含度	37
2.2 模糊逻辑算子	40
2.3 模糊集上的包含度	44
2.4 基于模糊逻辑算子的模糊粗糙近似与模糊信任函数	46
参考文献	54
第 3 章 直觉模糊包含度及广义模糊粗糙集上的包含度	56
3.1 直觉模糊包含度	56
3.2 直觉模糊集上基于包含度的相似度	63
3.3 广义模糊粗糙集之间的混合单调包含度	67
3.4 广义模糊粗糙集之间基于混合单调包含度的相似度	76
参考文献	79
第 4 章 基于包含度的变精度序信息系统及应用	80
4.1 优势关系下基于包含度的变精度粗糙集	80
4.2 粗糙度、依赖度和重要度	87
4.3 序信息系统变精度粗糙集的约简理论与算法	89
4.4 案例研究	95
参考文献	99
第 5 章 基于包含度的区间值模糊粗糙集理论与方法	102
5.1 模糊集的包含度	102
5.2 区间数和区间值模糊集的包含度	106
5.3 双论域上的区间值模糊粗糙集	115
5.4 区间值信息系统上基于包含度的变精度粗糙集理论与方法	130
参考文献	132

第 6 章 基于包含度的结构粗糙集近似方法	135
6.1 结构粗糙集	135
6.2 结构概率粗糙集近似	138
6.3 基于包含度的粗糙集近似	141
6.4 基于包含度的结构粗糙集近似	143
参考文献	144
第 7 章 基于包含度的概念格	146
7.1 形式背景上基于包含度的 Galois 连接	146
7.2 基于包含度的概念格的构造	149
7.3 邻域系统上的多尺度概念格	154
7.4 多尺度概念格的应用实例	157
参考文献	162
第 8 章 单边模糊概念格知识约简方法	163
8.1 单边模糊概念格	164
8.2 单边模糊概念格的格保持约简	170
8.3 单边模糊概念格的粒度约简	178
8.4 单边模糊概念格的粒度约简与其他约简之间的关系	182
参考文献	184
第 9 章 三元概念分析	187
9.1 三元概念分析的基本概念	187
9.2 三元背景的蕴涵及规则	194
9.3 三元因子分析	196
9.4 模糊三元概念分析	201
参考文献	204
第 10 章 基于模糊软集的集成预测方法及应用	207
10.1 模糊软集	207
10.2 支持向量机	209
10.3 集成预测模型	214
10.4 实证分析	217
参考文献	227
索引	229

第1章 包含度与粒计算的数学基础^①

从多视角、多层次观察、分析与解决问题,是人类智能的特点之一。作为知识获取和数据挖掘的重要工具,粒计算(Granular Computing, GrC)是在解决大规模复杂问题时模拟人类思考问题自然模式的一个新的理论、技术和方法,其基本思想是在问题求解过程中使用信息粒,从不同角度、不同层次上对现实问题进行描述、推理与求解。信息粒是指人类在认识、推理和作决策中,将大量复杂的信息按其各自的特征和性能划分成若干较简单的块、类、群或组等,而每个为此划分出来的块、类、群或组被称为一个粒。这种处理信息的过程就被称为信息粒化(information granulation)。1979年Zadeh在世界上首次提出并讨论了模糊信息粒化问题^[1],从此信息粒化的基本思想被应用到了许多领域中。在众多的粒计算模型中,Rough集方法^[2,3]、模糊集方法^[4]、概念格方法^[1]以及证据理论^[5]是公认的经典的粒计算分类方法,对于粒计算研究的发展和推动起着非常重要的作用,并得到了越来越多的学者关注^[6,7]。1985年,Hobbs提出了粒度(granularity)的概念^[8],接着Zadeh在1996-1997年期间第一次提出粒计算(granular computing)的概念^[9,10],此时粒计算在软计算、知识发现、数据挖掘中扮演的角色越来越重要,而且取得了很好的成果^[11-17]。

1.1 模糊集合与模糊测度

在人工智能中遇到大量的经典集合。例如,在专家诊断系统中,有专家的集合、病例的集合、症状的集合;在汉字识别系统中,有汉字的集合、笔画的集合、拼音字母的集合等。我们把在一个实际问题中考虑的对象全体称为一个基本集合或论域,通常以 X 记之。像26个英文字母集合,平面上所有点的集合等都是论域。论域 X 中的一部分称为 X 的子集合,通常以 A, B, C 表示。 X 中的对象称为元素,以小写字母 x, y 记之。如果 x 属于 A ,记为 $x \in A$;若 x 不属于 A ,记为 $x \notin A$ 。以 $\emptyset$ 表示空集合,以 X 表示全集合。任给性质 $P_1, P_2, \dots, P_m$,具有性质 $P_1, P_2, \dots, P_m$ 的全体 x 构成一个经典集合:

$$A = \{x \mid P_1(x), P_2(x), \dots, P_m(x)\}$$

^①本章由张文修(西安交通大学)和徐伟华(重庆理工大学)执笔。本章工作获得国家自然科学基金项目(项目编号:61105041, 61472463, 61402064)资助。

对于 X 中的两个子集 A 和 B , 当 $x \in A$ 时必有 $x \in B$, 称 A 含于 B , 或 B 包含 A , 记为 $A \subseteq B$ 。当 $A \subseteq B$ 且 $B \subseteq A$ 同时成立, 称 A 等于 B , 并记为 $A = B$ 。

设 X 为论域, 记

$$\mathcal{P}(X) = \{A \mid A \subseteq X\}$$

称 $\mathcal{P}(X)$ 为 X 的幂集。如果 X 中有 n 个元素, 则 $\mathcal{P}(X)$ 中有 2^n 个元素。对于 X 的两个子集 A 和 B 可定义运算:

$$A \cup B = \{x \mid x \in A \text{ 或 } x \in B\}$$

$$A \cap B = \{x \mid x \in A \text{ 且 } x \in B\}$$

$$A^c = \{x \mid x \notin A\}$$

称 $A \cup B$ 与 $A \cap B$ 分别为 A 与 B 的并集合与交集集合, A^c 称为 A 的补集合。如果定义 $A \setminus B = A \cap B^c$ 为 A 与 B 的差集合, 则 $A^c = X \setminus A$ 。

设 $D \subseteq \mathcal{P}(X)$, 称 D 为集代数, 若满足以下条件:

- (1) $X \in D$;
- (2) 若 $X \in D, Y \in D$, 则 $X \cup Y \in D$;
- (3) $X \in D$ 时 $X^c \in D$ 。

对于 X 中的子集可以引进特征函数:

$$A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

则特征函数完全刻画了经典集合。特征函数是从 X 到 $\{0, 1\}$ 的映射, X 中的一个经典子集合对应着 X 上的一个特征函数, X 上的一个特征函数完全确定了 X 中的一个经典子集合

$$A = \{x \mid A(x) = 1\}$$

从特征函数的角度来看, 经典集合是一个分明集合, 它对应着二值逻辑。从集合的角度来讲, 一个元素或属于这个集合, 或不属于这个集合, 二者必居其一。从逻辑的角度来讲, 一个命题或真或假, 不可能是又真又假。经典集合反映了一种“非此即彼”的分明逻辑。

从实际问题来讲, 二值逻辑并不能完全反映实际情况。例如, “张三 是年轻人, 李四是老年人” 就不能反映在二值逻辑中。张三目前是 30 岁, 是否为年轻人? 李四目前是 60 岁, 是否为老年人? 如果张三是年轻人, 那么再过一年是否还是年轻人? 如果仍为年轻人, 一直下去就可能导致李四也是年轻人的荒唐结论。在这里“年轻人”和“非年轻人”、“老年人”和“非老年人”之间没有明确的界限, 在一定意义下它是一种过渡状态。为了描述这种不分明的状态, 我们需要扩充特征函

数为隶属函数。所谓 X 上的隶属函数是指 X 到 $[0, 1]$ 的映射, 它代表的是一个模糊集^[18]。

例 1.1 设 $X = [0, 100]$ 表示年龄的某个集合, $\tilde{Y}$ 及 $\tilde{O}$ 表示“年轻”与“年老”。它的隶属函数分别为

$$\tilde{Y}(x) = \begin{cases} 1, & 0 \leq x \leq 25 \\ \left[1 + \left(\frac{x-25}{5}\right)^2\right]^{-1}, & 25 < x \leq 100 \end{cases}$$

$$\tilde{O}(x) = \begin{cases} 0, & 0 < x \leq 50 \\ \left[1 + \left(\frac{x-50}{5}\right)^{-2}\right]^{-1}, & 50 < x \leq 100 \end{cases}$$

对于模糊集合 $\tilde{A}$, 必须用隶属函数 $\tilde{A}(x)$ 描述。 $\tilde{A}(x)$ 表示 x 隶属于模糊集合 $\tilde{A}$ 的程度。若 $\tilde{A}(x_1) = 0.8$, $\tilde{A}(x_2) = 0.6$, 那么 x_1 相对 x_2 更属于 $\tilde{A}$ 。一般地, 一个模糊集可以表示为图 1.1。

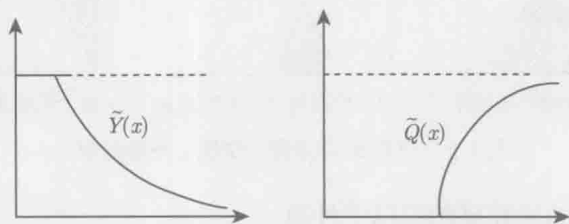


图 1.1 “年轻”与“年老”的隶属函数

对于一个模糊集, 可以表示为

$$\tilde{A} = \{(x, \tilde{A}(x)) \mid x \in X\}$$

如果 X 是有限集合或可数集合, 可表示为

$$\tilde{A} = \Sigma \tilde{A}(x_i)/x_i$$

如果 X 是无限不可数集合, 可表示为

$$\tilde{A} = \int \tilde{A}(x)/x$$

用 $\mathcal{F}(X)$ 表示 X 上的模糊集合的全体, 即

$$\mathcal{F}(X) = \{\tilde{A} \mid \tilde{A}: X \rightarrow [0, 1]\}$$