浙商院文库

不完备信息系统知识发现和规则提取的粗糙集方法研究

RESEARCH ON ROUGH SET APPROACHES OF KNOWLEDGE DISCOVERY AND RULE EXTRACTING IN INCOMPLETE INFORMATION SYSTEM

魏利华　著

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

不完备信息系统知识发现和规则提取的粗糙集方法研究

魏利华　著

阜阳师范学院图书馆
★
藏书

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

内容简介

粗糙集理论是处理不确定性问题的新型数学工具。该理论不需要太多先验知识，只需要解决问题所需的数据集，与概率理论、模糊集理论、证据理论不同，因这些理论还需要其他的先验知识。本著作主要研究拓展粗糙集模型，探讨这些模型的性质，深入探索基于这些粗糙集模型的约简算法和规则提取问题，主要研究以下方面：粗糙集的历史及模型、基于差异关系的不完备信息系统粗糙集方法、不完备区间值信息系统的知识获取、不完备模糊信息系统的粗糙集方法与知识约简、不完备信息系统的扩展优势粗糙集和对未来工作的一些展望。

版权专有　侵权必究

图书在版编目（CIP）数据

不完备信息系统知识发现和规则提取的粗糙集方法研究/魏利华著．—北京：北京理工大学出版社，2015.3

ISBN 978－7－5682－0023－3

Ⅰ.①不…　Ⅱ.①魏…　Ⅲ.①集论－研究　Ⅳ.①O144

中国版本图书馆 CIP 数据核字（2014）第 298267 号

出版发行／北京理工大学出版社有限责任公司
社　　址／北京市海淀区中关村南大街 5 号
邮　　编／100081
电　　话／（010）68914775（总编室）
　　　　　82562903（教材售后服务热线）
　　　　　68948351（其他图书服务热线）
网　　址／http：//www.bitpress.com.cn
经　　销／全国各地新华书店
印　　刷／保定市中画美凯印刷有限公司
开　　本／787 毫米×1092 毫米　1/16
印　　张／7.5
字　　数／175 千字
版　　次／2015 年 3 月第 1 版　2015 年 3 月第 1 次印刷
定　　价／39.00 元

责任编辑／封　雪
责任校对／孟祥敬
责任印制／李志强

图书出现印装质量问题，请拨打售后服务热线，本社负责调换

前　言

处理不确定性问题的新型数学工具粗糙集理论是由波兰数学家华沙理工大学教授 Z. Pawlak 于 1982 年提出的。该理论只需要解决问题所需的数据集，不需要太多先验知识，与概率理论、模糊集理论、证据理论不同，因这些理论还需要其他的先验知识。当今对粗糙集的研究主要集中在模型的推广、知识获取与约简、规则提取以及不确定度量研究。本书主要研究拓展粗糙集模型，探讨这些模型的性质，深入探索基于这些粗糙集模型的约简算法和规则提取问题，其主要研究内容和创新点包含以下方面：

（1）在不完备目标信息系统中，建立了基于差异关系的粗糙集模型，相对于等价关系来说，差异关系是从另一个角度来描述对象。同时，探讨了上、下近似算子的相关重要性质和该系统中知识的不确定度量问题。为了提取简化的规则，研究了相对上近似约简和相对下近似约简的相关问题。

（2）把序邻域关系引入了不完备区间值信息系统，建立了新的粗糙集模型，同时对下、上近似算子的性质进行了详细的证明。尽管基于不同粗细的关系，得到的粗糙集合的精度和粗糙度可能一样，为了克服这个问题，系统地研究了该系统中知识的不确定度量问题。本书也研究了不完备区间值信息系统的约简、下近似约简和上近似约简。

（3）在不完备模糊信息系统中，建立了基于优势关系的粗糙集模型，并证明了其上下近似算子的重要性质。同时把知识粒度、分辨度等概念引入该系统中，系统地探讨了优势关系粗糙集系统中知识的不确定度量问题。通过给出相对上近似和相对下近似约简，提取简化的“至多”和“至少”决策规则。

（4）由于现实生活中数据的复杂性，一个不完备信息系统中可能同时包含递增、递减的偏序关系。为此，在不完备信息系统中，构造了能处理较复杂问题的优势粗糙集模型，研究了其相关性质。同时，在该系统中建立了带值的优势粗糙集模型，该模型使用概率理论考虑了一个对象优于另一个对象的程度。最后通过区分属性集合、区分矩阵和函数设计了简便

的约简算法，进一步丰富了粗糙集理论。

本书试图在总结过去研究工作的基础上阐述作者在粗糙集方面的新研究成果。第 1 章简要介绍了粗糙集的历史及模型，第 2 章研究基于差异关系的不完备信息系统粗糙集方法，第 3 章研究不完备区间值信息系统的知识获取，第 4 章研究不完备模糊信息系统的粗糙集方法与知识约简，第 5 章研究不完备信息系统的扩展优势粗糙集，第 6 章是对研究成果的总结及对未来工作的一些展望。

本书的研究工作得到了南京理工大学计算机学院师生的大力支持与帮助，也得到了浙江商业职业技术学院的支持与资助。特此向支持和关心作者研究工作的所有单位和个人表示感谢。作者还要感谢教育作者多年的导师唐振民教授，以及学长和同人的帮助与支持；感谢出版社同人为本书出版付出的辛勤劳动。书中有部分内容参考了有关单位和个人的研究成果，均已在参考文献中列出，在此一并致谢。

由于本书追求的目标是介绍作者本人的研究新成果，这给撰写本书增添了难度。加上作者水平所限，虽然几经改稿，书中错误和缺点在所难免，希望广大读者不吝赐教。

魏利华

2014 年 10 月

此为试读，需要完整PDF请访问：www.ertongbook.com

主要符号说明

U	非空有限论域
$P(U)$	集合 U 的幂集
AT	信息系统的条件属性集
$\{d\}$	目标信息系统的决策属性集
$\sim X$	集合 X 的补集
$\|X\|$	集合 X 的基数
IS	信息系统
IIS	不完备信息系统
$IIIS$	不完备区间值信息系统
$IFIS$	不完备模糊信息系统
R_A^{*D}	不完备信息系统中的差异关系
R_A^{*I}	不完备区间值信息系统中的序邻域关系
R_A^{*F}	不完备模糊信息系统中的优势关系
$R_A^{*\geqslant}$	不完备信息系统中的扩展优势关系
$GK(R_A^{*D})$	知识 R_A^{*D} 的粒度
$Dis(R_A^{*D})$	知识 R_A^{*D} 的分辨度
$E(R_A^{*D})$	知识 R_A^{*D} 的信息熵
$E_r(R_A^{*D})$	知识 R_A^{*D} 的粗糙熵
$G(R_A^{*D})$	知识 R_A^{*D} 的不确定度量
$\boldsymbol{M}_{AT}^{*I}$	不完备区间值信息系统的区分矩阵

Δ^{*I}	不完备区间值信息系统的区分函数
M_{AT}^{*IL}	不完备区间值信息系统的下近似区分矩阵
M_{AT}^{*IH}	不完备区间值信息系统的上近似区分矩阵
Δ^{*IL}	不完备区间值信息系统的下近似区分函数
Δ^{*IH}	不完备区间值信息系统的上近似区分函数

目　录

第 1 章
绪 论

由于现代网络信息和计算机技术的快速发展及大量的电子设备的使用，人们收集数据的能力越来越强，数据不论是量的角度，还是维数都在急剧增加，从而导致数据、信息的不确定性越来越突出，其关系也变得越来越复杂. 要想从海量的信息中挖掘出有用的知识，必须采用适当的技术方法，建立相关模型，最终给决策系统提供支持，知识发现和约简则专门研究这些方法和技术.

粗糙集理论最早是由波兰数学家华沙理工大学教授 Z. Pawlak 于 1982 年提出的，是处理不确定、不精确知识的数学工具和数据分析理论，同时对处理非常复杂的系统也是一种非常有效的方法. 但是该理论与其他处理不确定问题的理论有一定的区别，它只需要问题所需处理的数据集合，而不需要其他的任何先验信息，而证据理论、模糊集理论、概率理论则不行. 粗糙集已经成为智能计算当中一个较新的研究热点，现已成功地应用到了数据的分析与决策、股票数据分析、医疗诊断、决策支持、过程控制、工业制造、商业经济、数据挖掘、模式识别、机器学习、故障检测、地震预报知识发现等多个领域.

第一届粗糙集国际研讨会是 1992 年在波兰召开的，从那时开始，几乎每年都召开一次关于粗糙集主题的国际会议，包括“粗糙集、模糊集、数据挖掘与粒度计算国际会议”“粗糙集进展与计算国际会议”等. 两年后，国际上又成立了粗糙集研究学会. 1995 年，粗糙集又被列为了新出现的计算机科学的研究课题. 1998 年，关于粗糙集的专辑在国际信息科学期刊出版了，而且其他国际高级刊物也开辟了粗糙集的专栏. 这些重要的会议和杂志，推动粗糙集不断快速发展. 相对国外来说，粗糙集在国内发展虽然稍晚但是较快，很多学者发表了关于粗糙集理论的论文和出版了关于粗糙集的专著. 同时举办了多次关于粗糙集的国际学术

会议，推动了粗糙集的发展，而且很快得到了国际同行的高度认可.

粗糙集理论用于数据分析、知识发现，与其他处理不确定问题的证据理论、模糊集理论、概率理论等相比，有如下突出特点：

（1）不需要提供除所需处理数据之外的任何先验信息；

（2）能处理、表达不完备的信息，能在保留所需信息不变的情况下进行属性约简，从而方便规则提取；

（3）能挖掘出数据表中的隐含信息；

（4）粗糙集理论和其他处理不确定问题的理论具有较好的互补性，取长补短，从而丰富了粗糙集理论，使其功能更强大和更优良.

众所周知，Z. Pawlak 提出的粗糙集是建立在等价关系（不可区分关系）的基础之上的，它所处理的分类必须是正确的或者完全肯定的，该要求限制了粗糙集在现实生活中的应用范围. 因此，很多学者们对粗糙集的方法进行了拓展：一是把被近似的集合从经典的集合推广到模糊集、直觉模糊集；二是把研究的论域从一个推广到两个，进而推广到多个；三是把关系从等价关系推广到优势关系、模糊关系、直觉模糊关系、相容关系（相似关系）、一般二元关系、邻域关系等；四是把信息系统的属性值，推广到了模糊值、区间值、集值等. 近年来，为了粗糙集应用的广泛性，学者们结合多种理论，目前已经取得的成果有：变精度粗糙集模型、程度粗糙集模型、模糊粗糙集模型、粗糙模糊集模型、直觉模糊粗糙集、粗糙直觉模糊集、相容关系粗糙集模型、优势关系粗糙集模型、多粒度粗糙集模型等.

粗糙集理论的一个前提就是在保持分类能力不变的情况下，它的分类主要通过等价关系来诱导，因为等价关系对论域能构成一个划分. 该理论将划分的数据集理解为知识或者概念，通过两个近似算子，下近似算子和上近似算子对未知的知识、概念进行刻画. 如果得到的下近似集合和上近似集合是相等的，则称被近似的集合是可定义的或者是精确的；如果得到的下近似集合和上近似集合是不相等的则称被近似的集合是不可定义的或者是粗糙的. 一般地，粗糙集的不确定性主要是由于下、上近似集不等，即边界域的存在而导致的，所以不确定性研究也是当前的一个研究热点，当前，这方面取得的成果详见文献［1，9，31，34，75，82－83，86，97］. 除此之外，知识约简一直是粗糙集研究的一个核心问题，约简能去除冗余的知识，保留有用信息，对提取简化规则奠定好的基础.

信息系统是具有对象和属性关系的一个数据库，属性一般分为条件属性和决策属性，该数据库由数据隐含着对象与属性之间的关系，最终用属性来表达知识模式，具有明确的意义，是可以很好理解的. 不同的数据表规模不一样，属性的取值也不一样，有的是定性值、有的是定量值、有的是离散值、有的是区间值、有的是连续值、有的是缺省值、有的是集合

此为试读，需要完整PDF请访问：www.ertongbook.com

值等.

本书重点以不完备的信息系统为基本的研究对象，以粗糙集、模糊集等为基本工具，目的是知识发现，深入研究不完备信息系统的知识约简问题，有非常明确的应用目标和研究结构.

1.1 Pawlak 粗糙集模型

设论域 U 是所有对象构成的有限集合，论域 U 上的一族划分为一个知识库，记为 (U, R)，有时也称 (U, R) 为 Pawlak 近似空间，其中 R 是论域 $U \times U$ 上的一个二元等价关系，即满足自反性、对称性和传递性. 对象 x 的等价类记为 $[x]_R$，

$$[x]_R = \{y \in U \mid (x, y) \in R\}.$$

由等价关系 R 诱导出的论域 U 的划分记为 U/R，

$$U/R = \{[x]_R \mid x \in U\}.$$

定义 1.1.1 设 (U, R) 为 Pawlak 近似空间，其中 R 是论域 $U \times U$ 上的一个等价关系，对于任意的子集 $X \subseteq U$，则 X 的下近似集和上近似集分别为：

$$\underline{R}(X) = \cup \{Y \in U/R \mid Y \subseteq X\};$$

$$\bar{R}(X) = \cup \{Y \in U/R \mid Y \cap X \neq \varnothing\}.$$

由于二元关系 R 是等价关系，因此下近似集和上近似集也可以表示成如下形式：

$$\underline{R}(X) = \{x \in U \mid [x]_R \subseteq X\};$$

$$\bar{R}(X) = \{x \in U \mid [x]_R \cap X \neq \varnothing\}.$$

如果 $\underline{R}(X) = \bar{R}(X)$，则称集合 X 是可定义集或者是精确集；如果 $\underline{R}(X) \neq \bar{R}(X)$，则称集合 X 是不可定义集或者是粗糙集.

集合 $\underline{R}(X)$ 称为 X 的正域，记为 $Pos_R(X)$，表示根据知识 R 判断一定属于集合 X 的元素组成的集合. 集合 $U - \bar{R}(X)$ 称为 X 的负域，记为 $Neg_R(X)$，表示根据知识 R 判断一定不属于集合 X 的元素组成的集合. 集合 $\bar{R}(X) - \underline{R}(X)$ 称为 X 的边界域，记为 $Bn_R(X)$，表示根据知识 R 不能肯定判断属于集合 X 也不能肯定判断属于集合 X 的补集 $\sim X$.

显然，集合 X 的正域、边界域、负域构成了论域 U 的一个划分，即：

$$U = Pos_R(X) \cup Neg_R(X) \cup Bn_R(X).$$

定理 1.1.1 设 (U, R) 为 Pawlak 近似空间，对于任意的子集 X，$Y \subseteq U$，则如下性质成立：

(1) $\underline{R}(X) \subseteq X \subseteq \bar{R}(X)$；

(2) $\underline{R}(\sim X) = \sim \bar{R}(X), \bar{R}(\sim X) = \sim \underline{R}(X)$；

(3) $\underline{R}(\varnothing) = \bar{R}(\varnothing) = \varnothing$；

(4) $\underline{R}(U) = \bar{R}(U) = U$；

(5) $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$；

(6) $\bar{R}(X \cup Y) = \bar{R}(X) \cup \bar{R}(Y)$；

(7) $\underline{R}(X \cup Y) \supseteq \underline{R}(X) \cup \underline{R}(Y)$；

(8) $\bar{R}(X \cap Y) \subseteq \bar{R}(X) \cap \bar{R}(Y)$；

(9) $X \subseteq Y \Rightarrow \underline{R}(X) \subseteq \underline{R}(Y)$；

(10) $X \subseteq Y \Rightarrow \bar{R}(X) \subseteq \bar{R}(Y)$；

(11) $\underline{R}(\underline{R}(X)) = \underline{R}(X)$；

(12) $\bar{R}(\bar{R}(X)) = \bar{R}(X)$；

(13) $\underline{R}(\sim \underline{R}(X)) = \sim \underline{R}(X)$；

(14) $\bar{R}(\sim \bar{R}(X)) = \sim \bar{R}(X)$；

(15) 任意 $K \in c$，$\underline{R}(K) = K$；

(16) 任意 $K \in c$，$\bar{R}(K) = K$.

其中，c 表示论域 U 在基于等价关系下的所有可定义集的全体.

定义 1.1.2 设 (U, R) 为 Pawlak 近似空间，对于任意的子集 $X \subseteq U$，其下近似和上近似分别为 $\underline{R}(X)$，$\bar{R}(X)$，则集合 X 的精度为：

$$\alpha_R(X) = \frac{|\underline{R}(X)|}{|\bar{R}(X)|}.$$

显然，精度的取值为 0 到 1，它从正面刻画了集合 X 在等价关系 R 下精确的程度，同时，也可以从反面来刻画集合 X 的精确程度，即粗糙度：

$$\rho_R(X) = 1 - \alpha_R(X) = 1 - \frac{|\underline{R}(X)|}{|\bar{R}(X)|}.$$

显然，粗糙度的取值也为 0 到 1. 当 $\alpha_R(X) = 1$，$\rho_R(X) = 0$，则集合 X 是可定义集；当 $\alpha_R(X) < 1$，$\rho_R(X) > 0$，则集合 X 是不可定义集.

1.2 基于容差关系的不完备信息系统粗糙集

本节将简要介绍不完备信息系统的相关知识.

定义 1.2.1 设 $IS=(U, AT, F)$ 为信息系统，其中，U 为对象集或论域：

$$U=\{x_1, x_2, \cdots, x_n\}.$$

论域 U 中的每一个 $x_i(i=1, 2, \cdots, n)$ 称为一个对象. AT 为属性集：

$$AT=\{a_1, a_2, \cdots, a_m\}.$$

属性集 AT 中的每一个 $a_j(j=1, 2, \cdots, m)$ 称为一个属性. F 是论域 U 和属性集 AT 间的关系集，即：

$$F=\{f_j | j\leqslant m\}, f_j: U\rightarrow V_{a_j}.$$

其中，V_{a_j} 为对象在属性 a_j 的取值.

在信息系统中，论域 U 和属性集 AT 的关系一定是存在的，这是知识发现的重要信息基础. 一般地，$f(x, a)=v$ 表示对象 x 在属性 a 下的取值为 v. 所有属性的值域记为 V：

$$V=\bigcup_{j=1}^{m} V_{a_j}.$$

在信息系统 IS 中，一个属性可以诱导一个等价关系，一个表可以看作是定义的一族等价关系. 此时，可以称 $R_A=\{(x, y)\in U\times U | f(x, a)=f(y, a), \ \forall a\in A\}$ 为关于属性集 A 的诱导的等价关系. 一般说来，描述不同对象的属性集是比较大的，但是对于信息系统的知识发现来说有些属性并不总是必要的. 有些属性是绝对不必要的，去掉这种属性并不影响分类的知识发现；而有些属性是绝对必要的，去掉这种属性必然会影响分类的知识发现. 还有一些属性是相对必要的属性，它可能与其他属性联合起来确定分类的知识发现. 而知识约简就是要在属性集中寻找一个最小的属性集，使得这个最小属性集确定的分类知识与用全体属性集确定的分类知识是相同的.

定义 1.2.2 设 $IS=(U, AT, F)$ 为信息系统，对于任意的属性子集 $A\subseteq AT$，其满足如下两个条件：

(1) $R_A=R_{AT}$；

(2) 对于任意的属性 $a\in A$，$R_{A-\{a\}}\neq R_{AT}$.

则称属性集 A 是该信息系统 IS 的约简.

如果所有约简的交非空时，则称此交集为该信息系统 IS 的核，核中的属性是绝对必要属性.

有时，由于数据毁坏或者丢失，导致某些对象的某些属性值缺失，为了研究的方便，通常把“ * ”这个空值给这些属性.

定义 1.2.3 在信息系统 $IS=(U, AT, F)$ 中，如果至少存在一个属性 $a \in AT$，使得 V_a 含有空值，则称该信息系统 IS 为不完备信息系统，记为 IIS.

定义 1.2.4 设 $IIS=(U, AT, F)$ 是一个不完备信息系统，任意的属性子集 $A \subseteq AT$，$a \in A$，则容差关系为:

$$R_A^* = \{(x,y) \in U \times U \mid f(x,a) = f(y,a) \vee f(x,a) = * \vee f(y,a) = *\}.$$

由以上定义可知，性质 $R_A^* = \bigcap_{a \in A} R_{\{a\}}^*$成立. 令 $U/R_A^* = \{[x]_{R_A^*} \mid x \in U\}$ 表示分类，其中，$[x]_{R_A^*}$称为对象 x 的容差类. 一般来说，U/R_A^* 中的容差类一般不是论域 U 的划分，而是论域 U 的一个覆盖.

定义 1.2.5 设 $IIS=(U, AT, F)$ 是一个不完备信息系统，任意的属性子集 $A \subseteq AT$，$X \subseteq U$，则集合 X 基于容差关系的下近似集和上近似集分别为:

$$\underline{R}_A^*(X) = \{x \in U \mid [x]_{R_A^*} \subseteq X\};$$

$$\bar{R}_A^*(X) = \{x \in U \mid [x]_{R_A^*} \cap X \neq \varnothing\}.$$

其中，$\underline{R}_A^*(X)$，$\bar{R}_A^*(X)$ 分别为集合 X 的下近似和上近似，如果 $\underline{R}_A^*(X) = \bar{R}_A^*(X)$，则称集合 X 在该不完备信息系统中是可定义的或者是精确的；否则，称集合 X 在该不完备信息系统中是不可定义的或者是粗糙的.

定义 1.2.6 设 $IIS=(U, AT, F)$ 是一个不完备信息系统，任意的属性子集 $A \subseteq AT$，

(1) 如果 $R_A^* = R_{AT}^*$，则属性子集 A 是系统 IIS 的容差协调集；

(2) 如果属性子集 A 是系统 IIS 的容差协调集，而任意的属性集 $B \subset A$，属性集 B 不是容差协调集，则属性子集 A 是系统 IIS 的容差约简.

按照定义去寻找不完备信息系统 IIS 的容差约简，计算量相当大，所以有必要找更简捷的方法，通常借助于容差区分矩阵.

$$D_{AT}^*(x, y) = \begin{cases} \{a \in AT \mid (x, y) \notin R_{\{a\}}^*\} & (x, y) \notin R_{AT}^*, \\ \varnothing & \text{其他}. \end{cases}$$

则 $D_{AT}^*(x, y)$ 为不完备信息系统 IIS 对象序对 (x, y) 的容差关系区分属性集. 再者，矩阵

$$\boldsymbol{M}_{AT}^* = \{D_{AT}^*(x, y) \mid (x, y) \notin R_{AT}^*\}$$

为不完备信息系统 IIS 的容差关系区分矩阵.

定义 1.2.7 设 $IIS=(U, AT, F)$ 是一个不完备信息系统，则该系统中容差关系区分函数为:

$$\Delta^* = \bigwedge_{D^*_{AT}(x,\ y) \in M^*_{AT}} \bigvee D^*_{AT}(x,\ y).$$

定理 1.2.1 设 $IIS=(U,\ AT,\ F)$ 是一个不完备信息系统，对于任意的属性子集 $A \subseteq AT$，属性集 A 是容差关系协调约简，当且仅当属性集 A 是 Δ^* 的基本蕴含项.

其他的关于不完备信息系统中的约简理论详见文献 [90－91].

1.3 模糊粗糙集

模糊集理论是美国 Zadeh 教授于 1965 年提出的、专门处理不确定问题的又一种数学工具，但是该理论与粗糙集理论不同，它主要是针对现实生活中的亦此亦彼的模糊现象提出来的.

定义 1.3.1 设 U 是一个非空有限论域，模糊集 $\tilde{A}$ 是论域 U 到区间 [0, 1] 上的映射，即：

$$\tilde{A}: U \to [0,\ 1],$$

$$x \to \tilde{A}(x).$$

则称 $\tilde{A}(x)$ 为对象 x 在模糊集 $\tilde{A}$ 上的隶属度.

论域 U 上的全体模糊集记为 $F(U)$.

定义 1.3.2 设 U 是一个非空有限论域，如果二元关系 $\tilde{R}$ 是 $U \times U$ 到区间 [0, 1] 上的映射，即：

$$\tilde{R}: U \times U \to [0,\ 1],$$

$$(x,\ y) \to \tilde{R}(x,\ y).$$

则称关系 $\tilde{R}$ 为 $U \times U$ 上的模糊二元关系，其中，$\tilde{R}(x,\ y)$ 为对象 x，y 的隶属度.

定义 1.3.3 设 $(U,\ R)$ 是 Pawlak 近似空间，对于论域 U 上的模糊集合 $\tilde{A}$，则模糊集 $\tilde{A}$ 基于 Pawlak 近似空间的下近似集和上近似集分别为：

$$\underline{R}(\tilde{A})(x) = \wedge \{ \tilde{A}(y) \mid y \in [x]_R \};$$

$$\overline{R}(\tilde{A})(x) = \vee \{ \tilde{A}(y) \mid y \in [x]_R \}.$$

模糊集 $\tilde{A}$ 被近似之后依然是模糊集. 事实上, 上述是在经典的等价关系下对论域上的模糊集合的近似描述, 下面将介绍在模糊二元关系下, 对论域上的经典笛卡儿集合的近似描述.

定义 1.3.4 设近似空间 (U, $\tilde{R}$) 为模糊近似空间, 如果 $U \times U$ 上的模糊二元关系 $\tilde{R}$ 是自反的, 则模糊近似空间 (U, $\tilde{R}$) 是模糊等价近似空间, 当且仅当 $\tilde{R}$ 是一个模糊等价关系.

定义 1.3.5 设 (U, $\tilde{R}$) 是模糊近似空间, 对于任意的对象 $x \in U$, 有

$$[x]: U \to [0, 1],$$

$$y \to \tilde{R}(x, y).$$

称 $[x]$ 为对象 x 的模糊邻域算子.

定义 1.3.6 设 (U, $\tilde{R}$) 是模糊近似空间, 对于论域 U 上的任意经典子集 X, 则集合 X 在模糊近似空间中的下近似集和上近似集分别为:

$$\underline{\tilde{R}}(X)(y) = \bigwedge_{x \notin X}(1 - \tilde{R}(x, y));$$

$$\overline{\tilde{R}}(X)(y) = \bigvee_{x \in X}(\tilde{R}(x, y)).$$

接下来, 将介绍在模糊近似空间 (U, $\tilde{R}$) 对模糊集合 $\tilde{A}$ 的近似刻画.

定义 1.3.7 设 (U, $\tilde{R}$) 是模糊近似空间, 对于论域 U 上的任意模糊集 $\tilde{A}$, 则模糊集合 $\tilde{A}$ 在模糊近似空间中的下近似集和上近似集分别为:

$$\underline{\tilde{R}}(\tilde{A})(x) = \wedge\{\tilde{A}(y) \vee (1 - \tilde{R}(x, y)) \mid y \in U\};$$

$$\overline{\tilde{R}}(\tilde{A})(x) = \vee\{\tilde{A}(y) \wedge (\tilde{R}(x, y)) \mid y \in U\}.$$

1.4 优势关系粗糙集

现实生活中, 有时需要根据属性值的值域进行排序, 比如学生的成绩. 在一个信息系统中, 可由属性的值域得到递增的偏序关系或递减的偏序关系, 再由偏序关系, 可得到优势关系. 如果属性的值域是递增的偏序关系或递减的偏序关系, 则这个属性称为一个准则, 由准

此为试读,需要完整PDF请访问: www.ertongbook.com