粗糙集理论与方法

西安交通大学数学研究生教学丛书

张文修 吴伟志
梁吉业 李德玉 编著

科学出版社

责任编辑／林 鹏 杨 波 · 封面设计／高海英 · 责任校对／曹锐军
（O-1282.0101）

粗糙集理论与方法

西安交通大学数学研究生教学丛书

计算力学	张 武等
计算智能中的仿生理论和算法	徐宗本等
数值计算的算法与分析	张可村等
统计推断导引	范金城等
抽象代数	盛德成
偏微分方程现代数值方法	黄艾香等
常微分方程定性与稳定性方法	马知恩等
随机逼近及自适应算法	聂赞坎等
近代优化方法	徐成贤等
应用非线性分析	陈红斌等
动力系统基础及其方法	陈绥阳等
粗糙集理论与方法	张文修等
计算机数学	陈志平等

ISBN 7-03-008798-4
9 787030 087980 >

ISBN 7-03-008798-4/O · 1282
定 价：22.00 元

西安交通大学数学研究生教学丛书

粗糙集理论与方法

张文修　吴伟志
梁吉业　李德玉　编著

科学出版社

2001

内 容 简 介

本书系统地介绍了粗糙集理论的基本内容与方法，力图概括国内外最新成果．主要内容有：粗糙集的基本概念，粗糙计算方法，粗糙集的代数性质与粗糙逻辑，粗糙集的各种推广模型，粗糙集与其他处理不确定或不精确问题理论的联系以及不完备信息系统下的粗糙集方法．

本书可作为计算机科学、应用数学、自动控制、信息科学和管理工程等专业的高年级学生及研究生的教材，也可作为研究粗糙集理论与方法的科技人员的参考书．

图书在版编目（CIP）数据

粗糙集理论与方法/张文修等编著．－北京：科学出版社，2001

（西安交通大学数学研究生教学丛书）

ISBN 7-03-008798-4

Ⅰ．粗… Ⅱ．张… Ⅲ．粗糙集 Ⅳ．O144

中国版本图书馆 CIP 数据核字(2000)第 69236 号

科 学 出 版 社 出版

北京东黄城根北街16号

邮政编码：100717

新 蕾 印 刷 厂 印刷

科学出版社发行 各地新华书店经销

*

2001 年 7 月第 一 版 开本：B5(720×1000)

2001 年 7 月第一次印刷 印张：14 3/4

印数：1-3 000 字数：261 000

定价：22.00 元

（如有印装质量问题，我社负责调换〈新欣〉）

前　　言

当今，社会已经进入了网络信息时代，计算机与网络信息技术的飞速发展使得各个领域的数据和信息急剧增加（信息爆炸），并且由于人类的参与使数据与信息系统中的不确定性更加显著（复杂系统）. 如何从大量的、杂乱无章的、强干扰的数据（海量数据）中挖掘潜在的、有利用价值的信息（有用知识），这给人类的智能信息处理能力提出了前所未有的挑战. 由此产生了人工智能研究的一个崭新领域——数据挖掘（DM）和数据库知识发现（KDD）.

在 DM 和 KDD 诸多方法中，粗糙集理论与方法对于处理复杂系统不失为一种较为有效的方法，因为它与概率方法、模糊集方法和证据理论方法等其他处理不确定性问题理论的最显著的区别是它无需提供问题所需处理的数据集合之外的任何先验信息. 当然，由于该理论未能包含处理不精确或不确定原始数据的机制，所以与其他处理不确定性问题的理论有很强的互补性.

粗糙集理论是波兰数学家 Z. Pawlak 于 1982 年提出的一种数据分析理论. 由于最初关于粗糙集理论的研究主要集中在波兰，因此当时并没有引起国际计算机界和数学界的重视，研究地域仅局限于东欧一些国家. 直到 1990 年前后，由于该理论在数据的决策与分析、模式识别、机器学习与知识发现等方面的成功应用，才逐渐引起了世界各国学者的广泛关注. 1991 年 Z. Pawlak 的专著《粗糙集——关于数据推理的理论》(Rough Sets——Theoretical Aspects of Reasoning about Data)的问世，标志着粗糙集理论及其应用的研究进入了活跃时期. 1992 年在波兰召开了关于粗糙集理论的第一届国际学术会议. 1995 年 ACM Communication 将粗糙集列为新浮现的计算机科学的研究课题. 目前，粗糙集理论已成为信息科学最为活跃的研究领域之一. 同时，该理论还在医学、化学、材料学、地理学、管理科学和金融等其他学科得到了成功的应用.

本书的目的是介绍粗糙集的基本理论与方法以及这一理论的研究发展状况. 为了阅读方便，本书对国内外已发表的文章进行了系统化处理，规范了数学概念与符号，在统一的框架下叙述了粗糙集理论的最新研究成果，同时也包含了作者的某些新成果，期望为从事粗糙集理论研究人员和研究生进入这一新领域提供捷径.

鉴于我们从事该领域的研究工作时间较短，加之自身知识的局限性，错误与不妥之处在所难免，热忱欢迎广大同仁批评、指正.

作　　者

2000 年 8 月

目　　录

第一章　粗糙集理论的基本概念

粗糙集理论是一种新的处理模糊和不确定性知识的数学工具.其主要思想就是在保持分类能力不变的前提下,通过知识约简,导出问题的决策或分类规则.目前,粗糙集理论已被成功地应用于机器学习、决策分析、过程控制、模式识别与数据挖掘等领域.本章介绍标准粗糙集理论(Pawlak 粗糙集模型)的基本概念,作为后面各章节的基础.

§1.1　知识与知识库

设 $U\neq\varnothing$ 是我们感兴趣的对象组成的有限集合,称为论域.任何子集 $X\subseteq U$,称为 U 中的一个概念或范畴.为规范化起见,我们认为空集也是一个概念.U 中的任何概念族称为关于 U 的抽象知识,简称知识.本书主要是对在 U 上能形成划分的那些知识感兴趣.一个划分 $\mathscr{C}$ 定义为:$\mathscr{C}=\{X_1,X_2,\cdots,X_n\}$;$X_i\subset U$,$X_i\neq\varnothing$,$X_i\cap X_j=\varnothing$,对于 $i\neq j$,$i,j=1,2,\cdots,n$;$\bigcup\limits_{i=1}^{n}X_i=U$.

U 上的一族划分称为关于 U 的一个知识库(knowledge base).

设 R 是 U 上的一个等价关系,U/R 表示 R 的所有等价类(或者 U 上的分类)构成的集合,$[x]_R$ 表示包含元素 $x\in U$ 的 R 等价类.一个知识库就是一个关系系统 $K=(U,\mathbf{R})$,其中 U 为非空有限集,称为论域,$\mathbf{R}$ 是 U 上的一族等价关系.

若 $\mathbf{P}\subseteq\mathbf{R}$,且 $\mathbf{P}\neq\varnothing$,则 $\bigcap\mathbf{P}$($\mathbf{P}$ 中所有等价关系的交集)也是一个等价关系,称为 $\mathbf{P}$ 上的不可区分(indiscernibility)关系,记为 $\mathrm{ind}(\mathbf{P})$,且有

$$[x]_{\mathrm{ind}(\mathbf{P})}=\bigcap_{R\in\mathbf{P}}[x]_R. \tag{1.1}$$

这样,$U/\mathrm{ind}(\mathbf{P})$(即等价关系 $\mathrm{ind}(\mathbf{P})$ 的所有等价类)表示与等价关系族 $\mathbf{P}$ 相关的知识,称为 K 中关于 U 的 $\mathbf{P}$ 基本知识($\mathbf{P}$ 基本集).为简单起见,我们用 $U/\mathbf{P}$ 代替 $U/\mathrm{ind}(\mathbf{P})$,$\mathrm{ind}(\mathbf{P})$ 的等价类称为知识 $\mathbf{P}$ 的基本概念或基本范畴.特别地,如果 $Q\in\mathbf{R}$,则称 Q 为 K 中关于 U 的 Q 初等知识,Q 的等价类为知识 $\mathbf{R}$ 的 Q 初等概念或 Q 初等范畴.

事实上,$\mathbf{P}$ 基本范畴是拥有知识 $\mathbf{P}$ 的论域的基本特性.换句话说,它们是知识的基本模块.

同样,我们也可定义:当 $K=(U,\mathbf{R})$ 为一个知识库,$\mathrm{ind}(K)$ 定义为 K 中

所有等价关系的族，记作 $\mathrm{ind}(K)=\{\mathrm{ind}(\mathbf{P})\mid\varnothing\neq\mathbf{P}\subseteq\mathbf{R}\}$.

例 1.1 给定一玩具积木的集合 $U=\{x_1,x_2,\cdots,x_8\}$，并假设这些积木有不同的颜色(红、黄、蓝)，形状(方、圆、三角)，体积(小，大). 因此，这些积木都可以用颜色、形状、体积这些知识来描述. 例如一块积木可以是红色、小而圆的，或黄色、大而方的等. 如果我们根据某一属性描述这些积木的情况，就可以按颜色、形状、体积分类.

按颜色分类：

x_1,x_3,x_7——红；

x_2,x_4——蓝；

x_5,x_6,x_8——黄.

按形状分类：

x_1,x_5——圆；

x_2,x_6——方；

x_3,x_4,x_7,x_8——三角.

按体积分类：

x_2,x_7,x_8——大；

x_1,x_3,x_4,x_5,x_6——小.

换言之，我们定义三个等价关系(即属性)：颜色 R_1，形状 R_2 和体积 R_3，通过这些等价关系，可以得到下面三个等价类：

$$U/R_1=\{\{x_1,x_3,x_7\},\{x_2,x_4\},\{x_5,x_6,x_8\}\},$$
$$U/R_2=\{\{x_1,x_5\},\{x_2,x_6\},\{x_3,x_4,x_7,x_8\}\},$$
$$U/R_3=\{\{x_2,x_7,x_8\},\{x_1,x_3,x_4,x_5,x_6\}\}.$$

这些等价类是由知识库 $K=(U,\{R_1,R_2,R_3\})$ 中的初等概念(初等范畴)构成的.

基本范畴是初等范畴的交集构成的，例如下列集合：

$$\{x_1,x_3,x_7\}\cap\{x_3,x_4,x_7,x_8\}=\{x_3,x_7\},$$
$$\{x_2,x_4\}\cap\{x_2,x_6\}=\{x_2\},$$
$$\{x_5,x_6,x_8\}\cap\{x_3,x_4,x_7,x_8\}=\{x_8\}.$$

它们分别为 $\{R_1,R_2\}$ 的基本范畴，即：红色三角形，蓝色方形，黄色三角形.

下列集合：

$$\{x_1,x_3,x_7\}\cap\{x_3,x_4,x_7,x_8\}\cap\{x_2,x_7,x_8\}=\{x_7\},$$
$$\{x_2,x_4\}\cap\{x_2,x_6\}\cap\{x_2,x_7,x_8\}=\{x_2\},$$
$$\{x_5,x_6,x_8\}\cap\{x_3,x_4,x_7,x_8\}\cap\{x_2,x_7,x_8\}=\{x_8\}.$$

它们分别为 $\{R_1,R_2,R_3\}$ 的基本范畴，即：红色大三角形，蓝色大方形，黄色大

三角形.

下列集合:

$$\{x_1,x_3,x_7\}\cup\{x_2,x_4\}=\{x_1,x_2,x_3,x_4,x_7\},$$
$$\{x_2,x_4\}\cup\{x_5,x_6,x_8\}=\{x_2,x_4,x_5,x_6,x_8\},$$
$$\{x_1,x_3,x_7\}\cup\{x_5,x_6,x_8\}=\{x_1,x_3,x_5,x_6,x_7,x_8\}.$$

它们分别为$\{R_1\}$的范畴,即:红或蓝(非黄),蓝或黄(非红),红或黄(非蓝).

注 有些范畴在这个知识库中是无法得到的,例如集合:

$$\{x_2,x_4\}\cap\{x_1,x_5\}=\varnothing,$$
$$\{x_1,x_3,x_7\}\cap\{x_2,x_6\}=\varnothing$$

为空集,也就是说,在我们的知识库中不存在蓝色圆形和红色方形的范畴,即为空范畴.

下面讨论两个知识库之间的关系.

令 $K=(U,\mathbf{P})$和 $K'=(U,\mathbf{Q})$为两个知识库.若 $\mathrm{ind}(\mathbf{P})=\mathrm{ind}(\mathbf{Q})$,即$U/\mathbf{P}=U/\mathbf{Q}$,则称 K 和K'($\mathbf{P}$ 和 $\mathbf{Q}$)是等价的,记作 $K\simeq K'$($\mathbf{P}\simeq\mathbf{Q}$).因此,当 K 和K'有同样的基本范畴集时,知识库 K 和K'中的知识都能使我们确切地表达关于论域的完全相同的事实.

这个概念意味着可以用不同的属性集对对象进行描述,以表达关于论域的完全相同的事实.

对于 $K=(U,\mathbf{P})$和 $K'=(U,\mathbf{Q})$两个知识库,当 $\mathrm{ind}(\mathbf{P})\subset\mathrm{ind}(\mathbf{Q})$时,我们称知识 $\mathbf{P}$(知识库 K)比知识 $\mathbf{Q}$(知识库 K')更精细,或者说 $\mathbf{Q}$ 比 $\mathbf{P}$ 更粗糙.当 $\mathbf{P}$ 比 $\mathbf{Q}$ 更精细时,我们也称 $\mathbf{P}$ 为 $\mathbf{Q}$ 的特化,$\mathbf{Q}$ 为 $\mathbf{P}$ 的推广.这意味着,推广是将某些范畴组合在一起,而特化则是将范畴分割成更小的单元.

§1.2 不精确范畴,近似与粗糙集

令 $X\subseteq U$,R 为 U 上的一个等价关系.当 X 能表达成某些R 基本范畴的并时,称 X 是R 可定义的;否则称 X 为R 不可定义的.

R 可定义集是论域的子集,它可在知识库 K 中精确地定义,而 R 不可定义集不能在这个知识库中定义.R 可定义集也称作R 精确集,而 R 不可定义集也称为R 非精确集或R 粗糙集(rough set).

当存在等价关系 $R\in\mathrm{ind}(K)$且 X 为R 精确集时,集合 $X\subseteq U$ 称为K 中的精确集;当对于任何 $R\in\mathrm{ind}(K)$,X 都为R 粗糙集,则 X 称为K 中的粗糙集.

对于粗糙集可以近似地定义,我们使用两个精确集,即粗糙集的上近似

(upper approximation)和下近似(lower approximation)来描述.

给定知识库 $K=(U,\mathbf{R})$,对于每个子集 $X\subseteq U$ 和一个等价关系 $R\in \mathrm{ind}(K)$,定义两个子集:

$$\underline{R}X=\bigcup\{Y\in U/R \mid Y\subseteq X\}, \tag{1.2}$$

$$\overline{R}X=\bigcup\{Y\in U/R \mid Y\cap X\neq\varnothing\}. \tag{1.3}$$

分别称它们为 X 的 R 下近似集和 R 上近似集.

下近似、上近似也可用下面的等式表达:

$$\underline{R}X=\{x\in U \mid [x]_R\subseteq X\}, \tag{1.4}$$

$$\overline{R}X=\{x\in U \mid [x]_R\cap X\neq\varnothing\}. \tag{1.5}$$

集合 $\mathrm{bn}_R(X)=\overline{R}X-\underline{R}X$ 称为 X 的 R 边界域;$\mathrm{pos}_R(X)=\underline{R}X$ 称为 X 的 R 正域;$\mathrm{neg}_R(X)=U-\overline{R}X$ 称为 X 的 R 负域. 显然:$\overline{R}X=\mathrm{pos}_R(X)\cup\mathrm{bn}_R(X)$.

$\underline{R}X$ 或 $\mathrm{pos}_R(X)$是由那些根据知识 R 判断肯定属于 X 的 U 中元素组成的集合;$\overline{R}X$ 是那些根据知识 R 判断可能属于 X 的 U 中元素组成的集合;$\mathrm{bn}_R(X)$是那些根据知识 R 既不能判断肯定属于 X 又不能判断肯定属于 $\sim X$(即 $U-X$)的 U 中元素组成的集合;$\mathrm{neg}_R(X)$是那些根据知识 R 判断肯定不属于 X 的 U 中元素组成的集合.

下列性质是显而易见的.

定理 1.1

(1) X 为 R 可定义集当且仅当 $\underline{R}X=\overline{R}X$.

(2) X 为 R 粗糙集当且仅当 $\underline{R}X\neq\overline{R}X$.

我们也可将 $\underline{R}X$ 描述为 X 中的最大可定义集,将 $\overline{R}X$ 描述为含有 X 的最小可定义集.

这样,范畴就是可以用已知知识表达的信息项. 换句话说,范畴就是用我们的知识可表达的具有相同性质的对象的子集. 一般地说,在一给定的知识库中,并不是所有对象子集都可以构成范畴(即用知识表达的概念). 因此,这样的子集可以看作为粗范畴(即不精确或近似范畴),它只能用知识通过两个精确范畴,即上、下近似集,粗略地定义.

从近似的定义,我们可以直接得到 R 下近似集和上近似集的下列性质.

定理 1.2

(1) $\underline{R}X\subseteq X\subseteq\overline{R}X$.

(2) $\underline{R}\varnothing=\overline{R}\varnothing=\varnothing$, 　$\underline{R}U=\overline{R}U=U$.

(3) $\overline{R}(X\cup Y)=\overline{R}X\cup\overline{R}Y$.

(4) $\underline{R}(X\cap Y)=\underline{R}X\cap\underline{R}Y$.

(5) $X\subseteq Y\Rightarrow\underline{R}X\subseteq\underline{R}Y$.

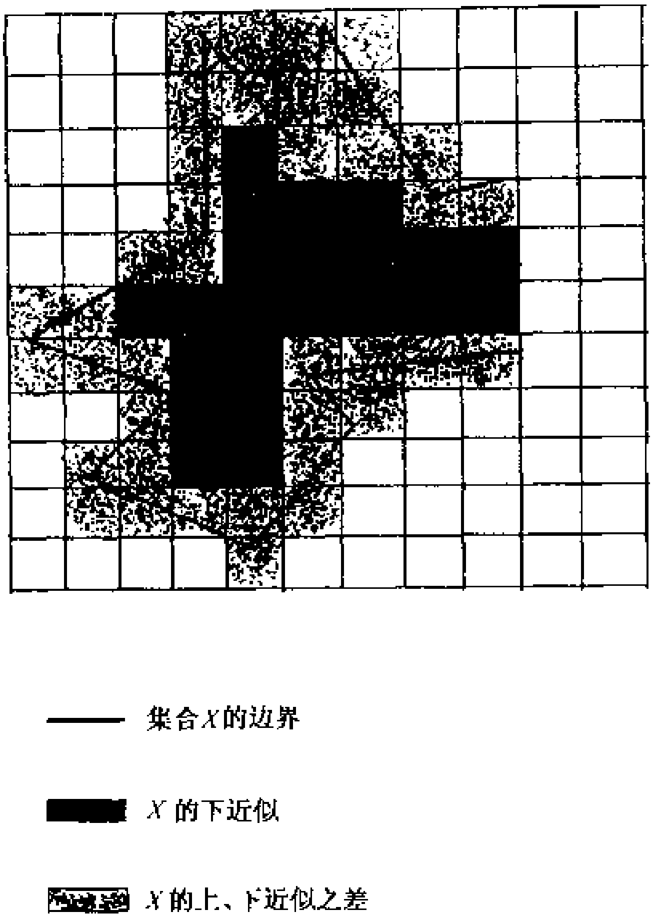

图 1.1 粗糙近似

(6) $X \subseteq Y \Rightarrow \overline{R}X \subseteq \overline{R}Y$.

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

(8) $\overline{R}(X \cap Y) \subseteq \overline{R}X \cap \overline{R}Y$.

(9) $\underline{R}(\sim X) = \sim \overline{R}X$.

(10) $\overline{R}(\sim X) = \sim \underline{R}X$.

(11) $\underline{R}(\underline{R}X) = \overline{R}(\underline{R}X) = \underline{R}X$.

(12) $\overline{R}(\overline{R}X) = \underline{R}(\overline{R}X) = \overline{R}X$.

证

(1a)设 $x \in \underline{R}X$，则有$[x] \subseteq X$；而 $x \in [x]$，所以 $x \in X$. 因此，$\underline{R}X \subseteq X$.

(1b)设 $x \in X$，则$[x] \cap X \neq \varnothing$，所以，$x \in \overline{R}X$. 因此 $X \subseteq \overline{R}X$.

(2a)由(1)知，$\underline{R}\varnothing \subseteq \varnothing$，而$\varnothing \subseteq \underline{R}\varnothing$，因此 $\underline{R}\varnothing = \varnothing$.

(2b)假设 $\overline{R}\varnothing \neq \varnothing$，则存在 x 使得$x \in \overline{R}\varnothing$，即$[x] \cap \varnothing \neq \varnothing$，而$[x] \cap \varnothing = \varnothing$，与假设矛盾. 因此，$\overline{R}\varnothing = \varnothing$.

(2c)由(1)知，$\underline{R}U \subseteq U$. 又因为当 $x \in U$，有$[x] \subseteq U$，所以 $x \in \underline{R}U$，即

$U \subseteq \underline{R}U$，因此，$\underline{R}U=U$.

(2d)由(1)知 $\overline{R}U \supseteq U$，但 $\overline{R}U \subseteq U$，因此 $\overline{R}U=U$.

(3)

$$\begin{aligned} x \in \overline{R}(X \cup Y) &\Leftrightarrow [x] \cap (X \cup Y) \neq \varnothing \\ &\Leftrightarrow ([x] \cap X) \cup ([x] \cap Y) \neq \varnothing \\ &\Leftrightarrow [x] \cap X \neq \varnothing \vee [x] \cap Y \neq \varnothing \\ &\Leftrightarrow x \in \overline{R}X \vee x \in \overline{R}Y \\ &\Leftrightarrow x \in \overline{R}X \cup \overline{R}Y, \end{aligned}$$

因此 $\overline{R}(X \cup Y)=\overline{R}X \cup \overline{R}Y$.

(4)

$$\begin{aligned} & x \in \underline{R}(X \cap Y) \\ &\Leftrightarrow [x] \subseteq X \cap Y \\ &\Leftrightarrow [x] \subseteq X \wedge [x] \subseteq Y \\ &\Leftrightarrow x \in \underline{R}X \cap \underline{R}Y. \end{aligned}$$

因此 $\underline{R}(X \cap Y)=\underline{R}X \cap \underline{R}Y$.

(5)设 $X \subseteq Y$，则 $X \cap Y=X$，所以 $\underline{R}(X \cap Y)=\underline{R}X$. 由(4)知，$\underline{R}X \cap \underline{R}Y=\underline{R}X$，因此 $\underline{R}X \subseteq \underline{R}Y$.

(6)设 $X \subseteq Y$，则 $X \cup Y=Y$，所以 $\overline{R}(X \cup Y)=\overline{R}Y$. 由(3)知，$\overline{R}X \cup \overline{R}Y=\overline{R}Y$，因此 $\overline{R}X \subseteq \overline{R}Y$.

(7)因为 $X \subseteq X \cup Y$，$Y \subseteq X \cup Y$，所以 $\underline{R}X \subseteq \underline{R}(X \cup Y)$，$\underline{R}Y \subseteq \underline{R}(X \cup Y)$，故 $\underline{R}X \cup \underline{R}Y \subseteq \underline{R}(X \cup Y)$.

(8)因为 $X \cap Y \subseteq X$，$X \cap Y \subseteq Y$，所以 $\overline{R}(X \cap Y) \subseteq \overline{R}X$，$\overline{R}(X \cap Y) \subseteq \overline{R}Y$，故 $\overline{R}(X \cap Y) \subseteq \overline{R}X \cap \overline{R}Y$.

(9)因为 $x \in \underline{R}X \Leftrightarrow [x] \subseteq X \Leftrightarrow [x] \cap \sim X=\varnothing$

$$\Leftrightarrow x \notin \overline{R}(\sim X) \Leftrightarrow x \in \sim \overline{R}(\sim X),$$

所以 $\underline{R}X=\sim \overline{R}(\sim X)$.

(10)在(9)中用 $\sim X$ 代替 X，即得 $\overline{R}X=\sim \underline{R}(\sim X)$.

(11a)由(1)知，$\underline{R}(\underline{R}X) \subseteq \underline{R}X$. 又当 $x \in \underline{R}X$ 时，有 $[x] \subseteq X$，因此，$\underline{R}[x] \subseteq \underline{R}X$. 但 $\underline{R}[x]=[x]$，于是 $[x] \subseteq \underline{R}X$，所以 $x \in \underline{R}(\underline{R}X)$，即 $\underline{R}X \subseteq \underline{R}(\underline{R}(X)$. 故 $\underline{R}(\underline{R}X)=\underline{R}X$.

(11b)由(1)知，$\underline{R}X \subseteq \overline{R}(\underline{R}X)$. 又当 $x \in \overline{R}(\underline{R}X)$，有 $[x] \cap \underline{R}X \neq \varnothing$，即存在 $y \in [x]$ 使得 $y \in \underline{R}X$，所以 $[y] \subseteq X$，但 $[x]=[y]$，这样就有 $[x] \subseteq X$，即 $x \in \underline{R}X$，因此，$\underline{R}X \supseteq \overline{R}(\underline{R}X)$. 故 $\overline{R}(\underline{R}X)=\underline{R}X$.

(12a)由(1)知，$\overline{R}X \subseteq \overline{R}(\overline{R}X)$，又当 $x \in \overline{R}(\overline{R}X)$ 时，有 $[x] \cap \overline{R}X \neq \varnothing$. 因此存在 $y \in [x]$ 且 $y \in \overline{R}X$，所以 $[y] \cap X \neq \varnothing$. 但 $[x]=[y]$，所以 $[x] \cap X \neq \varnothing$，即 $x \in \overline{R}X$，这样就有 $\overline{R}X \supseteq \overline{R}(\overline{R}X)$. 故 $\overline{R}(\overline{R}X)=\overline{R}X$.

(12b)由(1)知,$\underline{R}(\overline{R}X)\subseteq\overline{R}X$. 又当 $x\in\overline{R}X$ 时有$[x]\cap X\neq\varnothing$,所以$[x]\subseteq\overline{R}X$ (因为任取 $y\in[x]$时,则有$[y]\cap X=[x]\cap X\neq\varnothing$,即 $y\in\overline{R}X$),即 $x\in\underline{R}(\overline{R}X)$,这样就有 $\underline{R}(\overline{R}X)\supseteq\overline{R}X$. 故 $\underline{R}(\overline{R}X)=\overline{R}X$. □

集合近似的概念导致了一个新的概念——成员关系. 因为在我们的方法中,对于一个集合的定义是与集合的知识相联系的,所以成员关系一定也和知识有关,可以形式地定义为

$$x\ \underline{\in}_R X \quad \text{当且仅当} \quad x\in\underline{R}X;$$

$$x\ \overline{\in}_R Y \quad \text{当且仅当} \quad x\in\overline{R}X;$$

这里$\underline{\in}_R$ 表示根据R,x 肯定地属于X;$\overline{\in}_R$ 表示根据R,x 可能属于X. 分别称$\underline{\in}_R$ 和$\overline{\in}_R$ 为下和上成员关系.

成员关系依赖于我们的知识,即一个对象是否属于一个集合依赖于我们的知识,并且这不是绝对特性.

由定理 1.2 我们可以得到成员关系的性质:

定理 1.3

(1) $x\underline{\in}X$ 蕴涵$x\in X$ 蕴涵$x\overline{\in}X$;

(2) $X\subseteq Y$ 蕴涵($x\underline{\in}X$ 蕴涵$x\underline{\in}Y$ 且$x\overline{\in}X$ 蕴涵$x\overline{\in}Y$);

(3) $x\overline{\in}(X\cup Y)$当且仅当 $x\overline{\in}X$ 或$x\in Y$;

(4) $x\underline{\in}(X\cap Y)$当且仅当 $x\underline{\in}X$ 且$x\underline{\in}Y$;

(5) $x\underline{\in}X$ 或$x\underline{\in}Y$ 蕴涵$x\underline{\in}(X\cup Y)$;

(6) $x\overline{\in}(X\cap Y)$蕴涵 $x\overline{\in}X$ 且$x\overline{\in}Y$;

(7) $x\underline{\in}(\sim X)$当且仅当 $x\overline{\in}X$ 不成立;

(8) $x\overline{\in}(\sim X)$当且仅当 $x\underline{\in}X$ 不成立.

为简单起见,我们省略了以上式子中的下标 R.

集合(范畴)的不精确性是由于边界域的存在而引起的. 集合的边界域越大,其精确性则越低. 为了更准确地表达这一点,我们引入精度的概念. 由等价关系 R 定义的集合X 的近似精度为

$$\alpha_R(X)=\frac{|\underline{R}X|}{|\overline{R}X|}, \tag{1.6}$$

其中 $X\neq\varnothing$,$|X|$表示集合 X 的基数.

精度 $\alpha_R(X)$用来反映我们对于了解集合 X 的知识的完全程度. 显然,对每一个 R 和$X\subseteq U$ 有 $0\leqslant\alpha_R(X)\leqslant 1$. 当 $\alpha_R(X)=1$ 时,X 的R 边界域为空集,集合 X 为R 可定义的;当 $\alpha_R(X)<1$ 时,集合 X 有非空R 边界域,集合 X

为 R 不可定义的.

当然,其他一些量度同样可用来定义集合 X 的不精确程度.

例如,可用 $\alpha_R(X)$的一种变形,即 X 的 R 粗糙度 $\rho_R(X)$来定义:

$$\rho_R(X) = 1 - \alpha_R(X).$$

X 的 R 粗糙度与精度恰恰相反,它表示的是集合 X 的知识的不完全程度.

我们可以看到,与概率论和模糊集合论不同,不精确性的数值不是事先假定的,而是通过表达知识不精确性的概念近似计算得到的,这样不精确性的数值表示的是有限知识(对象分类能力)的结果,这里我们不需要用一个机构来指定精确的数值去表达不精确的知识,而是采用量化概念(分类)来处理.不精确的数值特征用来表示概念的精确度.

除了用数值(近似程度的精度)来表示粗糙集的特征外,也可根据上、下近似的定义来表达粗糙集的另一个有用的特征,即拓扑特征.

下面定义四种不同的重要粗糙集:

(1)如果 $\underline{R}X \neq \varnothing$ 且 $\overline{R}X \neq U$,则称 X 为 R 粗糙可定义.

(2)如果 $\underline{R}X = \varnothing$ 且 $\overline{R}X \neq U$,则称 X 为 R 内不可定义.

(3)如果 $\underline{R}X \neq \varnothing$ 且 $\overline{R}X = U$,则称 X 为 R 外不可定义.

(4)如果 $\underline{R}X = \varnothing$ 且 $\overline{R}X = U$,则称 X 为 R 全不可定义.

这个划分的直观意义是这样的:

如果集合 X 为 R 粗糙可定义,则意味着我们可以确定 U 中某些元素属于 X 或 $\sim X$.

如果 X 为 R 内不可定义,则意味着我们可以确定 U 中某些元素是否属于 $\sim X$,但不能确定 U 中的任一元素是否属于 X.

如果 X 为 R 外不可定义的,则意味着我们可以确定 U 中某些元素是否属于 X,但不能确定 U 中的任一元素是否属于 $\sim X$.

如果 X 为 R 全不可定义,则意味着我们不能确定 U 中任一元素是否属于 X 或 $\sim X$.

下面我们给出集合拓扑划分的一个有用性质.

定理 1.4

(1)集合 X 为 R 粗糙可定义(或 R 全不可定义)当且仅当 $\sim X$ 为 R 粗糙可定义(或 R 全不可定义);

(2)集合 X 为 R 外(内)不可定义当且仅当 $\sim X$ 为 R 内(外)不可定义.

证 (1)设 X 为 R 粗糙可定义,则 $\underline{R}X \neq \varnothing$, $\overline{R}X \neq U$.

$\underline{R}X \neq \varnothing \Leftrightarrow$ 存在 $x \in X$,使得 $[x]_R \subseteq X$

$\Leftrightarrow [x]_R \cap \sim X = \varnothing \Leftrightarrow \overline{R}(\sim X) \neq U$;

类似地，

$$\overline{R}X \neq U \Leftrightarrow \text{存在 } y \in U\text{，使得}[y]_R \cap \overline{R}X = \varnothing$$
$$\Leftrightarrow [y]_R \subseteq \sim \overline{R}X \Leftrightarrow [y]_R \subseteq \underline{R}(\sim X) \Leftrightarrow \underline{R}(\sim X) \neq \varnothing.$$

同理可证其余情况. □

至此，我们已经介绍了两种刻画粗糙集的方法. 其一为用近似程度的精度来表示粗糙集的数字特征；其二为用粗糙集的分类表示粗糙集的拓扑特征. 粗糙集的数字特征表示了集合边界域的大小，但没有说明边界域的结构；而粗糙集的拓扑特征没有给出边界域大小的信息，它提供的是边界域的结构.

此外，粗糙集的数字特征和粗糙集的拓扑特征之间存在一种关系. 首先，如果集合为内不可定义或全不可定义，则其精度为 0；其次，当集合为外不可定义或全不可定义，则它的补集的精度为 0. 这样即使知道了集合的精度，我们也不能确定它的拓扑结构；反过来，集合的拓扑结构也不具备精度的信息.

因此在粗糙集的实际应用中，我们需要将边界域的两种信息结合起来，既要考虑精度因素，又要考虑到集合的拓扑结构.

例 1.2 给定一知识库 $K=(U,\mathbf{R})$ 和一个等价关系 $R \in \mathrm{ind}(K)$，其中 $U=\{x_0, x_1, \cdots, x_{10}\}$，且有 R 的下列等价类：

$$E_1 = \{x_0, x_1\},$$
$$E_2 = \{x_2, x_6, x_9\},$$
$$E_3 = \{x_3, x_5\},$$
$$E_4 = \{x_4, x_8\},$$
$$E_5 = \{x_7, x_{10}\}.$$

集合 $X_1=\{x_0, x_1, x_4, x_8\}$ 为 R 可定义集，因为

$$\underline{R}X_1 = \overline{R}X_1 = E_1 \cup E_4.$$

集合 $X_2=\{x_0, x_3, x_4, x_5, x_8, x_{10}\}$ 为 R 粗糙可定义集.

X_2 的近似集，边界和精度为

$\underline{R}X_2 = E_3 \cup E_4 = \{x_3, x_4, x_5, x_8\}$,

$\overline{R}X_2 = E_1 \cup E_3 \cup E_4 \cup E_5 = \{x_0, x_1, x_3, x_4, x_5, x_7, x_8, x_{10}\}$,

$\mathrm{bn}_R(X_2) = E_1 \cup E_5 = \{x_0, x_1, x_7, x_{10}\}$,

$\alpha_R(X_2) = 4/8 = 1/2$.

集合 $X_3=\{x_0, x_2, x_3\}$ 为 R 内不可定义. 因为

$$\underline{R}X_3 = \varnothing,$$
$$\overline{R}X_3 = E_1 \cup E_2 \cup E_3 = \{x_0, x_1, x_2, x_3, x_5, x_6, x_9\} \neq U.$$

集合 $X_4=\{x_0, x_1, x_2, x_3, x_4, x_7\}$ 为 R 外不可定义.

X_4 的近似集,边界和精度为

$$\underline{R}X_4=E_1=\{x_0,x_1\},$$
$$\overline{R}X_4=U,$$
$$\mathrm{bn}_R(X_4)=E_2\cup E_3\cup E_4\cup E_5$$
$$=\{x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\},$$
$$\alpha_R(X_3)=2/11.$$

集合 $X_5=\{x_0,x_2,x_3,x_4,x_7\}$ 为 R 全不可定义,因为

$$\underline{R}X_1=\varnothing,\qquad \overline{R}X_5=U.$$

下面讨论近似分类问题.

令 $\mathscr{F}=\{X_1,X_2,\cdots,X_n\}$ 是 U 的一个分类或划分,这个分类独立于知识 R.例如,它可能由一个专家为解决一个分类问题所给出.子集 $X_i(i=1,2,\cdots,n)$ 是划分 $\mathscr{F}$ 的类.$\mathscr{F}$ 的 R 下近似和上近似分别定义为:$\underline{R}\mathscr{F}=\{\underline{R}X_1,\underline{R}X_2,\cdots,\underline{R}X_n\}$ 和 $\overline{R}\mathscr{F}=\{\overline{R}X_1,\overline{R}X_2,\cdots,\overline{R}X_n\}$.

我们定义两个量度来描述近似分类的不精确性.

第一个量度为根据 R,$\mathscr{F}$ 的近似分类精度:

$$\alpha_R(\mathscr{F})=\frac{\sum_{i=1}^{n}|\underline{R}X_i|}{\sum_{i=1}^{n}|\overline{R}X_i|}.\tag{1.7}$$

第二个量度为根据 R,$\mathscr{F}$ 的近似分类质量:

$$\gamma_R(\mathscr{F})=\frac{\sum_{i=1}^{n}|\underline{R}X_i|}{|U|}.\tag{1.8}$$

近似分类的精度描述的是当使用知识 R 对对象分类时,可能的决策中正确决策的百分比;分类的质量表示的是应用知识 R 能确切地划入 $\mathscr{F}$ 类的对象的百分比.

粗糙集概念和通常的集合有着本质的区别.在集合的相等上也有一个重要的区别.在普通集合里,如果两个集合有完全相同的元素,则这两个集合相等.在粗糙集理论中,我们需要另一个集合相等的概念,即近似(粗糙)相等.两个集合在普通集合里不是相等的,但在粗糙集里可能是近似相等的.因为两个集合是否近似相等是根据我们得到的知识判断的.

现在我们介绍集合的三种近似相等的定义.

令 $K=(U,\mathbf{R})$ 是一个知识库,$X,Y\subseteq U$ 且 $R\in\mathrm{ind}(K)$.

(1)若 $\underline{R}X=\underline{R}Y$,则称集合 X 和 Y 为 R 下粗相等,记作 $X\eqsim_R Y$.

(2)若 $\overline{R}X=\overline{R}Y$,则称集合 X 和 Y 为 R 上粗相等,记作 $X\simeq_R Y$.

(3)若 $X \eqsim_R Y$ 且 $X \simeq_R Y$，则称集合 X 和 Y 为 R 相等，记作 $X \approx_R Y$.

易知，对于任何不可区分关系 R，$\eqsim_R$，$\simeq_R$ 和 $\approx_R$ 为等价关系.

由定义可以直接得到关系 $\eqsim_R$，$\simeq_R$ 和 $\approx_R$ 的下列基本性质：

定理 1.5 对于任何等价关系，有下列性质：

(1) $X \eqsim Y$ 当且仅当 $(X \cap Y) \eqsim X$ 且 $(X \cap Y) \eqsim Y$.

(2) $X \simeq Y$ 当且仅当 $(X \cup Y) \simeq X$ 且 $(X \cup Y) \simeq Y$.

(3) 若 $X \simeq X'$ 且 $Y \simeq Y'$，则 $X \cup Y \simeq X' \cup Y'$.

(4) 若 $X \eqsim X'$ 且 $Y \eqsim Y'$，则 $X \cap Y \eqsim X' \cap Y'$.

(5) 若 $X \simeq Y$；则 $X \cup \sim Y \simeq U$.

(6) 若 $X \eqsim Y$，则 $X \cap \sim Y \eqsim \varnothing$.

(7) 若 $X \subseteq Y$ 且 $Y \simeq \varnothing$，则 $X \simeq \varnothing$.

(8) 若 $X \subseteq Y$ 且 $X \simeq U$，则 $Y \simeq U$.

(9) $X \simeq Y$ 当且仅当 $\sim X \eqsim \sim Y$.

(10)若 $X \eqsim \varnothing$ 或 $Y \eqsim \varnothing$，则 $X \cap Y \eqsim \varnothing$.

(11)若 $X \simeq U$ 或 $Y \simeq U$，则 $X \cup Y \simeq U$.

为简单起见，我们省略了以上式子中的下标 R.

注 用 $\simeq$ 代替 $\eqsim$(或者反过来)上述性质不成立.

现在我们可以用集合的粗相等来表达集合的下近似和上近似.

定理 1.6 对于任何等价关系 R

(1) $\underline{R}X$ 是所有满足 $X \eqsim_R Y$ 的集合 $Y \subseteq U$ 的交集.

(2) $\overline{R}X$ 是所有满足 $X \simeq_R Y$ 的集合 $Y \subseteq U$ 的并集.

集合论的基本概念之一是包含关系. 在粗糙集框架下也可以引入类似的定义，这里集合的粗包含可用与集合的粗相等同样的方法定义如下：

令 $K=(U,\mathbf{R})$ 为一知识库；$X, Y \subseteq U$ 且 $R \in \mathrm{ind}(K)$.

(1) 若 $\underline{R}X \subseteq \underline{R}Y$，则称集合 X 为 R 下包含于 Y，记作 $X \underset{\sim}{\subset}_R Y$.

(2) 若 $\overline{R}X \subseteq \overline{R}Y$，则称集合 X 为 R 上包含于 Y，记作 $X \overset{\sim}{\subset}_R Y$.

(3) 若 $X \overset{\sim}{\subset}_R Y$ 且 $X \underset{\sim}{\subset}_R Y$，则称集合 X 为 R 包含于 Y，记作 $X \underset{\sim}{\overset{\sim}{\subset}}_R Y$.

易知，$\underset{\sim}{\subset}_R$，$\overset{\sim}{\subset}_R$ 和 $\underset{\sim}{\overset{\sim}{\subset}}_R$ 为拟序关系.

从定义可以直接推导出下列简单性质.

定理 1.7

(1) 若 $X \subseteq Y$，则 $X \underset{\sim}{\subset} Y$，$X \overset{\sim}{\subset} Y$ 且 $X \underset{\sim}{\overset{\sim}{\subset}} Y$.

(2) 若 $X \underset{\sim}{\subset} Y$ 且 $Y \underset{\sim}{\subset} X$，则 $X \eqsim Y$.

(3) 若 $X \overset{\sim}{\subset} Y$ 且 $Y \overset{\sim}{\subset} X$，则 $X \simeq Y$.

(4) 若 $X\bar{\underline{\subset}}Y$ 且 $Y\bar{\underline{\subset}}X$,则 $X\approx Y$.

(5) $X\bar{\subset}Y$ 当且仅当 $X\cup Y\simeq Y$.

(6) $X\underline{\subset}Y$ 当且仅当 $X\cap Y\eqsim X$.

(7) 若 $X\subseteq Y$,$X\eqsim X'$ 且 $Y\eqsim Y'$,则 $X'\underline{\subset}Y'$.

(8) 若 $X\subset Y$,$X\simeq X'$ 且 $Y\simeq Y'$,则 $X'\bar{\subset}Y'$.

(9) 若 $X\subset Y$,$X\approx X'$ 且 $Y\approx Y'$,则 $X'\bar{\underline{\subset}}Y'$.

(10)若 $X'\bar{\subset}X$ 且 $Y'\bar{\subset}Y$,则 $X'\cup Y'\bar{\subset}X\cup Y$.

(11)若 $X'\underline{\subset}X$ 且 $Y'\underline{\subset}Y$,则 $X'\cap Y'\underline{\subset}X\cap Y$.

(12)若 $X\underline{\subset}Y$ 且 $X\eqsim Z$,则 $Z\underline{\subset}Y$.

(13)若 $X\bar{\subset}Y$ 且 $X\simeq Z$,则 $Z\bar{\subset}Y$.

(14)若 $X\bar{\underline{\subset}}Y$ 且 $X\approx Z$,则 $Z\bar{\underline{\subset}}Y$.

为简单起见,我们省略了以上式子中的下标 R.

将粗糙集的概念与普通集合论相比较,可以看出粗糙集的基本性质,如元素的成员关系,集合的等价和包含,都与不可区分关系所表示的论域的知识有关.因此,一个元素是否属于某一集合,不是该元素的客观的性质,即取决于我们对它的了解程度.同样,集合的相等和包含没有绝对的意义,即取决于我们对所研究问题中的集合的了解程度.

§1.3 知识约简

知识约简是粗糙集理论的核心内容之一.众所周知,知识库中知识(属性)并不是同等重要的,甚至其中某些知识是冗余的.所谓知识约简,就是在保持知识库分类能力不变的条件下,删除其中不相关或不重要的知识.

知识约简中有两个基本概念:约简(reduct)和核(core).

在对约简和核进行讨论之前,我们先作如下定义:

令 $\mathbf{R}$ 为一族等价关系,$R\in\mathbf{R}$,如果

$$\mathrm{ind}(\mathbf{R})=\mathrm{ind}(\mathbf{R}-\{R\}),\tag{1.9}$$

则称 R 为 $\mathbf{R}$ 中不必要的;否则称 R 为 $\mathbf{R}$ 中必要的.

如果每一个 $R\in\mathbf{R}$ 都为 $\mathbf{R}$ 中必要的,则称 $\mathbf{R}$ 为独立的;否则称 $\mathbf{R}$ 为依赖的.

定理 1.8 如果 $\mathbf{R}$ 是独立的,$\mathbf{P}\subseteq\mathbf{R}$,则 $\mathbf{P}$ 也是独立的.

证 用反证法.假设 $\mathbf{P}\subseteq\mathbf{R}$ 且 $\mathbf{P}$ 是依赖的,则存在 $\mathbf{S}\subset\mathbf{P}$,使得

$$\mathrm{ind}(\mathbf{S}) = \mathrm{ind}(\mathbf{P}),$$

这意味着

$$\mathrm{ind}(\mathbf{S} \cup (\mathbf{R} - \boldsymbol{P})) = \mathrm{ind}(\mathbf{R}), \quad \mathbf{S} \cup (\mathbf{R} - \boldsymbol{P}) \subset \boldsymbol{R}.$$

因此 **R** 为依赖的，与已知条件矛盾，故定理得证. □

设 $\mathbf{Q} \subseteq \mathbf{P}$. 如果 **Q** 是独立的，且 $\mathrm{ind}(\mathbf{Q}) = \mathrm{ind}(\mathbf{P})$，则称 **Q** 为 **P** 的一个约简. 显然，**P** 可以有多种约简. **P** 中所有必要关系组成的集合称为 **P** 的核，记作 $\mathrm{core}(\mathbf{P})$.

核与约简有如下关系.

定理 1.9 $\mathrm{core}(\mathbf{P}) = \bigcap \mathrm{red}(\mathbf{P})$，其中 $\mathrm{red}(\mathbf{P})$ 表示 **P** 的所有约简.

证 (i) $\mathrm{core}(\mathbf{P}) \subseteq \bigcap \mathrm{red}(\mathbf{P})$. 设 $R \in \mathrm{core}(\mathbf{P})$，我们需要证明 $R \in \bigcap \mathrm{red}(\mathbf{P})$. 若 $R \notin \bigcap \mathrm{red}(\mathbf{P})$，则存在某个 $\mathbf{Q} \in \mathrm{red}(\mathbf{P})$ 使得 $R \notin \mathbf{Q}$. 从而有 $\mathbf{Q} \subseteq \mathbf{P} - \{R\}$，$\mathrm{ind}(\mathbf{Q}) \supseteq \mathrm{ind}(\mathbf{P} - \{R\}) \supseteq \mathrm{ind}(\mathbf{P})$.

另一方面，$\mathrm{ind}(\mathbf{P}) = \mathrm{ind}(\mathbf{Q})$，因为 **Q** 是 **P** 的一个约简. 这样，我们有 $\mathrm{ind}(\mathbf{P} - \{R\}) = \mathrm{ind}(\mathbf{P})$，这说明 R 在 **P** 中是不必要的，即 $R \notin \mathrm{core}(\mathbf{P})$，矛盾.

(ii) $\mathrm{core}(\mathbf{P}) \supseteq \bigcap \mathrm{red}(\mathbf{P})$. 设 $R \in \bigcap \mathrm{red}(\mathbf{P})$，我们需要证明 $R \in \mathrm{core}(\mathbf{P})$. 若 $R \notin \mathrm{core}(\mathbf{P})$，则 $\boldsymbol{R}$ 在 **P** 中是不必要的，于是有 $\mathrm{ind}(\mathbf{P}) = \mathrm{ind}(\mathbf{P} - \{R\})$. 另一方面，因为约简总是存在的，我们知道 $\mathbf{P} - \{R\}$ 有一个约简 $\mathbf{P}_0$. 从 $\mathbf{P}_0 \subseteq \mathbf{P} - \{R\}$ 知 $R \notin \mathbf{P}_0$. $\mathbf{P} - \{R\}$ 的约简 $\mathbf{P}_0$ 满足

(1) $\mathrm{ind}(\mathbf{P}_0) = \mathrm{ind}(\mathbf{P} - \{R\})$.

(2) 如果 $\mathbf{P}' \subset \mathbf{P}_0$，则 $\mathrm{ind}(\mathbf{P} - \{R\}) \subset \mathrm{ind}(\mathbf{P}')$.

这两个条件可以重写为

(1′) $\mathrm{ind}(\mathbf{P}_0) = \mathrm{ind}(\mathbf{P})$.

(2′) 如果 $\mathbf{P}' \subset \mathbf{P}_0$，则 $\mathrm{ind}(\mathbf{P}) \subset \mathrm{ind}(\mathbf{P}')$. 注意到 $\mathbf{P}_0 \subseteq \mathbf{P} - \{R\} \subset \mathbf{P}$. 我们知道 $\mathbf{P}_0$ 是 **P** 的一个约简，因此存在某个 $\mathbf{Q} \in \mathrm{red}(\mathbf{P})$，使得 $\mathbf{P}_0 = \mathbf{Q}$. 因此 $R \notin \mathbf{Q}$，从而 $R \notin \bigcap \mathrm{red}(\mathbf{P})$，矛盾. □

可以看出，核这个概念的用处有两个方面：首先它可以作为所有约简的计算基础，因为核包含在所有的约简之中，并且计算可以直接进行；其次可解释为在知识约简时它是不能消去的知识特征集合.

例 1.3 设 $K = (U, \mathbf{R})$ 是一个知识库，其中 $U = \{x_1, x_2, \cdots, x_8\}$，$\mathbf{R} = \{R_1, R_2, R_3\}$，等价关系 R_1，R_2 和 R_3 有下列等价类：

$U/R_1 = \{\{x_1, x_4, x_5\}, \{x_2, x_8\}, \{x_3\}, \{x_6, x_7\}\}$,

$U/R_2 = \{\{x_1, x_3, x_5\}, \{x_6\}, \{x_2, x_4, x_7, x_8\}\}$,

$U/R_3 = \{\{x_1, x_5\}, \{x_6\}, \{x_2, x_7, x_8\}, \{x_3, x_4\}\}$.

关系 $\mathrm{ind}(\mathbf{R})$ 有下列等价类：

$U/\mathrm{ind}(\mathbf{R}) = \{\{x_1, x_5\}, \{x_2, x_8\}, \{x_3\}, \{x_4\}, \{x_6\}, \{x_7\}\}$.

关系 R_1 为 $\mathbf{R}$ 中必要的,因为

$$U/\mathrm{ind}(\mathbf{R}-\{R_1\})=\{\{x_1,x_5\},\{x_2,x_7,x_8\},\{x_3\},\{x_4\},\{x_6\}\}\neq U/\mathrm{ind}(\mathbf{R}).$$

对于关系 R_2,我们有

$$U/\mathrm{ind}(\mathbf{R}-\{R_2\})=\{\{x_1,x_5\},\{x_2,x_8\},\{x_3\},\{x_4\},\{x_6\},\{x_7\}\}=U/\mathrm{ind}(\mathbf{R}),$$

故关系 R_2 是 $\mathbf{R}$ 中不必要的.

同样,对于关系 R_3,

$$U/\mathrm{ind}(\mathbf{R}-\{R_3\})=\{\{x_1,x_5\},\{x_2,x_8\},\{x_3\},\{x_4\},\{x_6\},\{x_7\}\}=U/\mathrm{ind}(\mathbf{R}),$$

因此关系 R_3 也是 $\mathbf{R}$ 中不必要的.

这表明通过等价关系 R_1,R_2 和 R_3 的集合定义的分类与根据 R_1 和 R_2 或 R_1 和 R_3 定义的分类相同,即表明该系统的知识可以通过 $U/\mathrm{ind}(\{R_1,R_2\})$或 $U/\mathrm{ind}(\{R_1,R_3\})$来表达.

为了得到 $\mathbf{R}=\{R_1,R_2,R_3\}$的约简,我们检验$\{R_1,R_2\}$和$\{R_1,R_3\}$是否为独立的,因为 $U/\mathrm{ind}(\{R_1,R_2\})\neq U/\mathrm{ind}(R_1)$,且 $U/\mathrm{ind}(\{R_1,R_2\})\neq U/\mathrm{ind}(R_2)$,因此$\{R_1,R_2\}$为独立的,且$\{R_1,R_2\}$为 $\mathbf{R}$ 的一个约简.同理,$\{R_1,R_3\}$也是 $\mathbf{R}$ 的一个约简.

这样 $\mathbf{R}$ 有两个约简,即$\{R_1,R_2\}$和$\{R_1,R_3\}$,一个核 $\mathrm{core}(\mathbf{R})=\{R_1,R_2\}\cap\{R_1,R_3\}=\{R_1\}$.

在应用中,一个分类相对于另一个分类的关系十分重要,因此我们将介绍知识的相对约简(relative reduct)和相对核(relative core)的概念.首先我们定义一个分类相对于另一个分类的正域.

令 P 和 Q 为 U 中的等价关系,Q 的 P 正域记为 $\mathrm{pos}_P(Q)$,即

$$\mathrm{pos}_P(Q)=\bigcup_{X\in U/Q}\underline{P}X.$$

Q 的 P 正域是 U 中所有根据分类 U/P 的信息可以准确地划分到关系 Q 的等价类中去的对象集合.

令 $\mathbf{P}$ 和 $\mathbf{Q}$ 为等价关系族,$R\in\mathbf{P}$,如果

$$\mathrm{pos}_{\mathrm{ind}(\mathbf{P})}(\mathrm{ind}(\mathbf{Q}))=\mathrm{pos}_{\mathrm{ind}(\mathbf{P}-\{R\})}(\mathrm{ind}(\mathbf{Q})),\tag{1.10}$$

则称 R 为 $\mathbf{P}$ 中 $\mathbf{Q}$ 不必要的;否则 R 为 $\mathbf{P}$ 中 $\mathbf{Q}$ 必要的.

为简单起见,我们也用 $\mathrm{pos}_{\mathbf{P}}(\mathbf{Q})$代替 $\mathrm{pos}_{\mathrm{ind}(\mathbf{P})}(\mathrm{ind}(\mathbf{Q}))$.

如果 $\mathbf{P}$ 中的每个 R 都为 $\mathbf{Q}$ 必要的,则称 $\mathbf{P}$ 为 $\mathbf{Q}$ 独立的(或 $\mathbf{P}$ 相对于 $\mathbf{Q}$ 独立).

设 $\mathbf{S}\subseteq\mathbf{P}$, $\mathbf{S}$ 为 $\mathbf{P}$ 的 $\mathbf{Q}$ 约简当且仅当 $\mathbf{S}$ 是 $\mathbf{P}$ 的 $\mathbf{Q}$ 独立子族且 $\mathrm{pos}_{\mathbf{S}}(\mathbf{Q})=$

$\mathrm{pos}_{\mathbf{P}}(\mathbf{Q})$. $\mathbf{P}$ 的 $\mathbf{Q}$ 约简简称为相对约简.

$\mathbf{P}$ 中所有 $\mathbf{Q}$ 必要的原始关系构成的集合称为 $\mathbf{P}$ 的 $\mathbf{Q}$ 核,简称为相对核,记为 $\mathrm{core}_{\mathbf{Q}}(\mathbf{P})$.

相对核与相对约简的关系见下述定理.

定理 1.10 $\mathrm{core}_{\mathbf{Q}}(\mathbf{P})=\bigcap\mathrm{red}_{\mathbf{Q}}(\mathbf{P})$,其中 $\mathrm{red}_{\mathbf{Q}}(\mathbf{P})$是所有 $\mathbf{P}$ 的 $\mathbf{Q}$ 约简构成的集合.

这个定理的证明与定理 1.9 的证明类似. □

例 1.4 设 $K=(U,\mathbf{P})$是一个知识库,其中 $U=\{x_1,x_2,\cdots,x_8\}$,$\mathbf{P}=\{R_1,R_2,R_3\}$,等价关系 R_1,R_2 和 R_3 有下列等价类:

$$U/R_1=\{\{x_1,x_3,x_4,x_5,x_6,x_7\},\{x_2,x_8\}\},$$
$$U/R_2=\{\{x_1,x_3,x_4,x_5\},\{x_2,x_6,x_7,x_8\}\},$$
$$U/R_3=\{\{x_1,x_5,x_6\},\{x_2,x_7,x_8\},\{x_3,x_4\}\}.$$

由 P 导出的分类为

$$U/\mathrm{ind}(\mathbf{P})=\{\{x_1,x_5\},\{x_3,x_4\},\{x_2,x_8\},\{x_6\},\{x_7\}\}.$$

进一步假设等价关系 Q 有下列等价类:

$$U/Q=\{\{x_1,x_5,x_6\},\{x_3,x_4\},\{x_2,x_7\},\{x_8\}\}.$$

Q 的 $\mathbf{P}$ 正域为

$$\mathrm{pos}_{\mathbf{P}}(Q)=\{x_1,x_5\}\cup\{x_3,x_4\}\cup\{x_6\}\cup\{x_7\}=\{x_1,x_3,x_4,x_5,x_6,x_7\}.$$

现在从 P 中去掉R_1 得到

$$\begin{aligned}U/(\mathbf{P}-\{R_1\})&=U/\{R_2,R_3\}\\&=\{\{x_1,x_5\},\{x_3,x_4\},\{x_2,x_7,x_8\},\{x_6\}\}.\end{aligned}$$

因为

$$\begin{aligned}\mathrm{pos}_{(\mathbf{P}-\{R_1\})}(Q)&=\{x_1,x_5\}\cup\{x_3,x_4\}\cup\{x_6\}\\&=\{x_1,x_3,x_4,x_5,x_6\}\neq\mathrm{pos}_{\mathbf{P}}(Q),\end{aligned}$$

故 R_1 是 $\mathbf{P}$ 中 Q 必要的.

从 $\mathbf{P}$ 中去掉 R_2 得到

$$\begin{aligned}U/(\mathbf{P}-\{R_2\})&=U/\{R_1,R_3\}\\&=\{\{x_1,x_5,x_6\},\{x_3,x_4\},\{x_2,x_8\},\{x_7\}\}.\end{aligned}$$

由此导出正域

$$\begin{aligned}\mathrm{pos}_{(\mathbf{P}-\{R_2\})}(Q)&=\{x_1,x_5,x_6\}\cup\{x_3,x_4\}\cup\{x_7\}\\&=\{x_1,x_3,x_4,x_5,x_6,x_7\}=\mathrm{pos}_{\mathbf{P}}(Q),\end{aligned}$$

因此 R_2 为 $\mathbf{P}$ 中 Q 不必要的.

最后省略 $\mathbf{P}$ 中的 R_3 得到

$$U/(\mathbf{P}-\{R_3\})=U/\{R_1,R_2\}$$

$$=\{\{x_1,x_3,x_4,x_5\},\{x_2,x_8\},\{x_6,x_7\}\}.$$

正域为

$$\mathrm{pos}_{(\mathbf{P}-\{R_3\})}(Q)=\varnothing\neq\mathrm{pos}_{\mathbf{P}}(Q),$$

因此 R_3 为 **P** 中 Q 必要的.

这样 **P** 的 Q 核为 $\{R_1,R_3\}$,它也是 **P** 的 Q 约简.

§1.4 知识的依赖性

知识的依赖性可形式化地定义如下:令 $K=(U,\mathbf{R})$ 是一个知识库,$\mathbf{P},\mathbf{Q}\subseteq\mathbf{R}$.

(1)知识 **Q** 依赖于知识 **P**(记作 $\mathbf{P}\Rightarrow\mathbf{Q}$)当且仅当 $\mathrm{ind}(\mathbf{P})\subseteq\mathrm{ind}(\mathbf{Q})$.

(2)知识 **P** 与知识 **Q** 等价(记作 $\mathbf{P}\equiv\mathbf{Q}$)当且仅当 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{Q}\Rightarrow\mathbf{P}$.

(3)知识 **P** 与知识 **Q** 独立(记作 $\mathbf{P}\not\equiv\mathbf{Q}$)当且仅当 $\mathbf{P}\Rightarrow\mathbf{Q}$ 与 $\mathbf{Q}\Rightarrow\mathbf{P}$ 均不成立.

显然,$\mathbf{P}\equiv\mathbf{Q}$ 当且仅当 $\mathrm{ind}(\mathbf{P})=\mathrm{ind}(\mathbf{Q})$.

当知识 **Q** 依赖于知识 **P** 时,我们也说知识 **Q** 是由知识 **P** 导出的.

我们用下面的例子说明知识的依赖性.

例 1.5 假设知识 **P** 和知识 **Q** 有下列分类:

$$U/\mathbf{P}=\{\{x_1,x_5\},\{x_2,x_8\},\{x_3\},\{x_4\},\{x_6\},\{x_7\}\},$$
$$U/\mathbf{Q}=\{\{x_1,x_5\},\{x_2,x_7,x_8\},\{x_3,x_4,x_6\}\}.$$

可见,$\mathrm{ind}(\mathbf{P})\subseteq\mathrm{ind}(\mathbf{Q})$,因此有 $\mathbf{P}\Rightarrow\mathbf{Q}$.

通过简单推导可得下列性质:

定理 1.11 下列条件是等价的:

(1) $\mathbf{P}\Rightarrow\mathbf{Q}$.

(2) $\mathrm{ind}(\mathbf{P}\cup\mathbf{Q})=\mathrm{ind}(\mathbf{P})$.

(3) $\mathrm{pos}_{\mathbf{P}}(\mathbf{Q})=U$.

(4)对于所有 $X\in U/\mathbf{Q}$,有 $\underline{\mathbf{P}}X=X$,其中 $\underline{\mathbf{P}}X$ 表示 $\underline{\mathrm{ind}(\mathbf{P})}X$.

定理 1.11 表明,如果知识 **Q** 依赖于知识 **P**,则在知识库中,知识 **Q** 是多余的.在这种情况下,知识 $\mathbf{P}\cup\mathbf{Q}$ 与知识 **P** 提供同样的对象特征.

下面是依赖性的一些重要性质.

定理 1.12

(1)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{Q}\Rightarrow\mathbf{R}$,则 $\mathbf{P}\Rightarrow\mathbf{R}$.

(2)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{Q}\Rightarrow\mathbf{R}$,则 $\mathbf{P}\cup\mathbf{Q}\Rightarrow\mathbf{R}$.

(3)如果 $\mathbf{P}\Rightarrow\mathbf{R}\cup\mathbf{Q}$,则 $\mathbf{P}\Rightarrow\mathbf{R}$ 且 $\mathbf{P}\Rightarrow\mathbf{Q}$.

(4)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{Q}\cup\mathbf{R}\Rightarrow\mathbf{T}$,则 $\mathbf{P}\cup\mathbf{R}\Rightarrow\mathbf{T}$.

(5)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{R}\Rightarrow\mathbf{T}$,则 $\mathbf{P}\cup\mathbf{R}\Rightarrow\mathbf{Q}\cup\mathbf{T}$.

(6)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{P}'\supset\mathbf{P}$,则 $\mathbf{P}'\Rightarrow\mathbf{Q}$.

(7)如果 $\mathbf{P}\Rightarrow\mathbf{Q}$ 且 $\mathbf{Q}'\subset\mathbf{Q}$,则 $\mathbf{P}\Rightarrow\mathbf{Q}'$.

有时候知识的依赖性可能是部分的,这意味着知识 Q 仅有部分是由知识 P 导出的,部分可导出可由知识的正域来定义.

令 $K=(U,\mathbf{R})$为一知识库,且 $\mathbf{P},\mathbf{Q}\subseteq\mathbf{R}$.当

$$k = \gamma_{\mathbf{P}}(\mathbf{Q}) = |\mathrm{pos}_{\mathbf{P}}(\mathbf{Q})| / |U| \tag{1.11}$$

时,我们称知识 Q 是 $k(0\leqslant k\leqslant 1)$度依赖于知识 P 的,记作 $\mathbf{P}\Rightarrow_k\mathbf{Q}$.

当 $k=1$ 时,我们称 Q 完全依赖于 P;当 $0<k<1$ 时,称 Q 粗糙(部分)依赖于 P;当 $k=0$ 时,称 Q 完全独立于 P.如果 $\mathbf{P}\Rightarrow_1\mathbf{Q}$,也记为 $\mathbf{P}\Rightarrow\mathbf{Q}$.

由依赖性的定义可知,当 $\mathbf{P}\Rightarrow_k\mathbf{Q}$ 时,则由 Q 导出的分类 $U/\mathbf{Q}$ 的正域覆盖了知识库的 $k\times100\%$ 个元素;另一方面,只有属于分类正域的元素能被惟一的分类,即对象的 $k\times100\%$ 可以通过知识 P 划入分类 $U/\mathbf{Q}$ 的模块中.

系数 $\gamma_{\mathbf{P}}(\mathbf{Q})$可以看作 Q 和 P 间的依赖度.

当然,还可以使用粗糙依赖的另外量度,但这里的量度看来更适合于各种应用,并且也易于计算和解释.

部分依赖性 $\mathbf{P}\Rightarrow_k\mathbf{Q}$ 的量度 k 不能够完全反映 $U/\mathbf{Q}$ 中类之间的分布情况.例如,一些决策类可能完全由 P 描述,但另一些可能仅仅由 P 部分描述.因此,我们需要使用一个系数 $\gamma_{\mathbf{P}}(X)=|\underline{\mathbf{P}}X|/|X|,(X\in U/\mathbf{Q})$来表明通过知识 P 能将 $U/\mathbf{Q}$ 中每个类的多少个元素正确划分.

这样,两个值 $\gamma_{\mathbf{P}}(\mathbf{Q})$和 $\gamma_{\mathbf{P}}(X),X\in U/\mathbf{Q}$ 给出了知识 P 的"分类能力"关于分类 $U/\mathbf{Q}$ 的全部信息.

例 1.6 计算知识 Q 和知识 P 间的依赖性量度.假定 $U=\{1,2,\cdots,8\}$;$U/\mathbf{Q}=\{X_1,X_2,X_3,X_4,X_5\}$,其中 $X_1=\{1\},X_2=\{2,7\},X_3=\{3,6\},X_4=\{4\},X_5=\{5,8\}$;$U/\mathbf{P}=\{Y_1,Y_2,Y_3,Y_4,Y_5,Y_6\}$,其中 $Y_1=\{1,5\},Y_2=\{2,8\},Y_3=\{3\},Y_4=\{4\},Y_5=\{6\},Y_6=\{7\}$.

因为 $\underline{\mathbf{P}}X_1=\varnothing,\underline{\mathbf{P}}X_2=Y_6,\underline{\mathbf{P}}X_3=Y_3\cup Y_5,\underline{\mathbf{P}}X_4=Y_4,\underline{\mathbf{P}}X_5=\varnothing$,因此,$\mathrm{pos}_{\mathbf{P}}(\mathbf{Q})=Y_3\cup Y_4\cup Y_5\cup Y_6=\{3,4,6,7\}$.这就是说,只有这些元素可以通过知识 P 划入到分类 $U/\mathbf{Q}$ 的模块.因此 Q 和 P 间的依赖度为

$$r_{\mathbf{P}}(\mathbf{Q}) = 4/8 = 0.5.$$

§1.5 知识表达系统

知识表达在智能数据处理中占有十分重要的地位.

形式上,四元组 $S=(U,A,V,f)$是一个知识表达系统,其中

U:对象的非空有限集合,称为论域;

A:属性的非空有限集合;

$V=\bigcup_{a\in A}V_a$, V_a 是属性 a 的值域;

$f:U\times A\rightarrow V$ 是一个信息函数,它为每个对象的每个属性赋予一个信息值,即 $\forall a\in A, x\in U, f(x,a)\in V_a$.

知识表达系统也称为信息系统. 通常也用 $S=(U,A)$来代替 $S=(U,A,V,f)$.

知识表达系统的数据以关系表的形式表示. 关系表的行对应要研究的对象,列对应对象的属性,对象的信息是通过指定对象的各属性值来表达.

容易看出,一个属性对应一个等价关系,一个表可以看作是定义的一族等价关系,即知识库. 前几节讨论的问题都可以用属性及属性值引入的分类来表示,知识约简可转化为属性约简.

例 1.7 一个关于某些病人的知识表达系统如表 1.1,其中

表 1.1

病人	头痛	肌肉痛	体温
e_1	是	是	正常
e_2	是	是	高
e_3	是	是	很高
e_4	否	是	正常
e_5	否	否	高
e_6	否	是	很高

$U=\{e_1,e_2,e_3,e_4,e_5,e_6\}$, $A=\{$头痛,肌肉痛,体温$\}$.

令 $P\subseteq A$,定义属性集 P 的不可区分关系 $\mathrm{ind}(P)$为

$$\mathrm{ind}(P)=\{(x,y)\in U\times U \mid \forall a\in P, f(x,a)=f(y,a)\}.$$

如果$(x,y)\in\mathrm{ind}(P)$,则称 x 和 y 是 P 不可区分的. 容易证明 $\forall P\subseteq A$,不可区分关系 $\mathrm{ind}(P)$是 U 上的等价关系,符号 $U/\mathrm{ind}(P)$(简记为 U/P)表示不可区分关系 $\mathrm{ind}(P)$在 U 上导出的划分,$\mathrm{ind}(P)$中的等价类称为 P 基本集. 符号$[x]_P$ 表示包含 $x\in U$ 的 P 等价类.

在不产生混淆情况下,我们也用 P 来代替 $\mathrm{ind}(P)$.

在例 1.7 中,若取属性集 $P=\{$头痛,肌肉痛$\}$, $X=\{e_2,e_4,e_6\}$,则有

$$U/P=\{\{e_1,e_2,e_3\},\{e_4,e_6\},\{e_5\}\}.$$

P 基本集为$\{e_1,e_2,e_3\}$,$\{e_4,e_6\}$,$\{e_5\}$.

$$\underline{P}X=\mathrm{pos}_P(X)=\{e_4,e_6\},\overline{P}X=\{e_1,e_2,e_3,e_4,e_6\},$$

$$\mathrm{neg}_P(X)=U-\overline{P}X=\{e_5\},\mathrm{bn}_P(X)=\overline{P}X-\underline{P}X=\{e_1,e_2,e_3\}.$$

属性集{头痛,肌肉痛,体温}有一个约简{头痛,体温},{头痛,体温}亦为核.

§1.6 决策表

决策表是一类特殊而重要的知识表达系统.多数决策问题都可以用决策表形式来表达,这一工具在决策应用中起着重要的作用.

决策表可以根据知识表达系统定义如下.

表 1.2

条件属性				决策属性
病人	头痛	肌肉痛	体温	流感
e_1	是	是	正常	否
e_2	是	是	高	是
e_3	是	是	很高	是
e_4	否	是	正常	否
e_5	否	否	高	否
e_6	否	是	很高	是
e_7	否	否	高	是
e_8	否	是	很高	否

设 $S=(U,A,V,f)$为一知识表达系统,$A=C\cup D$,$C\cap D=\varnothing$,C 称为条件属性集,D 称为决策属性集.具有条件属性和决策属性的知识表达系统称为决策表.

例 1.8 一个关于某些病人的决策如表 1.2.其中 $U=\{e_1,e_2,\cdots,e_8\}$,$C=\{$头痛,肌肉痛,体温$\}$,$D=\{$流感$\}$.

令 C_1=头痛,C_2=肌肉痛,C_3=体温,则

$$U/\{C_1\}=\{\{e_1,e_2,e_3\},\{e_4,e_5,e_6,e_7,e_8\}\},$$

$$U/\{C_2\}=\{\{e_1,e_2,e_3,e_4,e_6,e_8\},\{e_5,e_7\}\},$$

$$U/\{C_3\} = \{\{e_1,e_4\},\{e_2,e_5,e_7\},\{e_3,e_6,e_8\}\};$$
$$U/\{C_1,C_2\} = \{\{e_1,e_2,e_3\},\{e_4,e_6,e_8\},\{e_5,e_7\}\},$$
$$U/\{C_1,C_3\} = \{\{e_1\},\{e_2\},\{e_3\},\{e_4\},\{e_5,e_7\},\{e_6,e_8\}\},$$
$$U/\{C_2,C_3\} = \{\{e_1,e_4\},\{e_2\},\{e_5,e_7\},\{e_3,e_6,e_8\}\};$$
$$U/C = \{\{e_1\},\{e_2\},\{e_3\},\{e_4\},\{e_5,e_7\},\{e_6,e_8\}\},$$
$$U/D = \{\{e_2,e_3,e_6,e_7\},\{e_1,e_4,e_5,e_8\}\}.$$

因为

$$\mathrm{pos}_C(D) = \{e_1\} \cup \{e_2\} \cup \{e_3\} \cup \{e_4\} = \{e_1,e_2,e_3,e_4\},$$
$$k = \gamma_C(D) = |\mathrm{pos}_C(D)| / |U| = 4/8 = 0.5,$$

所以 D 部分依赖(依赖度为 0.5)于 C. 又因为

$$\mathrm{pos}_{(C-\{C_1\})}(D) = \{e_1,e_2,e_4\} \neq \mathrm{pos}_C(D),$$
$$\mathrm{pos}_{(C-\{C_2\})}(D) = \{e_1,e_2,e_3,e_4\} = \mathrm{pos}_C(D),$$
$$\mathrm{pos}_{(C-\{C_3\})}(D) = \varnothing \neq \mathrm{pos}_C(D);$$
$$\mathrm{pos}_{(C-\{C_2,C_1\})}(D) = \{e_1,e_4\} \neq \mathrm{pos}_C(D),$$
$$\mathrm{pos}_{(C-\{C_2,C_3\})}(D) = \varnothing \neq \mathrm{pos}_C(D).$$

所以 C 的 D 约简(相对约简)为 $C-\{C_2\}=\{C_1,C_3\}$, C 的 D 核(相对核)也为 $\{C_1,C_3\}$.

在决策表中,不同的属性可能具有不同的重要性. 例如,当由症状描述病人的情况时,对于识别病人的身体状况有些症状具有更重要的意义.

为了找出某些属性(或属性集)的重要性(significance),我们的方法是从表中去掉一些属性,再来考察没有该属性后分类会怎样变化. 若去掉该属性相应分类变化较大,则说明该属性的强度大,即重要性高;反之,说明该属性的强度小,即重要性低.

令 C 和 D 分别为条件属性集和决策属性集,属性子集 $C'\subseteq C$ 关于 D 的重要性定义为

$$\sigma_{CD}(C') = \gamma_C(D) - \gamma_{C-C'}(D).$$

特别当 $C'=\{a\}$ 时,属性 $a\in C$ 关于 D 的重要性为

$$\sigma_{CD}(a) = \gamma_C(D) - \gamma_{C-\{a\}}(D).$$

例 1.9 在例 1.8 中,我们有

$$\sigma_{CD}(\text{头痛}) = 4/8 - 3/8 = 1/8,$$
$$\sigma_{CD}(\text{肌肉痛}) = 4/8 - 4/8 = 0,$$
$$\sigma_{CD}(\text{体温}) = 4/8 - 0 = 4/8.$$

由此可见，在决策表 1.2 中，{体温}最重要，其次是{头痛}，{肌肉痛}是不重要的.

在决策表中，最重要的是决策规则的产生.

设 $S=(U,A,V,f)$是一个决策表，$A=C\cup D$，$C\cap D=\varnothing$，其中 C 为条件属性集，D 为决策属性集. 令 X_i 和 Y_j 分别代表 U/C 与 U/D 中的各个等价类，$\mathrm{des}(X_i)$表示对等价类 X_i 的描述，即等价类 X_i 对于各条件属性值的特定取值；$\mathrm{des}(Y_j)$表示对等价类 Y_j 的描述，即等价类 Y_j 对于各决策属性值的特定取值.

决策规则定义如下：

$$r_{ij}:\mathrm{des}(X_i)\rightarrow \mathrm{des}(Y_j),\quad Y_j\cap X_i\neq\varnothing,$$

规则的确定性因子 $\mu(X_i,Y_j)=|Y_j\cap X_i|/|X_i|$，$0<\mu(X_i,Y_j)\leqslant 1$.

当 $\mu(X_i,Y_j)=1$ 时，r_{ij}是确定的；当 $0<\mu(X_i,Y_j)<1$ 时，r_{ij}是不确定的.

注 在产生决策规则之前，可首先对决策表中的属性进行约简.

例 1.10 在例 1.8 中，我们对表 1.2 进行属性约简可得表 1.3. 这里 $U=\{e_1,e_2,\cdots,e_8\}$，$C=\{$头痛，体温$\}$，$D=\{$流感$\}$.

$$U/C=\{X_1,X_2,X_3,X_4,X_5,X_6\},$$

其中 $X_1=\{e_1\}$，$X_2=\{e_2\}$，$X_3=\{e_3\}$，$X_4=\{e_4\}$，$X_5=\{e_5,e_7\}$，$X_6=\{e_6,e_8\}$.

$$U/D=\{Y_1,Y_2\},$$

其中 $Y_1=\{e_2,e_3,e_6,e_7\}$，$Y_2=\{e_1,e_4,e_5,e_8\}$.

确定性规则有

r_{12}：(头痛，是)且(体温，正常)→(流感，否).

r_{21}：(头痛，是)且(体温，高)→(流感，是).

r_{31}：(头痛，是)且(体温，很高)→(流感，是).

r_{42}：(头痛，否)且(体温，正常)→(流感，否).

不确定性规则有

r_{51}：(头痛，否)且(体温，高)→(流感，是)，规则的确定性因子为 0.5.

r_{52}：(头痛，否)且(体温，高)→(流感，否)，规则的确定性因子为 0.5.

r_{61}：(头痛，否)且(体温，很高)→(流感，是)，规则的确定性因子为 0.5.

r_{62}：(头痛，否)且(体温，很高)→(流感，否)，规则的确定性因子为 0.5.

表 1.3

	条件属性		决策属性
病人	头痛	体温	流感
e_1	是	正常	否
e_2	是	高	是
e_3	是	很高	是
e_4	否	正常	否
e_5	否	高	否
e_6	否	很高	是
e_7	否	高	是
e_8	否	很高	否

§1.7 区分矩阵与区分函数

利用区分矩阵(discernibility matrix)来表达知识有许多优点,特别是它能容易地计算约简和核.

令 $S=(U,A,V,f)$是一个知识表达系统,$|U|=n$. S 的区分矩阵是一个 $n\times n$ 矩阵,其任一元素为

$$\alpha(x,y) = \{a \in A \mid f(x,a) \neq f(y,a)\}. \tag{1.12}$$

因此,$\alpha(x,y)$是区别对象 x 和 y 的所有属性的集合.

下面我们引入一个布尔函数,称其为区分函数(discernibility function),用△表示. 对每个属性 $a\in A$,我们指定一个布尔变量"a". 若 $\alpha(x,y)=\{a_1, a_2,\cdots,a_k\}\neq\varnothing$,则指定一个布尔函数 $a_1\vee a_2\vee\cdots\vee a_k$,用$\sum\alpha(x,y)$来表示;若 $\alpha(x,y)=\varnothing$,则指定布尔常量 1.(布尔)区分函数△可定义如下:

$$\triangle = \prod_{(x,y)\in U\times U} \sum\alpha(x,y). \tag{1.13}$$

现在介绍几个与布尔函数有关的概念.

(1)布尔表达式 E 是一个范式,如果 E 仅仅由布尔变量和常量通过析取与合取运算来表达.

(2)布尔表达式 E 是一个合取范式,如果 E 是由一些析取式组成的合取所构成的范式.

(3)布尔表达式 E 是一个析取范式,如果 E 是由一些合取式组成的析取所构成的范式.

(4)布尔表达式 E 是一个极小析取范式,如果 E 是一个析取范式且包含最小数目的合取式.

区分函数△有如下性质:函数△的极小析取范式中的所有合取式是属性

集 A 的所有约简.

换句话说,约简是满足能区别由整个属性集区别的所有对象的属性极小子集.

易知,如果 $B \subset A$ 是满足条件

$$B \cap \alpha(x,y) \neq \varnothing, \quad \forall \alpha(x,y) \neq \varnothing \tag{1.14}$$

的极小子集(关于包含),则 B 是 A 的一个约简.

核是区分矩阵中所有单个元素组成的集合,即

$$\mathrm{core}(A) = \{a \in A \mid \alpha(x,y) = \{a\}, \text{其中 } x,y \in U\}. \tag{1.15}$$

例 1.11 考虑表 1.4 给出的知识表达系统. 表 1.4 对应的区分矩阵由表 1.5 给出.

表 1.4

U	a	b	c	d
1	0	1	2	0
2	1	2	0	2
3	1	0	1	0
4	2	1	0	1
5	1	1	0	2

注 区分矩阵是对称矩阵,因此,我们仅需计算矩阵的一半元素.

表 1.5 对应的区分函数为

$$\begin{aligned}\triangle &= (a \vee b \vee c \vee d)(a \vee b \vee c)(a \vee c \vee d)(a \vee c \vee d)\\ &\quad \cdot (b \vee c \vee d)(a \vee b \vee d)b(a \vee b \vee c \vee d)\\ &\quad \cdot (b \vee c \vee d)(a \vee d)\\ &= ab \vee bd.\end{aligned}$$

因此,这个知识表达系统有两个约简 $\{a,b\}$ 和 $\{b,d\}$,核是 $\{b\}$.

对于决策表 $S=(U,A,V,f)$,$A = C \cup D$,$C \cap D = \varnothing$,C 为条件属性集,D 为决策属性集,我们可用类似的方法计算其相对约简和相对核.

表 1.5

	1	2	3	4	5
1					
2	a,b,c,d				
3	a,b,c	b,c,d			
4	a,c,d	a,b,d	a,b,c,d		
5	a,c,d	b	b,c,d	a,d	

决策表 S 的区分矩阵是一个 $n\times n$ 矩阵，其任一元素为

$$\alpha^*(x,y)=\{a\in C\mid f(x,a)\neq f(y,a)\text{ 且 }w(x,y)\}$$

对于 $x,y\in U$，$w(x,y)$满足

$$x\in \mathrm{pos}_C(D)\text{ 且 }y\notin \mathrm{pos}_C(D),$$

或者

$$x\notin \mathrm{pos}_C(D)\text{ 且 }y\in \mathrm{pos}_C(D),$$

或者

$$x,y\in \mathrm{pos}_C(D)\text{ 且}(x,y)\notin \mathrm{ind}(D).$$

决策表 S 的区分函数$\triangle^*$定义为

$$\triangle^*=\prod_{(x,y)\in U\times U}\sum\alpha^*(x,y).$$

区分函数$\triangle^*$有如下性质：函数$\triangle^*$的极小析取范式中的所有合取式是 C 的所有 D 约简.

易知，如果 $C'\subseteq C$ 是满足条件

$$C'\cap\alpha^*(x,y)\neq\varnothing,\qquad \forall\alpha^*(x,y)\neq\varnothing$$

的极小子集(关于包含)，则 C'是C 的D 约简(相对约简).

D 核(相对核)是决策表 S 的区分矩阵中所有单个元素组成的集合，即

$$\mathrm{core}_D(C)=\{a\in C\mid\alpha^*(x,y)=\{a\},\text{其中 }x,y\in U\}.\quad(1.16)$$

例 1.12　考虑表 1.6 给出的决策表. 这里 $C=\{a,b,c\}$，$D=\{d\}$. 决策表 1.6 对应的区分矩阵由表 1.7 给出. 表 1.7 对应的区分函数为

$$\begin{aligned}\triangle^*&=a(a\vee b)(a\vee b\vee c)(a\vee b)(a\vee b\vee c)\\&\quad\cdot(a\vee b)(b\vee c)(a\vee c)a\\&=ab\vee ac.\end{aligned}$$

因此，这个决策表有两个 D 约简$\{a,b\}$和$\{a,c\}$，D 核是$\{a\}$.

表 1.6

U	a	b	c	d
1	2	2	0	1
2	1	2	0	0
3	1	2	0	1
4	0	0	0	0
5	1	0	1	0
6	2	0	1	1

表 1.7

	1	2	3	4	5	6
1						
2	a					
3	a					
4	a,b	a,b	a,b			
5	a,b,c	b,c	b,c			
6		a,b,c	a,b,c	a,c	a	

第二章　粗糙集模型的算法

粗糙集方法已被成功地应用于分类问题和决策分析(见第六章). 对于粗糙集理论的应用而言,设计有效的算法是非常重要的. 本章主要介绍基于决策表的粗糙集模型的算法,并分析每种算法的时间复杂性.

§2.1　信息系统和决策表

一个信息系统 S 是一个系统(U,A),其中 $U=\{u_1,u_2,\cdots,u_{|U|}\}$是有限非空集,称为论域或对象空间,$U$ 中的元素称为对象;$A=\{a_1,a_2,\cdots a_{|A|}\}$也是一个有限非空集,$A$ 中的元素称为属性;对于每个 $a\in A$,有一个映射 a: $U\to a(U)$,且 $a(U)=\{a(u)\mid u\in U\}$称为属性 a 的值域.

一个信息系统可以用一个信息表来表示,当没有重复元组时,信息表是一个关系.

如果 $A=C\cup D, C\cap D=\varnothing$,则称信息系统$(U,A)$为一个决策表,其中 C 中的属性称为条件属性,D 中的属性称为决策属性.

例 2.1　下列信息系统是发表在"Popular Science"上的 CTR(Car Test Results)数据库.

其中属性是:

x_1——size, overall length, x_2——number of cylinders,

x_3——presence of a turbocharger, x_4——type of fuel system,

x_5——engine displacement, x_6——compression ratio,

x_7——power, x_8——type of transmission,

x_9——weight, y——mileage.

属性值如下:

c——compact, s——subcompact, sm——small, y——yes,

n——no, E——EFI, B——2-BBL, m——medium,

ma——manual, h——high, he——heavy, l——light,

lo——low, a——auto.

表 2.1

U/A	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	y
u_1	c	6	y	E	m	h	h	a	m	m
u_2	c	6	n	E	m	m	h	ma	m	m
u_3	c	6	n	E	m	h	h	ma	m	m
u_4	c	4	y	E	m	h	h	ma	1	h
u_5	c	6	n	E	m	m	m	ma	m	m
u_6	c	6	n	B	m	m	m	a	he	lo
u_7	c	6	n	E	m	m	h	ma	he	lo
u_8	s	4	n	B	sm	h	lo	ma	1	h
u_9	c	4	n	B	sm	h	lo	ma	m	m
u_{10}	c	4	n	B	sm	h	m	a	m	m
u_{11}	s	4	n	E	sm	h	lo	ma	1	h
u_{12}	s	4	n	E	m	m	m	ma	m	h
u_{13}	c	4	n	B	m	m	m	ma	m	m
u_{14}	s	4	y	E	sm	h	h	ma	m	h
u_{15}	s	4	n	B	sm	m	lo	ma	m	h
u_{16}	c	4	y	E	m	m	h	ma	m	m
u_{17}	c	6	n	E	m	m	h	a	m	m
u_{18}	c	4	n	E	m	m	h	a	m	m
u_{19}	s	4	n	E	sm	h	m	ma	m	h
u_{20}	c	4	n	E	sm	h	m	ma	m	h
u_{21}	c	4	n	B	sm	h	m	ma	m	m

这是一个信息系统,其中 $U=\{u_1,u_2,\cdots,u_{21}\}$,$A=\{x_1,x_2,\cdots,x_9,y\}$.事实上,它也是一个数据库关系.现在,我们假设 $C=\{x_1,x_2,\cdots,x_9\}$,$D=\{y\}$,则信息系统(关系)变成了一个决策表.

§2.2 简单分类

设(U,A)是一个信息系统,对于每一个属性 $a\in A$,我们引入一个 U 中

的划分 U/a：两个对象 $u,v\in U$ 在同一类中当且仅当 $a(u)=a(v)$.

算法 P

设 (U,A) 是一个信息系统，$U=\{u_1,u_2,\cdots,u_{|U|}\}$. 这个算法对于 $a\in A$ 给出了分类. 使用下面的指针：i 指向当前的输入对象 u_i；s 记录已经找到的 s 个类 $V_1,V_2,\cdots,V_s$；j 取值 $1,2,\cdots,s$，用来检验当前的输入对象 u_i 是否有 $a(V_j)=a(u_i)$.

如果对于某个 j 有 $a(u_i)=a(V_j)$，则令 u_i 在 V_j 中：$u_i\in V_j$；否则，建立一个新类，$s+1\rightarrow s$，$V_s=\{u_i\}$. 当算法结束时，我们有划分 $U/a=\{V_1,V_2,\cdots,V_s\}$.

P_1. [*Initialize*] Set $1\rightarrow i, 1\rightarrow j, 1\rightarrow s$. $V_1=\{u_1\}$.

P_2. [*Is* $i=|U|$?] If $i=|U|$, then the classification is completed, and we have $U/a=\{V_1,V_2,\cdots,V_s\}$. If $i<|U|$, then go to P_3.

P_3. [*Increase* i] $i+1\rightarrow i$, $1\rightarrow j$, go to P_4.

P_4. [*Is* $j=s$?] If $j=s$, then establish a new class $s+1\rightarrow s$, $V_s=\{u_i\}$ and go to P_2 to input the next object (if any). If $j<s$, then go to P_5.

P_5. [*Increase* j] $j+1\rightarrow j$, go to P_6.

P_6. [*Is* $a(u_i)=a(V_j)$?] If $a(u_i)=a(V_j)$, then $u_i\in V_j$, go to P_2. Otherwise, go to P_4 to check the next V_j (if any).

在最坏情况下，需要对于 $i=2,3,\cdots,|U|$ 检验 $a(V_j)=a(u_i)$ 是否成立，共需 $1+2+\cdots+(|U|-1)=(|U|)(|U|-1)/2=O(|U|^2)$ 次检验. 因此，这个算法的时间复杂性为 $O(|U|^2)$.

这个算法可以以并行方式计算所有的分类 $U/a_1,U/a_2,\cdots,U/a_{|A|}$.

例 2.2 应用这个算法到 CTR 数据库，我们有下面划分：

$$U/x_1=\{V_{11},V_{12}\},$$

其中

$$V_{11}=\{u_1,u_2,u_3,u_4,u_5,u_6,u_7,u_9,u_{10},u_{13},u_{16},u_{17},u_{18},u_{20},u_{21}\},$$
$$V_{12}=\{u_8,u_{11},u_{12},u_{14},u_{15},u_{19}\};$$

$$U/x_2=\{V_{21},V_{22}\},$$

其中

$$V_{21}=\{u_1,u_2,u_3,u_5,u_6,u_7,u_{17}\},$$
$$V_{22}=\{u_4,u_8,u_9,u_{10},u_{11},u_{12},u_{13},u_{14},u_{15},u_{16},u_{18},u_{19},u_{20},u_{21}\};$$

$$U/x_3=\{V_{31},V_{32}\},$$

其中

$$V_{31}=\{u_1,u_4,u_{14},u_{16}\},$$
$$V_{32}=\{u_2,u_3,u_5,u_6,u_7,u_8,u_9,u_{10},u_{11},u_{12},u_{13},u_{15},u_{17},u_{18},u_{19},u_{20},u_{21}\};$$

$$U/x_4=\{V_{41},V_{42}\},$$

其中

$$V_{41}=\{u_1,u_2,u_3,u_4,u_5,u_7,u_{11},u_{12},u_{14},u_{16},u_{17},u_{18},u_{19},u_{20}\},$$
$$V_{42}=\{u_6,u_8,u_9,u_{10},u_{13},u_{15},u_{21}\};$$

$$U/x_5=\{V_{51},V_{52}\},$$

其中

$$V_{51}=\{u_1,u_2,u_3,u_4,u_5,u_6,u_7,u_{12},u_{13},u_{16},u_{17},u_{18}\},$$
$$V_{52}=\{u_8,u_9,u_{10},u_{11},u_{14},u_{15},u_{19},u_{20},u_{21}\};$$

$$U/x_6=\{V_{61},V_{62}\},$$

其中

$$V_{61}=\{u_1,u_3,u_4,u_8,u_9,u_{10},u_{11},u_{14},u_{19},u_{20},u_{21}\},$$
$$V_{62}=\{u_2,u_5,u_6,u_7,u_{12},u_{13},u_{15},u_{16},u_{17},u_{18}\};$$

$$U/x_7=\{V_{71},V_{72},V_{73}\},$$

其中

$$V_{71}=\{u_1,u_2,u_3,u_4,u_7,u_{14},u_{16},u_{17},u_{18}\},$$
$$V_{72}=\{u_5,u_6,u_{10},u_{12},u_{13},u_{19},u_{20},u_{21}\},$$
$$V_{73}=\{u_8,u_9,u_{11},u_{15}\};$$

$$U/x_8=\{V_{81},V_{82}\},$$

其中

$$V_{81}=\{u_1,u_6,u_{10},u_{17},u_{18}\};$$
$$V_{82}=\{u_2,u_3,u_4,u_5,u_7,u_8,u_9,u_{11},u_{12},u_{13},u_{14},u_{15},u_{16},u_{19},u_{20},u_{21}\};$$

$$U/x_9=\{V_{91},V_{92},V_{93}\},$$

其中

$$V_{91}=\{u_1,u_2,u_3,u_5,u_9,u_{10},u_{12},u_{13},u_{14},u_{15},u_{16},u_{17},u_{18},u_{19},u_{20},u_{21}\},$$
$$V_{92}=\{u_4,u_8,u_{11}\},$$
$$V_{93}=\{u_6,u_7\};$$

$$U/y=\{W_1,W_2,W_3\},$$

其中

$$W_1=\{u_1,u_2,u_3,u_5,u_9,u_{10},u_{13},u_{16},u_{17},u_{18},u_{21}\},$$

$$W_2=\{u_4,u_8,u_{11},u_{12},u_{14},u_{15},u_{19},u_{20}\},\quad W_3=\{u_6,u_7\}.$$

§2.3 支持子集

设 $W\subseteq U$. 对于分类 U/a,定义 W 的下近似为 $W^{(U/a)^-}=\bigcup_{V\in U/a,V\subseteq W}V$,也用 $S_a(W)$来表示. 子集 $S_a(W)$称为 W 关于属性 a 的支持子集(support subset),$\mathrm{spt}_a(W)=|S_a(W)|/|U|$称为 W 关于属性 a 的支持度(support degree);对于分类 U/a,定义 W 的上近似为 $W^{(U/a)^+}=\bigcup_{V\in U/a,V\cap W\neq\varnothing}V$;$acc_a(W)=|W^{(U/a)^-}|/|W^{(U/a)^+}|$称为 W 关于属性 a 的近似精度.

例 2.3 对于 CTR 数据库,我们有 $S_{x_1}(W_2)=V_{12}=\{u_8,u_{11},u_{12},u_{14},u_{15},u_{19}\}$, $S_{x_9}(W_2)=V_{92}=\{u_4,u_8,u_{11}\}$. 这意味着下面:考虑 W_2 中的所有元组 u.

元组 $u=u_8,u_{11},u_{12},u_{14},u_{15},u_{19}$ 支持一个规则,这个规则表明条件 $x_1(u)=s$ 蕴涵决策 $y(u)=h$,其强度为 $\mathrm{spt}_{x_1}(W_2)=6/21$.

元组 $u=u_4,u_8,u_{11}$ 支持一个规则,这个规则表明条件 $x_9(u)=l$ 蕴涵决策 $y(u)=h$,其强度为 $\mathrm{spt}_{x_9}(W_2)=3/21$.

简单地, $S_x(W)$意味着规则"x 蕴涵 $y=y(W)$"有强度 $\mathrm{spt}_x(W)=|S_x(W)|/|U|$.

现在我们给出一个计算下近似 $W^{(U/a)^-}=\bigcup_{V\in U/a,V\subseteq W}V$ 的算法,其中 $a\in A$, $W\subseteq U$.

算法 L

设(U,A)是一个信息系统,令 $U/a=\{V_1,V_2,\cdots,V_s\}$, $1\leqslant s\leqslant|U|$, $W\subseteq U$. 这个算法给出了 W 关于U/a 的下近似

$$W^{(U/a)^-}=\bigcup_{V=V_1,V_2,\cdots,V_s,V\subseteq W}V.$$

L_1 [*Initialize*] Set $1\to j,\varnothing\to L$.

L_2 [*Is* $V_j\subseteq W$?] If $V_j\subseteq W$, then $L\cup V_j\to L$. Otherwise, go to L_3 to check the next V_j (if any).

L_3 [*Is* $j=s$?] If $j=s$, then the algorithm is completed, and we have $W^{(U/a)^-}=L$. Otherwise, go to L_4 to check the next V_j.

L_4 [*Increase j*] Set $j+1 \to j$. go to L_2.

算法 L 需要检验 s 个子集 $V_1, V_2, \cdots, V_s, s \leqslant |U|$，因此它的时间复杂性是 $O(|U|)$.

类似地，我们能够设计一个计算上近似的算法.

算法 R

设(U, A)是一个信息系统，$U/a = \{V_1, V_2, \cdots, V_s\}, 1 \leqslant s \leqslant |U|$. 令 $W \subseteq U$. 这个算法给出了 W 关于 U/a 的上近似 $W^{(U/a)^*} = \bigcup_{V = V_1, V_2, \cdots, V_s, V \cap W \neq \varnothing} V$.

R_1 [*Initialize*] Set $1 \to j, \varnothing \to R$.

R_2 [*Is* $V_j \cap W \neq \varnothing$?] If $V_j \cap W \neq \varnothing$, then $R \cup V_j \to R$. Otherwise, go to R_3 to check the next V_j (if any).

R_3 [*Is* $j = s$?] If $j = s$, then the algorithm is completed, and we have $W^{(U/a)^*} = R$. Otherwise, go to R_4 to check the next V_j.

R_4 [*Increase j*] Set $j+1 \to j$. go to R_2.

算法 R 需要检验 s 个子集 $V_1, V_2, \cdots, V_s, s \leqslant |U|$，因此它的时间复杂性是 $O(|U|)$.

例 2.4 对于 CTR 数据库，使用算法 L，我们可以计算不同支持子集的支持度：

$$\mathrm{spt}_{x_1}(W_1) = |S_{x_1}(W_1)|/|U| = |\bigcup_{V \in U/x_1, V \subseteq W_1} V|/|U| = |\varnothing|/21 = 0,$$

$$\mathrm{spt}_{x_1}(W_2) = |S_{x_1}(W_2)|/|U| = |\bigcup_{V \in U/x_1, V \subseteq W_2} V|/|U| = |V_{12}|/21 = 6/21,$$

$$\mathrm{spt}_{x_1}(W_3) = |S_{x_1}(W_3)|/|U| = |\bigcup_{V \in U/x_1, V \subseteq W_3} V|/|U| = |\varnothing|/21 = 0;$$

$$\mathrm{spt}_{x_j}(W_i) = |S_{x_j}(W_i)|/|U| = |\varnothing|/21 = 0,$$

$$i = 1, 2, 3, \quad j = 2, 3, 4, 5, 6, 7, 8;$$

$$\mathrm{spt}_{x_9}(W_1) = |S_{x_9}(W_1)|/|U| = |\bigcup_{V \in U/x_9, V \subseteq W_1} V|/|U| = |\varnothing|/21 = 0,$$

$$\mathrm{spt}_{x_9}(W_2) = |S_{x_9}(W_2)|/|U| = |\bigcup_{V \in U/x_9, V \subseteq W_2} V|/|U| = |V_{92}|/21 = 3/21,$$

$$\mathrm{spt}_{x_9}(W_3) = |S_{x_9}(W_3)|/|U| = |\bigcup_{V \in U/x_9, V \subseteq W_3} V|/|U| = |V_{93}|/21 = |W_3|/21 = 2/21.$$

§2.4 决策属性的支持度

设 $y \in D$ 是决策表(U,A)中的一个决策属性，$A=C\cup D, C\cap D=\varnothing$. 现在我们考虑决策属性 $y\in D$ 的整体决策而不是对于"决策子集"$W\in U/y$ 的一个局部决策.

决策属性 $y\in D$ 关于条件属性 $a\in C$ 的支持子集是子集

$$S_a(y)=\bigcup_{W\in U/y}W^{(U/a)^-}=\bigcup_{W\in U/y}(\bigcup_{V\in U/a, V\subseteq W}V),$$

$$\mathrm{spt}_a(y)=|\bigcup_{W\in U/y}W^{(U)/a)^-}|/|U|$$

称为 y 关于 a 的支持度.

如果 $U/y=U/\delta$，其中 δ 是"全体"划分 $U/\delta=\{U\}$，则对于 $a\in C$，我们有 $S_a(y)=U, \mathrm{spt}_a(y)=1$.

例 2.5 对于 CTR 数据库，我们有 $S_{x_1}(y)=V_{12}=\{u_8,u_{11},u_{12},u_{14},u_{15},u_{19}\}$，$S_{x_9}(y)=V_{92}\cup V_{93}=\{u_4,u_8,u_{11}\}\cup\{u_6,u_7\}$. 这意味着下面：考虑包含在一个决策中的所有元组 $u\in U$.

元组 $u=u_8,u_{11},u_{12},u_{14},u_{15},u_{19}$ 支持一个规则，这个规则表明条件 $x_1(u)=s$ 蕴涵决策 $y(u)=h$，其强度为 $\mathrm{spt}_{x_1}(W_2)=6/21$.

元组 $u=u_4,u_8,u_{11}$ 支持一个规则，这个规则表明条件 $x_9(u)=l$ 蕴涵决策 $y(u)=h$，其强度为 $\mathrm{spt}_{x_9}(W_2)=3/21$.

元组 $u=u_6,u_7$ 支持一个规则，这个规则表明条件 $x_9(u)=he$ 蕴涵决策 $y(u)=lo$，其强度为 $\mathrm{spt}_{x_9}(W_3)=2/21$.

简而言之，$S_x(y)$ 意味着规则"x 蕴涵 y"有强度 $\mathrm{spt}_x(y)=|S_x(y)|/|U|$.

假设 $U/y=\{W_1,W_2,\cdots,W_t\}, 1\leqslant t\leqslant|U|$. 我们可以使用算法 L 找 $W_1^{(U/a)^-}, W_2^{(U/a)^-},\cdots,W_t^{(U/a)^-}$. 因为算法 L 找一个 $W^{(U/a)^-}$ 的时间复杂性是 $O(|U|)$，因此找所有 $t(\leqslant|U|)$个的时间复杂性是 $O(|U|^2)$.

例 2.6 对于 CTR 数据库，使用算法 L，我们有

$$\mathrm{spt}_{x_1}(y)=|S_{x_1}(y)|/21=|V_{12}|/21=6/21,$$

$$\mathrm{spt}_{x_9}(y)=|S_{x_9}(y)|/21=|V_{92}\cup V_{93}|/21=5/21,$$

$$\mathrm{spt}_{x_j}(y)=|S_{x_j}(y)|/21=0,$$

其中 $j=2,3,4,5,6,7,8$.

§2.5 交的计算

设(U,A)是一个信息系统,对于两个属性 $a,b\in A$,我们需要计算论域 U 的下面分类 U/ab:两个对象 $u,v\in U$ 在同一类当且仅当$a(u)=a(v)$且$b(u)=b(v)$.

对于属性集 $X\subseteq A$,我们按下面定义论域 U 的分类:两个对象 $u,v\in U$ 在同一类当且仅当对每个$a\in X$ 有$a(u)=a(v)$.

我们也定义 $U/\varnothing=U/\delta$,其中 δ 是"全划分":$U/\delta=\{U\}$,$|U/\delta|=1$.

算法 I

设(U,A)是一个信息系统.对于 $a,b\in A$,令 $U/a=\{V_{11},V_{12},\cdots,V_{1s}\}$,$1\leqslant s\leqslant|U|$;$U/b=\{V_{21},V_{22},\cdots,V_{2t}\}$,$1\leqslant t\leqslant|U|$.这个算法给出分类 $U/ab=\{V_1,V_2,\cdots,V_r\}$,$1\leqslant r\leqslant st$.我们使用下面指针:$i=1,2,\cdots,s$ 指向 V_{1i};$j=1,2,\cdots,t$ 指向 V_{2j};变量 r 记录我们已经找到的U/ab 的r 个类 $V_1,V_2,\cdots,V_r$.

对于每个 i 和每个j,我们检验 $V_{1i}\cap V_{2j}=\varnothing$是否成立.如果 $V_{1i}\cap V_{2j}=\varnothing$,则忽略它;否则建立一个新类:$r+1\rightarrow r$,$V_r=V_{1i}\cap V_{2j}$.

I_1 [*Initialize*] Set $1\rightarrow i,1\rightarrow j,0\rightarrow r$.

I_2 [*Is* $V_{1i}\cap V_{2j}=\varnothing$?] If intersection $V_{1i}\cap V_{2j}=\varnothing$, then go to I_3 to check the next intersection. Otherwise, set $r+1\rightarrow r$ and establish a new class $V_r=V_{1i}\cap V_{2j}$ for U/ab. go to I_3 check the next intersection.

I_3 [*Is* $j=t$?] If $j=t$, then go to I_5 to check next i. Otherwise, go to next step I_4 to see next j.

I_4 [*Increase* j] Set $j+1\rightarrow j$. go to I_2.

I_5 [*Is* $i=s$?] If $i=s$, then the classification is completed, and we have $U/ab=\{V_1,V_2,\cdots,V_r\}$. Otherwise, go to I_6.

I_6 [*Increase* i] Set $i+1\rightarrow i,1\rightarrow j$. go to I_2.

算法 I 需要检验 $st\leqslant|U|^2$ 个交 $V_{1i}\cap V_{2j}$,因此它的时间复杂性为 $O(|U|^2)$.

例 2.7 对于 CTR 数据库,使用算法 I,我们得到

(1) U/x_1x_9 是 $\{u_1,u_2,u_3,u_5,u_9,u_{10},u_{13},u_{16},u_{17},u_{18},u_{20},u_{21}\}$, $\{u_{12},u_{14},u_{15},u_{19}\}$, $\{u_4\}$, $\{u_8,u_{11}\}$, $\{u_6,u_7\}$.

(2) $U/x_1x_9x_5$ 是 $\{u_1,u_2,u_3,u_5,u_{13},u_{16},u_{17},u_{18}\}$, $\{u_9,u_{10},u_{20},u_{21}\}$, $\{u_{12}\}$, $\{u_{14},u_{15},u_{19}\}$, $\{u_4\}$, $\{u_8,u_{11}\}$. $\{u_6,u_7\}$.

(3) $U/x_1x_9x_5x_2$ 是 $\{u_1,u_2,u_3,u_5,u_{17}\}$, $\{u_{13},u_{16},u_{18}\}$, $\{u_9,u_{10},u_{20},u_{21}\}$, $\{u_{14},u_{15},u_{19}\}$, $\{u_{12}\}$, $\{u_4\}$, $\{u_8,u_{11}\}$, $\{u_6,u_7\}$.

§2.6 多个条件的支持度

令 $W\subsetneq U$ 是 U 的子集,我们考虑子集 W 关于条件属性集的支持度.

对于决策表$(U,C\cup D)$中的条件属性集 $X\subseteq C$, W 关于X 的支持子集是 $S_X(W)=W^{(U/X)} = \bigcup_{V\in U/X,V\subseteq W}V$, $\mathrm{spt}_X(W)=|W^{(U/X)^-}|/|U|$ 称为 W 关于X 的支持度.

例2.8 对于 CTR 数据库,我们有 $S_{x_1x_9}(W_2)=\{u_4\}\cup\{u_8,u_{11}\}\cup\{u_{12},u_{14},u_{15},u_{19}\}$. 这意味着下面:考虑所有元组 $u\in W_2$.

元组 $u=u_4$ 支持一个规则,这个规则表明条件 $x_1(u)=c$ 和 $x_9(u)=l$ 蕴涵决策 $y(u)=h$.

元组 $u=u_8,u_{11}$支持一个规则,这个规则表明条件 $x_1(u)=s$ 和 $x_9(u)=l$ 蕴涵决策 $y(u)=h$.

元组 $u=u_{12},u_{14},u_{15},u_{19}$支持一个规则,这个规则表明条件 $x_1(u)=s$ 和 $x_9(u)=m$ 蕴涵决策 $y(u)=h$.

简单地, $S_X(W)$意味着规则"X 蕴涵 $Y=Y(W)$"有强度 $\mathrm{spt}_X(W)=|S_X(W)|/|U|$.

我们知道下面:(1)令 $X_1,X_2\subseteq C$ 是两个条件属性子集. 如果 $X_1\subseteq X_2$, 则 $S_{X_1}(W)\subseteq S_{X_2}(W)$, $\mathrm{spt}_{X_1}(W)\leqslant \mathrm{spt}_{X_2}(W)$. (2)令 $X_1,X_2\subseteq C$ 是两个条件属性子集,则 $S_{X_1\cup X_2}(W)\supseteq S_{X_1}(W)$, $S_{X_2}(W)$; $\mathrm{spt}_{X_1\cup X_2}(W)\geqslant \mathrm{spt}_{X_1}(W)$, $\mathrm{spt}_{X_2}(W)$.

为了计算支持子集,我们需要下面计算:(1)计算 U/X 需要$|X|-1$ 个交 U/x,其中 $x\in X$. 计算一个交的时间复杂性是 $O(|U|^2)$. 因此这一步的代价是 $O(|X||U|^2)$. (2)使用算法 L 找 $W^{(U/X)^-}$ 的时间复杂性是 $O(|U|)$. 因此计算支持子集 $S_X(W)$的时间复杂性是 $O(|X||U|^2)=O(|A|\cdot|U|^2)$.

令 $Y\subseteq D$ 是$(U,C\cup D)$中的一些决策属性子集. 现在我们考虑 Y 关于 $X\subseteq C$ 的支持. Y 关于 X 的支持子集是 $S_X(Y)=\bigcup_{W\in U/Y}W^{(U/X)^-}=$

$\bigcup_{W\in U/Y}(\bigcup_{V\in U/X,V\subseteq W}V)$，$\mathrm{spt}_X(Y)=|\bigcup_{W\in U/Y}W^{(U/X)^-}|/|U|$称为 Y 关于X 的支持度.

例 2.9 对于 CTR 数据库，我们有 $S_{x_1x_9}(y)=(\{u_4\}\cup\{u_8,u_{11}\}\cup\{u_{12},u_{14},u_{15},u_{19}\})\cup\{u_6,u_7\}$. 这意味着下面：考虑所有元组 $u\in U$，并且做出整个决策.

元组 $u=u_4$ 支持一个规则，这个规则表明条件 $x_1(u)=c$ 和 $x_9(u)=l$ 蕴涵决策 $y(u)=h$.

元组 $u=u_8,u_{11}$支持一个规则，这个规则表明条件 $x_1(u)=s$ 和 $x_9(u)=l$ 蕴涵决策 $y(u)=h$.

元组 $u=u_{12},u_{14},u_{15},u_{19}$支持一个规则，这个规则表明条件 $x_1(u)=s$ 和$x_9(u)=m$ 蕴涵决策 $y(u)=h$.

元组 $u=u_6,u_7$ 支持一个规则，这个规则表明条件 $x_1(u)=c$ 和$x_9(u)=he$ 蕴涵决策 $y(u)=lo$.

简单地，$S_X(Y)$意味着规则"X 蕴涵 Y"有强度

$$\mathrm{spt}_X(Y)=|S_X(Y)|/|U|.$$

我们知道下面：(1)如果 $U/Y=U/\delta$，其中 δ 是"全体"划分 $U/\delta=\{U\}$，则对于所有 $X\subseteq C$ 有 $S_X(Y)=U$，$\mathrm{spt}_X(Y)=1$. (2)令 $X_1,X_2\subseteq C$ 是两个条件属性子集. 如果 $X_1\subseteq X_2$，则 $S_{X_1}(Y)\subseteq S_{X_2}(Y)$，$\mathrm{spt}_{X_1}(Y)\leqslant \mathrm{spt}_{X_2}(Y)$. (3)令 $X_1,X_2\subseteq C$ 是两个条件属性子集，则 $S_{X_1\cup X_2}(Y)\supseteq S_{X_1}(Y),S_{X_2}(Y)$；$\mathrm{spt}_{X_1\cup X_2}(Y)\geqslant \mathrm{spt}_{X_1}(Y),\mathrm{spt}_{X_2}(Y)$.

为了计算支持子集，我们需要进行下面计算：(1)计算 U/Y 需要$|Y|-1$个交 $U/y(y\in Y)$. 计算一个交的时间复杂性是 $O(|U|^2)$. 因此这一步的代价是 $O(|Y||U|^2)$. (2)计算 U/X 需要$|X|-1$ 个交 $U/x(x\in X)$. 因此这一步的代价是 $O(|X||U|^2)$. (3)设 $U/Y=\{W_1,W_2,\cdots,W_t\}$，$1\leqslant t\leqslant|U|$. 我们可以使用算法 L 找 $W_1^{(U/X)^-},W_2^{(U/X)^-},\cdots,W_t^{(U/X)^-}$. 因为算法 L 找一个 $W^{(U/X)^-}$ 的时间复杂性是 $O(|U|)$，因此找所有 $t(\leqslant|U|)$个 $W_i^{(U/X)^-}$ 的时间复杂性是 $O(|U|^2)$. 由上可知，计算支持子集 $S_X(Y)$的时间复杂性是 $O((|X|+|Y|)|U|^2)=O(|A||U|^2)$.

§2.7 函数依赖

令(U,A)是一个决策表，$A=C\cup D$，$C\cap D=\varnothing$，其中 C 是条件属性集，D 是决策属性集. 两个属性子集 $X_1,X_2\subseteq C$ 之间的函数依赖(functional dependency)(对于决策属性子集 $Y \subseteq D$ 而言)是一个描述,用 $X_1 \to X_2(Y)$ 来表示.在决策表中,$X_1 \to X_2(Y)$ 成立当且仅当 $S_{X_1}(Y) \supseteq S_{X_2}(Y)$,即对于每个 $W \in U/Y$,我们有 $S_{X_1}(W) \supseteq S_{X_2}(W)$.

我们知道,如果 $X_1 \supseteq X_2$,则 $X_1 \to X_2(Y)$.

例 2.10 对于 CTR 数据库,我们有 $x_1x_9 \to x_1, x_9(y)$,因为 $S_{x_1x_9}(y) = (\{u_4\} \cup \{u_8, u_{11}\} \cup \{u_{12}, u_{14}, u_{15}, u_{19}\}) \cup \{u_6, u_7\} \supseteq S_{x_1}(y) = \{u_8, u_{11}, u_{12}, u_{14}, u_{15}, u_{19}\}$, $S_{x_9}(y) = \{u_4, u_8, u_{11}\} \cup \{u_6, u_7\}$.

$X_1 \to X_2(Y)$ 意味着规则"X_1 蕴涵 Y"是比规则"X_2 蕴涵 Y"更强.

函数依赖 $X_1 \to X_2(Y)$ 又可以定义为:对于每个 $W \in U/Y$,有 $W^{(U/X_1)^-} \supseteq W^{(U/X_2)^-}$,即对于每个 $W \in U/Y$,有 $\bigcup_{V \in U/X_1, V \subseteq W} V \supseteq \bigcup_{V \in U/X_2, V \subseteq W} V$.

对于 $X_1, X_2 \subseteq C$, $Y \subseteq D$,检验 $X_1 \to X_2(Y)$ 的时间复杂性是 $O(|A| \cdot |U|^2)$,因为它是计算 $S_{X_1}(Y)$ 和 $S_{X_2}(Y)$ 的代价.

§2.8 恒等依赖

令 (U, A) 是一个决策表,$A = C \cup D$, $C \cap D = \varnothing$,其中 C 是条件属性集,D 是决策属性集.两个属性子集 $X_1, X_2 \subseteq C$ 之间的恒等依赖(identity dependency)(对于决策属性子集 $Y \subseteq D$ 而言)是一个描述,用 $X_1 \leftrightarrow X_2(Y)$ 来表示.在决策表 (U, A) 中,$X_1 \leftrightarrow X_2(Y)$ 成立的充分必要条件是 $S_{X_1}(Y) = S_{X_2}(Y)$;即对于每一个 $W \in U/Y$,我们有 $S_{X_1}(W) = S_{X_2}(W)$.

简单地,$X_1 \leftrightarrow X_2(Y)$ 意味着规则"X_1 蕴涵 Y"是和规则"X_2 蕴涵 Y"同样强.

恒等依赖 $X_1 \leftrightarrow X_2(Y)$ 又可以定义为:对于每个 $W \in U/Y$,有 $W^{(U/X_1)^-} = W^{(U/X_2)^-}$,即对于每个 $W \in U/Y$,有 $\bigcup_{V \in U/X_1, V \subseteq W} V = \bigcup_{V \in U/X_2, V \subseteq W} V$.

对于 $X_1, X_2 \subseteq C$, $Y \subseteq D$,检验 $X_1 \leftrightarrow X_2(Y)$ 的时间复杂性是 $O(|A| \cdot |U|^2)$,因为它是计算 $S_{X_1}(Y)$ 和 $S_{X_2}(Y)$ 的代价.

§2.9 重要性和核

令 (U, A) 是一个决策表,$A = C \cup D$, $C \cap D = \varnothing$,其中 C 是条件属性集,D 是决策属性集.我们知道 $X_1 \supseteq X_2$ 蕴涵 $S_{X_1}(Y) \supseteq S_{X_2}(Y)$(即 $X_1 \supseteq X_2$ 蕴涵 $X_1 \to X_2(Y)$),其中 $X_1, X_2 \subseteq C$, $Y \subseteq D$.因此我们知道

(1) $S_C(Y)\supseteq S_X(Y)$,即对于所有的 $X\subseteq C, Y\subseteq D$,有 $C\rightarrow X(Y)$.

(2) $S_X(Y)\supseteq S_{X-\{x\}}(Y)$,即对于非空集 $X\subseteq C, x\in X$ 和 $Y\subseteq D$,有 $X\rightarrow(X-\{x\})(Y)$.

在下面一节,对于 $Y\subseteq D$,我们研究一系列条件属性子集,$X\rightarrow X-\{x\}\rightarrow X-\{x,x'\}\rightarrow X-\{x,x',x''\}\rightarrow\varnothing(Y)$,其中 $X\subseteq C, x, x', x'', \cdots\in X$;$S_X(Y)\supseteq S_{X-\{x\}}(Y)\supseteq S_{X-\{x,x'\}}(Y)\supseteq S_{X-\{x,x',x''\}}(Y)\supseteq\cdots\supseteq S_{\varnothing}(Y)=S_{\delta}(Y)=\bigcup_{W\in U/Y}(\bigcup_{V=U,V\subseteq W}V)=\varnothing$,其中 $U/Y\neq U/\delta=\{U\}$.

令 $\varnothing\subset X\subseteq C, \varnothing\subset Y\subseteq D, U/Y\neq U/\delta=\{U\}$.给定 $x\in X$,如果 $S_X(Y)\supset S_{X-\{x\}}(Y)$,则称 x 在 X 中是重要的(对于 Y 而言);如果 $S_X(Y)=S_{X-\{x\}}(Y)$,则称 x 在 X 中是不重要的(对于 Y 而言).

利用恒等依赖来表示属性的重要性,我们有 $x\in X$ 在 X 中是重要的(对于 Y 而言)当且仅当 $X\leftrightarrow(X-\{x\})(Y)$ 不成立;$x\in X$ 在 X 中是不重要的(对于 Y 而言)当且仅当 $X\leftrightarrow(X-\{x\})(Y)$ 成立.

例 2.11 令 $X\subseteq C, X=\{x\}, Y\subseteq D, U/Y\neq\{U\}$.注意到:$U/X=U/x, U/(X-\{x\})=U/\varnothing=U/\delta=\{U\}$.因此 $S_{X-\{x\}}(Y)=S_{\varnothing}(Y)=\bigcup_{W\in U/Y}(\bigcup_{V=U,V\subseteq W}V)=\varnothing, S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)$.

我们有下面:(1)如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)\neq\varnothing$,则 x 在 X 中是重要的(对于 Y 而言).(2)如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)=\varnothing$,则 x 在 X 中是不重要的(对于 Y 而言).

令 $\varnothing\subset X\subseteq C, \varnothing\subset Y\subseteq D, U/Y\neq U/\delta=\{U\}$.给定 $x\in X$,我们定义 x 在 X 中的重要性(对于 Y 而言)为 $\mathrm{sig}^Y_{X-\{x\}}(x)=(|S_X(Y)|-|S_{X-\{x\}}(Y)|)/|U|$.

在特别情形下,当 $X=\{x\}$ 时,我们用 $\mathrm{sig}^Y(x)$ 来表示 $\mathrm{sig}^Y_{\varnothing}(x)$,即

$$\begin{aligned}\mathrm{sig}^Y(x)=\mathrm{sig}^Y_{\varnothing}(x)&=(|S_x(Y)|-|S_{\delta}(Y)|)/|U|\\&=|S_x(Y)|/|U|.\end{aligned}$$

因此我们总是有 $\mathrm{sig}^Y(x)>0$,除非 $S_x(Y)=\varnothing$.

计算一个属性重要性 $\mathrm{sig}^Y_{X-\{x\}}(x)$ 的时间复杂性为 $O(|A||U|^2)$,因为它是计算 $S^1_X(Y)$ 和 $S_{X-\{x\}}(Y)$ 的代价.

我们知道(1) $0\leqslant\mathrm{sig}^Y_{X-\{x\}}(x)\leqslant 1$.(2)属性 x 在 X 中是重要的(对于 Y 而言)当且仅当 $\mathrm{sig}^Y_{X-\{x\}}(x)>0$.

例 2.12 令 $X\subseteq C, X=\{x\}, Y\subseteq D, U/Y\neq\{U\}$.注意到:$U/X=U/x, U/(X-\{x\})=U/\varnothing=U/\delta=\{U\}$.因此 $S_{X-\{x\}}(Y)=S_{\varnothing}(Y)=\bigcup_{W\in U/Y}(\bigcup_{V=U,V\subseteq W}V)=\varnothing, S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)$.

我们有

(1) 如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)\neq\varnothing$,则 $\mathrm{sig}^Y_{X-\{x\}}(x)=\mathrm{sig}^Y_{\varnothing}(x)>0$.

(2) 如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)=\varnothing$,则 $\mathrm{sig}^Y_{X-\{x\}}(x)=\mathrm{sig}^Y_{\varnothing}(x)=0$.

例 2.13 对于 CTR 数据库,我们有

(1) $\mathrm{sig}^y(x_1)=6/21$;$\mathrm{sig}^y(x_j)=0, j=2,3,4,5,6,7,8$;$\mathrm{sig}^y(x_9)=5/21$.

(2) $\mathrm{sig}^y_{x_1x_9}(x_2)=5/21$,$\mathrm{sig}^y_{x_1x_9}(x_3)=2/21$,$\mathrm{sig}^y_{x_1x_9}(x_4)=4/21$,$\mathrm{sig}^y_{x_1x_9}(x_5)-8/21$,$\mathrm{sig}^y_{x_1x_9}(x_6)=6/21$,$\mathrm{sig}^y_{x_1x_9}(x_7)=7/21$,$\mathrm{sig}^y_{x_1x_9}(x_8)=4/21$.

(3) $\mathrm{sig}^y_{x_1x_9x_5}(x_2)=\mathrm{sig}^y_{x_1x_9x_5}(x_3)=\mathrm{sig}^y_{x_1x_9x_5}(x_6)=0$,$\mathrm{sig}^y_{x_1x_9x_5}(x_4)=4/21$,$\mathrm{sig}^y_{x_1x_9x_5}(x_7)=\mathrm{sig}^y_{x_1x_9x_5}(x_8)=1/21$.

令$\varnothing\subset X\subseteq C$,$\varnothing\subset Y\subseteq D$,$U/Y\neq U/\delta=\{U\}$.所有在 X 中是重要的属性$x\in X$ 组成的集合(对于 Y 而言)称为 X 的核(对于 Y 而言),用 C^Y_X 来表示.即 $C^Y_X=\{x\in X\mid \mathrm{sig}^Y_{X-\{x\}}(x)>0\}$.

我们也定义 $C^Y_{\varnothing}=\varnothing$.

例 2.14 令 $X\subseteq C, X=\{x\}, Y\subseteq D, U/Y\neq\{U\}$.

从例 2.11 我们有

(1) 如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)\neq\varnothing$,则 $C^Y_{\{x\}}=\{x\}$,因为 x 在 X 中是重要的(对于 Y 而言).

(2) 如果 $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)=\varnothing$,则 $C^Y_{\{x\}}=\varnothing$,因为 x 在 X 中是不重要的(对于 Y 而言).

算法 C

令(U,A)是一个决策表,$A=C\cup D$,$C\cap D=\varnothing$,其中 C 是条件属性集,D 是决策属性集.令$\varnothing\subset X\subseteq C$,$\varnothing\subset Y\subseteq D$,$U/Y\neq U/\delta=\{U\}$.这个算法计算 X 的核C^Y_X(对于 Y 而言).令 $X=\{x_1,x_2,\cdots,x_{|X|}\}$,我们使用 $x\in C^Y_X$ 来表示x 包含在C^Y_X 中(即 x 在X 中是重要的(对于 Y 而言)).

C_1.[*Initialize*]Set $1\rightarrow i$.

C_2.[*Is* $\mathrm{sig}^Y_{X-\{x_i\}}(x_i)>0$?]If $\mathrm{sig}^Y_{X-\{x_i\}}(x_i)>0$, go to C_3. If $\mathrm{sig}^Y_{X-\{x_i\}}(x_i)=0$, go to C_4.

C_3.[*Set* $x_i\in C^Y_X$]Set $x_i\in C^Y_X$, go to C_4.

C_4.[*Is* $i=|X|$?]If $i=|X|$, then algorithm is completed and C^Y_X is the core of X(for Y). If $i<|X|$, then go to C_5.

C_5.[*Increase* i] Set $i+1\rightarrow i$. Go to C_2.

这个算法对于 $i=1,2,\cdots,|X|$ 计算 $|X|$ 个重要性 $\mathrm{sig}_{X-\{x_i\}}^{Y}(x_i)$. 计算一个重要性的时间复杂性是 $O(|A||U|^2)$. 因此算法 C 的时间复杂性是 $O(|X||A||U|^2)$.

例 2.15 对于 CTR 数据库,使用算法 C 我们可以得到

$$C_{\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9\}}^{Y}=\{x_4,x_9\}.$$

§2.10 属性依赖性

令 (U,A) 是一个决策表, $A=C\cup D$, $C\cap D=\varnothing$,其中 C 是条件属性集, D 是决策属性集. 令 $\varnothing\subset X\subseteq C$, $\varnothing\subset Y\subseteq D$, $U/Y\neq U/\delta=\{U\}$. 如果每个 $x\in X$ 在 X 中是重要的(对于 Y 而言),则称非空子集 X 是独立的(对于 Y 而言);否则 X 是依赖的(对于 Y 而言). 空集 $\varnothing$ 称为独立的(对于 Y 而言).

为了检验 X 是否独立,我们需要计算 $|X|$ 个重要性 $\mathrm{sig}_{X-\{x_i\}}^{Y}(x_i)$ $(i=1,2,\cdots,|X|)$,并判断其是否大于 0. 计算一个重要性的时间复杂性是 $O(|A|\cdot|U|^2)$. 因此检验 X 的依赖性的时间复杂性是 $O(|X||A||U|^2)$.

例 2.16 令 $X\subseteq C$, $X=\{x\}$, $Y\subseteq D$, $U/Y\neq\{U\}$ 从例 2.11,我们有下面:

情形 A $S_x(Y)=\bigcup_{W\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)\neq\varnothing$. X 是独立的(对于 Y 而言),因为 x 在 X 中是重要的(对于 Y 而言).

情形 B $S_x(Y)=\bigcup_{V\in U/Y}(\bigcup_{V\in U/x,V\subseteq W}V)=\varnothing$. X 是依赖的(对于 Y 而言),因为 x 在 X 中不是重要的(对于 Y 而言).

例 2.17 对于 CTR 数据库,我们知道 $X=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9\}$ 的下列子集的依赖性(对于 Y 而言). (1) $\varnothing$ 是独立的. (2) 单元素集 $\{x_1\}$, $\{x_9\}$ 是独立的;所有别的单元素集 $\{x_j\}$ $(j=2,3,4,5,6,7,8)$ 是依赖的,因为

$$S_{x_1}(y)=\bigcup_{i=1,2,3}(\bigcup_{V\in U/x_1,V\subseteq W_i}V)=V_{12}\neq\varnothing,$$

$$S_{x_9}(y)=\bigcup_{i=1,2,3}(\bigcup_{V\in U/x_9,V\subseteq W_i}V)=V_{92}\cup V_{93}\neq\varnothing,$$

$$S_{x_j}(y)=\bigcup_{i=1,2,3}(\bigcup_{V\in U/X_j,V\subseteq W_i}V)=\varnothing,j=2,3,4,5,6,7,8.$$

§2.11 约 简

令 (U,A) 是一个决策表, $A=C\cup D$, $C\cap D=\varnothing$,其中 C 是条件属性集, D 是决策属性集. 令 $\varnothing\subset X\subseteq C$, $\varnothing\subset Y\subseteq D$, $U/Y\neq U/\delta=\{U\}$. 我们总是可

以找到 X 的一个极小子集 X_0 使得 $S_{X_0}(Y)=S_X(Y)$.

也就是说,X_0 是 X 的一个子集使得 $X\leftrightarrow\cdots\leftrightarrow X_0\rightarrow\cdots\rightarrow\varnothing(Y)$,其中 $X\supset\cdots\supset X_0\supset\cdots\supset\varnothing$, $S_X(Y)=\cdots=S_{X_0}(Y)\supset\cdots\supset S_{\varnothing}(Y)=S_\delta(Y)=\bigcup_{W\in U/Y}(\bigcup_{V=U,V\subseteq W}V)=\varnothing$,其中 $U/Y\neq U/\delta=\{U\}$.我们称 X 的这种子集 X_0 为 X 的约简(对于 Y 而言).注意约简可能不是惟一的.我们引进下述约简的概念.

令(U,A)是一个决策表,$A=C\cup D$,$C\cap D=\varnothing$,其中 C 是条件属性集,D 是决策属性集.令$\varnothing\subset X\subseteq C$,$\varnothing\subset Y\subseteq D$,$U/Y\neq U/\delta=\{U\}$.如果 $X_0\subseteq X$ 满足

(1)$S_{X_0}(Y)=S_X(Y)$,即 $X_0\leftrightarrow X(Y)$.

(2)如果 $X'\subset X_0$,则 $S_X(Y)\supset S_{X'}(Y)$,即如果 $X'\subset X_0$,则 $X'\leftrightarrow X(Y)$ 不成立.

则称 X_0 是 X 的一个约简(对于 Y 而言).

空子集$\varnothing$的约简为$\varnothing$(对于 Y 而言).

这就是说,X_0 是 X 的极小恒等依赖子集(对于 Y 而言),即对于 $X\supseteq\cdots\supseteq X_0\supset X'\supseteq\cdots$,我们有 $X\leftrightarrow\cdots\leftrightarrow X_0\rightarrow X'\rightarrow\cdots(Y)$.

从这个定义,我们知道找所有约简的时间复杂性是指数的.首先,我们需要考虑 X 的所有 $|2^X|=2^{|X|}$ 个子集.其次,对于每个子集 X_0,我们需要计算 $S_{X_0}(Y)$.这样,对于每个子集 X_0,计算 $S_{X_0}(Y)$的时间复杂性是 $O(|A|\cdot|U|^2)$.因此,整个代价是 $O(2^{|X|}|A|\,|U|^2)$.

对于知识发现,X 的一个约简 X_0(对于 Y 而言)意味着下面:(1)规则"X_0 蕴涵 Y"和规则"X 蕴涵 Y"同样强.(2)如果 $X'\subset X_0$,则规则"X' 蕴涵 Y"严格弱于规则"X 蕴涵 Y".

第三章　一般关系下的粗糙集模型

在 Pawlak 粗糙集模型中，论域上的等价关系起着至关重要的作用．但在很多实际问题中，论域上的二元关系不是等价的，这时 Pawlak 粗糙集模型的应用受到了限制，为此必须将 Pawlak 粗糙集模型进行推广．本章讨论在一般二元关系下的粗糙集模型．

§3.1　二元关系与邻域算子

设 U 是非空有限论域，对于任意的 $x\in U$，对应一个 U 的子集 $n(x)$ 称为 x 的一个邻域．x 的一个邻域可以包含 x，也可以不包含 x．这样，$n:U\to P(U)$ 是一个算子，称为邻域算子．对于 $X\subseteq U$，记 $n(X)=\bigcup\limits_{x\in X}n(x)$，称 $n(X)$ 为集合 X 的邻域．

定义 3.1　设 $n:U\to P(U)$ 是一个邻域算子，称 n 是串行的(serial)，若对于任意 $x\in U$ 存在 $y\in U$ 使得 $y\in n(x)$，即 $\forall x\in U,n(x)\neq\varnothing$．称 n 是逆串行的(inverse serial)，若对于任意 $x\in U$ 存在 $y\in U$ 使得 $x\in n(y)$，即 $n(U)=U$．称 n 是自反的，若对于任意 $x\in U,x\in n(x)$．称 n 是对称的，若对于任意 $x,y\in U,x\in n(y)$ 蕴涵 $y\in n(x)$．称 n 是传递的，若对于任意 $x,y,z\in U,y\in n(x)$ 和 $z\in n(y)$ 蕴涵 $z\in n(x)$．称 n 是欧几里得的(Euclidean)，若对于任意 $x,y,z\in U,y\in n(x)$ 和 $z\in n(x)$ 蕴涵 $z\in n(y)$．

定义 3.2　设 R 和 S 是 U 上的两个二元关系，我们定义

$$R\cap S=\{(x,y)\mid xRy\wedge xSy\},$$
$$R\cup S=\{(x,y)\mid xRy\vee xSy\},$$
$$R^{-1}=\{(y,x)\mid xRy\},$$
$$\sim R=\{(x,y)\mid\neg xRy\}.$$

它们分别称为 R 和 S 的交，R 和 S 的并，R 的逆和 R 的补．

显然，R 是自反的 $\Leftrightarrow R^{-1}$ 是自反的；R 是传递的 $\Leftrightarrow R^{-1}$ 是传递的；R 是对称的 $\Leftrightarrow R=R^{-1}$；对于 U 上的任意二元关系 R，$R\cap R^{-1}$ 和 $R\cup R^{-1}$ 都是对称的．

定义 3.3　设 R 是 U 上的二元关系，对于 $x,y\in U$，若 xRy，即 $(x,y)\in R$，则称 x 是 y 的前继(predecessor)，y 是 x 的后继(successor)，记

$$R_s(x)=\{y\in U\mid xRy\},$$

$$R_p(x) = \{y \in U \mid yRx\},$$
$$R_{p\wedge s}(x) = \{y \in U \mid xRy \wedge yRx\} = R_p(x) \cap R_s(x),$$
$$R_{p\vee s}(x) = \{y \in U \mid xRy \vee yRx\} = R_p(x) \cup R_s(x).$$

它们分别称为 x 的后继邻域,前继邻域,下继邻域和上继邻域.

显然这些邻域有以下关系:

$$R_{p\wedge s}(x) \subseteq R_p(x) \subseteq R_{p\vee s}(x),$$
$$R_{p\wedge s}(x) \subseteq R_s(x) \subseteq R_{p\vee s}(x),$$
$$R_p(x) = R_s^{-1}(x).$$

二元关系 R 和邻域算子 R_s,R_p 之间能被相互惟一确定

$$xRy \Leftrightarrow x \in R_p(y) \Leftrightarrow y \in R_s(x).$$

定理 3.1 若 R 和 S 都是 U 上的二元关系,则

(1) $R \subseteq S \Leftrightarrow \forall x \in U, R_s(x) \subseteq S_s(x)$;

(2) $(\sim R)_s(x) = \sim R_s(x)$;

(3) $(R \cap S)_s(x) = R_s(x) \cap S_s(x)$,

特别地,$R_{p\wedge s}(x) = (R \cap R^{-1})_s(x)$;

(4) $(R \cup S)_s(x) = R_s(x) \cup S_s(x)$,

特别地 $R_{p\vee s}(x) = (R \cup R^{-1})_s(x)$.

证 显然. □

二元关系 R 的性质可以由其邻域算子来描述:

R 是串行的 $\Leftrightarrow \forall x \in U, R_s(x) \neq \varnothing \Leftrightarrow R_p(U) = U$,

R 是逆串行的 $\Leftrightarrow \forall x \in U, R_p(x) \neq \varnothing \Leftrightarrow R_s(U) = U$,

R 是自反的 $\Leftrightarrow \forall x \in U, x \in R_s(x)$,

R 是对称的 $\Leftrightarrow \forall x, y \in U, x \in R_s(y)$ 蕴涵 $y \in R_s(x)$,

R 是传递的 $\Leftrightarrow \forall x, y, z \in U, y \in R_s(x)$ 和 $z \in R_s(y)$ 蕴涵 $z \in R_s(x)$,

$\Leftrightarrow \forall x, y \in U, y \in R_s(x)$ 蕴涵 $R_s(y) \subseteq R_s(x)$,

R 是欧几里得的 $\Leftrightarrow \forall x, y, z \in U, y \in R_s(x)$ 和 $z \in R_s(x)$ 蕴涵 $z \in R_s(y)$,

$\Leftrightarrow \forall x, y \in U, y \in R_s(x)$ 蕴涵 $R_s(x) \subseteq R_s(y)$.

表 3.1 给出了二元关系的性质和它导出的邻域算子的性质之间的联系.

表 3.1 二元关系 R 及其导出的邻域算子

R	R_s	R_p	$R_{p\wedge s}$	$R_{p\vee s}$
任意			对称	对称
串行	串行	逆串行		(逆)串行
逆串行	逆串行	串行		(逆)串行
自反	自反	自反	自反	自反
对称	对称	对称	对称	对称
传递	传递	传递	传递	
欧几里得	欧几里得		欧几里得	

§3.2 二元关系与粗糙近似算子

Pawlak 粗糙集模型中的近似空间上的二元关系是一个等价关系的，下面我们将粗糙集模型推广到任意的二元关系，并讨论一些特殊的二元关系与近似算子的特性刻画.

定义 3.4 设 U 是有限非空的论域，$R\subseteq U\times U$ 为 U 上一个任意的二元关系，称 $A=(U,R)$为广义近似空间. 对于任意 $X\subseteq U$，X 关于近似空间 $A=(U,R)$的下近似$\underline{\mathrm{apr}}_AX$ 和上近似$\overline{\mathrm{apr}}_AX$ 分别定义为

$$\underline{\mathrm{apr}}_AX=\{x\in U\mid R_s(x)\subseteq X\},$$
$$\overline{\mathrm{apr}}_AX=\{x\in U\mid R_s(x)\cap X\neq\varnothing\}.$$

X 关于近似空间 A 的正域 $\mathrm{pos}_A(X)$，负域 $\mathrm{neg}_A(X)$和边界 $\mathrm{bn}_A(X)$分别定义为

$$\mathrm{pos}_A(X)=\underline{\mathrm{apr}}_AX,$$
$$\mathrm{neg}_A(X)=\sim\overline{\mathrm{apr}}_AX=\{x\in U\mid R_s(x)\cap X=\varnothing\},$$
$$\mathrm{bn}_A(X)=\overline{\mathrm{apr}}_AX\setminus\underline{\mathrm{apr}}_AX.$$

当$\underline{\mathrm{apr}}_AX=\overline{\mathrm{apr}}_AX$ 时，称 X 关于近似空间 A 是可定义的，否则称 X 关于近似空间 A 是粗糙的.

注 如果 R 是等价关系，则 $R_s(x)$就是对象 x 所在的 R 等价类$[x]_R$，也就是说 x 所在的等价类可以看成 x 的邻域，这时所得到的下近似$\underline{\mathrm{apr}}_AX$ 和上近似$\overline{\mathrm{apr}}_AX$ 就是在 Pawlak 意义下的下近似 $\underline{R}X$ 和上近似 $\overline{R}X$. 以后若 R 比较明确我们将近似算子中的下标 A 省略.

显然，近似算子$\underline{\mathrm{apr}}$和$\overline{\mathrm{apr}}$是 $P(U)\to P(U)$的算子，我们称系统$(P(U),\cap,\cup,\sim,\underline{\mathrm{apr}},\overline{\mathrm{apr}})$为广义粗糙集代数.

定理 3.2 设 R 是论域 U 上的一个任意的二元关系，则按定义 3.4 给出的下近似和上近似满足下列对偶性质：

(1)$\underline{\mathrm{apr}}X=\sim(\overline{\mathrm{apr}}(\sim X))$，$\overline{\mathrm{apr}}X=\sim(\underline{\mathrm{apr}}(\sim X))$.

(2)$\underline{\mathrm{apr}}U=U$，$\overline{\mathrm{apr}}\varnothing=\varnothing$.

(3)$\underline{\mathrm{apr}}(X\cap Y)=\underline{\mathrm{apr}}X\cap\underline{\mathrm{apr}}Y$，$\overline{\mathrm{apr}}(X\cup Y)=\overline{\mathrm{apr}}X\cup\overline{\mathrm{apr}}Y$.

(4)$X\subseteq Y\Rightarrow\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}Y$，$\overline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}Y$.

(5)$\underline{\mathrm{apr}}(X\cup Y)\supseteq\underline{\mathrm{apr}}X\cup\underline{\mathrm{apr}}Y$，$\overline{\mathrm{apr}}(X\cap Y)\subseteq\overline{\mathrm{apr}}X\cap\overline{\mathrm{apr}}Y$.

证 (1)由于

$$x\in\underline{\mathrm{apr}}X\Leftrightarrow R_s(x)\subseteq X\Leftrightarrow R_s(x)\cap(\sim X)=\varnothing$$
$$\Leftrightarrow x\in\sim(\overline{\mathrm{apr}}(\sim X)),$$

所以$\underline{\mathrm{apr}}X=\sim(\overline{\mathrm{apr}}(\sim X))$. 同理，$\overline{\mathrm{apr}}X=\sim(\underline{\mathrm{apr}}(\sim X))$.

(2)显然.

(3)由于

$$\begin{aligned}x\in\underline{\mathrm{apr}}(X\cap Y)&\Leftrightarrow R_s(x)\subseteq X\cap Y\\&\Leftrightarrow R_s(x)\subseteq X\text{ 且 }R_s(x)\subseteq Y\\&\Leftrightarrow x\in\underline{\mathrm{apr}}X\text{ 且 }x\in\underline{\mathrm{apr}}Y\\&\Leftrightarrow x\in\underline{\mathrm{apr}}X\cap\underline{\mathrm{apr}}Y,\end{aligned}$$

所以$\underline{\mathrm{apr}}(X\cap Y)=\underline{\mathrm{apr}}X\cap\underline{\mathrm{apr}}Y$. 同理，$\overline{\mathrm{apr}}(X\cup Y)=\overline{\mathrm{apr}}X\cup\overline{\mathrm{apr}}Y$.

(4)由定义直接可得.

(5)由性质(4)即得. □

注 上述定理中(5)一般情况下等式不成立.

例 3.1 设 $U=\{a,b,c\}$，$R=\{(a,a),(a,b),(b,b),(b,a),(c,b)\}$，$X=\{a\}$，$Y=\{b,c\}$，则可以验证

$$\underline{\mathrm{apr}}X\cup\underline{\mathrm{apr}}Y=\{c\}\text{，但}\underline{\mathrm{apr}}(X\cup Y)=U.$$

又若取 $A=\{a,c\}$，$B=\{b,c\}$，则可以验证

$$\overline{\mathrm{apr}}(A\cap B)=\overline{\mathrm{apr}}\{c\}=\varnothing,$$

而

$$\overline{\mathrm{apr}}A\cap\overline{\mathrm{apr}}B=\{a,b\}\cap U=\{a,b\}.$$

下面我们讨论论域上的二元关系与近似算子的特征联系.

定理 3.3 设 R 是 U 上二元关系，则以下等价：

(1) R 是串行的.

(2) $\underline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}X$，$\forall X\subseteq U$.

(3) $\underline{\mathrm{apr}}\varnothing=\varnothing$.

(4) $\overline{\mathrm{apr}}U=U$.

证 "(1)⇒(2)"对于$\forall x\in\underline{\mathrm{apr}}X$，由定义有 $R_s(x)\subseteq X$，由 R 是串行的可知$R_s(x)\neq\varnothing$，因此 $R_s(x)\cap X\neq\varnothing$，从而 $x\in\overline{\mathrm{apr}}X$，即$\underline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}X$.

"(2)⇒(1)"反证,若∃$x\in U$使$R_s(x)=\varnothing$,则对于$\forall X\subseteq U$,$R_s(x)\subseteq X$,即$x\in\underline{\mathrm{apr}}X$,但是$R_s(x)\cap X=\varnothing$,于是$x\notin\overline{\mathrm{apr}}X$,这与(2)成立矛盾!

"(2)⇔(3)"事实上,由定理3.2的对偶性质(1)可得

$$\begin{aligned}\underline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}X&\Leftrightarrow\underline{\mathrm{apr}}X\cap\sim\overline{\mathrm{apr}}X=\varnothing\\&\Leftrightarrow\underline{\mathrm{apr}}X\cap\underline{\mathrm{apr}}(\sim X)=\varnothing\\&\Leftrightarrow\underline{\mathrm{apr}}(X\cap\sim X)=\varnothing\\&\Leftrightarrow\underline{\mathrm{apr}}\varnothing=\varnothing.\end{aligned}$$

"(3)⇔(4)"由定理3.2的对偶性质(1)可得

$$\underline{\mathrm{apr}}\varnothing=\varnothing\Leftrightarrow\sim\underline{\mathrm{apr}}\varnothing=\sim\underline{\mathrm{apr}}(\sim U)=\sim\varnothing=U\Leftrightarrow\overline{\mathrm{apr}}U=U.$$

□

引理3.1 设R是U上任意一个二元关系,则对于$\forall x\in U$,有

$$\overline{\mathrm{apr}}\{x\}=R_s^{-1}(x)=R_p(x)=\{y\in U\mid x\in R_s(y)\}.\tag{3.1}$$

证 对于$\forall y\in\overline{\mathrm{apr}}\{x\}$,由定义$R_s(y)\cap\{x\}\neq\varnothing$,即$x\in R_s(y)$,也就是说$y\in R_p(x)$,从而$\overline{\mathrm{apr}}\{x\}\subseteq\{y\in U\mid x\in R_s(y)\}$,反之亦然,故(3.1)式成立.

□

定理3.4 设R是U上任意一个二元关系,则以下等价:

(1) R是逆串行的.

(2) $\overline{\mathrm{apr}}\{x\}\neq\varnothing$, $\forall x\in U$.

(3) $R_s(U)=U$.

证 由于R是逆串行的当且仅当对于$\forall x\in U$,有$R_s^{-1}(x)=R_p(x)\neq\varnothing$,从而可由(3.1)式以及定义直接证得. □

定理3.5 设R是U上任意一个二元关系,则以下等价:

(1) R是自反的.

(2) $\underline{\mathrm{apr}}X\subseteq X$, $\forall X\subseteq U$.

(3) $X\subseteq\overline{\mathrm{apr}}X$, $\forall X\subseteq U$.

证 "(1)⇒(2)"设$x\in\underline{\mathrm{apr}}X$,由定义得$R_s(x)\subseteq X$,由于$R$是自反的,因此$x\in R_s(x)$,从而$x\in X$,即$\underline{\mathrm{apr}}X\subseteq X$.

"(2)⇒(3)"对于$\forall X\subseteq U$,由(2)可得$\underline{\mathrm{apr}}(\sim X)\subseteq\sim X$,从而由定理3.2的对偶性质(1)知$X=\sim(\sim X)\subseteq\sim\underline{\mathrm{apr}}(\sim X)=\overline{\mathrm{apr}}X$.

"(3)⇒(1)"若(3)成立,则对于任意$x\in U$,有$x\in\overline{\mathrm{apr}}\{x\}$,由引理3.1可得$x\in R_p(x)$,从而$x\in R_s(x)$,即$R$自反的. □

定理3.6 设R是U上任意一个二元关系,则以下等价:

(1) R是传递的.

(2) $\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}\ \underline{\mathrm{apr}}X$, $\forall X\subseteq U$.

(3) $\overline{\mathrm{apr}}\,\overline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}X$, $\forall X\subseteq U$.

证　"(1)⇒(2)"对于 $\forall x\in\underline{\mathrm{apr}}X$，由定义可得 $R_s(x)\subseteq X$. 对于 $\forall y\in R_s(x)$，由 R 的传递性可得 $R_s(y)\subseteq R_s(x)$，从而 $R_s(y)\subseteq X$，由下近似定义可知 $y\in\underline{\mathrm{apr}}X$，于是由 $y\in R_s(x)$ 的任意性得 $R_s(x)\subseteq\underline{\mathrm{apr}}X$，这样就有 $x\in\underline{\mathrm{apr}}\,\underline{\mathrm{apr}}X$，因此(2)成立.

"(2)⇒(3)"可由定理 3.2 的对偶性质(1)证得.

"(3)⇒(1)"设 $y\in R_s(x)$ 并且 $z\in R_s(y)$，则由 $z\in R_s(y)$ 和引理 3.1 得 $y\in\overline{\mathrm{apr}}\{z\}$，结合 $y\in R_s(x)$ 得 $y\in R_s(x)\cap\overline{\mathrm{apr}}\{z\}$，这说明 $R_s(x)\cap\overline{\mathrm{apr}}\{z\}\neq\varnothing$，从而由上近似的定义得 $x\in\overline{\mathrm{apr}}\,\overline{\mathrm{apr}}\{z\}$，由(3)成立推得 $x\in\overline{\mathrm{apr}}\{z\}$，于是由上近似定义有 $z\in R_s(x)$. 这样我们从 $y\in R_s(x)$ 和 $z\in R_s(y)$ 得到了 $z\in R_s(x)$，即 R 是传递的. □

定理 3.7　设 R 是 U 上任意一个二元关系，则以下等价：

(1) R 是对称的.

(2) $X\subseteq\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$, $\forall X\subseteq U$.

(3) $\overline{\mathrm{apr}}\,\underline{\mathrm{apr}}X\subseteq X$, $\forall X\subseteq U$.

证　"(1)⇒(2)"对于 $\forall x\in X$ 和 $y\in R_s(x)$，由于 R 是对称的，因此 $x\in R_s(y)$，于是 $x\in R_s(y)\cap X$，这说明 $R_s(y)\cap X\neq\varnothing$，由上近似的定义得 $y\in\overline{\mathrm{apr}}X$. 由 $y\in R_s(x)$ 的任意性得 $R_s(x)\subseteq\overline{\mathrm{apr}}X$，再由下近似的定义得 $x\in\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$，这样就证明了 $X\subseteq\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$.

"(2)⇔(3)"可由定理 3.2 的性质(1)证得.

"(2)⇒(1)"设 $x, y\in U$，且 $y\in R_s(x)$. 由于(2)成立，因此 $x\in\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}\{x\}$，由下近似定义得 $R_s(x)\subseteq\overline{\mathrm{apr}}\{x\}$，结合 $y\in R_s(x)$ 得 $y\in\overline{\mathrm{apr}}\{x\}$，从而由上近似的定义得 $R_s(y)\cap\{x\}\neq\varnothing$，这样就有 $x\in R_s(y)$，所以 R 是对称的. □

定理 3.8　设 R 是 U 上任意一个二元关系，则以下等价：

(1) R 是欧几里得的.

(2) $\overline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$, $\forall X\subseteq U$.

(3) $\overline{\mathrm{apr}}\,\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}X$, $\forall X\subseteq U$.

证　"(1)⇒(2)"设 $x\in\overline{\mathrm{apr}}X$，由上近似定义得 $R_s(x)\cap X\neq\varnothing$. 对于 $\forall y\in R_s(x)$，由 R 的欧几里得性得 $R_s(x)\subseteq R_s(y)$，结合 $R_s(x)\cap X\neq\varnothing$ 可得 $R_s(y)\cap X\neq\varnothing$，由上近似的定义知 $y\in\overline{\mathrm{apr}}X$，再由 $y\in R_s(x)$ 的任意性得 $R_s(x)\subseteq\overline{\mathrm{apr}}X$，从而由下近似的定义可知 $x\in\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$，这样我们就证明了 $\overline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}\,\overline{\mathrm{apr}}X$.

"(2)⇒(1)"设 $y\in R_s(x)$，并且 $z\in R_s(x)$，由 $z\in R_s(x)$ 和引理 3.1 可得

$x\in\overline{\mathrm{apr}}\{z\}$由于(2)成立,因此 $x\in\underline{\mathrm{apr}}\ \overline{\mathrm{apr}}\{z\}$,由下近似的定义得 $R_s(x)\subseteq\overline{\mathrm{apr}}\{z\}$,结合 $y\in R_s(x)$得 $y\in\overline{\mathrm{apr}}\{z\}$,于是由引理 3.1 得 $z\in R_s(y)$. 这样我们就从 $y\in R_s(x)$和 $z\in R_s(x)$得到了 $z\in R_s(y)$,即 R 是欧几里得的.

"(2)$\Leftrightarrow$(3)"可由定理 3.2 的性质(1)证得. □

给定 U 上任意一个二元关系 R,它可以导出邻域算子 $R_s, R_p, R_{p\wedge s}$和 $R_{p\vee s}$,由这些邻域算子可以定义下近似和上近似

(D1) $\begin{cases}\underline{\mathrm{apr}}_pX=\{x\in U\mid R_p(x)\subseteq X\},\\ \overline{\mathrm{apr}}_pX=\{x\in U\mid R_p(x)\cap X\neq\varnothing\}.\end{cases}$

(D2) $\begin{cases}\underline{\mathrm{apr}}_sX=\{x\in U\mid R_s(x)\subseteq X\},\\ \overline{\mathrm{apr}}_sX=\{x\in U\mid R_s(x)\cap X\neq\varnothing\}.\end{cases}$

(D3) $\begin{cases}\underline{\mathrm{apr}}_{p\wedge s}X=\{x\in U\mid R_{p\wedge s}(x)\subseteq X\},\\ \overline{\mathrm{apr}}_{p\wedge s}X=\{x\in U\mid R_{p\wedge s}(x)\cap X\neq\varnothing\}.\end{cases}$

(D4) $\begin{cases}\underline{\mathrm{apr}}_{p\vee s}X=\{x\in U\mid R_{p\vee s}(x)\subseteq X\},\\ \overline{\mathrm{apr}}_{p\vee s}X=\{x\in U\mid R_{p\vee s}(x)\cap X\neq\varnothing\}.\end{cases}$

对于 Pawlak 粗糙集模型,由于 R 是等价关系,因此上述定义(D1)~(D4)都是等价的.

定理 3.9 (D1)~(D4)中任意两对是等价的当且仅当 R 是对称的.

证 "$\Leftarrow$"显然.

"$\Rightarrow$"取单点集 $X=\{x\}$,由引理 3.1 得

$$\overline{\mathrm{apr}}_p\{x\} = \{y\in U\mid x\in R_p(y)\} = R_s(x). \tag{3.2}$$

同理

$$\overline{\mathrm{apr}}_s\{x\} = R_p(x), \tag{3.3}$$

$$\overline{\mathrm{apr}}_{p\wedge s}\{x\} = R_{p\wedge s}(x), \tag{3.4}$$

$$\overline{\mathrm{apr}}_{p\vee s}\{x\} = R_{p\vee s}(x). \tag{3.5}$$

由于对于任意 $x\in U$,(3.2)~(3.5)中任意两个式子等式的右边是相等的当且仅当 R 是对称的,必要性得证. □

注 由(3.2)~(3.5)可知上近似可以用集合的邻域算子来表示,即

$$\overline{\mathrm{apr}}_pX = R_s(X),$$

$$\overline{\mathrm{apr}}_sX = R_p(X),$$

$$\overline{\mathrm{apr}}_{p\wedge s}X = R_{p\wedge s}(X), \tag{3.6}$$

$$\overline{\mathrm{apr}}_{p\vee s}X = R_{p\vee s}(X). \tag{3.7}$$

这样我们发现由前继邻域算子定义的 X 的上近似恰好是 X 的后继邻域;由后继邻域算子定义的 X 的上近似恰好是 X 的前继邻域;由上继邻域算子和下继邻域算子定义的 X 的上近似分别恰好是 X 的上继邻域和下继邻域. 特别地由

(3.6)和(3.7)我们可以得到

$$\underline{\mathrm{apr}}_{p\wedge s}X = \{x \in U \mid \overline{\mathrm{apr}}_{p\wedge s}\{x\} \subseteq X\},$$
$$\underline{\mathrm{apr}}_{p\vee s}X = \{x \in U \mid \overline{\mathrm{apr}}_{p\vee s}\{x\} \subseteq X\}.$$

但近似算子$\underline{\mathrm{apr}}_p$ 与$\overline{\mathrm{apr}}_p$ 以及$\underline{\mathrm{apr}}_s$ 与$\overline{\mathrm{apr}}_s$ 之间没有类似上面的关系式.

§3.3 近似算子的其他定义形式与比较

我们知道 Pawlak 粗糙集的近似算子的定义有两种主要形式,即

$$\text{(i)}\begin{cases}\underline{R}X = \{x \in U \mid [x] \subseteq X\},\\ \overline{R}X = \{x \in U \mid [x] \cap X \neq \varnothing\}.\end{cases}$$

$$\text{(ii)}\begin{cases}\underline{R}X = \bigcup\{[x] \mid [x] \subseteq X\},\\ \overline{R}X = \bigcup\{[x] \mid [x] \cap X \neq \varnothing\}.\end{cases}$$

由于 R 是等价关系,这两种定义是等价的,并且定义中的每一对下近似和上近似构成了对偶.那么当 R 是 U 上的任意的二元关系时,用 $R_s(x)$代替$[x]$并按(i)和(ii)方式去定义下、上近似,即

$$\text{(Ⅰ)}\begin{cases}\underline{\mathrm{apr}}X = \{x \in U \mid R_s(x) \subseteq X\},\\ \overline{\mathrm{apr}}X = \{x \in U \mid R_s(x) \cap X \neq \varnothing\}.\end{cases}$$

$$\text{(Ⅱ)}\begin{cases}\underline{\mathrm{apr}}X = \bigcup\{R_s(x) \mid R_s(x) \subseteq X\},\\ \overline{\mathrm{apr}}X = \bigcup\{R_s(x) \mid R_s(x) \cap X \neq \varnothing\}.\end{cases}$$

这时(Ⅰ)和(Ⅱ)是否等价?每一对近似是否还能构成对偶?我们说一般情况下(Ⅰ)和(Ⅱ)不是等价的,由上一节知(Ⅰ)中的下近似与上近似可以构成对偶,但是(Ⅱ)中的下近似与上近似不能构成对偶,见下例

例3.2 设 $U=\{a,b,c\}$,$R=\{(a,a),(a,b),(b,c),(c,b)\}$,这时 $R_s(a)=\{a,b\}$,$R_s(b)=\{c\}$,$R_s(c)=\{b\}$,令 $X=\{a,c\}$,可以验证

$$\underline{\mathrm{apr}}X = \{b\} \neq \underline{\mathrm{apr}}X = \{c\},$$
$$\overline{\mathrm{apr}}X = \{a,b\} \neq \overline{\mathrm{apr}}X = \{a,b,c\},$$
$$\sim\underline{\mathrm{apr}}(\sim X) = \sim\{b\} = \{a,c\} \neq \overline{\mathrm{apr}}X = \{a,b,c\}.$$

近似算子的对偶性质很重要,能简化近似的计算,为了解决(Ⅱ)中定义的近似算子不能构成对偶的缺陷,我们用两种方式去推广(Ⅱ)的定义.

$$\begin{aligned}(\text{Ⅱ}')\ \underline{\mathrm{apr}}'X &= \bigcup\{R_s(x) \mid R_s(x) \subseteq X, x \in U\}\\ &= \{x \in U \mid \exists y[x \in R_s(y), R_s(y) \subseteq X]\},\\ \overline{\mathrm{apr}}'X &= \sim\underline{\mathrm{apr}}'(\sim X)\\ &= \sim\{x \in U \mid \exists y[x \in R_s(y), R_s(y) \subseteq \sim X]\}\end{aligned}$$

$$= \{x \in U \mid \forall y[x \in R_s(y) \Rightarrow R_s(y) \cap X \neq \varnothing]\} \cup (\sim R_s(U)).$$

$$\begin{aligned}
(\text{II}'')\underline{\mathrm{apr}}''X &= \sim \overline{\mathrm{apr}}''(\sim X)\\
&= \sim \{x \in U \mid \exists y[x \in R_s(y), R_s(y) \cap \sim X \neq \varnothing]\}\\
&= \{x \in U \mid \forall y[x \in R_s(y) \Rightarrow R_s(y) \cap \sim X = \varnothing]\} \cup (\sim R_s(U))\\
&= \{x \in U \mid \forall y[x \in R_s(y) \Rightarrow R_s(y) \subseteq X]\} \cup (\sim R_s(U))\\
\overline{\mathrm{apr}}''X &= \bigcup \{R_s(x) \mid R_s(x) \cap X \neq \varnothing\}\\
&= \{x \in U \mid \exists y[x \in R_s(y), R_s(y) \cap X \neq \varnothing]\}.
\end{aligned}$$

定理 3.10 设 R 是论域 U 上的一个任意的二元关系，则下近似$\underline{\mathrm{apr}}'$和上近似$\overline{\mathrm{apr}}'$满足下列性质：

(1) $\underline{\mathrm{apr}}'X = \sim(\overline{\mathrm{apr}}'(\sim X))$，$\overline{\mathrm{apr}}'X = \sim(\underline{\mathrm{apr}}'(\sim X))$；

(2) $X \subseteq Y \Rightarrow \underline{\mathrm{apr}}'X \subseteq \underline{\mathrm{apr}}'Y$，$\overline{\mathrm{apr}}'X \subseteq \overline{\mathrm{apr}}'Y$；

(3) $\underline{\mathrm{apr}}'(X \cap Y) \subseteq \underline{\mathrm{apr}}'X \cap \underline{\mathrm{apr}}'Y$，$\overline{\mathrm{apr}}'(X \cup Y) \supseteq \overline{\mathrm{apr}}'X \cup \overline{\mathrm{apr}}'Y$；

(4) $\underline{\mathrm{apr}}'(X \cup Y) \supseteq \underline{\mathrm{apr}}'X \cup \underline{\mathrm{apr}}'Y$，$\overline{\mathrm{apr}}'(X \cap Y) \subseteq \overline{\mathrm{apr}}'X \cap \overline{\mathrm{apr}}'Y$；

(5) $\underline{\mathrm{apr}}'X \subseteq X \subseteq \overline{\mathrm{apr}}'X$；

(6) $\underline{\mathrm{apr}}'X = \underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X$，$\overline{\mathrm{apr}}'\,\overline{\mathrm{apr}}'X = \overline{\mathrm{apr}}'X$；

(7) 当 R 是逆串行时，有$\underline{\mathrm{apr}}'U = U$，$\overline{\mathrm{apr}}'\varnothing = \varnothing$.

证 由对偶性每一部分只证一半就可以了.(1)～(5)由定义直接可得.

(6)设 $x \in \underline{\mathrm{apr}}'X$，由定义存在 $y \in U$，$x \in R_s(y)$且 $R_s(y) \subseteq X$，显然对于任意 $z \in R_s(y)$，有 $z \in \underline{\mathrm{apr}}'X$，从而 $R_s(y) \subseteq \underline{\mathrm{apr}}'X$，结合 $x \in R_s(y)$可得 $x \in \underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X$，于是

$$\underline{\mathrm{apr}}'X \subseteq \underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X. \tag{3.8}$$

其次，由(2) 和(5) 可得

$$\underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X \subseteq \underline{\mathrm{apr}}'X, \tag{3.9}$$

结合(3.8) 和(3.9) 得

$$\underline{\mathrm{apr}}'X = \underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X.$$

(7)可由逆串行的定义直接证得. □

定理 3.11 设 R 是论域 U 上的一个任意的二元关系，则下近似$\underline{\mathrm{apr}}''$和上近似$\overline{\mathrm{apr}}''$满足下列性质：

(1) $\underline{\mathrm{apr}}''X = \sim(\overline{\mathrm{apr}}''(\sim X))$，$\overline{\mathrm{apr}}''X = \sim(\underline{\mathrm{apr}}''(\sim X))$；

(2) $\underline{\mathrm{apr}}''U = U$，$\overline{\mathrm{apr}}''\varnothing = \varnothing$；

(3) $X \subseteq Y \Rightarrow \underline{\mathrm{apr}}''X \subseteq \underline{\mathrm{apr}}''Y$，$\overline{\mathrm{apr}}''X \subseteq \overline{\mathrm{apr}}''Y$；

(4) $\underline{\mathrm{apr}}''(X \cap Y) = \underline{\mathrm{apr}}''X \cap \underline{\mathrm{apr}}''Y$，$\overline{\mathrm{apr}}''(X \cup Y) = \overline{\mathrm{apr}}''X \cup \overline{\mathrm{apr}}''Y$；

(5) $\underline{\mathrm{apr}}''(X\cup Y)\supseteq\underline{\mathrm{apr}}''X\cup\underline{\mathrm{apr}}''Y$，$\overline{\mathrm{apr}}''(X\cap Y)\subseteq\overline{\mathrm{apr}}''X\cap\overline{\mathrm{apr}}''Y$；

(6) $\underline{\mathrm{apr}}''\underline{\mathrm{apr}}''X\subseteq\underline{\mathrm{apr}}''X$，$\overline{\mathrm{apr}}''X\subseteq\overline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$；

(7) $X\subseteq\underline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$，$\overline{\mathrm{apr}}''\underline{\mathrm{apr}}''X\subseteq X$；

(8) 当 R 是逆串行时，有$\underline{\mathrm{apr}}''X\subseteq X\subseteq\overline{\mathrm{apr}}''X$.

证　由对偶性每一部分只证一半就可以了.(1)～(5)可由定义直接证得.

(6)我们只证对于任意 $X\subseteq U$，$\overline{\mathrm{apr}}''X\subseteq\overline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$ 成立.

事实上，对于任意 $x\in\overline{\mathrm{apr}}''X$，由定义 $\exists y\in U$，使 $x\in R_s(y)$ 且 $R_s(y)\cap X\neq\varnothing$. 对于 $\forall z\in R_s(y)$，由于 $R_s(y)\cap X\neq\varnothing$，因此 $z\in\overline{\mathrm{apr}}''X$，于是 $R_s(y)\subseteq\overline{\mathrm{apr}}''X$. 由 $x\in R_s(y)$ 得 $x\in R_s(y)\cap\overline{\mathrm{apr}}''X\neq\varnothing$，从而 $x\in\overline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$，所以$\overline{\mathrm{apr}}''X\subseteq\overline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$ 成立.

(7)对于任意 $X\subseteq U$ 和 $x\in X$，若 $x\notin R_s(U)$，则由定义知 $x\in\underline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$. 假如 $x\in R_s(U)$，则对于 $y\in U$ 若满足 $x\in R_s(y)$，则有 $x\in R_s(y)\cap X\neq\varnothing$，由上近似定义可以看出 $R_s(y)\subseteq\overline{\mathrm{apr}}''X$，因而由下近似的定义知 $x\in\underline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$. 这样我们就证明了 $X\subseteq\underline{\mathrm{apr}}''\overline{\mathrm{apr}}''X$.

(8)当 R 是逆串行时有 $R_s(U)=U$，对于任意 $X\subseteq U$ 和 $x\in X$，则存在 $y\in U$ 满足 $x\in R_s(y)$，从而 $x\in R_s(y)\cap X\neq\varnothing$，由定义得 $x\in\overline{\mathrm{apr}}''X$，这样就证明了 $X\subseteq\overline{\mathrm{apr}}''X$. □

定理 3.12　设 R 是论域 U 上的一个任意的二元关系，则以下成立：

(1)若 R 是逆串行的，则
$$\underline{\mathrm{apr}}''X\subseteq\underline{\mathrm{apr}}'X\subseteq X\subseteq\overline{\mathrm{apr}}'X\subseteq\overline{\mathrm{apr}}''X;$$

(2)若 R 是自反的，则
$$\underline{\mathrm{apr}}''X\subseteq\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}'X\subseteq X\subseteq\overline{\mathrm{apr}}'X\subseteq\overline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}''X.$$

证　由对偶性只须证一半即可.

(1)设 R 是逆串行的，由定理 3.10 和定理 3.11 知
$$\underline{\mathrm{apr}}'X\subseteq X\subseteq\overline{\mathrm{apr}}'X,\ \text{且}\ \underline{\mathrm{apr}}''X\subseteq X\subseteq\overline{\mathrm{apr}}''X.$$
由对偶性我们只须证
$$\underline{\mathrm{apr}}''X\subseteq\underline{\mathrm{apr}}'X. \tag{3.10}$$
设 $x\in\underline{\mathrm{apr}}''X$，由于 R 是逆串行的，所以 $R_s(U)=U$ 且 $\exists y\in U$ 使 $x\in R_s(y)$，由 $x\in\underline{\mathrm{apr}}''X$ 可知 $R_s(y)\subseteq X$，从而由定义知 $x\in\underline{\mathrm{apr}}'X$，因此(3.10)成立.

(2)对于 $\forall x\in\underline{\mathrm{apr}}''X$，由 R 的自反性可得 $x\in R_s(x)$，从而由$\underline{\mathrm{apr}}''X$ 的定义可知 $R_s(x)\subseteq X$，于是 $x\in\underline{\mathrm{apr}}X$，因此
$$\underline{\mathrm{apr}}''X\subseteq\underline{\mathrm{apr}}X. \tag{3.11}$$
其次，对于 $\forall x\in\underline{\mathrm{apr}}X$，由定义有 $R_s(x)\subseteq X$，从而由 $x\in R_s(x)$ 和$\underline{\mathrm{apr}}'X$ 的定义得 $x\in\underline{\mathrm{apr}}'X$，因此

$$\underline{\mathrm{apr}}X \subseteq \underline{\mathrm{apr}}'X. \tag{3.12}$$

又由 R 的自反性可得 R 是逆串行的，这样由(1)，(3.11)和(3.12)定理得证.

□

定理 3.13 设 R 是 U 上的二元关系，则由(Ⅱ′)和(Ⅱ″)定义的两对近似算子是等价的当且仅当集合类 $\{R_s(x)\neq\varnothing \mid x\in U\}$ 构成了 U 的一个划分.

证 "⇒"先证 $R_s(U)=U$. 反证，若不然，则存在 $x\in U$ 且 $x\notin R_s(U)$，因此 $x\in\underline{\mathrm{apr}}''U$，但显然 $x\notin\underline{\mathrm{apr}}'U$，这与 $\underline{\mathrm{apr}}''X=\underline{\mathrm{apr}}'X$，对任意 $X\subseteq U$ 成立相矛盾，从而

$$R_s(U)=U. \tag{3.13}$$

其次，再证

$$R_s(x)\cap R_s(y)\neq\varnothing \Rightarrow R_s(x)=R_s(y). \tag{3.14}$$

事实上，若 $\exists x,y\in U$，使 $R_s(x)\cap R_s(y)\neq\varnothing$，但 $R_s(x)\neq R_s(y)$，不妨设 $\exists x_0\in R_s(x)\setminus R_s(y)$，取 $z_0\in R_s(x)\cap R_s(y)$，则由定义可知 $z_0\in\overline{\mathrm{apr}}''\{x_0\}$. 但由 $z_0\in R_s(y)$ 和 $x_0\notin R_s(y)$ 可得 $z_0\notin\overline{\mathrm{apr}}'\{x_0\}$，这与 $\overline{\mathrm{apr}}''\{x_0\}=\overline{\mathrm{apr}}'\{x_0\}$ 矛盾！因此(3.14)成立.

由(3.13)和(3.14)可知 $\{R_s(x)\neq\varnothing \mid x\in U\}$ 构成了 U 的一个划分.

"⇐"由于 $\{R_s(x)\neq\varnothing \mid x\in U\}$ 构成了 U 的一个划分，因此 R 是逆串行的且 $R_s(U)=U$，由定理 3.12 可知对于任意 $X\subseteq U$ 有 $X\subseteq\overline{\mathrm{apr}}'X\subseteq\overline{\mathrm{apr}}''X$. 往证 $\overline{\mathrm{apr}}'X=\overline{\mathrm{apr}}''X$. 反证，若不然，存在 $X\subseteq U$ 使 $\overline{\mathrm{apr}}'X\subset\overline{\mathrm{apr}}''X$，于是存在 $x\in\overline{\mathrm{apr}}''X\setminus\overline{\mathrm{apr}}'X$. 由 $x\in\overline{\mathrm{apr}}''X$ 可知 $\exists y\in U$ 使

$$x\in R_s(y) \text{ 且 } R_s(y)\cap X\neq\varnothing, \tag{3.15}$$

又由 $x\notin\overline{\mathrm{apr}}'X$ 可知 $\exists z\in U$ 使

$$x\in R_s(z) \text{ 但 } R_s(z)\cap X=\varnothing. \tag{3.16}$$

因为 $x\in R_s(y)\cap R_s(z)\neq\varnothing$，由条件得 $R_s(y)=R_s(z)$，这样(3.15)和(3.16)是矛盾的.

□

定理 3.14 设 R 是 U 上的二元关系，则由(Ⅰ)和(Ⅱ′)定义的两对近似算子是等价的当且仅当 R 是自反和传递的.

证 "⇒"设对于任意 $X\subseteq U$，有 $\overline{\mathrm{apr}}X=\overline{\mathrm{apr}}'X$ 且 $\underline{\mathrm{apr}}X=\underline{\mathrm{apr}}'X$.

首先，对于 $\forall x\in U$，由定理 3.10 知 $x\in\overline{\mathrm{apr}}'\{x\}$. 由于 $\overline{\mathrm{apr}}'\{x\}=\overline{\mathrm{apr}}\{x\}=R_p(x)$，从而 $x\in R_p(x)$，即 xRx，所以 R 是自反的.

其次，由定理 3.10 的证明知

$$\underline{\mathrm{apr}}'X \subseteq \underline{\mathrm{apr}}'\,\underline{\mathrm{apr}}'X,$$

于是由条件有

$$\underline{\mathrm{apr}}X \subseteq \underline{\mathrm{apr}}\,\underline{\mathrm{apr}}X.$$

根据定理 3.6 知 R 是传递的.

"$\Leftarrow$"设 R 自反且传递,则由自反性可知 $x\in R_s(x)$,从而易见对于任意 $X\subseteq U$,有

$$\underline{\mathrm{apr}}X\subseteq \underline{\mathrm{apr}}'X. \tag{3.17}$$

另外对于 $\forall x\in \underline{\mathrm{apr}}'X$,由定义存在 $y\in U$ 使 $x\in R_s(y)$且 $R_s(y)\subseteq X$,由 R 的传递性得 $R_s(x)\subseteq R_s(y)\subseteq X$,从而由定义得 $x\in \underline{\mathrm{apr}}X$,这样就有

$$\underline{\mathrm{apr}}'X\subseteq \underline{\mathrm{apr}}X, \tag{3.18}$$

由(3.17) 和(3.18) 得

$$\underline{\mathrm{apr}}'X=\underline{\mathrm{apr}}X,\ \forall X\subseteq U.$$

由对偶性得 $\overline{\mathrm{apr}}X=\overline{\mathrm{apr}}'X$. □

定理 3.15 设 R 是 U 上的二元关系,则由(Ⅰ)和(Ⅱ″)定义的两对近似算子是等价的当且仅当 R 是对称且传递的.

证 "$\Rightarrow$"设对于任意 $X\subseteq U$,有 $\overline{\mathrm{apr}}X=\overline{\mathrm{apr}}''X$ 且 $\underline{\mathrm{apr}}X=\underline{\mathrm{apr}}''X$. 首先,由定理 3.11 的(7)知对于任意 $X\subseteq U$,有

$$\overline{\mathrm{apr}}''\,\underline{\mathrm{apr}}''X\subseteq X\subseteq \underline{\mathrm{apr}}''\,\overline{\mathrm{apr}}''X.$$

从而由定理 3.7 知 R 是对称的. 其次,假设 $x,y,z\in U$ 满足 $y\in R_s(x)$和 $z\in R_s(y)$,由定义显然有 $x\in \underline{\mathrm{apr}}R_s(x)$,从而由前提条件有 $x\in \underline{\mathrm{apr}}''R_s(x)$. 又由 R 的对称性知 $x\in R_s(y)$,这说明 $x\in R_s(U)$,这样对于 $\forall w\in U$ 若满足 $x\in R_s(w)$,则由 $x\in \underline{\mathrm{apr}}''R_s(x)$的定义知 $R_s(w)\subseteq R_s(x)$,于是由 $x\in R_s(y)$得 $R_s(y)\subseteq R_s(x)$,从而结合 $z\in R_s(y)$得 $z\in R_s(x)$,所以是传递的.

"$\Leftarrow$"假设 R 是对称和传递的,对于任意 $X\subseteq U$,设 $x\in \underline{\mathrm{apr}}''X$,若 $x\notin R_s(U)$,则由 R 的对称性知 $R_s(x)=\varnothing$,于是 $R_s(x)\subseteq X$,由定义知 $x\in \underline{\mathrm{apr}}X$. 若 $x\in R_s(U)$,由定义对于任意 $y\in U$,$x\in R_s(y)$蕴涵 $R_s(y)\subseteq X$,由 R 的传递性得 $R_s(x)\subseteq R_s(y)\subseteq X$,从而 $x\in \underline{\mathrm{apr}}X$. 这样就有

$$\underline{\mathrm{apr}}''X\subseteq \underline{\mathrm{apr}}X. \tag{3.19}$$

反之,对于 $x\in \underline{\mathrm{apr}}X$,由定义知 $R_s(x)\subseteq X$,若 $x\notin R_s(U)$,则直接由定义可以看出 $x\in \underline{\mathrm{apr}}''X$. 若 $x\in R_s(U)$,对于任意 $y\in U$,若满足 $x\in R_s(y)$,由 R 的对称性知 $y\in R_s(x)$,从而由 R 的传递性得 $R_s(y)\subseteq R_s(x)$,结合 $R_s(x)\subseteq X$ 得 $R_s(y)\subseteq X$,于是就有 $x\in \underline{\mathrm{apr}}''X$. 这样就证明了

$$\underline{\mathrm{apr}}X\subseteq \underline{\mathrm{apr}}''X. \tag{3.20}$$

由(3.19)和(3.20)得 $\underline{\mathrm{apr}}X=\underline{\mathrm{apr}}''X$,由对偶性可得 $\overline{\mathrm{apr}}X=\overline{\mathrm{apr}}''X$. □

定理 3.16 设 R 是 U 上的二元关系,则由(Ⅰ),(Ⅱ′)和(Ⅱ″)定义的三对近似算子是等价的当且仅当 R 是等价的.

证 由定理 3.13～3.15 即得. □

§3.4 近似算子的表示

这一节我们从一般的二元关系出发，引进关系划分函数，继而在论域里导出一个等价关系，使得在这个新的等价关系下定义的下近似和上近似对应原来一般关系下定义的下近似和上近似.

设(U,R)是广义近似空间，$|U|=n$，$\underline{\text{apr}}$和$\overline{\text{apr}}$是由定义3.4给出的近似算子，即

$$\underline{\text{apr}}X=\{x\in U\mid R_s(x)\subseteq X\},$$
$$\overline{\text{apr}}X=\{x\in U\mid R_s(x)\cap X\neq\varnothing\}.$$

定理3.2启示我们集合X的上近似$\overline{\text{apr}}X$可以通过单点集来计算，即

$$\overline{\text{apr}}X=\bigcup_{x\in X}\overline{\text{apr}}\{x\},$$

约定$\overline{\text{apr}}\varnothing=\varnothing$，我们引进函数$h:U\to 2^U$

$$h(x)=\overline{\text{apr}}\{x\},$$

集值映射h称为上近似分布算子. 这样，$\overline{\text{apr}}X=\bigcup_{x\in X}h(x)$，上近似分布算子作用类似于概率论中的概率分布函数，其形式又类似于概率推理中的可能性分布，显然上近似分布算子是一个可加算子(关于集合并). 由引理3.1可知

$$h(x)=R_p(x)=\{y\in U\mid x\in R_s(y)\}. \tag{3.21}$$

即$h(x)$是恰好是x关于关系R的逆元全体(前继). 由(3.21)可以看出$y\in h(x)$当且仅当$x\in R_s(y)$，因而二元关系反过来也可以通过上近似分布算子来表示

$$R_s(x)=\{y\in U\mid x\in h(y)\}. \tag{3.22}$$

公式(3.22)启示我们可以从一个集值映射出发去定义二元关系，从而导出基于集值映射的粗糙集理论，详细请见第十章.

对于任意给定的二元关系R，我们定义关系划分函数$j:2^U\to 2^U$

$$j(X)=\{x\in U\mid R_s(x)=X\},X\in 2^U.$$

显然，$x\in j(X)\Leftrightarrow R_s(x)=X$. $j(X)\neq\varnothing$当且仅当X属于R_s的值域. 由于论域U中的每一个元素的后继是惟一的，因此显然有以下

定理3.17 算子j满足以下性质：

(1) $\bigcup_{X\subseteq U}j(X)=U$，

(2) $X\neq Y\Rightarrow j(X)\cap j(Y)=\varnothing$.

注 集合类$\{j(X)\mid X\subseteq U\}$中有很多是空集，事实上$\{j(X)\mid X\subseteq U\}$中最多只有$n$个非空集，而集类$\{j(X)\mid X\subseteq U\}$中集合个数为$2^n$，所以$\{j(X)\mid X\subseteq U\}$只是构成了$U$的覆盖但不构成$U$的划分.

定理 3.18 近似算子可以由关系划分函数表示：

(1) $\underline{\mathrm{apr}}X=\bigcup_{Y\subseteq X} j(Y)$.

(2) $\overline{\mathrm{apr}}X=\bigcup_{Y\cap X\neq\varnothing} j(Y)$.

证 (1)设 $x\in\underline{\mathrm{apr}}X$，由定义 $R_s(x)\subseteq X$，从而存在惟一的 $A\subseteq X$ 使 $R_s(x)=A$，于是 $x\in j(A)$，即 $x\in\bigcup_{Y\subseteq X} j(Y)$，因此

$$\underline{\mathrm{apr}}X\subseteq\bigcup_{Y\subseteq X} j(Y). \tag{3.23}$$

反过来，对于任意 $y\in\bigcup_{Y\subseteq X} j(Y)$，存在 $B\subseteq X$ 使 $y\in j(B)$，于是 $R_s(y)=B$，由 $B\subseteq X$ 得 $R_s(y)\subseteq X$，从而由下近似定义得 $y\in\underline{\mathrm{apr}}X$，因此

$$\bigcup_{Y\subseteq X} j(Y)\subseteq\underline{\mathrm{apr}}X. \tag{3.24}$$

由(3.23)和(3.24)知(1)成立.

(2)由近似算子的对偶性，(1)以及定理 3.17 的(1)得

$$\begin{aligned}\overline{\mathrm{apr}}X&=\sim\underline{\mathrm{apr}}(\sim X)=\sim\bigcup_{Y\subseteq\sim X} j(Y)\\&=\sim\bigcup_{Y\cap X=\varnothing} j(Y)=\bigcup_{Y\cap X\neq\varnothing} j(Y).\end{aligned}$$ □

注 由上述定理可以看出 j 的作用类似于证据理论中生成信任函数和似然函数的概率分配(mass)函数(见第十章)，因此有时 j 也称为基本集分配(basic set assignment)函数.

定理 3.19 关系划分函数 j 可以用近似算子来表示：即对于任意 $X\subseteq U$，

$$j(X)=\underline{\mathrm{apr}}X\setminus\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y. \tag{3.25}$$

证 对于任意 $X\subseteq U$，若 $x\in j(X)$，则由定理 3.18 知 $x\in\underline{\mathrm{apr}}X$. 由于 $R_s(x)=X$，由定理 3.17 可知对于 X 的任意真子集 Y，有 $x\notin\underline{\mathrm{apr}}Y$，这是因为若 $x\in\underline{\mathrm{apr}}Y$，由定理 3.18 知存在 Y 的子集 Z 使 $x\in j(Z)$，即 $R_s(x)=Z$，这与 $R_s(x)$的值是惟一的矛盾. 这样我们就得到了 $x\in\underline{\mathrm{apr}}X\setminus\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y$，因此

$$j(X)\subseteq\underline{\mathrm{apr}}X\setminus\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y. \tag{3.26}$$

反之，对于任意 $y\in\underline{\mathrm{apr}}X\setminus\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y$，由 $y\in\underline{\mathrm{apr}}X$ 和定理 3.18 知存在惟一 $A\subseteq X$使$y\in j(A)$，如果 $A\subset X$，那么由定理 3.18 知 $y\in\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y$，因此只能有 $A=X$，即 $y\in j(X)$，从而

$$\underline{\mathrm{apr}}X\setminus\bigcup_{Y\subset X}\underline{\mathrm{apr}}Y\subseteq j(X). \tag{3.27}$$

由(3.26)和(3.27)得(3.25). □

对于 U 上任意一个二元关系 R，我们定义 U 上一个新的二元关系 D

$$xDy\Leftrightarrow R_s(x)=R_s(y).$$

显然,D 是等价关系. D 称为由 R 导出的关系,记 $[x]_D=\{y\in U\mid xDy\}=D_s(x)$,则

$$x\in j(X)\Leftrightarrow R_s(x)=X=R_s([x]_D)\Leftrightarrow[x]_D=j(X).$$

当 $j(X)\neq\varnothing$ 时,称 X 为焦点集(focal set),令 $\mathscr{F}=\{X\subseteq U\mid j(X)\neq\varnothing\}$,记 $\mathscr{M}=\{j(X)\mid X\in\mathscr{F}\}$. 由定理 3.17 知集类 $\mathscr{M}$ 构成了 U 的一个划分,显然它就是 U 关于关系 D 的商集,即 $U/D=\mathscr{M}$. $\mathscr{M}$ 中每一个集称为基本集,空集以及有限个基本集的并集称为可观察集或可定义集,其全体就是由 $\mathscr{M}$ 生成的 σ 代数 $\sigma(\mathscr{M})$. 利用 $\sigma(\mathscr{M})$,近似算子 $\underline{\mathrm{apr}}_R$ 和 $\overline{\mathrm{apr}}_R$ 有类似于 Pawlak 近似算子的表达式

$$\underline{\mathrm{apr}}_R X=\bigcup\{Y\subseteq U\mid Y\in\sigma(\mathscr{M}),R_s(Y)\subseteq X\},$$

$$\overline{\mathrm{apr}}_R X=\sim\underline{\mathrm{apr}}_R(\sim X)=\bigcap\{Y\subseteq U\mid Y\in\sigma(\mathscr{M}),\\ \sim R_s(\sim Y)\supseteq X\}.$$

同样,利用等价关系 D,近似算子又可表示为

$$\underline{\mathrm{apr}}_R X=\bigcup\{[x]_D\mid R_s([x]_D)\subseteq X\},$$

$$\overline{\mathrm{apr}}_R X=\bigcup\{[x]_D\mid R_s([x]_D)\cap X\neq\varnothing\}.$$

§3.5 程度粗糙集模型

在§3.2 中给出的一般关系下粗糙集模型是完全通过元素 x 的后继 $R_s(x)$ 与论域 U 中集合 X 的简单的定性关系来定义近似算子的,即当元素 x 的所有后继都属于 X 时,元素 x 就属于 X 的下近似 $\underline{\mathrm{apr}}X$,当元素 x 至少有一个后继属于 X 时,元素 x 就属于 X 的上近似 $\overline{\mathrm{apr}}X$,这个定义有一个缺陷,它没有考虑到 $R_s(x)$ 与 X 重叠部分的定量信息,因此如果人们要依据 $R_s(x)$ 与 X 重叠的多少来刻画或近似概念 X,这个模型就无能为力了,为了解决这个问题,我们可以引进程度粗糙集(graded rough set)概念.

定义 3.5 设 $A=(U,R)$ 为广义近似空间,$X\subseteq U$,k 为非负整数,定义 X 关于近似空间 A 依程度 k 的下近似和上近似分别为

$$\underline{\mathrm{apr}}_k X=\{x\in U\mid |R_s(x)|-|R_s(x)\cap X|\leqslant k\},$$

$$\overline{\mathrm{apr}}_k X=\{x\in U\mid |R_s(x)\cap X|>k\}.$$

即当元素 x 关于 R 的后继个数最多只有 k 个不属于 X 时它就属于 $\underline{\mathrm{apr}}_k X$,而当元素 x 关于 R 的后继属于 X 的个数多于 k 时它就属于 X 的上近似 $\overline{\mathrm{apr}}_k X$.

当 $\underline{\mathrm{apr}}_k X=\overline{\mathrm{apr}}_k X$ 时称 X 关于近似空间 A 依程度 k 是可定义的,否则称 X 关于近似空间 A 依程度 k 是粗糙的.

显然当 $k=0$ 时,程度粗糙集模型退化为一般关系下的粗糙集模型.

定理 3.20 设 $A=(U,R)$为广义近似空间，k 为非负整数，则程度近似算子满足下列性质：

(1) $\underline{\mathrm{apr}}_k X = \sim(\overline{\mathrm{apr}}_k(\sim X))$，$\overline{\mathrm{apr}}_k X = \sim(\underline{\mathrm{apr}}_k(\sim X))$.

(2) $\underline{\mathrm{apr}}_k U = U$，$\overline{\mathrm{apr}}_k \varnothing = \varnothing$.

(3) $\underline{\mathrm{apr}}_k(X \cap Y) \subseteq \underline{\mathrm{apr}}_k X \cap \underline{\mathrm{apr}}_k Y$，$\overline{\mathrm{apr}}_k(X \cup Y) \supseteq \overline{\mathrm{apr}} X \cup \overline{\mathrm{apr}}\, Y$.

(4) $X \subseteq Y \Rightarrow \underline{\mathrm{apr}}_k X \subseteq \underline{\mathrm{apr}}_k Y$，$\overline{\mathrm{apr}}_k X \subseteq \overline{\mathrm{apr}}_k Y$.

(5) $\underline{\mathrm{apr}}_k(X \cup Y) \supseteq \underline{\mathrm{apr}}_k X \cup \underline{\mathrm{apr}}_k Y$，$\overline{\mathrm{apr}}_k(X \cap Y) \subseteq \overline{\mathrm{apr}}_k X \cap \overline{\mathrm{apr}}_k Y$.

(6) $\underline{\mathrm{apr}} X = \underline{\mathrm{apr}}_0 X$，$\overline{\mathrm{apr}} X = \overline{\mathrm{apr}}_0 X$.

(7) $k \geqslant l \Rightarrow \underline{\mathrm{apr}}_k X \supseteq \underline{\mathrm{apr}}_l X$，$\overline{\mathrm{apr}}_k X \subseteq \overline{\mathrm{apr}}_l X$.

证 由定义直接得. □

第四章　粗糙集代数的公理化方法

在 Pawlak 粗糙集理论中，下近似和上近似并非最基本的概念，它们被论域 U 上的二元关系，论域的划分或覆盖，对象的邻域（见第三章）及偏序、格、布尔代数等等更为基本的概念所刻画，我们称沿着这条技术路线研究粗糙集的方法为构造性方法．另一方面，通过把下近似和上近似视为最基本的概念，将注意力放在研究粗糙集理论中所产生的代数系统上，利用一个公理集来刻画上、下近似算子，揭示公理集与此条件下所产生的代数系统之间的联系的粗糙集理论研究方法被称为公理化方法．

§4.1　粗糙集理论的构造性方法

设 U 是一个有限、非空论域，R 是 U 上的一个二元关系，令 $R_s(x)=\{y\mid xRy\}$，$R_p(x)=\{y\mid yRx\}$．对任意集 $X\subseteq U$，我们可以按下面两种方式定义 X 的下、上近似．

(C_1) $\underline{R}X=\{x\mid R_s(x)\subseteq X\}$，

$\overline{R}X=\{x\mid R_s(x)\cap X\neq\varnothing\}$．

(C_2) $\underline{R}X=\bigcup\{R_s(x)\mid R_s(x)\subseteq X\}$，

$\overline{R}X=\bigcup\{R_s(x)\mid R_s(x)\cap X\neq\varnothing\}$．

由第三章知道，当 R 是 U 上的等价关系时，(C_1)和(C_2)等价．但对一般的二元关系，情况并非如此．

显然，下、上近似可以被视为定义在 U 的幂集 2^U 上的一对一元算子．

定义 4.1　算子 $L,H:2^U\to2^U$ 称为是对偶的，如果对 $\forall A\subseteq U$，它们满足

(L_0) $LA=\sim H\sim A$．

(H_0) $HA=\sim L\sim A$．

由 §3.2 可知，若采用(C_1)定义下、上近似，$\underline{R}$ 和 $\overline{R}$ 恰为一对对偶算子．在下面的讨论中，我们选择(C_1)作为下、上近似的定义．

由(C_1)可知，一个元属于 X 的下近似，如果它的所有后继都属于 X；一个元属于 X 的上近似，如果至少有一个它的后继属于 X．所以(C_1)又可被等价地表示为如下形式：

(C_1) $\underline{R}X=\{x\mid\forall y[xRy\Rightarrow y\in X]\}$，

$\overline{R}X=\{x\mid\exists y[xRy,y\in X]\}$.

注 在这种解释下,下、上近似的概念被紧密地联系于模态逻辑中的必然性和可能性算子.也就是说,如果一个元等价于(即此时 R 是等价关系)X 的下近似中的一个元,则该元必定属于 X,同时一个元如果等价于 X 的上近似中的一个元,则该元可能属于 X.关于这一点读者可以在第6章中有更深的体会.

定义 4.2 设 R 是 U 上的二元关系,下面的(4.1)式定义了 $2^U\to 2^U$ 的一对对偶近似算子 $\underline{R}$ 和 $\overline{R}$.

$$\begin{aligned}\underline{R}X&=\{x\mid\forall y[xRy\Rightarrow y\in X]\}=\{x\mid R_s(x)\subseteq X\},\\ \overline{R}X&=\{x\mid\exists y[xRy,y\in X]\}=\{x\mid R_s(x)\cap X\neq\varnothing\}.\end{aligned}\tag{4.1}$$

称系统$(2^U,\cap,\cup,\sim,\underline{R},\overline{R})$为一个粗糙集代数,这里$\cap$,$\cup$和$\sim$分别指集合的交、并和补.

考虑单元集$\{y\}$的上近似,由引理3.1,我们有

$$\overline{R}\{y\}=\{x\mid R_s(x)\cap\{y\}\neq\varnothing\}=\{x\mid y\in R_s(x)\}=R_p(y),\tag{4.2}$$

因此

$$y\in R_s(x)\Leftrightarrow x\in\overline{R}\{y\}.\tag{4.3}$$

由定义4.2及(4.3)式,我们可以建立单元集和任意集之上、下近似之间的联系.

(C_3) $\underline{R}X=\{x\mid\forall y[x\in\overline{R}\{y\}\Rightarrow y\in X]\}$,

$\overline{R}X=\{x\mid\exists y[x\in\overline{R}\{y\},y\in X]\}$.

由定理3.2,$\underline{R}$ 和 $\overline{R}$ 有如下性质:

(L_1) $\underline{R}U=U$.

(H_1) $\overline{R}\varnothing=\varnothing$.

(L_2) $\underline{R}(X\cap Y)=\underline{R}X\cap\underline{R}Y$.

(H_2) $\overline{R}(X\cup Y)=\overline{R}X\cup\overline{R}Y$.

注意到这四个性质中标号相同者是对偶的.性质(L_2)说明 $\underline{R}$ 对$\cap$运算是分配的;性质(H_2)说明 $\overline{R}$ 对$\cup$运算是分配的.它们蕴涵如下的性质:

(L_3) $\underline{R}(X\cup Y)\supseteq\underline{R}X\cup\underline{R}Y$.

(H_3) $\overline{R}(X\cap Y)\subseteq\overline{R}X\cap\overline{R}Y$.

像集合的补一样,下、上近似算子 $\underline{R}$ 和 $\overline{R}$ 是两个一元算子,所以一个粗糙集代数$(2^U,\cap,\cup,\sim,\underline{R},\overline{R})$是通过幂集布尔代数$(2^U,\cap,\cup,\sim)$中追加两个一元算子 $\underline{R}$ 和 $\overline{R}$ 的一个扩张.为方便起见,我们假定算子 $\underline{R}$ 和 $\overline{R}$ 与$\sim$有相同的优先级,并从右到左结合.如式 $\underline{R}\overline{R}\sim X\cap Y$ 等价于$\underline{R}(\overline{R}(\sim X))\cap Y$.

§4.2 粗糙集理论的公理化方法

在粗糙集的公理化方法中，基本概念是系统$(2^U,\cap,\cup,\sim,L,H)$，这里$(2^U,\cap,\cup,\sim)$是集代数，$L,H:2^U\to 2^U$是两个一元算子，它们被一些公理所定义而非联系于二元关系．我们称L,H是近似算子，在本章中，我们假定L，H是一对对偶的一元算子，即L,H满足定义4.1．

由L,H的对偶性，我们可以用其中的一个来定义另一个．比如$\sim L\sim$就是H，$\sim H\sim$就是L．但为方便计，通常仍用两个．

重复使用一元算子，可以被认为是算子的复合．如LH是L和H的复合．由L,H的对偶性，

$$
\begin{aligned}
L\sim X &= \sim HX, \\
\sim LX &= H\sim X.
\end{aligned}
\tag{4.4}
$$

当然，我们可以考虑任意有限个一元算子的复合，但利用L和H的对偶性，均可得到该算子的对偶形式，如$LHLHHL=\sim HLHLLH\sim$．

定理4.1 设$L,H:2^U\to 2^U$是一对对偶的一元算子．则存在一个U上的二元关系R使得$LX=\underline{R}X$，$HX=\overline{R}X$对任意集$X\subseteq U$都成立的充分必要条件是L和H满足性质：

(L_1) $LU=U$．

(L_2) $L(X\cap Y)=LX\cap LY$．

(H_1) $H\varnothing=\varnothing$．

(H_2) $H(X\cup Y)=HX\cup HY$．

证 定理的必要性在上一节中已经有了，所以我们只需证明定理的充分性．设H满足(H_1)和(H_2)，利用H我们可以构造U上的一个二元关系R如下：

$$xRy\Leftrightarrow x\in H(\{y\}).$$

所以，$R_s(x)=\{y\mid x\in H(\{y\})\}$．等价地，这个式子可以写成$H(\{y\})=\{x\mid y\in R_s(x)\}$．由$\overline{R}$的定义(即定义4.2可知，$\overline{R}(\varnothing)=\varnothing=H(\varnothing)$，对单元集有

$$
\begin{aligned}
\overline{R}(\{y\}) &= \{x\mid \exists t[t\in R_s(x),t\in\{y\}]\} \\
&= \{x\mid y\in R_s(x)\} = H(\{y\}).
\end{aligned}
$$

由公理(H_2)及论域U的有限性，对$\forall X\subseteq U$，我们有

$$\overline{R}X=\bigcup_{y\in X}\overline{R}(\{y\})=\bigcup_{y\in X}H(\{y\})=HX.$$

故 $\overline{R}=H$. 由对偶性易证 $\underline{R}=L$.　□

由定理 4.1,我们可以获得定义 4.2 在代数框架下的一个翻版.

定义 4.3　设 $L,H:2^U\to 2^U$ 是一对对偶算子,如果 L 满足公理(L_1)和(L_2),或者等价地,H 满足(H_1),(H_2)则系统($2^U,\cap,\cup,\sim,L,H$)称为一个粗糙集代数;L,H 称为近似算子.

在给定(L_1)的条件下,(L_2)等价于下面的公理(K),

(K) $L(\sim X\cup Y)\subseteq \sim LX\cup LY$.

利用对偶性易知(K)等价于(K′)

(K′) $\sim HX\cap HY\subseteq H(\sim X\cap Y)$.

在论域为有限集情况下,(H_2)等价于

(S) $HX=\bigcup\limits_{x\in X}H\{x\}$.

显然,若近似算子 H 满足(S),则 H 被其在论域上所有单元集上的值惟一决定.因此,我们可以建立 U 的任一子集的近似与单元子集的上近似之间的联系.

显然(S)可以等价地表达为 $HX=\{x\mid \exists y[x\in H\{y\},y\in X]\}$再由对偶性 $LX=\sim H\sim X$ 可得$LX=\{x\mid\forall y[x\in H\{y\}\Rightarrow y\in X]\}$由于 LU 的条件命题$\forall y[x\in H\{y\}\Rightarrow y\in U]$是一个必然命题,故 $LU=U$.由此又可得到 $H\varnothing=\varnothing$.

(A_5) $LX=\{x\mid\forall y[x\in H\{y\}\Rightarrow y\in X]\}$,

$HX=\{x\mid\exists y[x\in H\{y\},y\in X]\}$.

我们所讨论的满足(L_1),(L_2)和(H_1),(H_2)的一对对偶算子 L,H 中,L 是可乘的(即对交运算可分配);H 是可加的(即对并运算是可分配的),但 L 不是可加的,H 不是可乘的.关于这一点,我们可以很容易地从最简单的 Pawlak 粗糙集模型中找到反例,只需将 Pawlak 粗糙集模型($2^U,\cap,\cup,\sim,\underline{R},\overline{R}$)(这里 R 为等价关系)中的$\underline{R}$ 视为L,$\overline{R}$ 视为H.当然 L,H 满足对偶条件及(L_1),(L_2)和(H_1),(H_2).但我们知道,在这个粗糙集模型中,L 不是可加的,H 也不是可乘的.

尽管 L 不可加,H 不可乘,但它们却满足较弱的条件:

(L_3) $L(X\cup Y)\supseteq LX\cup LY$.

(H_3) $H(X\cap Y)\subseteq HX\cap HY$.

显然,(L_3)和(H_3)可以等价地表述为 L 和 H 对集合包含关系的单调性,即

(L_4) $X\subseteq Y\Rightarrow LX\subseteq LY$.

(H_4) $X\subseteq Y\Rightarrow HX\subseteq XY$.

容易看出,(L_2)蕴涵(L_3)和(L_4),(H_2)蕴涵(H_3)和(H_4).但反之不对,实事上,它们等价于

(L_5) $L(X\cap Y)\subseteq LX\cap LX$.

(H_5) $H(X\cup Y)\supseteq HX\cup HY$.

上面所列的各公理(性质)之间的关系,可以归纳为下面的定理.

定理 4.2 对近似算子 L,H,下面的蕴涵和等价关系成立:

(a) (L_1),(L_2)$\Leftrightarrow$(L_1),(K), (H_1),(H_2)$\Leftrightarrow$(H_1),(K′).

(b) (L_2)$\Rightarrow$(L_3),(L_4),(L_5), (H_2)$\Rightarrow$(H_3),(H_4),(H_5).

(c) (H_2)$\Leftrightarrow$(S).

(d) (L_3)$\Leftrightarrow$(L_4)$\Leftrightarrow$(L_5), (H_3)$\Leftrightarrow$(H_4)$\Leftrightarrow$(H_5).

在一个粗糙集代数$(2^U,\cap,\cup,\sim,L,H)$中,两个一元算子复合可以产生六个一元算子,即 $LL,HH,HL,LH,\sim L=H\sim$ 及 $\sim H=L\sim$.一般来讲,它们不同于 $L,H,\sim$ 中的任何一个.这就意味着重复使用一元算子一般是不可简化的.但上面的六个复合一元算于恰好构成三对对偶算子,即 LL 和 HH,LH 和 HL,以及 $\sim L$ 和 $\sim H$.

由 L、H 满足(L_1),(L_2),(H_1)及(H_2),我们有

$$LLU=U,$$
$$HH\varnothing=\varnothing,$$
$$LL(X\cap Y)=LLX\cap LLY,$$
$$HH(X\cup Y)=HHX\cup HHY.$$

所以,系统$(2^U,\cap,\cup,\sim,LL,HH)$也是一个粗糙集代数.易知,对任意自然数 n,$(2^U,\cap,\cup,\sim,\underbrace{LL\cdots L}_{n个},\underbrace{HH\cdots H}_{n个})$也是粗糙集代数.

对对偶算子 HL 和 LH,由于

$$HL(X\cap Y)=H(LX\cap LY)\subseteq HLX\cap HLY,$$
$$LH(X\cup Y)=L(HX\cup HY)\supseteq LHX\cup LHY.$$

故 HL 和 LH 满足(H_3)和(L_3)而不满足(L_1),(L_2),(H_1)和(H_2),所以系统$(2^U,\cap,\cup,\sim,LH,HL)$是比粗糙集代数弱的一个系统.

由 de Morgan 律及公理(L_1),(L_2),(H_1)和(H_2),另一对对偶算子 $\sim L$ 和 $\sim H$ 满足

($\sim L_1$) $\sim LU=\varnothing$.

($\sim H_1$) $\sim H\varnothing=U$.

($\sim L_2$) $\sim L(X\cap Y)=\sim LX\cup\sim LY$.

$(\sim H_2)$ $\sim H(X\cup Y)=\sim HX\cap\sim HY$.

容易证明$\sim L$,$\sim H$满足$(\sim L_1)$,$(\sim L_2)$,$(\sim H_1)$和$(\sim H_2)$当且仅当L,H满足(L_1),(L_2),(H_1)和(H_2).所以系统$(2^U,\cap,\cup,\sim,\sim L,\sim H)$可以被认为是粗糙集代数的另一种表现形式.

§4.3 构造性方法与公理化方法的关系

任意给定论域U上的一个二元关系R,按定义4.2,R可导出一个粗糙集代数$(2^U,\cap,\cup,\sim,\underline{R},\overline{R})$.又由定理4.1可知构造性方法与公理化方法在某种意义上是统一的.当然,具有不同性质的二元关系应该诱导出具有不同性质的(或者说满足不同公理的)近似算子.本节的目的就是要寻找二元关系的特殊性质与近似算子的不同公理之间的对应关系.

通常用来刻画二元关系的性质有以下五种:

串行性(serial)$\Leftrightarrow\forall x\,\exists y[xRy]$

$\Leftrightarrow\forall x\,\exists y[y\in R_s(x)]$

$\Leftrightarrow\forall x[R_s(x)\neq\varnothing]$.

自反性(reflexive)$\Leftrightarrow\forall x[xRx]$

$\Leftrightarrow\forall x[x\in R_s(x)]$.

对称性(symmetric)$\Leftrightarrow\forall x,y[xRy\Rightarrow yRx]$

$\Leftrightarrow\forall x,y[y\in R_s(x)\Rightarrow x\in R_s(y)]$.

传递性(transitive)$\Leftrightarrow\forall x,y,z[(xRy,yRz)\Rightarrow xRz]$

$\forall x,y[y\in R_s(x)\Rightarrow R_s(y)\subseteq R_s(x)]$.

欧几里得性(Euclidean)$\Leftrightarrow\forall x,y,z[(xRy,xRz)\Rightarrow yRz]$

$\Leftrightarrow\forall x,y[y\in R_s(x)\Rightarrow R_s(x)\subseteq R_s(y)]$.

正像我们在§4.1的注中所提到的,一对粗糙集代数中的近似算子所起的作用类似于模态逻辑中必然性算子和可能性算子所起的作用.在模态逻辑中,可能世界集上的不同二元关系蕴涵必然性和可能性算子的不同性质,对应于在模态逻辑中常用的公理,我们用L来代替必然性算子[ ],用H代替可能性算子$\langle\rangle$,用$\sim$代替$\neg$,用$\cap$代替合取$\wedge$,用$\cup$代替析取$\vee$,用$\subseteq$代替蕴涵$\rightarrow$,相对应地,可获得粗糙集代数中的若干公理.

(D) $LX\subseteq HX$.

(T) $LX\subseteq X$.

(T′) $X\subseteq HX$.

(B) $X\subseteq LHX$.

(B′) $HLX\subseteq X$.

(4) $LX \subseteq LLX$.

(4′) $HHX \subseteq HX$.

(5) $HX \subseteq LHX$.

(5′) $HLX \subseteq LX$.

具有相同字母或数字编号的公理是对偶的，一个近似算子满足一个公理时，则该算子的对偶算子一定满足对偶的公理. 本节中，我们想要说明的是在公理化方法中这些公理恰好是构造性方法中二元关系特殊性质的翻版.

定理 4.3 设 $L, H: 2^U \to 2^U$ 是一对近似算子，即 L 满足(L_1)，(L_2)，H 满足(H_1)和(H_2). 且 $H = \sim L \sim$，则存在

(a) 一个 U 上的串行关系，

(b) 一个 U 上的自反关系，

(c) 一个 U 上的对称关系，

(d) 一个 U 上的传递关系，

(e) 一个 U 上的欧几里得关系，

R 使得 $LX = \underline{R}X$，$HX = \overline{R}X$ 对所有的 $X \subseteq U$ 成立的充分必要条件是 L 和 H 分别满足

(a) (D)，

(b) (T′)，

(c) (B′)，

(d) (4′)，

(e) (5′).

证 (必要性)(1)设 R 是 U 上的串行关系. 令 $L = \underline{R}$，$H = \overline{R}$，对 $\forall X \subseteq U$. 若 $x_0 \in LX = \underline{R}X = \{x \mid \forall y[y \in R_s(x) \Rightarrow y \in X]\}$，因为 R 是串行的，所以存在 $y_0 \in R_s(x_0)$. 又由集 $\underline{R}X$ 的定义条件，知 $y_0 \in X$，故 $x_0 \in \overline{R}X = \{x \mid \exists y [y \in R_s(x), y \in X]\}$，所以有 $LX \subseteq HX$.

(2)设 R 是 U 上的自反关系，令 $L = \underline{R}$，$H = \overline{R}$. 对 $\forall X \subseteq U$，若 $x_0 \in X$，由 R 的自反性，$x_0 \in R_s(x_0)$. 由 $\overline{R}X$ 的定义，知 $x_0 \in \overline{R}X$. 故(T′) $X \subseteq HX$ 成立.

(3)设 R 是 U 上的对称关系，令 $L = \underline{R}$，$H = \overline{R}$. 对 $\forall X \subseteq U$，若 $x_0 \in X$，则对 $\forall y \in U$，如果 $y \in R_s(x_0)$，由 R 的对称性可得 $x_0 \in R_s(y)$，故 y 满足 HX 定义即 $y \in HX$. 所以 $x_0 \in LHX$，即(B) $X \subseteq LHX$ 成立.

(4)设 R 是 U 上的传递关系，令 $L = \underline{R}$，$H = \overline{R}$. 对 $\forall X \subseteq U$，若 $x_0 \in LX$，来证明 $x_0 \in LLX$. 为此需证对 $\forall y \in U$，为 $y \in R_s(x_0) \Rightarrow y \in LX$.

事实上，对 $\forall z \in U$，若 $z \in R_s(y)$，因 R 是传递的，故 $z \in R_s(y) \subseteq$

$R_s(x_0)$,又因为 $x_0\in LX$,所以 $z\in X$,所以 $y\in LX$,这表明 $x_0\in LLX$,故(4) $LX\subseteq LLX$ 成立.

(5)设 R 是 U 上的欧几里得关系,令 $L=\underline{R}$,$H=\overline{R}$.对 $\forall X\subseteq U$,若 $x_0\in HX$,来证明 $x_0\in LHX$,即对 $\forall z\in U$,若 $z\in R_s(x_0)\Rightarrow z\in HX$.因为 $x_0\in HX$,故存在 $y_0\in R_s(x_0)$且 $y_0\in X$,又因 R 是欧几里得的,故 $z\in R_s(y_0)$,由 HX 的定义条件可知 $z\in HX$.故(5) $HX\subseteq LHX$ 成立.

(充分性)由定理 4.1,对偶公理及公理(L_1),(L_2),(H_1)及(H_2)保证了 U 上二元关系 R 的存在性,且使得 $\underline{R}=L$,$\overline{R}=H$.更进一步,如果

(1) L,H 满足(D) $LX\subseteq HX$,由 $LU=U\Rightarrow\bigcup_{y\in U}H\{y\}=HU=U$,故对 $\forall x\in U$,$\exists y\in U$ 使得 $x\in H\{y\}$.由 R 的构造(见定理 4.1 的证明),xRy.故 R 是串行的.

(2) L,H 满足(T′) $X\subseteq HX$,由 X 的任意性$\to$对 $\forall x\in U$,$\{x\}\subseteq H\{x\}$,即 $x\in H\{x\}\Rightarrow xRx$,故 R 是自反的.

(3) L,H 满足(B) $X\subseteq LHX$,对 $\forall x,y\in U$,设 xRy 即 $x\in H\{y\}$.由公理(B),$\{x\}\subseteq LH\{x\}\Leftrightarrow\{x\}\subseteq\sim H(\sim H\{x\})\Leftrightarrow x\in\sim H(\sim H\{x\})\Leftrightarrow x\notin H(\sim H\{x\})\Leftrightarrow R_s(x)\cap(\sim H\{x\})=\varnothing\Leftrightarrow R_s(x)\subseteq H\{x\}$ 这意味着 $y\in H\{x\}$,即 yRx,故 R 是对称的.

(4) L,H 满足(4′) $HHX\subseteq HX$.对 $\forall x,y\in U$,设 $y\in R_s(x)$,进一步假设 $z\in R_s(y)$,则 $y\in H\{z\}\Leftrightarrow\{y\}\subseteq H\{z\}$,而公理($H_2$)蕴涵算子 H 是单调的,故 $H\{y\}\subseteq HH\{z\}$.由公理(4′)知 $H\{y\}\subseteq H\{z\}$.又由 $y\in R_s(x)\Leftrightarrow x\in H\{y\}$ 知 $x\in H\{z\}$,故 $z\in R_s(x)$.所以 $R_s(y)\subseteq R_s(x)$,即 R 传递.

(5) L,H 满足 $HX\subseteq LHX$.对 $\forall x,y\in U$,设 $y\in R_s(x)$,进一步假设 $z\in R_s(x)$,则 $x\in H\{z\}\Leftrightarrow\{x\}\subseteq H\{z\}$.由公理(5) $HX\subseteq LHX$,得 $\{x\}\subseteq H\{z\}\subseteq LH\{z\}\Rightarrow x\in LH\{z\}=\sim H(\sim H\{z\})\Leftrightarrow x\notin H(\sim H\{z\})\Leftrightarrow R_s(x)\cap(\sim H\{z\})=\varnothing\Leftrightarrow R_s(x)\subseteq H\{z\}$.又因为 $y\in R_s(x)$ 故 $y\in H\{z\}\Leftrightarrow z\in R_s(y)$ 故 $R_s(x)\subseteq R_s(y)$,所以 R 是欧几里得的. □

§4.4 特殊类型的粗糙集代数

在本节中,我们选择近似算子公理来刻画近似算子.这些公理揭示了如下三个系统:

$$S_L=(2^U,\cap,\cup,\sim,L,H),$$
$$S_{LH}=(2^U,\cap,\cup,\sim,HL,LH),$$
$$S_{LL}=(2^U,\cap,\cup,\sim,LL,HH).$$

之间的联系.通过选择不同的公理集,我们可以获得不同的粗糙集代数,并且在给定的公理集下,讨论上述三个系统之间的关系.

4.4.1 串行粗糙集代数

公理(D)表明 LX 是 HX 的子集,这诱发我们区别近似算子 L,H 为下、上近似算子.

如果一个粗糙集代数是由在基本公理(L_1),(L_2)基础上仅附加公理(D)而得到,则我们记之为 KD,称为串行粗糙集代数.由定理 4.2 知道,粗糙集代数 KD 等价于 U 上的一个串行二元关系诱导的粗糙集代数,它是允许我们在集合包含意义下定义下,上近似的最弱的一类粗糙集代数系统.在 KD 粗糙集代数系统中,下面的性质是显然成立的.

(L_6) $L\varnothing=\varnothing$.

(H_6) $HU=U$.

(L_7) $LX\cap L\sim X=\varnothing$.

(H_7) $HX\cup H\sim X=U$.

在给定(L_2)的条件下,(L_6)和(L_7)与(D)是等价的.当然在给定(H_2)的条件下,(H_6)和(H_7)与(D)也是等价的.事实上,在(L_2)的条件下,(D)⇒(L_6),(L_7)是容易的,至于反过来的结论,令 R 是 U 上由 L,H 决定的二元关系,$X_0=\{x\in U\mid R_s(x)=\varnothing\}$.我们只需证明 $X_0=\varnothing$,若不然 $X_0\neq\varnothing$,由 $L=\underline{R}$ 的定义可知,对 $\forall\, Y\subseteq U$,$X_0\subseteq LY$,故有 $X_0\subseteq L(X_0)\cap L(\sim X_0)=\varnothing$,矛盾.这个矛盾说明 $X_0=\varnothing$,即 R 是串行的,即 L,H 满足(D).值得一提的是,由(L_1),(L_2)和(L_6)构成的公理集通常被用来公理化区间结构,所以 KD 粗糙集代数是一个区间结构.

在一个 KD 粗糙集代数中,近似算子 LL 和 HH 服从公理(D),所以系统 S_{LL} 仍然是一个 KD 粗糙集代数,而算子 HL 满足公理(L_1)和(L_6),LH 满足公理(H_1)和(H_6),但它们不满足(L_2)和(H_2),所以系统 S_{HL} 不是一个粗糙集代数.由公理(D)及 L,H 的单调性,我们所考虑的六个近似算子有如下性质:

$$\begin{aligned}&LX\subseteq HX,\\&LLX\subseteq LHX\subseteq HHX,\\&LLX\subseteq HLX\subseteq HHX.\end{aligned}\tag{4.5}$$

K 粗糙集代数没有这些性质.

4.4.2 自反粗糙集代数

通过在基本公理中追加公理(T),我们可以得到 KT 粗糙集代数.在 KT 粗糙集代数中,对 $\forall\, X\subseteq U$ 有 $LX\subseteq X\subseteq HX$,因此公理(D)成立,故一个 KT

粗糙集代数自然是一个 KD 粗糙集代数，这一点也可以直接由二元关系的自反性蕴涵串行性而得到，因此也叫自反粗糙集代数.

在一个 KT 粗糙集代数中，LL 满足(T)，HH 满足(T′). 所以系统 S_{LL} 也是一个 KT 粗糙集代数. 但系统 S_{HL} 仍然不是一个粗糙集代数，因为 HL 和 LH 不满足(L_2)和(H_2). 对我们所考虑的三个系统 S_L，S_{HL} 和 S_{LL}，可以建立较(4.5)式更强的关系式

$$\begin{gathered}LLX \subseteq LX \subseteq X \subseteq HX \subseteq HHX, \\ LLX \subseteq LX \subseteq LHX \subseteq HX \subseteq HHX, \\ LLX \subseteq LX \subseteq HLX \subseteq HX \subseteq HHX.\end{gathered} \tag{4.6}$$

由此可知，算子 L，H 会产生出较之 LL 和 HH 产生的更紧的近似.

公理(T)和(T′)还蕴涵着下面一些性质：

(L_8) $X \subseteq LY \Rightarrow LX \subseteq LY$.

(H_8) $HX \subseteq Y \Rightarrow HX \subseteq HY$.

(L_9) $LX \subseteq Y \Rightarrow LX \subseteq HY$.

(H_9) $X \subseteq HY \Rightarrow LX \subseteq HY$.

(L_{10}) $LX \subseteq LY \Rightarrow LX \subseteq Y$.

(H_{10}) $HX \subseteq HY \Rightarrow X \subseteq HY$.

(L_{11}) $HX \subseteq LY \Rightarrow HX \subseteq Y$.

(H_{11}) $HX \subseteq LY \Rightarrow X \subseteq LY$.

4.4.3 对称粗糙集代数

通过增加公理(B)，我们可以得到 KB 粗糙集代数，也称对称粗糙集代数. 由定理 4.2，满足基本公理及公理(B)的算子 L，H 所诱导的 U 上的二元关系 R 是对称的，即 $\forall x, y \in U, xRy \Leftrightarrow yRx$，所以公理(B)可以被等价地表示为

(b′) $\forall x, y[x \in H\{y\} \Rightarrow y \in H\{x\}]$.

所以我们有

$$\begin{aligned}x \in HX &\Leftrightarrow \exists y[x \in H\{y\}, y \in X] \\ &\Leftrightarrow \exists y[y \in H\{x\}, y \in X] \\ &\Leftrightarrow X \cap H\{x\} \neq \varnothing.\end{aligned}$$

又由 H，L 的对偶性，有 $x \in LX \Leftrightarrow H\{x\} \subseteq X$，因此，上、下近似算子可以被表示为

(A_3) $LX = \{x \mid H\{x\} \subseteq X\}$,

$HX = \{x \mid H\{x\} \cap X \neq \varnothing\}$.

在 KB 粗糙集代数中，下面两个性质成立：

(L_{12}) $X \subseteq LY \Rightarrow HX \subseteq Y$.

(H_{12}) $HX \subseteq Y \Rightarrow X \subseteq LY$.

它们等价于 $X \subseteq LY \Rightarrow HX \subseteq Y$.

定理 4.4 设 $L, H: 2^U \rightarrow 2^U$ 是一对对偶算子,则(A_3)成立的充分必要条件是公理(H_2)及(B')成立.

证 公理(H_2)及(B')蕴涵(H_1). 因为由(H_2)及(B')可知,对 $\forall X \subseteq U$, $H(LX) = H(LX \cup \varnothing) = HLX \cup H\varnothing \subseteq X$ 成立. 由 X 的任意性可知, $H\varnothing = \varnothing$,即$(H_1)$成立. 再根据本子节开始的论述即知$(A_3)$成立.

反之,若(A_3)成立,则由(A_3)的第二式立即可知(H_2)成立. 下面我们来证明(B')的对偶式(B)成立. 因为(H_2)成立,故知 H 是单调的,对 $\forall x \in X$,立即有 $H\{x\} \subseteq HX$,根据(A_3)的第一式可知 $x \in LHX$,故 $X \subseteq LHX$. 所以(B)成立. □

由定理 4.4 可知一个 KB 粗糙集代数也可以由对偶公理及(A_3)或由对偶公理,(H_2)及(B')来刻画.

在一个 KB 粗糙集代数中,由(B)有, $HX \subseteq LH(HX)$,再由 L 的单调性有 $LHX \subseteq LLHHX$. 再次利用(B),则有 $X \subseteq LLHHX$. 故系统 S_{LL} 仍然为一个粗糙集代数. 系统 S_{HL} 仍然不是粗糙集代数. 公理(B)及(B')描述了任意集 X 及它在算子 HL 及 LH 下的像之间的关系, $HLX \subseteq X \subseteq LHX$. 三个系统 S_L, S_{HL} 及 S_{LL} 之间没有什么关系可以建立.

4.4.4 传递粗糙集代数

传递粗糙集代数被基本公理附加公理(4)所刻画. 我们记之为 K4,在给定公理(H_2)的条件下,公理$(4')$可以被等价地表达为

(iv') $\forall x, y[y \in H\{x\} \Rightarrow H\{y\} \subseteq H\{x\}]$.

证 $(H_2), (4') \Rightarrow (iv')$

对 $\forall x, y \in U$,若 $y \in H\{x\}$,则由(H_2)导出的 H 的单调性有 $H\{y\} \subseteq HH\{x\}$,又由$(4')$ $HH\{x\} \subseteq H\{x\}$,故 $H\{y\} \subseteq H\{x\}$.

$(H_2)(iv') \Rightarrow (4')$

$\forall x \in HHX$,由定义 $\exists y$ 使得 $y \in H\{x\}$ 且 $y \in HX$. 由 $y \in HX$ 可知, $\exists z \in H\{y\}$ 且 $z \in X$. 而由(iv')可知 $H\{y\} \subseteq H\{x\}$,故 $z \in H\{x\}$. 所以对 x,有 z 使得 $z \in H\{x\}$ 且 $z \in X$,故 $x \in HX$,所以 $HHX \subseteq HX$. □

读者还可以证明公理$(4')$的另外两个等价表达形式:

(iv'') $\forall x[HH\{x\} \subseteq H\{x\}]$.

(iv''') $\forall x, y, z[(y \in H\{x\}, z \in H\{y\}) \Rightarrow z \in H\{x\}]$.

由算子的单调性及公理(4)及$(4')$,我们有如下性质:

(L_{13}) $LX \subseteq Y \Rightarrow LX \subseteq LY$.

(H_{13}) $X\subseteq HY \Rightarrow HX\subseteq HY$.

公理(4)及(4′)描述了算子 L 和 LL 及 H 和 HH 之间的联系,容易证明算子 L,H 满足公理(4)及(4′)时,算子 LL 和 HH 也满足(4)及(4′).故系统 S_{LL} 也是一个 K4 粗糙集代数.但系统 S_{HL} 依然不是粗糙集代数.

4.4.5 自反且对称的粗糙集代数

自反且对称的粗糙集代数被公理(L_2),(T)和(B)所刻画,记之为 KTB.由定理 4.4 可知,(L_1)蕴涵于公理(L_2)和(B).KTB 粗糙集代数对应于 U 上的相容关系所诱导的粗糙集代数.

由公理(T)及(B),我们有 $LX\subseteq LHX$.事实上,这个性质可以从公理(T′)及算子 L 的单调性而获得,换句话说,公理(T)及(B)的组合并不产生新的性质,即 KTB 的性质是 KT 的性质或者是 KB 的性质.由 4.4.2 节 KT 粗糙集代数的性质(4.6)式及公理(B)及(B′),我们有下列公式.

对 $\forall X\subseteq U$,在 KTB 中

$$LLX \subseteq LX \subsetneq HLX \subseteq X \subseteq LHX \subseteq HX \subseteq HHX. \tag{4.7}$$

(4.7)式表明,在 KTB 框架下,S_{LL},S_L 及 S_{HL} 将产生一个比一个更紧的上、下近似.

4.4.6 自反且传递的粗糙集代数

一个 KT4 粗糙集代数,即自反且传递的粗糙集代数,被公理(L_1),(L_2),(T)及(4)所定义.它对应于模态逻辑系统 S_4,从代数角度讲,KT4 粗糙集代数与拓扑布尔代数的概念相关联.

显然,给定公理(T)和(T′),则公理(4)及(4′)可以被下面的(L_{14})及(H_{14})等价地代替:

(L_{14}) $LX = LLX$.

(H_{14}) $HX = HHX$.

(L_{14})及(H_{14})是 KT_4 的一个特殊性质,它们在 KT 和 K4 中都不成立.这两个性质也可以被认为是约简律或算子的幂等律.它们使得用一个较短的算子序列代替一个较长的算子序列成为可能,换句话说,这时系统 S_L 和系统 S_{LL} 变成了同一个系统.

在 KT_4 粗糙集代数中,有

$$LLX = LX \subsetneq X \subseteq HX = HHX,$$
$$LLX = LX \subseteq LHX \subseteq HX = HHX,$$
$$LLX = LX \subsetneq HLX \subseteq HX = HHX.$$

在 KT_4 粗糙集代数中,L 满足公理(L_1),(L_2),(T)和(L_{14});H 满足公理

(H_1),(H_2),(T')和(H_{14}),恰好是拓扑空间的内部和闭包算子的 Kuratowski 公理,因此,一个 KT_4 粗糙集代数可以被认为是一个拓扑粗糙集代数.一个 U 的子集 X 被称为是开的,如果 $LX=X$.一个 U 的子集 X 是闭的,如果 $HX=X$.由(L_{14})及(H_{14})可知,对 $\forall X\subseteq U$,LX 总是开的,HX 总是闭的.而且,LX 是包含于 X 的最大开子集,HX 是包含 X 的最小的闭子集.

4.4.7 对称且传递的粗糙集代数

对称且传递的粗糙集代数可以由公理(L_2),(B)和(4)来刻画.由 4.4.3 节,(L_1)蕴涵于(L_2)和(B)中.对称且传递的粗糙集代数记为 KB4.由对称性及传递性可以推出关系的欧几里得性,故在对称且传递的粗糙集代数 KB4 中,公理(5)及(5′)是成立的.由单调性及公理(5)和(5′),我们有

(L_{15}) $X\subseteq LY\Rightarrow HX\subseteq LY$.

(H_{15}) $HX\subseteq Y\Rightarrow HX\subseteq LY$.

在给定公理(H_2)的条件下,(5′)可以等价地表示为

(v′) $\forall x,y,z[(x\in H\{y\},x\in H\{z\})\Rightarrow y\in H\{z\}]$.

在给定(b′)(见 4.4.3 节)的条件下,(v′)可以被表示为

(v″) $\forall x,y[y\in H\{x\}\Rightarrow H\{x\}\subseteq H\{y\}]$.

考虑 $\forall,x,y\in U$,如果 $x\in H\{y\}$,则由(b′)得 $y\in H\{x\}$.由(iv′),我们有 $H\{x\}=H\{y\}$,从而 $x\in H\{x\}$,$y\in H\{y\}$.所以 $\{H\{x\}\mid x\in U\}$形成一个两两不相交的子集族,但它们未必是 U 的一个划分.但若在该集族中追加一个元 $\{x\mid H\{x\}=\varnothing\}$,则可以构成 U 的一个划分.

由上述的讨论,我们有

$$\begin{aligned} x\in HX&\Leftrightarrow \exists y[x\in H\{y\},y\in X]\\ &\Leftrightarrow \exists y[x\in H\{y\},y\in H\{x\},y\in X]\\ &\Leftrightarrow \exists y[x\in H\{y\},y\in H\{y\}\cap X\neq\varnothing]\\ &\Leftrightarrow \exists y[x\in H\{y\},H\{y\}\cap X\neq\varnothing]. \end{aligned}\tag{4.8}$$

由(4.8)式可知,$HX\subseteq\bigcup\{H\{x\}\mid H\{x\}\cap X\neq\varnothing]$.反过来,若 $H\{y\}\cap X\neq\varnothing$ 则 $y\in X$,故 $H\{y\}\subseteq HX$,所以,$HX=\bigcup\{H\{x\}\mid H\{x\}\cap X\neq\varnothing\}$.由 L,H 的对偶性可知

$$LX-(\bigcup\{H\{x\}\mid H\{x\}\subseteq X\})\cup\{x\mid H\{x\}=\varnothing\},$$

故在 KB4 中,我们有

(A_4) $LX=(\bigcup\{H\{x\}\mid H\{x\}\subseteq X\})\cup\{x\mid H\{x\}=\varnothing\}$;

$HX=\bigcup\{H\{x\}\mid H\{x\}\cap X\neq\varnothing\}$.

若一对对偶算子 L,H 满足(A_4),则由(A_4)的第二式即知道 H 满足(H_2),而(H_1)是(H_2)的特殊情况.

注意到对 $\forall X \subseteq U$，HX 是由单元集的 H 像的并构成的，故若 $H\{x\} \subseteq HHX$，则 $H\{x\} \cap HX \neq \varnothing$，由算子 H 的定义可知 $H\{x\} \subseteq HX$，所以 $HHX \subseteq HX$，所以(4′)成立.

对 $\forall X \subseteq U$，令 $X_0 = \{x \in U \mid H\{x\} = \varnothing\}$. 故由($A_4$)的第一式可知：$LX = (\bigcup\{H\{x\} \mid H\{x\} \subseteq X\}) \cup X_0 \triangleq X_1 \cup X_0$，故 $HLX = HX_1 \cup HX_0$，而 $HX_0 = \bigcup_{x \in X_0} H\{x\} = \varnothing$，再由 $HX_1 \subseteq X_1 \subseteq X$，故 $HLX \subseteq X$，所以(B′)成立.

由上面的讨论，我们有如下定理.

定理 4.5　设 $L, H: 2^U \to 2^U$ 是一对对偶算子，则(A_4)成立当且仅当 H 满足(H_2)，(B′)及(4′).

定理 4.5 表明，KB4 是使(A_4)成立的最弱的粗糙集代数.

4.4.8　Pawlak 粗糙集代数

由前面的讨论，我们看到公理(T)，(B)，(4)和(5)并不形成独立的公理集，尽管它们两两之间是独立的. 其中的一个等价独立公理集可以由公理(T)，(B)和(4)所构成. 另一个等价独立公理集可以由公理(T)和(5)所构成. 附加这一些等价独立公理集于公理(L_1)和(L_2)可以导出 Pawlak 粗糙集代数，它是模态逻辑 S_5 的一个翻版. 所以一个 Pawlak 粗糙集代数可以被记作 KTB_4 或 KT_5 或 S_5. 公理(5)和(T)蕴涵公理(B)，公理(L_2)及(B)蕴涵公理(L_1). 公理(T′)蕴涵类 $\{H\{x\} \mid x \in U\}$ 是 U 的一个覆盖. 公理(B)和(4′)蕴涵任意两个不同的单元集的 H 像是不相交的(即若 $H\{x\} \neq H\{y\}$ 则 $H\{x\} \cap H\{y\} = \varnothing$)，所以在 KTB4 中类 $\{H\{x\} \mid x \in U\}$ 构成了 U 的一个划分.

一个 Pawlak 粗糙集代数是一个 KB 粗糙集代数，因此，(A_3)在 KTB_4 中成立，同时它也是一个 KB4 粗糙集代数，所以(A_4)成立. 又因为在 Pawlak 粗糙集代数中，对 $\forall x \in U$，$x \in H\{x\}$，故 $H\{x\} \neq \varnothing$，所以在 Pawlak 粗糙集代数中，我们有比(A_4)更强的(A_5)：

(A_5)　$LX = \bigcup\{H\{x\} \mid H\{x\} \subseteq X\}$；

$HX = \bigcup\{H\{x\} \mid H\{x\} \cap X \neq \varnothing\}$.

若一对对偶算子 L, H 满足(A_5)，则由 H 满足(A_4)第二式可知 L 满足(A_4)第一式. 由 L 同时满足(A_4)，(A_5)第一式可知，集合 $\{x \in U \mid H\{x\} = \varnothing\} = \varnothing$，故 H 满足(T′)，所以我们有如下定理.

定理 4.6　设 $L, H: 2^U \to 2^U$ 是一对对偶算子，则(A_5)成立的充分必要条件是 H 满足(H_2)，(T′)，(B′)和(4′).

因为一个 Pawlak 粗糙集代数是一个拓扑粗糙集代数，故 $HHX = HX$，$LLX = LX$，所以在 Pawlak 粗糙集代数中 $HH = H$，$LL = L$. 易证，在给定公

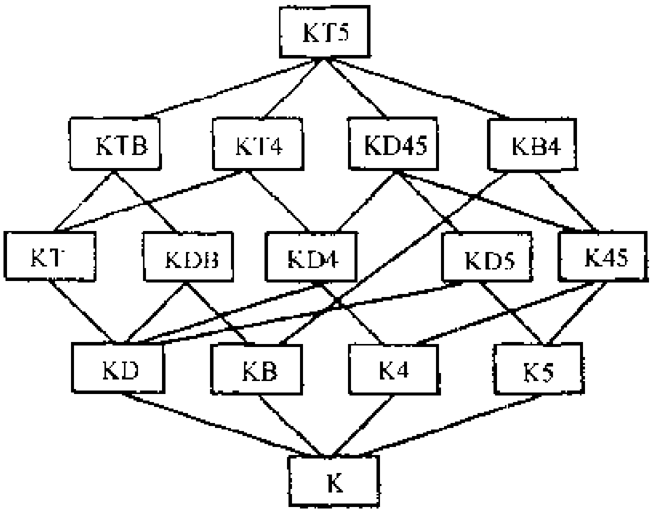

图 4.1

理(T)的情况下,(5)和(5′)等价于

(L_{16}) $HX = LHX$.

(H_{16}) $LX = HLX$.

因此,在 Pawlak 粗糙集代数中,任何子集的上、下近似都是开集也是闭集,且此时三个系统 S_L,S_{HL},S_{LL}变成了同一个系统.

至此,我们已经讨论了几种重要的特殊粗糙集代数类,定理 4.1 及定理 4.2 建立了粗糙集的构造性方法与公理化方法之间的联系.图 4.1 说明了我们所讨论的粗糙集代数类之间的强弱及蕴涵关系.图中的方框代表粗糙集代数类,方框中的字符代表该类的近似算子所满足的公理,而方框外面的注记对应于二元关系的性质.

注 需要指出的是,本章的讨论及结果均基于一个给定的构造算法:

$$\underline{R}X = \{x \mid \forall y[xRy \Rightarrow y \in X]\},$$
$$\overline{R}X = \{x \mid \exists y[xRy, y \in X]\}.$$

另外,我们在讨论特殊的粗糙集代数类时,对偶公理及(L_1),(L_2)及(H_1),(H_2)均作为基本公理被要求满足.显然,公理化方法,不应被限制于如此狭小的范围之内,所以本章的内容只能算作粗糙集公理化方法的一个侧面.

第五章　粗糙集系统的代数结构

粗糙集在数据库应用中的成功推动了粗糙集理论的研究. 粗糙集的代数分析是理论研究中最活跃的一个分支. 本章介绍粗糙集的 Stone 代数和 Nelson 代数及其性质,以及粗糙概念的代数刻画.

§5.1　粗糙集的 Stone 代数

在一个 Pawlak 粗糙集代数中,设 U 是论域,R 是 U 上的等价关系,对 $\forall X\subseteq U$,我们用 $\underline{R}(X)$和 $\overline{R}(X)$分别表示 X 在近似空间(U,R)中的下、上近似. $[x]_R$ 用来表示由 x 决定的 R 的等价类. 对 $\forall X,Y\subseteq U$,若 $\underline{R}(X)=\underline{R}(Y)$且 $\overline{R}(X)=\overline{R}(Y)$,则我们称 X,Y 粗相等,记作 $X\approx Y$. 显然$\approx$是 $P(U)$上的等价关系,$\approx$的等价类称为粗糙集. 所有粗糙集所构成的类记为 $\mathscr{R}$,即

$$\mathscr{R}=\{[X]_{\approx}\mid X\subseteq U\}. \tag{5.1}$$

设(U,R)是一个近似空间,$\mathscr{R}$是粗糙集类. 我们的目标是通过引进粗糙集的并,交及伪补运算,证明 $\mathscr{R}$是一个完备的 Stone 代数.

初看起来,利用集合的并、交运算来定义粗糙集的并、交运算是自然的,即令

$$[X]_{\approx}\cup[Y]_{\approx}=[X\cup Y]_{\approx},$$
$$[X]_{\approx}\cap[Y]_{\approx}=[X\cap Y]_{\approx}.$$

但下面的例子说明,这种貌似定义的形式其实并非定义,这是因为当我们选取等价类的不同代表元时,上述表达式可能产生不同的结果.

例 5.1　设有一个近似空间(U,R)满足,$\{x_1\},\{x_2\},\{x_3,x_4,x_5,x_6\},\{x_7\},\{x_8\}\in U/R$. $X=\{x_1,x_2,x_3\}$, $Y=\{x_4,x_7,x_8\}$, $X_1=\{x_1,x_2,x_3,x_4\}$, $Y_1=\{x_5,x_6,x_7,x_8\}$. 则有$[X_1]_{\approx}=[X]_{\approx}$,$[Y_1]_{\approx}=[Y]_{\approx}$. 但 $\underline{R}(X_1\cup Y_1)\neq\underline{R}(X\cup Y)$. 所以上面的表达式的计算依赖于等价类的代表元的选取,不能作为定义.

根据选择公理,对等价关系 R 的等价类所构成的集类,存在一个选择函数,我们记之为 f,以 R_gf 记 f 的值的集合. 我们定义 $P(U)$上的一个一元运算 S 如下:

$$S:P(U)\to P(U),$$

对$\forall X\subseteq U$,令 $S(X)=R_gf\cap X$.

显然 S 完全由选择函数 f 所决定.

为方便,我们以大写的花体字母来表示粗糙集,如 $\mathscr{X},\mathscr{Y},\mathscr{Z}$等.设 $\mathscr{X},\mathscr{Y},\mathscr{Z}\in\mathscr{R}$.如果 $\mathscr{X}=[X]_{\approx}$,则对$\forall Y\in\mathscr{X}$有 $\underline{R}(Y)=\underline{R}(X),\overline{R}(Y)=\overline{R}(X)$.所以我们可以采用下面的符号而不会引起误解,对$\forall X\in\mathscr{X},\underline{\mathscr{X}}=\underline{R}(X),\overline{\mathscr{X}}=\overline{R}(X)$.

下面我们定义粗糙并和粗糙交.

定义 5.1

$$\bigcup_{t\in T}\mathscr{X}_t=[\bigcup_{t\in T}\underline{\mathscr{X}_t}\cup S(\bigcup_{t\in T}\overline{\mathscr{X}_t})]_{\approx},\tag{5.2}$$

$$\bigcap_{t\in T}\mathscr{X}_t=[\bigcap_{t\in T}\underline{\mathscr{X}_t}\cup S(\bigcap_{t\in T}\overline{\mathscr{X}_t})]_{\approx}.\tag{5.3}$$

这里 T 是一个指标集,显然$\bigcup\mathscr{X}_t,\bigcap\mathscr{X}_t\in\mathscr{R}$.我们有如下的引理.

引理 5.1 (a)设 $A,B\subseteq U$ 且 $\underline{R}(A)=A$,则

$$\underline{R}(A\cup B)=\underline{R}(A)\cup\underline{R}(B)=A\cup\underline{R}(B),$$
$$\overline{R}(A\cup B)=\overline{R}(A)\cup\overline{R}(B)=A\cup\overline{R}(B).$$

(b)设 $z\in U,\mathscr{X}\in\mathscr{R}$ 且 $\mathrm{card}([z]_R)=1$,则

$$[z]_R\subseteq\overline{\mathscr{X}}\Leftrightarrow[z]_R\subseteq X.$$

引理 5.2 设 $X\subseteq U$ 是可定义的,则 $\overline{R}(S(X))=\overline{R}(X)$.

证 由 $\overline{R}$ 的单调性,立即有 $\overline{R}(S(X))\subseteq\overline{R}(X)$.又设$[z]_R\cap X\neq\varnothing$,由于 X 是可定义的,故$[z]_R\subseteq X$.这意味着$[z]_R\cap S(X)\neq\varnothing$,所以$[z]_R\subseteq\overline{R}(S(X))$.所以有 $\overline{R}(S(X))\supseteq\overline{R}(X)$. □

引理 5.3 设 S,S'是由选择函数 f 和 f'所决定的两个运算,$\{\mathscr{X}_t\mid t\in T\}$是任意一个粗糙集子类,则

$$[\bigcup\underline{\mathscr{X}_t}\cup S(\bigcup\mathscr{X}_t)]_{\approx}=[\bigcup\underline{\mathscr{X}_t}\cup S'(\bigcup\mathscr{X}_t)]_{\approx},$$
$$[\bigcap\underline{\mathscr{X}_t}\cup S(\bigcap\mathscr{X}_t)]_{\approx}=[\bigcap\underline{\mathscr{X}_t}\cup S'(\bigcap\mathscr{X}_t)]_{\approx}.$$

证 因为$\underline{\mathscr{X}_t}$是可定义的,所以$\bigcup_{t\in T}\underline{\mathscr{X}_t}$也是可定义的.由引理 5.1(a),有$\underline{R}(\bigcup\underline{\mathscr{X}_t}\cup S(\bigcup\overline{\mathscr{X}_t}))=\bigcup\underline{\mathscr{X}_t}\cup\underline{R}(S(\bigcup\overline{\mathscr{X}_t}))$.

由于对$\forall z\in U$,若$[z]_R\subseteq S(\bigcup\overline{\mathscr{X}_t})$,则 $\mathrm{card}[z]_R=1$,这蕴涵着$[z]_R\subseteq\bigcup\overline{\mathscr{X}_t}$,所以$[z]_R\subseteq S(\bigcup\underline{\mathscr{X}_t})$.故

(i) $\underline{R}(\bigcup\underline{\mathscr{X}_t}\cup S(\bigcup\overline{\mathscr{X}_t}))=\bigcup\underline{\mathscr{X}_t}$.

由引理 5.1(a),有 $\overline{R}(\bigcup\underline{\mathscr{X}_t}\cup S(\bigcup\overline{\mathscr{X}_t}))=\bigcup\underline{\mathscr{X}_t}\cup\overline{R}(S(\bigcup\overline{\mathscr{X}_t}))$.易知$\overline{R}(S(\bigcup\overline{\mathscr{X}_t}))=\bigcup\overline{\mathscr{X}_t}$,而$\bigcup\underline{\mathscr{X}_t}\subseteq\bigcup\overline{\mathscr{X}_t}$,故

(ii) $\overline{R}(\bigcup\underline{\mathscr{X}_t}\cup S(\bigcup\overline{\mathscr{X}_t}))=\bigcup\overline{\mathscr{X}_t}$.

同理可证(iii) $\underline{R}(\bigcup\underline{\mathscr{X}_t}\cup S'(\bigcup\overline{\mathscr{X}_t}))=\bigcup\underline{\mathscr{X}_t}$.

(iv)　$\overline{R}(\bigcup\underline{\mathscr{X}_t}\cup S'(\bigcup\overline{\mathscr{X}_t}))=\bigcup\overline{\mathscr{X}_t}$.

由粗糙集之定义,引理 5.3 中的第一式成立.对第二式可以类似地证明.

□

引理 5.3 说明:我们定义的粗糙集的交、并运算不依赖于选择函数 f.

推论 5.1　对任意的类$\{\mathscr{X}_t|t\in T\}$及粗糙集的交、并运算,有

$$\underline{\bigcup_{t\in T}\mathscr{X}_t}=\bigcup_{t\in T}\underline{\mathscr{X}_t},\qquad \overline{\bigcup_{t\in T}\mathscr{X}_t}=\bigcup_{t\in T}\overline{\mathscr{X}_t},$$

$$\underline{\bigcap_{t\in T}\mathscr{X}_t}=\bigcap_{t\in T}\underline{\mathscr{X}_t},\qquad \overline{\bigcap_{t\in T}\mathscr{X}_t}=\bigcap_{t\in T}\overline{\mathscr{X}_t}.$$

推论 5.2　对$\forall\mathscr{X}\in\mathscr{R}$,有$\mathscr{X}=[\underline{\mathscr{X}}\cup S(\overline{\mathscr{X}})]_{\approx}$.

定理 5.1　设(U,R)是一个近似空间,$\mathscr{R}$是由该近似空间决定的粗糙集类,则代数$(\mathscr{R},\cup,\cap)$是一个完备的分配格.

证　显然,粗糙集的交、并运算满足交换律、幂等律.下面我们以其中的一个为例证明它们满足结合律.比如:$\mathscr{X}\cap(\mathscr{Y}\cap\mathscr{Z})=(\mathscr{X}\cap\mathscr{Y})\cap\mathscr{Z}$

由推论 5.1 及推论 5.2,

$$\begin{aligned}\mathscr{X}\cap(\mathscr{Y}\cap\mathscr{Z})&=[\underline{\mathscr{X}\cap(\mathscr{Y}\cap\mathscr{Z})}\cup S(\overline{\mathscr{X}\cap(\mathscr{Y}\cap\mathscr{Z})})]_{\approx}\\&=[(\underline{\mathscr{X}}\cap(\underline{\mathscr{Y}}\cap\underline{\mathscr{Z}}))\cup S(\overline{\mathscr{X}}\cap(\overline{\mathscr{Y}}\cap\overline{\mathscr{Z}}))]_{\approx}\\&=[((\underline{\mathscr{X}}\cap\underline{\mathscr{Y}})\cap\underline{\mathscr{Z}})\cup S((\overline{\mathscr{X}}\cap\overline{\mathscr{Y}})\cap\overline{\mathscr{Z}})]_{\approx}\\&=[(\underline{\mathscr{X}\cap\mathscr{Y}}\cap\underline{\mathscr{Z}})\cup S(\overline{\mathscr{X}\cap\mathscr{Y}}\cap\overline{\mathscr{Z}})]_{\approx}\\&=(\mathscr{X}\cap\mathscr{Y})\cap\mathscr{Z}.\end{aligned}$$

下面我们证明吸收律中的一个,以$\mathscr{X}\cap(\mathscr{X}\cup\mathscr{Y})=\mathscr{X}$为例.

$$\begin{aligned}\mathscr{X}\cap(\mathscr{X}\cup\mathscr{Y})&=[(\underline{\mathscr{X}}\cap\underline{\mathscr{X}\cup\mathscr{Y}})\cup S(\overline{\mathscr{X}}\cap(\overline{\mathscr{X}\cup\mathscr{Y}}))]_{\approx}\\&=[(\underline{\mathscr{X}}\cap(\underline{\mathscr{X}}\cup\underline{\mathscr{Y}}))\cup S(\overline{\mathscr{X}}\cap(\overline{\mathscr{X}}\cup\overline{\mathscr{Y}}))]_{\approx}\\&=[\underline{\mathscr{X}}\cup S(\overline{\mathscr{X}})]_{\approx}=\mathscr{X}.\end{aligned}$$

现在我们证明$(\mathscr{R},\cup,\cap)$是一个分配格,即对$\forall\mathscr{X},\mathscr{Y},\mathscr{Z}\in\mathscr{R}$,有

$$\mathscr{X}\cap(\mathscr{Y}\cup\mathscr{Z})=(\mathscr{X}\cap\mathscr{Y})\cup(\mathscr{X}\cap\mathscr{Z}),$$

$$\mathscr{X}\cup(\mathscr{Y}\cap\mathscr{Z})=(\mathscr{X}\cup\mathscr{Y})\cap(\mathscr{X}\cup\mathscr{Z}).$$

事实上,我们有

$$\begin{aligned}\mathscr{X}\cap(\mathscr{Y}\cup\mathscr{Z})&=[(\underline{\mathscr{X}}\cap(\underline{\mathscr{Y}}\cup\underline{\mathscr{Z}}))\cup S(\overline{\mathscr{X}}\cap(\overline{\mathscr{Y}}\cup\overline{\mathscr{Z}}))]_{\approx}\\&=[((\underline{\mathscr{X}}\cap\underline{\mathscr{Y}})\cup(\underline{\mathscr{X}}\cap\underline{\mathscr{Z}}))\cup S((\overline{\mathscr{X}}\cap\overline{\mathscr{Y}})\cup(\overline{\mathscr{X}}\cap\overline{\mathscr{Z}}))]_{\approx}\\&=[(\underline{\mathscr{X}\cap\mathscr{Y}}\cup\underline{\mathscr{X}\cap\mathscr{Z}})\cup S(\overline{\mathscr{X}\cap\mathscr{Y}}\cup\overline{\mathscr{X}\cap\mathscr{Z}})]_{\approx}\\&=[(\underline{(\mathscr{X}\cap\mathscr{Y})\cup(\mathscr{X}\cap\mathscr{Z})})\cup S(\overline{(\mathscr{X}\cap\mathscr{Y})\cup(\mathscr{X}\cap\mathscr{Z})})]_{\approx}\\&=(\mathscr{X}\cap\mathscr{Y})\cup(\mathscr{X}\cap\mathscr{Z}),\end{aligned}$$

$$\begin{aligned}\mathscr{X}\cup(\mathscr{Y}\cap\mathscr{Z})&=[(\underline{\mathscr{X}}\cup(\underline{\mathscr{Y}}\cap\underline{\mathscr{Z}}))\cup S(\overline{\mathscr{X}}\cup(\overline{\mathscr{Y}}\cap\overline{\mathscr{Z}}))]_{\approx}\\&=[((\underline{\mathscr{X}}\cup\underline{\mathscr{Y}})\cap(\underline{\mathscr{X}}\cup\underline{\mathscr{Z}}))\cup S((\overline{\mathscr{X}}\cup\overline{\mathscr{Y}})\cap(\overline{\mathscr{X}}\cup\overline{\mathscr{Z}}))]_{\approx}\end{aligned}$$

$$
\begin{aligned}
&=[(\underline{\mathscr{X}\cup\mathscr{Y}}\cap\underline{\mathscr{X}\cup\mathscr{Z}})\cup S(\overline{\mathscr{X}\cup\mathscr{Y}}\cap\overline{\mathscr{X}\cup\mathscr{Z}})]_{\approx}\\
&=[\underline{(\mathscr{X}\cup\mathscr{Y})\cap(\mathscr{X}\cup\mathscr{Z})}\cup S(\overline{(\mathscr{X}\cup\mathscr{Y})\cap(\mathscr{X}\cup\mathscr{Z})})]_{\approx}\\
&=(\mathscr{X}\cup\mathscr{Y})\cap(\mathscr{X}\cup\mathscr{Z}).
\end{aligned}
$$
□

下面我们按照标准的方法在 $\mathscr{R}$ 上定义一个偏序，即 $\mathscr{X}\subseteq\mathscr{Y}$ 当且仅当 $\mathscr{X}\cap\mathscr{Y}=\mathscr{X}$. 显然地，$\mathscr{X}\subseteq\mathscr{Y}$ 当且仅当 $\underline{\mathscr{X}}\subseteq\underline{\mathscr{Y}}$ 且 $\overline{\mathscr{X}}\subseteq\overline{\mathscr{Y}}$.

在 $\mathscr{R}$ 中有两个元 $[\varnothing]_{\approx}$ 和 $[U]_{\approx}$. 我们下面来说明它们分别是格 $(\mathscr{R},\cup,\cap)$ 的最小元和最大元.

对 $\forall\mathscr{X}\in\mathscr{R}$,

$$
\begin{aligned}
\mathscr{X}\cap[U]_{\approx}&=[(\underline{\mathscr{X}}\cap\underline{R}(U))\cup S(\overline{\mathscr{X}}\cap\overline{R}(U))]_{\approx}\\
&=[(\underline{\mathscr{X}}\cap U)\cup S(\overline{\mathscr{X}}\cap U)]_{\approx}\\
&=[\underline{\mathscr{X}}\cup S(\overline{\mathscr{X}})]_{\approx}\\
&=\mathscr{X},
\end{aligned}
$$

故 $\mathscr{X}\subseteq[U]_{\approx}$. 又

$$
\begin{aligned}
\mathscr{X}\cap[\varnothing]_{\approx}&=[(\underline{\mathscr{X}}\cap\underline{R}(\varnothing))\cup S(\overline{\mathscr{X}}\cap\overline{R}(\varnothing))]_{\approx}\\
&=[(\underline{\mathscr{X}}\cap\varnothing)\cup S(\overline{\mathscr{X}}\cap\varnothing)]_{\approx}\\
&=[\varnothing\cup S(\varnothing)]_{\approx}\\
&=[\varnothing]_{\approx},
\end{aligned}
$$

故 $[\varnothing]_{\approx}$ 和 $[U]_{\approx}$ 分别是 $(\mathscr{R},\cup,\cap)$ 的最小、最大元，分别记为 0,1.

现在我们知道 $(\mathscr{R},\cup,\cap,0,1)$ 是一个有界分配格. 但它却不是一个补格. 事实上，在 $\mathscr{R}$ 中，只有满足 $\underline{\mathscr{X}}=\overline{\mathscr{X}}$ 的元 $\mathscr{X}$ 才有补元.

设 L 是一个有最小元 0 的格，若对 $\forall x\in L$ 存在元 $x^*\in L$ 使得 $x\wedge x^*=0$ 且 $x\wedge z=0$ 蕴涵 $z\leqslant x^*$，则称 x^* 是 x 的一个伪补. 若一个格 L 的每一个元都有伪补，则称 L 为一个伪补格.

尽管 $(\mathscr{R},\cup,\cap,0,1)$ 中不是每一个元都有补元，但我们可以为其中的每一个元定义一个伪补.

定义 5.2 对 $\forall\mathscr{X}\in\mathscr{R}$，令 $\mathscr{X}^*=[U-\overline{\mathscr{X}}]_{\approx}$.

引理 5.4 $(\mathscr{R}\cup,\cap,*)$ 是一个伪补格.

证 我们将证明，对 $\forall\mathscr{X}\in\mathscr{R}$，$\mathscr{X}^*$ 恰好是 $\mathscr{X}$ 的一个伪补.

因为 $\underline{\mathscr{X}^*}=\underline{R}(U-\overline{\mathscr{X}})=U-\overline{\mathscr{X}})$，有 $\overline{\mathscr{X}^*}=\overline{R}(U-\overline{\mathscr{X}})=U-\overline{\mathscr{X}}$，所以 $\mathscr{X}\cap\mathscr{X}^*=[(\underline{\mathscr{X}}\cap\underline{\mathscr{X}^*})\cup S(\overline{\mathscr{X}}\cap\overline{\mathscr{X}^*})]_{\approx}=[\varnothing\cap S(\varnothing)]_{\approx}=[\varnothing]_{\approx}=0$.

设对 $\forall\mathscr{Z}\in\mathscr{R}$，$\mathscr{X}\cap\mathscr{Z}=0$，即 $[(\underline{\mathscr{X}}\cap\underline{\mathscr{Z}})\cup S(\overline{\mathscr{X}}\cap\overline{\mathscr{Z}})]_{\approx}=0$. 所以 $\underline{\mathscr{X}}\cap\underline{\mathscr{Z}}=\varnothing$ 且 $\overline{\mathscr{X}}\cap\overline{\mathscr{Z}}=\varnothing$. 故有 $\overline{\mathscr{Z}}\subseteq U-\overline{\mathscr{X}}$，$\underline{\mathscr{Z}}\subseteq U-\underline{\mathscr{X}}$. 由于 $\underline{\mathscr{Z}}\subseteq\overline{\mathscr{Z}}$，所以 $\underline{\mathscr{Z}}\subseteq U-\overline{\mathscr{X}}$. 所以 $\mathscr{Z}\subseteq\mathscr{X}^*$. 故 $\mathscr{X}^*$ 是 X 的伪补. □

定义 5.3　设$(L,\vee,\wedge,*,0,1)$是一个有界分配的伪补格,若对$\forall a\in L$有$a^*\vee a^{**}=1$,则称L是一个 Stone 代数,$a^*\vee a^{**}=1$称为 Stone 恒等式.

引理 5.5　在代数$(\mathscr{R},\cup,\cap,*,0,1)$中 Stone 恒等式成立.

证　对$\forall \mathscr{X}\in\mathscr{R}$,$\mathscr{X}^*=[U-\overline{\mathscr{X}}]_\approx=[(U-\overline{\mathscr{X}})\cup S(U-\overline{\mathscr{X}})]_\approx$.由$\overline{\mathscr{X}^*}=U-\underline{\mathscr{X}}$,有$\mathscr{X}^{**}=[U-\overline{\mathscr{X}^*}]_\approx=[\mathscr{X}]_\approx=[\mathscr{X}\cup S(\mathscr{X})]_\approx$.所以$\mathscr{X}^*\cup\mathscr{X}^{**}=[\underline{R}(U-\overline{\mathscr{X}})\cup\overline{\mathscr{X}})\cup S(\overline{R}((U-\overline{\mathscr{X}})\cup\overline{\mathscr{X}})]_\approx=[\underline{U}\cup S(\overline{U})]_\approx=1$.　□

事实上,$(\mathscr{R},\cup,\cap,*,0,1)$还是原子的.$\mathscr{R}$的原子有如下的形式:

$$[S([x]_R)]_\approx.$$

现在我们可以陈述这一节的主要结论了.

定理 5.2　设$\mathscr{R}$是由近似空间(U,R)所决定的粗糙集类,则代数$(\mathscr{R},\cup,\cap,*,0,1)$是一个完备的,原子的 Stone 代数.

§5.2　粗糙近似空间

定义 5.4　设(U,R)是一个近似空间,定义$\mathscr{R}^*=\{(\underline{X},\overline{X})\mid X\subseteq U\}$为粗糙近似空间.

对$\forall X\subseteq U$,因为$X\in[\mathscr{X}]_\approx=\mathscr{X}$且我们有记号$\underline{\mathscr{X}}=\underline{X}$,$\overline{\mathscr{X}}=\overline{X}$.所以$\mathscr{R}^*$又可以表示为$\{(\underline{\mathscr{X}},\overline{\mathscr{X}})\mid \mathscr{X}\in\mathscr{R}\}$.

引理 5.6　设$\mathscr{R}$是由近似空间(U,R)所决定的粗糙集类,$\mathscr{R}^*$是粗糙近似空间,则存在一个$\mathscr{R}$到$\mathscr{R}^*$的一一对应g.

证　设$\mathscr{X}$是$\mathscr{R}$中的一个元,定义$g(\mathscr{X})=(\underline{\mathscr{X}},\overline{\mathscr{X}})$,则$g(\mathscr{X})\in\mathscr{R}^*$.若$g(\mathscr{X})=g(\mathscr{Y})$,则$(\underline{\mathscr{X}},\overline{\mathscr{X}})=(\underline{\mathscr{Y}},\overline{\mathscr{Y}})$当且仅当$\underline{\mathscr{X}}=\underline{\mathscr{Y}}$且$\overline{\mathscr{X}}=\overline{\mathscr{Y}}$.则有$\mathscr{X}=\mathscr{Y}$.对$\forall(\underline{\mathscr{X}},\overline{\mathscr{X}})\in\mathscr{R}^*$,则显然有$\mathscr{X}=[\mathscr{X}]_\approx\in\mathscr{R}$使得$g(\mathscr{X})=(\underline{\mathscr{X}},\overline{\mathscr{X}})$.故$g$是一个一一对应.　□

定义 5.5　设$\{R_t\mid t\in T\}$是$\mathscr{R}^*$的一个子集,即对$\forall t\in T$,$R_t=(\underline{\mathscr{X}_t},\overline{\mathscr{X}_t})\in\mathscr{R}^*$,定义$\cup R_t=(\cup\underline{\mathscr{X}_t},\cup\overline{\mathscr{X}_t})$,$\cap R_t=(\cap\underline{\mathscr{X}_t},\cap\overline{\mathscr{X}_t})$.

引理 5.7　在定义 5.5 中,$\cup R_t$,$\cap R_t\in\mathscr{R}^*$.

证　令$Z=(\cup\underline{\mathscr{X}_t})\cup S(\cup\overline{\mathscr{X}_t})$,则由上一节的结论$\underline{Z}=\cup\underline{X_t}$,$\overline{Z}=\cup\overline{X_t}$,故$\cup R_t=(\underline{Z},\overline{Z})\in\mathscr{R}^*$.

类似地,可以证明$\cap R_t=\in\mathscr{R}^*$.　□

引理 5.7 说明定义 5.7 的确定义了$\mathscr{R}^*$上的两个二元运算$\cup$,$\cap$.读者容易证明$(\mathscr{R}^*,\cup,\cap)$是一个格.由于$\forall\mathscr{X},\mathscr{Y}\in\mathscr{R}$,$g(\mathscr{X}\cup\mathscr{Y})=(\underline{\mathscr{X}\cup\mathscr{Y}},\overline{\mathscr{X}\cup\mathscr{Y}})=(\underline{\mathscr{X}}\cup\underline{\mathscr{Y}},\overline{\mathscr{X}}\cup\overline{\mathscr{Y}})=(\underline{\mathscr{X}},\overline{\mathscr{X}})\cup(\underline{\mathscr{Y}},\overline{\mathscr{Y}})=g(\mathscr{X})\cup g(\mathscr{Y})$.

同理有$g(\mathscr{X}\cap\mathscr{Y})=g(\mathscr{X})\cap g(\mathscr{Y})$.由此我们有下面的结论.

设 $\mathscr{R}^{\circ}$,$\mathscr{R}^{*}$ 是由近似空间(U,R)所决定的粗糙集类和粗糙集近似空间,则格$(\mathscr{R}^{\circ},\cup,\cap)$与$(\mathscr{R}^{*},\cup,\cap)$是同态的.

读者可以证明 $g([\varnothing]_{\approx})$和 $g([U]_{\approx})$是$(\mathscr{R}^{*},\cup,\cap)$的最小最大元.所以$(\mathscr{R}^{*},\cup,\cap,g(0),g(1))$是有界分配格.且可以证明 $g(\mathscr{X}^{*})$是 $g(\mathscr{X})$的一个伪补.在$(\mathscr{R}^{*},\cup,\cap,g(0),g(1))$中 Stone 恒等式成立.故我们有下面的定理 5.3.

定理 5.3 格$(\mathscr{R}^{*},\cup,\cap,*,0,1)$,这里 $0=(\varnothing,\varnothing)$,$1=(U,U)$,是一个完备的,原子的 Stone 代数,且同构于$(\mathscr{R}^{\circ},\cup,\cap,*,0,1)$.

§5.3 粗糙集和 Nelson 代数

在本节中,我们进一步将粗糙集抽象化,所采用的术语和符号大部分来自格论和简单的拓扑.在第一节中,我们称(U,R)是一个近似空间,这是一个泛化的、笼统的称谓,它本质表达的是能给 U 的任一子集提供在包含关系意义下的上、下近似的一个集类,我们看到所有的可定义集构成了布尔代数 $P(U)$ 的一个完备子代数.U 的任意子集恰好是被这个子代数中的两个元来刻画的,所以我们可以在格论下更广泛地讨论粗糙集系统的性质,对下面的定义读者不难找到它们的直观背景.在本节中,以下我们用大写的黑体英文字母表示结构而用大写的非黑体字母表示对应结构的载集.

定义 5.6 (1) 设 U 是一个有限集,$\boldsymbol{B}(U)=(P(U),\cup,\cap,-,\varnothing,U)$ 是通常的集代数,$\boldsymbol{B}(U)$的任意子代数 $\boldsymbol{A}(U)=(A,\cup,\cap,-,\varnothing,U)(A\subseteq P(U))$被称为近似代数.

(2) $\boldsymbol{A}$ 的原子被称为基本集,$\boldsymbol{A}$ 的原子集记为 Atom$(\boldsymbol{A})$.因为原子是不交的且所有原子的并是 U,所以 Atom$(\boldsymbol{A})$恰好是 U 上的一个等价关系所决定的等价类的类.

(3) 对 $\forall X\in P(U)$,$\overline{X}=\bigcup\{A'\in \mathrm{Atom}(A)\mid A'\cap X\neq\varnothing\}$,$\underline{X}=\bigcup\{A'\in \mathrm{Atom}(\boldsymbol{A})\mid A'\subseteq X\}$分别称为 X 关于近似空间 $\boldsymbol{A}$ 的上、下近似.

(4) 对 $\forall X,Y\in P(U)$,X,Y 称为是上粗相等的,当且仅当 $\overline{X}=\overline{Y}$,记为 $X\simeq Y$.

(5) 对 $\forall X,Y\in P(U)$,X,Y 称为是下粗相等的,当且仅当 $\underline{X}=\underline{Y}$,记为 $X\eqsim Y$.

(6) 对 $\forall X,Y\in P(U)$,X,Y 称为是粗相等的,当且仅当 $X\eqsim Y$,$X\simeq Y$,记为 $X\approx Y$.

(7) $X\in P(U)$称为是可定义的,如果 $\underline{X}=\overline{X}$.

(8) 我们可以将 $\boldsymbol{A}$ 视为一个拓扑空间$\langle U,\boldsymbol{A}\rangle$的开子集框架,用 $\mathscr{I}$和 $\mathscr{C}$记

A 所诱导的 Kuratowski 内部和闭包算子. 如果 $X\in P(U)$,我们用 $\mathscr{E}(X)$表示 $-\mathscr{C}(X)$. 因为 **A** 是一个布尔代数,故 **A** 的任一元既是开集也是闭集,即开闭集.

(9) 如果 $x\in A$, $\blacktriangle x=\{x'\in A\mid x'\geqslant x\}$是 **A** 上的一个滤子,称为由 x 生成的上滤子. 显然,我们有如下的引理.

引理 5.8 对 U 上的任一近似空间**A** 及 $\forall X\in P(U)$.

(1) $\mathscr{C}(X)=\bigcap\{A'\in \mathbf{A}\mid A'\supseteq X\}$.

(2) $\mathscr{I}(X)=\bigcup\{A'\in \mathbf{A}\mid A'\subseteq X\}$.

(3) $\mathscr{C}(X)=\overline{X}$.

(4) $\mathscr{I}(X)=\underline{X}$.

(5) 如果 X 是可定义的,则 $\mathscr{C}(X)\cup\mathscr{I}(X)=U$.

在§5.1 中,我们已经知道一个粗糙集是 U 的一个子集模关系≈的等价类,任意一个粗糙集可以被近似空间 **A** 中的一对元素所刻画,即所谓上近似和下近似. 显然,给定一个近似空间 **A**,下面的映射 $r^*:P(U)\to A\times A$ 的逆象决定$P(U)$的一个划分.

$$r^*:r^*(X)=(\mathscr{I}(X),\mathscr{C}(X)),\tag{5.4}$$

即对$\forall X\in P(U)$,$(r^*)^{-1}(\mathscr{I}(X),\mathscr{C}(X))$恰为$[X]_{\approx}$.

现在我们用 $\mathscr{E}(X)$代替$(\mathscr{I}(X),\mathscr{C}(X))$中的 $\mathscr{C}(X)$. 注意这种代替并没有本质的区别,因为集论中一个集合与其补集是相互惟一决定的. 这样我们可以得到映射 r^*的一个改造 $r:P(U)\to A\times A$,

$$r:r(X)=(\mathscr{I}(X),\mathscr{E}(X)).\tag{5.5}$$

对$\forall X\in P(U)$, $r(X)$称为粗糙集的一个不交表示. 因为对$\forall X,Y\in P(U)$, $X\approx Y\Leftrightarrow r^*(X)=r^*(Y)\Leftrightarrow r(X)=r(Y)$,所以粗糙集的不交表示是合理的. 我们以 $RS(\mathbf{A})$表示 $P(U)$在映射 r 下的像,即 $RS(\mathbf{A})=r(P(U))=\mathrm{Im}r$. 对粗糙集的不交表示来讲,任一个粗糙集是 **A** 中两个元所成的一个有序偶(X,Y)满足 $X\cap Y=\varnothing$,但并非所有满足 $X\cap Y=\varnothing$的对都是一个粗糙集(注意这里我们没有区分粗糙集和粗糙集的不交表示). 下面我们将讨论$(X,Y)\in A\times A$ 是一个粗糙集不交表示的条件.

$$S=\bigcup\{E\in \mathrm{Atom}(\mathbf{A})\mid \mathrm{card}(E)=1\}.\tag{5.6}$$

引理 5.9 对$\forall(X,Y)\in \mathbf{A}\times A$, (X,Y)是粗糙集的不交表示的条件是 $X\cap Y=\varnothing$且 $X\cup Y\supseteq S$.

证 若(X,Y)是粗糙集的一个不交表示,则 $X\cap Y=\varnothing$,且存在 $Z\in P(U)$,使得 $X=\mathscr{I}(Z)$, $Y=\mathscr{E}(Z)=-\mathscr{C}(Z)$. 对$\forall e\in S$, $[e]_R=E\in \mathrm{Atom}(\mathbf{A})$. 若 $E\not\subseteq X=\mathscr{I}(Z)$,因为 $\mathrm{card}(E)=1$,故 $E\not\subseteq\mathscr{C}(Z)$. 所以 $E\not\subseteq-\mathscr{C}(Z)$. 即 $E\subseteq\mathscr{E}(Z)=Y$. 所以 $S\subseteq X\cup Y$.

反过来,若(X,Y)满足 $X\cap Y=\varnothing$ 且 $X\cup Y\supseteq S$.则对 $\forall E\subseteq S(\mathrm{card}(E)=1)$,则要么 $E\subseteq X$,要么 $E\subseteq Y$.我们以 $E_X(E_Y)$表示 $\mathrm{Atom}(\boldsymbol{A})$中满足 $E\subseteq X(E\subseteq Y)$且 $\mathrm{card}(E)=1$ 的原子,则 $E_Y\not\subseteq -Y$,$E_X\subseteq X\subseteq -Y$.所以 $-Y-X$ 中没有基数为 1 的原子.设 f 是§5.1 中定义的一个选择函数,令 $X\cup(R_f\cap(-Y-X))=Z$,则 $\mathscr{I}(Z)=X$,$\mathscr{E}(Z)=Y$.所以(X,Y)是粗糙集的一个不交表示. □

下面我们将引理 5.9 中关于粗糙集的不交表示的第二个条件 $X\cup Y\supseteq S$ 给出一个代数解释.

由 $\boldsymbol{A}$ 的定义,$S\in A$(因为开集的并是开集),让我们考虑 $\boldsymbol{A}$ 中由 S 生成的主滤子 $\uparrow S$.根据格论的标准结果,下面的关系 Θ 是 $\boldsymbol{A}$ 的一个同余关系:$\forall X,Y\in A$,$X\Theta Y$ 当且仅当 $\exists Z\in S$ 使得 $X\cap Z=Y\cap Z$.

由 Θ 的定义,我们有,对 $\forall D\in A$,$D\Theta U$ 当且仅当 $D\supseteq S$.(事实上,若 $D\Theta U$,则 $\exists Z\in\uparrow S$ 满足 $D\cap Z=Z\cap U=Z\Rightarrow Z\subseteq D$,又由 $Z\in\uparrow S$,可得 $S\subseteq D$.反过来,若 $S\subseteq D$,则 $D\in\uparrow S$ 则 $D=D\cap D=D\cap U$,所以 $D\Theta U$).换句话说,对 $\forall D\in A$,D 与 U 属于同一个 Θ 同余类当且仅当 $S\subseteq D$.将这一结果与引理 5.9 结合,则有如下引理.

引理 5.10 对 $\forall(X,Y)\in A\times A$,$(X,Y)$是一个粗糙集的不交表示的充分必要条件是(1)$X\cap Y=\varnothing$;(2)$(X\cup Y)\Theta U$.

令

$$N_\Theta(\boldsymbol{A})=\{(X_1,X_2)\mid X_1,X_2\in A,X_1\cap X_2=\varnothing,(X_1\cup X_2)\Theta U\},$$

则有如下定理.

定理 5.4 对任一近似空间 $\boldsymbol{A}$,$N_\Theta(\boldsymbol{A})=\mathrm{RS}(\boldsymbol{A})$.

恰巧 $N_\Theta(\boldsymbol{A})$有一个很好的逻辑-代数结构.

定义 5.7 一个结构 $\boldsymbol{H}=\langle H,\vee,\wedge,\rightarrow,\neg,1,0\rangle$被称为是一个 Heyting 代数,如果

(1) $\langle H,\vee,\wedge,1,0\rangle$是一个分配格.

(2) 运算 $\vee$ 和$\rightarrow$满足附加性质 $\forall a,b\in H$,$a\wedge c\leqslant b$ 当且仅当 $c\leqslant a\rightarrow b$.

(3) $\neg x\overset{\wedge}{=}x\rightarrow 0$.

根据格中伪补的定义,在 Heyting 代数中$\neg$是伪补运算.运算$\rightarrow$称为相对伪补,$a\rightarrow b$ 即元 a 相对于 b 的伪补,$\neg a$ 即元 a 相对于 0 的伪补.

引理 5.11 任一个有限分配格是一个 Heyting 代数.

事实上,任一个有限分配格显然有界,且 $\vee\{x\mid x\wedge a\leqslant b\}$ 显然存在,故可令 $a\rightarrow b=\vee\{x\mid x\wedge a\leqslant b\}$.所以引理 5.11 成立.

定义 5.8 一个代数 $\boldsymbol{N}=(N,\wedge,\vee,\rightarrow,\neg,\sim,0,1)$称为一个 Nelson 代

数,如果对 $\forall a,b\in N$,

(1) $(N,\wedge,\vee,0,1)$是一个分界分配格.

(2) $\sim\sim a=a$.

(3) $\sim(a\vee b)=\sim a\wedge\sim b$.

(4) $a\wedge\sim a\leqslant b\vee\sim b$.

(5) $a\wedge x\leqslant(\sim a\vee b)$当且仅当 $x\leqslant a\to b$.

(6) $a\to(b\to c)=(a\wedge b)\to c$.

(7) $\neg a=a\to\sim a=a\to 0$.

定义 5.9 一个代数 $\boldsymbol{A}$ 是半单的,如果它是单代数的次直积.

对我们的主要结果,这里需要关于 Nelson 代数的一个熟知的结果.

引理 5.12 一个 Nelson 代数是半单的,如果对 $\forall a\in N, a\vee\neg a=1$.

引理 5.13 (1)设 $\boldsymbol{H}$ 是一个 Heyting 代数,Θ 是 $\boldsymbol{H}$ 上的一个布尔同余(即 $\boldsymbol{H}/\Theta$ 是一个布尔代数),则结构 $N_\Theta(\boldsymbol{H})=\langle N_\Theta(H),\vee,\wedge,\to,\neg,\sim,1,0\rangle$是一个 Nelson 代数.这里

(a) $1=\langle 1,0\rangle;0=\langle 0,1\rangle$.

(b) $\langle X_1,X_2\rangle\wedge\langle Y_1,Y_2\rangle=\langle X_1\wedge Y_1,X_2\vee Y_2\rangle$.

(c) $\langle X_1,X_2\rangle\vee\langle Y_1,Y_2\rangle=\langle X_1\vee Y_1,X_2\wedge Y_2\rangle$.

(d) $\langle X_1,X_2\rangle\to\langle Y_1,Y_2\rangle=\langle X_1\to Y_1,X_2\wedge Y_2\rangle$.

(e) $\sim\langle X_1,X_2\rangle=\langle X_2,X_1\rangle$.

(f) $\neg\langle X_1,X_2\rangle=\langle\neg X_1,X_1\rangle=\langle X_1,X_2\rangle\to\langle 0,1\rangle$.

(注意尽管 $N_\Theta(\boldsymbol{H})$和 H 中的运算选用了相同的运算符,但读者容易区分它们.)

(2)设 $\boldsymbol{H}$ 是一个布尔代数,Θ 是一个 $\boldsymbol{H}$ 上的同余(蕴涵 Θ 是布尔同余),则 $N_\Theta(\boldsymbol{H})$是一个半单的 Nelson 代数.

引理的证明是验证性的,略.

定理 5.5 对任一近似空间 $\boldsymbol{A}$,则粗糙集系统 RS($\boldsymbol{A}$)是一个半单的 Nelson 代数.

证 因为近似空间 $\boldsymbol{A}$ 是一个布尔代数,由定理 5.4,$N_\Theta(\boldsymbol{H})=\mathrm{RS}(\boldsymbol{A})$.再由引理 5.13,RS($\boldsymbol{A}$)是一个半单的 Nelson 代数. □

注 因为 $\boldsymbol{A}$ 是一个集布尔代数,故运算 $1,0,\wedge,\vee$ 和$\neg$实际上是集论算子 $U,\varnothing,\cap,\cup$和$-$.且 $X_1\to Y_1=(-X_1)\cup Y_1$.

下面我们通过一个例子来说明上面所讨论的结构.

例 5.2 设 $C=(U=\{a,b,c\},Q=\{q_1,q_2\},V=\{0,1,2\},f)$是一个信息系统,信息函数 f 由表 5.1 给定.

表 5.1

f	q_1	q_2
a	0	1
b	1	2
c	1	2

布尔代数 $\boldsymbol{B}(U)=(P(U)\cap,\cup,-,\varnothing,U)$ 及近似空间 $\boldsymbol{A}$(作为 $\boldsymbol{B}(U)$ 的子代数)由图 5.1 和图 5.2 表示.

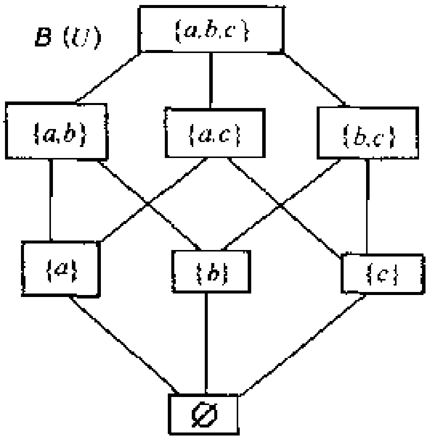

图 5.1

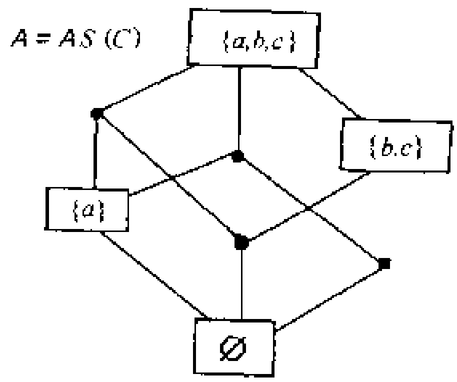

图 5.2

我们看到只有一个基数为 1 的基本集,即 $\{a\}$,所以 $S=\{a\}$, $\uparrow S=\{\{a\},\{a,b\},\{a,c\},\{a,b,c\}\}$. 所以同余关系 Θ 的同余类只有两个,即 $\uparrow S$ 和 $\{\varnothing,\{b\},\{c\},\{b,c\}\}$. 所以 $N_\Theta(\boldsymbol{A})=\{\langle\varnothing,\{a,b,c\}\rangle,\langle\{a,b,c\},\varnothing\rangle,\langle\varnothing,\{a\}\rangle,\langle\{a\},\varnothing\rangle,\langle\{a\},\{b,c\}\rangle,\langle\{b,c\},\{a\}\rangle\}$.

映射 r 如图 5.3 所示.

注 考察 Nelson 代数的公理,读者会发现,它们并没有直接保证伪补的存在性. 公理(1),(2)和(3)说明 Nelson 代数是一个 De Morgan 代数. 公理(4)称为正则性原理,它表明 Nelson 代数是一个 Kleene 代数. 公理(5)及公理(6)是 Nelson 代数的特有公理. 公理(5)要求 $a\rightarrow b$ 是满足 $a\wedge x\leqslant\sim a\vee b$ 的最大的元 x. 换句话说 $a\rightarrow b$ 是 a 相对于 $\sim a\vee b$ 的伪补. 公理(6)表明,这样的相对伪补满足所谓的 currying 性质(即将一个二元函数转化为两个一元函数),公理(7)是说 $\neg a$ 是 a 相对于其强非的伪补.

定义 5.10 在一个 Nelson 代数中,我们可以定义一个新的运算 $\Rightarrow$,对 $\forall a,b\in N$, $a\Rightarrow b\overset{\wedge}{=}(a\rightarrow b)\wedge(\sim b\rightarrow\sim a)$. $\Rightarrow$ 被称为强蕴涵.

在下面的内容中,我们用 $a,b\cdots,1,0$ 分别记 $N_\Theta(\boldsymbol{A})$ 中的元素即序偶

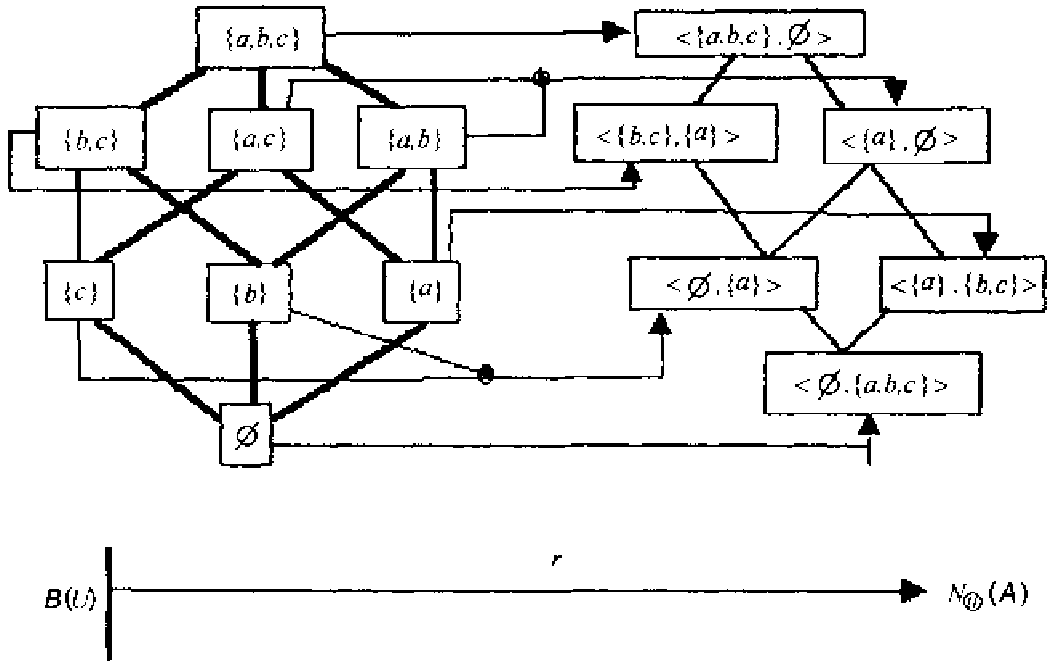

图 5.3

$\langle A_1,A_2\rangle,\langle B_1,B_2\rangle,\cdots,\langle U,\varnothing\rangle,\langle\varnothing,U\rangle$.

引理 5.14 对任一 Nelson 代数 $N_\Theta(\boldsymbol{H})$.

(1)由 $a\preccurlyeq b$ 当且仅当$a\to b=1$ 定义的关系$\preccurlyeq$是一个前序.

(2)由 $a\leqslant b$ 当且仅当$a\Rightarrow b=1$ 定义的关系$\leqslant$是一个偏序.

(3) $a\leqslant b$ 当且仅当 $a\wedge b=a$ 当且仅当 $a\vee b=b$ 当且仅当 $A_1\leqslant B_1$ 且 $B_2\leqslant A_2$.

证明是验证性的,略.

下面的例子说明$\to$,$\Rightarrow$都不满足定义 5.7(2).

例 5.3 近似空间 $\boldsymbol{A}$ 由例 5.2 给定.

(a) $\langle\varnothing,\{a\}\rangle\to\langle\{a\},\{b,c\}\rangle=\langle\{a,b,c\},\varnothing\rangle$,但$\langle\varnothing,\{a\}\rangle\wedge\langle\{a\},\{b,c\}\rangle=\langle\varnothing,\{a\}\rangle\not\leqslant\langle\{a\},\{b,c\}\rangle$. 所以$\langle\{a,b,c\},\varnothing\rangle$不是$\langle\varnothing,\{a\}\rangle$相对于$\langle\{a\},\{b,c\}\rangle$的伪补.

(b) $\langle\varnothing,\{a\}\rangle\Rightarrow\langle\{a\},\{b,c\}\rangle=(\langle\varnothing,\{a\}\rangle\to\langle\{a\},\{b,c\}\rangle)\wedge(\langle\{b,c\},\{a\}\rangle\to\langle\{a\},\varnothing\rangle)=\langle\{a,b,c\},\varnothing\rangle\wedge\langle\{a\},\varnothing\rangle=\langle\{a\},\varnothing\rangle$,但$\langle\{a\},\varnothing\rangle\wedge\langle\varnothing,\{a\}\rangle=\langle\varnothing,\{a\}\rangle\not\leqslant\langle\{a\},\{b,c\}\rangle$. 所以$\Rightarrow$也不是相对伪补运算.

下面让我们定义一个运算$\supset$满足定义 5.7(2).

定义 5.11 对 $\forall a,b\in N_\Theta(\boldsymbol{A})$,定义 $a\supset b$ 如下:对 $\forall c\in N_\Theta(A)$,$a\cap c\subseteq b$当且仅当$c\subseteq a\supset b$.

尽管我们形式地给出了相对伪补的定义,但这并不意味着定义是成功的.

下面的引理 5.15 说明定义是可行的.

引理 5.15(元 $a\supset b$ 的存在性) 对 $\forall a,b\in N_\Theta(\boldsymbol{A})$,元 $a\supset b$ 总是存在的.

证 因为$(N_\Theta(\boldsymbol{A}),\cap,\cup,0,1)$是有限分配格,若 c_1,c_2 满足 $a\cap c\subseteq b$,则 $c_1\cup c_2$ 也满足 $a\cap c\subseteq b$. 由 $N_\Theta(\boldsymbol{A})$的有限性,知$\{c\mid a\cap c\subseteq b\}$有最大元,故 $a\supset b$ 是惟一存在的. □

定理 5.6(元 $a\supset b$ 的刻画) 对 $\forall a,b\in N_\Theta(\boldsymbol{A})$, $a\supset b=(\neg a\wedge\neg\sim b)\vee\sim\neg\sim a\vee b$.

证 由引理 5.15,设 $a\supset b=\langle X,Y\rangle$,则$\langle X,Y\rangle$应满足:对 $\forall c=\langle C_1,C_2\rangle$,

(i) $a\cap c\subseteq b$.

(ii) $c\cap\langle X,Y\rangle=\langle C_1,C_2\rangle$.

(iii) $X\cap Y=\varnothing$.

$$\begin{aligned}
&\langle A_1,A_2\rangle\cap\langle C_1,C_2\rangle\subseteq\langle B_1,B_2\rangle\\
\Leftrightarrow\ &\langle A_1\cap C_1,A_2\cup C_2\rangle\subseteq\langle B_1,B_2\rangle\\
\Leftrightarrow\ &\langle A_1\cap C_1,A_2\cup C_2\rangle\cap\langle B_1,B_2\rangle=\langle A_1\cap C_1,A_2\cup C_2\rangle\\
\Leftrightarrow\ &\begin{cases}(A_1\cap C_1)\cap B_1=A_1\cap C_1,\\(A_2\cup C_2)\cup B_2=A_2\cup C_2,\end{cases}\\
\Leftrightarrow\ &\begin{cases}A_1\cap C_1\subseteq B_1,\\A_2\cup C_2\supseteq B_2,\end{cases}
\end{aligned}$$

$$c\cap\langle X,Y\rangle=\langle C_1,C_2\rangle\Leftrightarrow\begin{cases}C_1\subseteq X,\\C_2\supseteq Y,\end{cases}$$

所以 X 是满足$A_1\cap C_1\subseteq B_1$ 的$\subseteq$关系意义下的极大元,Y 是满足$A_2\cup C_2\supseteq B_2$ 的$\subseteq$关系意义下的极小元.

我们断言 $Y=B_2\cap-A_2$. 事实上,对 $\forall C_2$, $A_2\cup C_2\supseteq B_2\Leftrightarrow(A_2\cap(-C_2))\cup C_2\supseteq B_2$. $C_2\supseteq B_2-(A_2\cap(-C_2))=B_2\cap(-(A_2\cap(-C_2)))\supseteq B_2\cap(-A_2)$,所以 $Y=B_2\cap-A_2$.

同理可证$-A_1\cup B_1$是满足 $A_1\cap C_1\subseteq B_1$ 的极大元,但由于 X,Y 必须满足(iii)所以可令 $X=-A_1\cup B_1-(B_2\cap-A_2)=(-A_1\cup B_1)\cap-(B_2\cap-A_2)$.

$$\begin{aligned}
&(-A_1\cup B_1)\cap-(B_2\cap-A_2)\\
=&(-A_1\cup B_1)\cap(-B_2\cup A_2)\\
=&(-A_1\cap(-B_2\cup A_2))\cup(B_1\cap(-B_2\cup A_2))
\end{aligned}$$

$= ((-A_1 \cap -B_2) \cup (-A_1 \cap A_2)) \cup ((B_1 \cap -B_2) \cup (B_1 \cap A_2))$
$= (-A_1 \cap -B_2) \cup A_2 \cup B_1 \cup (B_1 \cap A_2)$
$= (-A_1 \cap -B_2) \cup (B_1 \cup A_2)$,

故有

$$X = (-A_1 \cap -B_2) \cup (B_1 \cup A_2);\ Y = B_2 \cap -A_2.$$

所以$\langle X, Y\rangle$是$N_\Theta(A)$中两个元$d = \langle D_1, D_2\rangle$和$e = \langle E_1, E_2\rangle$的析取. 使得$D_1 \cup E_1 = (-A_1 \cap -B_2) \cup (A_2 \cup B_1)$, $D_2 \cap E_2 = B_2 \cap -A_2$. 而$d$是两个元$d' = \langle D'_1, D'_2\rangle$和$d'' = \langle D''_1, D''_2\rangle$的合取，使得$D'_1 \cap D''_1 = -A_1 \cap -B_2$, $(D'_2 \cup D''_2) \cap E_2 = D_2 \cap E_2 = B_2 \cap -A_2$.

e是两个元$E' = \langle E'_1, E'_2\rangle$和$E'' = \langle E''_1, E''_2\rangle$的析取，使得$E'_1 \cup E''_1 = A_2 \cup B_1$, $(E'_2 \cap E''_2) \cap (D'_2 \cup D''_2) = (E'_2 \cap E''_2) \cap D_2 = D_2 \cap E_2 = B_2 \cap -A_2$.

现在我们寻求这组方程的具有最简表达复杂性的解.

我们断言$d' = \neg\langle A_1, A_2\rangle$, $d'' = \neg\sim\langle B_1, B_2\rangle$, $e' = \sim\neg\sim\langle A_1, A_2\rangle$, $e'' = \langle B_1, B_2\rangle$.

事实上，一方面$d' \wedge d'' = \neg\langle A_1, A_2\rangle \wedge \neg\sim\langle B_1, B_2\rangle = \langle -A_1, A_1\rangle \wedge \langle -B_2, B_2\rangle = \langle -A_1 \cap -B_2, A_1 \cup B_2\rangle$，另一方面$e' \vee e'' = \sim\neg\sim\langle A_1, A_2\rangle \vee \langle B_1, B_2\rangle = \langle A_2, -A_2\rangle \vee \langle B_1, B_2\rangle = \langle A_2 \cup B_1, -A_2 \cap B_2\rangle$.

而$d \vee e = \langle -A_1 \cap -B_2, A_1 \cup B_2\rangle \vee \langle A_2 \cup B_1, -A_2 \cap B_2\rangle = \langle (-A_1 \cap -B_2) \cup (A_2 \cup B_1), (A_1 \cup B_2) \cap (-A_2 \cap B_2)\rangle$. 因为$(A_1 \cup B_2) \supseteq B_2 \supseteq -A_2 \cap B_2$，所以$(A_1 \cup B_2) \cap (-A_2 \cap B_2) = -A_2 \cap B_2$. 故

$$d \vee e = \langle (-A_1 \cap -B_2) \cup (A_2 \cup B_1), -A_2 \cap B_2\rangle = \langle X, Y\rangle.$$

这正是我们所需要的. 故

$$\begin{aligned} a \supset b &= d \vee e = (d' \wedge d'') \vee (e' \vee e'') \\ &= (\neg a \wedge \neg\sim b) \vee \sim\neg\sim a \vee b. \end{aligned}$$ □

推论 5.3 对$\forall a \in N_\Theta(A)$, $a \supset 0 = \sim\neg\sim a$.

为了简化表达，我们记$\sim\neg\sim = \div$. 下面的例子说明如何使用运算$\supset$和$\div$.

例 5.4 （接例 5.2）

$\langle \varnothing, \{a\}\rangle \supset \langle \{a\}, \{b, c\}\rangle$
$= (\neg\langle \varnothing, \{a\}\rangle \wedge \neg\sim\langle \{a\}, \{b, c\}\rangle) \vee \sim\neg\sim\langle \varnothing, \{a\}\rangle \vee \langle \{a\}, \{b, c\}\rangle$
$= (\langle \{a, b, c\}, \varnothing\rangle \wedge \langle \{a\}, \{b, c\}\rangle) \vee \langle \{a\}, \{b, c\}\rangle \vee \langle \{a\}, \{b, c\}\rangle$
$= \langle \{a\}, \{b, c\}\rangle$

$=\langle\varnothing,\{a\}\rangle\supset\langle\varnothing,\{a,b,c\}\rangle$

$=\div\langle\varnothing,\{a\}\rangle\div\langle\{a\},\varnothing\rangle=\langle\varnothing,\{a,b,c\}\rangle$.

由上面的讨论可知,$\supset$是格 $N_\Theta(\boldsymbol{A})$作为 Heyting 代数的相对伪补,而$\div$则是伪补.

§5.4 粗糙概念的代数刻画

在本节中,读者将会看到运算$\div$在粗糙集系统的代数分析中的中心作用.事实上,我们将在本节研究$\div$和$\neg$在刻画可定义集、上粗相等和下粗相等中的作用.在此之前我们先就引入的新运算$\div$所能提供于 $N_\Theta(\boldsymbol{A})$的代数结论给出一个结论.在下面的内容中,给定 r 及$a\in N_\Theta(\boldsymbol{A})$,如果 a 在 r 下有惟一确定的逆像X,则用 $r^{-1}(a)$记之.

定义 5.12

(1)一个分配格 $\boldsymbol{L}=\langle L,\vee,\wedge,+,1,0\rangle$被称为是一个对偶伪补格,如果对$\forall x\in L$,$\exists x^+\in L$,使得对$\forall y\in L$,$x\vee y=1$ 当且仅当 $x^+\leqslant y$.元 x^+称为 x 的对偶伪补.

(2)一个分配格称为是一个双伪补格,如果它既是一个伪补格又是一个对偶伪补格.

(3)一个伪补格 $\boldsymbol{L}=\langle L,\vee,\wedge,*,1,0\rangle$称为是一个 Stone 代数,如果对$\forall x\in L$,$x^+\wedge x^{++}=0$.

(4)一个分配格称为一个双 Stone 代数,如果它既是一个 Stone 代数又是一个对偶 Stone 代数.

(5)一个双 Stone 代数 $\boldsymbol{L}$ 称为是正则的,如果对$\forall x,y\in L$,$x\wedge x^+\leqslant y\vee y^*$.

定理 5.7 对任意近似空间 $\boldsymbol{A}$,对伪补运算$\div$和对偶伪补$\neg$,$a\in N_\Theta(\boldsymbol{A})$是一个正则的双 Stone 代数.

证 运算$\div$是一个伪补运算,这一点在上一节中已经证明过了,由$\div$的定义,$\div a=\sim\neg\sim a=\langle A_2,-A_2\rangle$,$\div\div a=\langle -A_2,A_2\rangle$,所以$\div a\vee\div\div a=\langle A_2,-A_2\rangle\vee\langle -A_2,A_2\rangle=\langle U,\varnothing\rangle=1$,所以 $N_\Theta(\boldsymbol{A})$对运算$\div$是一个 Stone 代数.

又若$\langle X_1,X_2\rangle\vee\langle Y_1,Y_2\rangle=1$,则 $X_1\cup Y_1=U$,$X_2\cap Y_2=\varnothing$.所以有$-X_1=Y_1$,$Y_2\subseteq -Y_1=X_1$.故$\neg\langle X_1,X_2\rangle\wedge\langle Y_1,Y_2\rangle=\langle -X_1\cap Y_1,X_1\cup Y_2\rangle=\langle -X_1,X_1\rangle=\neg\langle X_1,X_2\rangle$,即$\neg x\leqslant y$,反过来也对.故$\neg$是 $N_\Theta(A)$中的一个对偶伪补.

由于$\neg x\wedge\neg\neg x=\langle -X_1,X_1\rangle\wedge\langle X_1,-X_1\rangle=\langle\varnothing,U\rangle=0$,故 $N_\Theta(\boldsymbol{A})$

是一个对偶的 Stone 代数(对运算$\neg$).

最后,对$\forall x,y\in N_\Theta(\boldsymbol{A})$,$x\wedge\neg x=\langle X_1,X_2\rangle\wedge\langle -X_1,X_1\rangle=\langle\varnothing,X_1\cup X_2\rangle$,$y\vee\div y=\langle Y_1,Y_2\rangle\vee\langle Y_2,-Y_2\rangle=\langle Y_1\cup Y_2,\varnothing\rangle$,$(x\wedge\neg x)\wedge(y\vee\div y)=\langle\varnothing,X_1\cup X_2\rangle=x\wedge\neg x$,所以$x\wedge\neg x\leqslant y\vee\div y$.故$\neg$,$\div$满足正则性条件.

由以上的讨论即知定理 5.7 真. □

定义 5.13 设$\boldsymbol{P}=\langle P,\leqslant\rangle$是一个偏序集,$g:P\to P$是一个保序映射,即对$\forall a,b\in P$,$a\leqslant b$蕴涵$g(a)\leqslant g(b)$.

(1)如果对$\forall a\in P$,$g(g(a))=a$,则称g是一个投射算子.

(2)如果g是一个$\boldsymbol{P}$上的投射算子且对$\forall a\in P$,$a\leqslant g(a)$,则称g是一个闭包算子.

(3)如果g是$\boldsymbol{P}$的一个投射算子且对$\forall a\in P$,$g(a)\leqslant a$,则称g是一个核算子.

引理 5.16 考虑格$N_\Theta(\boldsymbol{A})$,其中的偏序为$\leqslant$,则

(1)算子$\neg\neg$是一个核算子.

(2)算子$\div\div$是一个闭包算子.

证 (1)设$a\leqslant b$,即有$A_1\subseteq B_1$,$B_2\subseteq A_2$.而$\neg\neg a=\langle A_1,-A_1\rangle$,$\neg\neg b=\langle B_1,-B_1\rangle$.所以$\neg\neg a\leqslant\neg\neg b$.故$\neg\neg$是一个保序映射.$\neg\neg\neg\neg a=\neg\neg\langle A_1,-A_1\rangle=\langle A_1,-A_1\rangle$而$a=\langle A_1,A_2\rangle$.因为$-A_1\supseteq A_2$,故$\neg\neg\neg\neg a\leqslant a$.所以$\neg\neg$是核算子.

(2)设$a\leqslant b$,即有$A_1\subseteq B_1$,$B_2\subseteq A_2$.而$\div\div a=\sim\neg\sim\sim\neg\sim a=\sim\neg\neg\sim a=\sim\neg\neg\sim\langle A_1,A_2\rangle=\sim\neg\neg\langle A_2,A_1\rangle=\sim\langle A_2,-A_2\rangle=\langle -A_2,A_2\rangle$,$\div\div b=\langle -B_2,B_2\rangle$.因为$B_2\subseteq A_2$,故$-A_2\subseteq -B_2$,所以$\div\div a\leqslant\div\div b$.故$\div\div$是保序映射.易知$\div\div\div\div a=\langle -A_2,A_2\rangle$,而$a=\langle A_1,A_2\rangle$.由于$A_1\subseteq -A_2$,故$\div\div\div\div a\geqslant a$,所以$\div\div$是一个闭包算子. □

定义 5.14 我们称元$a\in N_\Theta(\boldsymbol{A})$是一个精确元,如果$A_1\cup A_2=U$.

精确元恰好反映了粗糙集系统中的可定义集.

引理 5.17 对$\forall a\in N_\Theta(\boldsymbol{A})$,

(1) a是精确的当且仅当$a\vee\sim a=1$.

(2) $\div a$是精确元.

(3) $\neg a$是精确元.

(4) a是精确元当且仅当$\sim a$是精确元.

(5)如果a是精确的,则$r^{-1}(a)$是U的可定义集.

引理 5.17 的证明是显然的,略.

定义 5.15 设$\boldsymbol{L}=\langle L,\wedge,\vee,0,1\rangle$是一个分配格,如果对一个元$x\in L$,

存在一个元 $x'\in L$，使得 $x'\vee x=1$，$x'\wedge x=0$，则称 x 是可补元，x'是 x 的补元. 格 $\boldsymbol{L}$ 的所有可补元所成的集合称为 $\boldsymbol{L}$ 的中心，记为 $C(\boldsymbol{L})$.

定理 5.8 $N_\Theta(\boldsymbol{A})$的所有精确元的集合是 $C(N_\Theta(\boldsymbol{A}))$.

证 由精确元的定义，元 $a=\langle A_1,A_2\rangle\in N_\Theta(\boldsymbol{A})$是精确元当且仅当 $A_1\cup A_2=U$ 当且仅当 $A_1=-A_2$ 当且仅当 $\sim a\vee a=1$ 且 $\sim a\wedge a=0$ 当且仅当 a 是可补元. □

直接地，如果 $a\in C(N_\Theta(\boldsymbol{A}))$，则 $\sim a=\neg a=\div a$. 下面让我们利用算子 $\div\div$，$\neg\neg$ 和强非 $\sim$ 定义一些 $N_\Theta(\boldsymbol{A})$上的等价关系，它们将被用于刻画集合的上粗、下粗相等.

定义 5.16 在 $N_\Theta(\boldsymbol{A})$中定义二元关系 $\eqsim$，$\simeq$，$\sim\eqsim$ 和 $\sim\simeq$，对 $\forall a,b\in N_\Theta(\boldsymbol{A})$，

(1) $a\eqsim b$ 当且仅当 $\div\div a=\div\div b$.

(2) $a\simeq b$ 当且仅当 $\neg\neg a=\neg\neg b$.

(3) $a\sim\eqsim b$ 当且仅当 $\sim a\eqsim\sim b$.

(4) $a\sim\simeq b$ 当且仅当 $\sim a\simeq\sim b$.

容易验证这四个二元关系都是 $N_\Theta(\boldsymbol{A})$上的等价关系，且我们分别用记号 $[a]_{\eqsim}$，$[a]_{\simeq}$，$[a]_{\sim\eqsim}$ 和 $[a]_{\sim\simeq}$ 表示 a 所在的各关系的等价类.

在下面的引理 5.18 中，我们列举了一些容易验证的结论，我们用括号括住了各种补所刻画的算子. 实际上，我们想要证明的是两个算子$(\div\div)$和$(\neg\neg)$之间的关系，以及强非 $\sim$ 是如何转化这些算子的. 读者可以体会到它与集论补 $-$ 在将闭包算子 $\mathscr{C}$ 和内部算子 $\mathscr{I}$ 相互转化中类似的作用. 这种类似并非偶然的，在稍后的内容中，我们将看到算子$(\div\div)$和$(\neg\neg)$分别联系于近似空间 $\boldsymbol{A}$ 作为 U 上的拓扑空间所诱导的闭包算子和内部算子，为了刻画上粗相等和下粗相等，我们将揭示这种关系.

引理 5.18 对 $\forall a\in N_\Theta(\boldsymbol{A})$，

(0) $a=b$ 当且仅当 $\sim a=\sim b$.

(1) $\sim(\div)a=(\neg)\sim a=\div\div a$.

(2) $\sim(\neg)a=(\div)\sim a=\neg\neg a$.

(3) $\sim(\div\div)a=(\neg\neg)\sim a=\div a$.

(4) $\sim(\neg\neg)a=(\div\div)\sim a=\neg a$.

(5) $a\simeq b$ 当且仅当 $\neg a=\neg b$.

(6) $a\eqsim b$ 当且仅当 $\div a=\div b$.

(7) $a\eqsim b$ 当且仅当 $a\sim\simeq b$.

(8) $a\simeq b$ 当且仅当 $a\sim\eqsim b$.

(9) $\div\div a$ 是$[a]_{\eqsim}$中的最大元.

(10) $\neg\neg a$ 是$[a]_{\simeq}$中的最小元.

(11) $a \simeq b$ 当且仅当$a \to b = b \to a = 1$.

(12) $a \eqsim b$ 当且仅当$\sim a \to \sim b = \sim b \to \sim a = 1$.

证 (0)~(4)的证明是平凡的,只需一些简单的推导.对(5),由定义 $a \simeq b$ 当且仅当$\div\div a = \div\div b$,由(0),当且仅当$\sim\div\div a = \sim\div\div b$ 由(4)当且仅当$\neg a = \neg b$.

(6)的证明与(5)的证明类似.对(7),由(6),$a \eqsim b$ 当且仅当$\div a = \div b$,由(0),当且仅当$\sim\div a = \sim\div b$,由(1),当且仅当$\neg\sim a = \neg\sim b$,由(5),当且仅当$\sim a \simeq \sim b$,由$\sim\simeq$的定义,当且仅当 $a \sim\simeq b$.

(8)的证明与(7)类似.对(9),设$\forall x \in [a]_{\eqsim}$,则$\div\div a = \div\div x$,$\div\div a \wedge x = \div\div x \wedge x = \sim\neg\neg\sim\langle X_1, X_2\rangle \wedge \langle X_1, X_2\rangle = \sim\neg\neg\langle X_2, X_1\rangle \wedge \langle X_1, X_2\rangle = \neg\langle X_2, X_1\rangle \wedge \langle X_1, X_2\rangle - \langle -X_2, X_2\rangle \wedge \langle X_1, X_2\rangle = \langle X_1, X_2\rangle = x$ 故$x \leqslant \div\div a$,由 x 的任意性,$\div\div a$ 是$[a]_{\eqsim}$中的最大元.

(10)的证明与(9)类似.对(11),因为 $a \simeq b$ 当且仅当$\neg\neg a = \neg\neg b$ 当且仅当$\langle A_1, -A_1\rangle = \langle B_1, -B_1\rangle$当且仅当 $A_1 = B_1$(当且仅当$-A_1 = -B_1$),则$a \to b = \langle -A_1 \cup B_1, A_1 \cap B_2\rangle = \langle U, \varnothing\rangle = 1$,$b \to a = \langle -B_1 \cup A_1, B_1 \cap A_2\rangle = \langle U, \varnothing\rangle = 1$.反之 $a \to b = 1$ 或 $b \to a = 1$,则 $A_1 = B_1$,则 $a \simeq b$.

(12)的证明与(11)的证明类似. □

定理 5.9 对$\forall X \subseteq U$,若 $r(X) = a \in N_\theta(A)$,则

(1) $r^{-1}(\neg a)$是使得 $\mathscr{I}(X) \cap Y = \varnothing$的最大可定义集 Y.

(2) $r^{-1}(\neg\neg a)$是使得 $Y \subseteq X$ 的最大可定义集Y.

(3) $r^{-1}(\div a)$是使得 $\mathscr{C}(X) \cap Y = \varnothing$的最大可定义集 Y.

(4) $r^{-1}(\div\div a)$是使得 $X \subseteq Y$ 的最小可定义集Y.

证 (1)由引理 5.17 可知$\neg a$ 是精确元.$\neg a = \langle -\mathscr{I}(X), \mathscr{I}(X)\rangle$.如果$Y = r^{-1}(\neg a)$,则$\mathscr{I}(Y) = -\mathscr{I}(X)$.但由引理 5.17 的(5),$Y$ 是可定义集,故$Y = \mathscr{I}(Y) = -\mathscr{I}(X)$.(1)得证.

(2)由引理 5.17,$\neg\neg a$ 是精确元.$\neg\neg a = \langle \mathscr{I}(X), -\mathscr{I}(X)\rangle$.如果 $Y = r^{-1}(\neg\neg a)$,则$\mathscr{I}(Y) = \mathscr{I}(X)$.因为 $r^{-1}(\neg\neg a)$是可定义集,$Y = \mathscr{I}(Y) = \mathscr{I}(X)$,故 $r^{-1}(\neg\neg a)$是满足 $Y \subseteq X$ 的最大可定义集.

(3),(4)可按同样的方式证明. □

我们从定理 5.9 中看到,算子$\div\div$和$\neg\neg$将任一粗糙元投射到一个特殊的,代表可定义集的精确元.

到现在为止,为了揭示粗糙集的不交表示在分析粗糙集系统的一般性质中的作用,我们已经研究了上粗相等和下粗相等的一些特征.下面让我们首先展示前面研究的逻辑代数算子是如何提供一种识别两个集是上粗或下粗相等

的.

定理 5.10(上粗相等的代数刻画) 对论域 U 上的任一个近似空间 $\mathbf{A}$，$\forall X, Y \subseteq U$，下面的陈述是等价的.

(1) $X \eqsim y$.

(2) $\div\div r(X) = \div\div r(Y)$.

(3) $\neg\div r(X) = \neg\div r(Y)$.

(4) $\neg\sim r(X) = \neg\sim r(Y)$.

(5) $\sim\div r(X) = \sim\div r(Y)$.

(6) $\div r(X) = \div r(Y)$.

证 (1)⇔(2) $X \eqsim Y$ 当且仅当 $\mathscr{C}(X) = \mathscr{C}(Y)$ 当且仅当 $-\mathscr{C}(X) = -\mathscr{C}(Y)$ 当且仅当 $\langle \mathscr{C}(X), -\mathscr{C}(X) \rangle = \langle \mathscr{C}(Y), -\mathscr{C}(Y) \rangle$ 当且仅当 $\div\div r(X) = \div\div r(Y)$，由引理 5.18(1)可知(1)⇔(2)⇔(4)⇔(5).

(1)⇔(3) $\neg\div r(X) = \neg\sim\neg\sim r(X) = \neg\sim\neg\sim\langle \mathscr{I}(X), -\mathscr{C}(X) \rangle = \langle \mathscr{C}(X), -\mathscr{C}(X) \rangle$. 所以 $\neg\div r(X) = \neg\div r(Y)$ 当且仅当 $\mathscr{C}(X) = \mathscr{C}(Y)$ 当且仅当(1).

(1)⇔(6) $\div r(X) = \sim\neg\sim r(X) = \sim\neg\sim\langle \mathscr{I}(X), -\mathscr{C}(X) \rangle = \langle -\mathscr{C}(X), \mathscr{C}(X) \rangle$. 所以 $\div r(X) = \div r(Y)$ 当且仅当 $\mathscr{C}(X) = \mathscr{C}(Y)$ 当且仅当(1). □

定理 5.11(下粗相等的代数刻画)

对论域 U 上的任一个近似空间 $\mathbf{A}$，$\forall X, Y \subseteq U$，下面的陈述等价.

(1) $X \simeq Y$.

(2) $\neg\neg r(X) = \neg\neg r(Y)$.

(3) $\sim\neg r(X) = \sim\neg r(Y)$.

(4) $\neg r(X) = \neg r(Y)$.

对定理 5.11，可按定理 5.10 的方式证明，略.

至此，我们有这样一个问题，即如何判定 $P(U)$ 上的等价关系是上粗相等，下粗相等关系，还是两者都不是.

引理 5.19 对 U 上的任一近似空间 $\mathbf{A}$，

(1) $\div\div r$ 是一个半格 $\langle P(U), \cup \rangle$ 到半格 $\langle N_\Theta(\mathbf{A}), \vee \rangle$ 的 0-1 同态.

(2) $\neg\neg r$ 是一个半格 $\langle P(U), \cap \rangle$ 到半格 $\langle N_\Theta(\mathbf{A}, \wedge \rangle$ 的 0-1 同态.

证 (1).

(a) $\div\div r(\varnothing) = \div\div\langle \varnothing, U \rangle = \langle \varnothing, U \rangle = 0$.

(b) $\div\div r(U) = \div\div\langle U, \varnothing \rangle = \langle U, \varnothing \rangle = 1$.

(c) $\div\div r(X \cup Y) = \div\div\langle \mathscr{I}(X \cup Y), -\mathscr{C}(X \cup Y) \rangle = \langle \mathscr{C}(X \cup Y), -\mathscr{C}(X \cup Y) \rangle = \langle \mathscr{C}(X) \cup \mathscr{C}(Y), -\mathscr{C}(X) \cap -\mathscr{C}(Y) \rangle = \langle \mathscr{C}(X), -\mathscr{C}(X) \rangle \vee$

$\langle \mathscr{I}(Y), -\mathscr{C}(Y)\rangle = \div\div\langle \mathscr{I}(X), -\mathscr{C}(X)\rangle \vee \div\div\langle \mathscr{I}(Y), -\mathscr{C}(Y)\rangle = \div\div r(X) \vee \div\div r(Y)$.

由(a),(b)和(c),(1)得证.

(2)的证明可由(1)的证明对偶地得到. □

引理 5.20 对论域 U 上的任一近似空间 $\boldsymbol{A}$,

(1)对 $\forall X, Y \in P(U)$,如果 $X \subseteq Y$,则 $r(X) \leqslant r(Y)$.

(2)对 $\forall a, b \in N_\Theta(\boldsymbol{A})$,如果 b 是精确元且 $a \leqslant b$,则 $r^{-1}(a) \leqslant r^{-1}(b)$.

证 (1)由拓扑空间的内部算子和闭包算子的单调性,若 $X \subseteq Y$ 则 $\mathscr{C}(X) \subseteq \mathscr{C}(Y)$, $\mathscr{I}(X) \subseteq \mathscr{I}(Y)$. 故 $r(X) \wedge r(Y) = \langle \mathscr{I}(X), -\mathscr{C}(X)\rangle \wedge \langle \mathscr{I}(Y), -\mathscr{C}(Y)\rangle = \langle \mathscr{I}(X) \cap \mathscr{I}(Y), -\mathscr{C}(X) \cup -\mathscr{C}(Y)\rangle = \langle \mathscr{I}(X), -\mathscr{C}(X)\rangle = r(X)$,所以 $r(X) \leqslant r(Y)$.

(2)设 $a = r(X), b = r(Y)$. 如果 $a \leqslant b$,则 $\mathscr{I}(X) \subseteq \mathscr{I}(Y)$ 且 $-\mathscr{C}(Y) \subseteq -\mathscr{C}(X)$,即 $\mathscr{C}(X) \subseteq \mathscr{C}(Y)$. 或者 $\mathscr{I}(X) \subset X \subset \mathscr{C}(X)$ 或者 $\mathscr{I}(X) = X = \mathscr{C}(X)$. 两种情况下均有 $X \subseteq \mathscr{C}(Y)$. 由 Y 是可定义集,故有 $X \subseteq \mathscr{C}(Y) = Y$. □

引理 5.21 对论域 U 上的任一个近似空间 $\boldsymbol{A}$,映射 $r^{-1}\div\div r: P(U) \to P(U)$(或等价地, $r^{-1}\neg\div r, r^{-1}\neg\sim r, r^{-1}\sim r$)是 $P(U)$ 上的拓扑闭包算子.

证 由定理 5.9(4), $r^{-1}\div\div r$ 是 $\langle P(U), \subseteq\rangle$ 的闭包算子,又因为 $r^{-1}\div\div r(X)$ 是满足 $X \subseteq Y$ 的最小可定义集,即 $r^{-1}\div\div r(X) = \mathscr{C}(X)$,故 $r^{-1}\div\div r$ 还是拓扑的. □

引理 5.22 对论域 U 上的任一个近似空间 $\boldsymbol{A}$,映射 $r^{-1}\neg\neg r: P(U) \to P(U)$(或等价地, $r^{-1}\sim\neg r, r^{-1}\neg\div r$)是 $P(U)$ 上的拓扑内部算子.

证明类似于引理 5.21 的证明,略.

引理 5.23 对 $N_\Theta(\boldsymbol{A})$ 的算子 $\div\div$ 和 $\neg\neg$,它们对 $\vee, \wedge$ 是可分配的.

证 我们证明其中的一个. 如 $\div\div(a \vee b) = \div\div a \vee \div\div b$. $\div\div(a \vee b) = \div\div\langle A_1 \cup B_1, A_2 \cap B_2\rangle = \langle -(A_2 \cap B_2), A_2 \cap B_2\rangle = \langle -A_2 \cup -B_2, A_2 \cap B_2\rangle = \langle -A_2, A_2\rangle \vee \langle -B_2, B_2\rangle = \div\div a \vee \div\div b$. 其余类似可证. □

定理 5.12 设 $\boldsymbol{A}$ 是论域 U 上的一个近似空间,则

(1) $P(U)$ 上的等价关系 $\eqsim$ 是半格 $\langle P(U), \cup\rangle$ 上的一个同余.

(2) $P(U)$ 上的等价关系 $\simeq$ 是半格 $\langle P(U), \cap\rangle$ 上的一个同余.

证 (1)由引理 5.19, $\div\div r$ 是 $\langle P(U), \cup\rangle$ 到 $\langle N_\Theta(\boldsymbol{A}), \vee\rangle$ 的同态映射. 对 $\forall a \in N_\Theta(\boldsymbol{A})$,令 $\Gamma(a) = (\div\div r)^{-1}(a)$. 则 $\{\Gamma(a) \mid a \in N_\Theta(\boldsymbol{A})\}$ 是 $P(U)$ 的划分. 用 $P(U)/\div\div r$ 表示由该划分所决定的商集. 由定理 5.10 可知 $P(U)/\div\div r = P(U)/\eqsim$. 所以 $\Gamma(a)$ 恰为 $\eqsim$ 的等价类. 定义 $[X]_\eqsim \cup [Y]_\eqsim = [X \cup Y]_\eqsim$,则对 $\forall X' \in [X]_\eqsim, Y' \in [Y]_\eqsim$,有 $\div\div r(X' \cup Y') = \div\div r(X')$

$\vee \div\div r(Y')=\div\div r(X)\vee\div\div r(Y)$. 这说明我们定义的$[X]_{\approx}\cup[Y]_{\approx}=[X\cup Y]_{\approx}$与等价类的代表元的选取无关. 故$\approx$是$\langle P(U),\cup\rangle$上的一个同余关系. 且若令$[X]_{\approx}=\Gamma(a)$, $[Y]_{\approx}=\Gamma(b)$, 则$[X\cup Y]_{\approx}=[X]_{\approx}\cup[Y]_{\approx}=\Gamma(a)\cup\Gamma(b)=\Gamma(a\vee b)$.

(2)可类似于(1)证明. □

注 这个结果中陈述的条件只是一个 $P(U)$ 上的等价关系为上粗(下粗)相等关系的必要条件, 下面的例子说明这个条件并非充分的.

例 5.5 设 $U=\{a,b,c\}$, $\boldsymbol{B}(U)=\langle P(U),\cap,\cup,-,0,1\rangle$是 U 的所有子集所构成的布尔代数. 令 R 是$P(U)$上的等价关系. $[\varnothing]_R=\{\varnothing,\{c\}\}$, $[\{a\}]_R=\{\{a\},\{a,c\}\}$, $[\{b\}]_R=\{\{b\},\{b,c\}\}$, $[\{a,b,c\}]_R=\{\{a,b,c\},\{a,b\}\}$. 可以验证 R 是$P(U)$上的$\cup$同余关系. 当然我们可以构造 U 上的一个近似空间 $\boldsymbol{A}$, 使得$\{b,c\}$, $\{b\}$有相同的上近似$\{b,c\}$, 于是 $c\in\mathscr{C}(\{b\})$. 但 $\boldsymbol{A}$ 中的任一元都是开闭集, 所以 $\mathscr{C}(c)=\mathscr{C}(\{b,c\})=\{b,c\}$, 故 $b\in\mathscr{C}(\{c\})$. 所以$\{b,c\}$也是$\{c\}$的上近似. 但$\{c\}$与$\{b,c\}$并不在同一个同余类中.

如果 K 是半格$\langle P(U),\cup\rangle$($\langle P(U)\cap\rangle$)上的一个同余关系, 令 $C^{\cup}(K)=\{\bigcup X\mid X\in K$(这里 X 指同余关系 K 的任一个同余类)$\}$, $C^{\cap}(K)=\{\bigcap X\mid X\in K\}$. 利用 $C^{\cup}(K)$和 $C^{\cap}(K)$, 我们有一个关于 $P(U)$上的同余关系是上(下)粗相等关系的一个刻画, 即定理 5.13.

定理 5.13 对于半格$\langle P(U),\cup\rangle$($\langle P(U),\cap\rangle$)上的同余关系 K, 下面的陈述是等价的.

(1) K 是上粗相等关系(下粗相等关系).

(2) $\langle C^{\cup}(K),\cap,\cup,-,\varnothing,U\rangle$($\langle C^{\cap}(K),\cap,\cup,-,\varnothing,U\rangle$)是一个$\boldsymbol{B}(U)$的布尔子代数.

证 (以上粗相等为例)

(a)若 K 是上粗相等关系, 则$[\varnothing]_K=\{\varnothing\}$, 所以$\bigcup[\varnothing]_K=\varnothing$. 显然, $\bigcup[U]_K=U$. 故 $\boldsymbol{B}(U)$的最小元和最大元均在 $C^{\cup}(K)$中, 而 $C^{\cup}(K)$中的元均为集合, 故其中的交、并运算与 $\boldsymbol{B}(U)$中的相同. 因为 $C^{\cup}(K)$的最大元即$\boldsymbol{B}(U)$的最大元. 故其中的补运算也与 $\boldsymbol{B}(U)$中的补运算相同, 故$\langle C^{\cup}(K),\cap,\cup,-,\varnothing,U\rangle$是 $\boldsymbol{B}(U)$的子代数.

(b)若$\langle C^{\cup}(K),\cap,\cup,-,\varnothing,U\rangle$是 $\boldsymbol{B}(U)$的子代数, 我们只需要证明任一 K 的同余类中的元有相同的上近似. 设 $Y\in[X]_K$, 则 $Y\subseteq\bigcup[X]_K$. 由上近似的单调性及对集合并运算的分配性可知

$$\mathscr{C}(Y)\subseteq\mathscr{C}(\bigcup[X]_K)=\bigcup_{Z\in[X]_K}\mathscr{C}(Z).$$

下面证明反向的包含关系. 设基本集 $E\subseteq\bigcup_{Z\in[X]_K}\mathscr{C}(Z)$, 则存在 $Z\in$

$[X]_K$ 使得 $Z\cap E\neq\varnothing$. 我们断言 $E\cap Y\neq\varnothing$. 在假设条件下,我们断言:除 $[\varnothing]_K$ 外,存在 K 的一个同余类 $[T]_K$ 使得其中的每一个元都不以 E 为其子集. 若不然,除 $[\varnothing]_K$ 外的每一个 K 的同余类的每一个元都含有 E.

考察任一非 $[\varnothing]_K$ 是同余类 $[W]_K$ 的补 $-[W]_K$. 可知该补在 $C^{\cup}(K)$ 中不存在. 这与 $\langle C^{\cup}(K),\cap,\cup,-,\varnothing,U\rangle$ 是 $\boldsymbol{B}(U)$ 的子布尔代数矛盾. 考察 $[X]_K\cup[T]_K$,由于 K 是 $\langle P(U),\cup\rangle$ 上的同余关系,一方面,$[X]_K\cup[T]_K=[Y\cup T]_K$,另一方面 $[X]_K\cup[T]_K=[Z\cup T]_K$. 但 $E\not\subseteq\bigcup[Y\cup T]_K$ 而 $E\subseteq\bigcup[Z\cup T]_K$,矛盾. 故 $Z\cap Y\neq\varnothing$,所以 $E\subseteq\mathscr{C}(Y)$. 这意味着 $\mathscr{C}(Y)=\mathscr{C}(\bigcup[X]_K)$. 由 Y 的任意性可知,$[X]_K$ 中任一元的上近似都相等,故 K 是上粗相等关系. □

接下来我们重新回到按我们前面进行的方式来讨论一个 $\langle P(U),\cup\rangle$ ($\langle P(U),\cap\rangle$)上的同余关系是一个上(下)粗相等关系的条件. 为此我们引用 Heyting 代数中的一个结论.

引理 5.24 设 $\boldsymbol{H}$ 是一个 Heyting 代数,Θ 是其上的一个布尔同余,$N_\Theta(\boldsymbol{H})$ 是由 $\boldsymbol{H}$ 的不交元所成的有序对构成的 Nelson 代数,定义 $N_\Theta(\boldsymbol{H})$ 上的等价关系 $\equiv$ 如下: $\forall a,b\in N_\Theta(\boldsymbol{H})$, $a\equiv b$ 当且仅当 $a\rightarrow b\cap b\rightarrow a=1$. 则 $\equiv$ 对运算 $\vee,\wedge,\neg,0,1$ 是同余关系,且商代数 $N_\Theta(\boldsymbol{H})/\equiv=\langle\{[x]_\equiv\mid x\in N_\Theta(\boldsymbol{H})\},\vee,\wedge,\neg,0,1\rangle$ 是与 $\boldsymbol{H}$ 同构的 Heyting 代数.

定理 5.14 设 $\boldsymbol{A}$ 是论域 U 上的一个近似空间. 令 $G(A)=\{r^{-1}\neg\neg r(X)\mid X\in P(U)\}$. 则 $G(\boldsymbol{A})=\langle G(A),\cap,\cup,-,\varnothing,U\rangle$ 是 $\boldsymbol{B}(U)$ 的一个布尔子代数.

证 因为 $N_\Theta(\boldsymbol{A})$ 是一个 Nelson 代数,故可在其中定义等价关系 $\equiv$,由引理 5.24,商代数 $N_\Theta(\boldsymbol{A})/\equiv$ 是一个同构于 $\boldsymbol{A}$ 的布尔代数. 而由引理 5.18 (11),$a\equiv b$ 当且仅当 $\neg\neg a=\neg\neg b$. 再由引理 5.18(10),$\neg\neg a$ 可以作为等价类 $[a]_\equiv$ 的一个代表元. 而 $\neg\neg r(X)$ 是精确元,故 $G(A)$ 与 $\boldsymbol{A}$ 存在一一对应,考虑到映射 $r^{-1}\neg\neg r$,可知 $G(\boldsymbol{A})$ 即是 $\boldsymbol{A}$,由引理 5.24,$N_\Theta(\boldsymbol{A})/\equiv\cong G(\boldsymbol{A})$. 所以 $\langle G(\boldsymbol{A}),\cap,\cup,-,\varnothing,U\rangle$ 是 $\boldsymbol{B}(U)$ 的一个子布尔代数. □

按同样的方式可以证明下面的定理 5.15.

定理 5.15 设是论域 U 上的近似空间,令 $F(A)=\{r^{-1}\div\div r(X)\mid X\in P(U)\}$. 则 $\boldsymbol{F}(\boldsymbol{A})=\langle F(A),\cap,\cup,-,\varnothing,U\rangle$ 是 $\boldsymbol{B}(U)$ 的子布尔代数.

注:注意到定理 5.12 可知 $G(A)=C^{\cup}(\approx)$, $F(A)=C^{\cup}(\simeq)$. 由定理 5.13 及定理 5.14 可知,若 K 是 $\langle P(U),\cup\rangle$ 上的同余关系,则当 $C^{\cup}(K)=$

$G(A)$时可知，K 是上粗相等关系，对下粗相等也有同样的注.

§5.5 半群中的粗理想

Z. Pawlak 提出粗糙集的概念以来，粗糙集理论与应用都在迅速发展.将粗糙集的思想引入到一个代数系统，可以按十分自然的方式导出所谓的粗糙代数的概念.如粗糙群，半群中的粗理想等.然而粗糙代数并未像粗糙集理论的其他方面那样获得长足的发展.究其原因可能在于粗糙代数的应用前景并不十分明朗，至今并未看到这方面的出色应用.在本节中我们简要介绍这方面的研究.

设 S 是一个半群，ρ 是 S 上的同余关系，即 ρ 是满足如下条件的 S 上的一个等价关系：对 $\forall x\in S,(a,b)\in\rho\Rightarrow(ax,bx),(xa,xb)\in\rho$ 我们以$[a]_\rho$ 记a 所在的ρ 同余类.

根据粗糙集的思想，自然地，同余关系 ρ 决定了 S 上的一个近似空间，因此就有了 S 的子集的近似.

设 A 是S 的任一子集，则 $\rho_-(A)=\{x\in S\mid[x]_\rho\subseteq A\}$，$\rho^-(A)=\{x\in S\mid[x]_\rho\cap A\neq\varnothing\}$，它们分别称为 A 的ρ 下近似和ρ 上近似. $\rho(A)=(\rho_-(A),\rho^-(A))$称为 ρ 粗糙集.

半群 S 上的一个同余关系ρ 称为是完备的，如果对$\forall a,b\in S,[a]_\rho[b]_\rho=[ab]_\rho$.

根据第一章的结论立刻有如下的定理.

定理 5.16 设 ρ,λ 是半群S 上的同余关系，$A,B\subseteq S$，则

(1) $\rho_-(A)\subseteq A\subseteq\rho^-(A)$.

(2) $\rho^-(A\cup B)=\rho^-(A)\cup\rho^-(B)$.

(3) $\rho_-(A\cap B)=\rho_-(A)\cap\rho_-(B)$.

(4) $A\subseteq B\Rightarrow\rho_-(A)\subseteq\rho_-(B)$.

(5) $A\subseteq B\Rightarrow\rho^-(A)\subseteq\rho^-(B)$.

(6) $\rho_-(A\cup B)\supseteq\rho_-(A)\cup\rho_-(B)$.

(7) $\rho^-(A\cap B)\subseteq\rho^-(A)\cap\rho^-(B)$.

(8) $\rho\subseteq\lambda\Rightarrow\rho_-(A)\supseteq\lambda_-(A)$.

(9) $\rho\subseteq\lambda\Rightarrow\rho^-(A)\subseteq\lambda^-(A)$.

定理 5.17 设 ρ 是半群S 上的同余关系，A,B 是S 的非空子集，则

$$\rho^-(A)\rho^-(B)\subseteq\rho^-(AB).$$

证 设 $c\in\rho^-(A)\rho^-(B)$，则存在 $a\in\rho^-(A),b\in\rho^-(A)$使得 $c=ab$.

故存在 $x\in[a]_\rho\cap A$，$y\in[b]_\rho\cap B$. 由 ρ 是同余关系，$xy\in[a]_\rho[b]_\rho\subseteq[ab]_\rho$. 因为 $xy\in AB$，故 $xy\in[ab]_\rho\cap AB$，所以 $ab\in\rho^-(AB)$. □

定理 5.18 设 ρ 是半群 S 上的完备同余，A，B 是 S 的非空子集，则

$$\rho_-(A)\rho_-(B)\subseteq\rho_-(AB).$$

证明类似于定理 5.17，略.

定理 5.19 设 ρ 和 λ 是半群 S 上的同余关系，A 是 S 的子集，则

$$(\rho\cap\lambda)^-(A)\subseteq\rho^-(A)\cap\lambda^-(A).$$

证 注意到 S 的同余关系的交仍是同余关系，则该结论是定理 5.16(9) 的直接推论. □

半群 S 的非空子集 A 称为是 S 的子半群，如果 $AA\subseteq A$；A 称为 S 的左理想，如果 $SA\subseteq A$；A 称为 S 的一个右理想，如果 $AS\subseteq A$；A 称为 S 的一个(双侧)理想，如果 A 既是左理想又是右理想.

设 ρ 是半群 S 的一个同余关系，S 的非空集 A 称为是一个上粗子半群，如果 $\rho^-(A)$ 是 S 的子半群；A 称为是 S 的一个上粗左(右，双侧)理想，如果 $\rho^-(A)$ 是 S 的左(右，双侧)理想.

定理 5.20 设 ρ 是半群 S 上的同余关系，则

(1)如果 A 是 S 的一个子半群，则 A 是 S 的一个上粗子半群.

(2)如果 A 是 S 的一个左(右，双侧)理想，则 A 是 S 的一个上粗左(右，双侧)理想.

证 (1)设 A 是 S 的子半群，则由定理 5.16(1)，有 $A\subseteq\rho^-(A)$. 由定理 5.17 及定理 5.16(5)，有 $\rho^-(A)\rho^-(A)\subseteq\rho^-(AA)\subseteq\rho^-(A)$. 这意味着 $\rho^-(A)$ 是 S 的子半群，故 A 是一个上粗子半群.

(2)设 A 是 S 的左理想，即 $SA\subseteq A$. 注意到 $\rho^-(S)=S$，则有 $S\rho^-(A)=\rho^-(S)\rho^-(A)\subseteq\rho^-(SA)\subseteq\rho^-(A)$. 这意味着 $\rho^-(A)$ 是 S 的左理想. 故 A 是一个上粗左理想，其余情况的证明类似. □

这个定理说明上粗子半群(左理想，右理想，双侧理想)是通常的子半群(左理想，右理想，双侧理想)等概念的扩张. 下面的例子说明定理 5.20 的逆不成立.

例 5.6 设 $S=\{a,b,c,d\}$ 是一个半群，其上的运算由下表 5.2 定义，S 上的同余关系 ρ 所决定的同余类为 $\{a\}$，$\{d\}$，$\{b,c\}$. 则对 $A=\{b\}$，$\rho^-(A)=\{b,c\}$. 且 $\{b,c\}S=S\{b,c\}=\{b,c\}$. 故 A 是一个上粗双侧理想，但 A 不是 S 的双侧理想.

表 5.2

	a	b	c	d
a	a	b	c	d
b	b	b	b	b
c	c	c	c	c
d	d	c	b	a

定理 5.21 设 ρ 是半群 S 上的一个完备同余关系，则

(1)如果 $A\subseteq S$ 是一个子半群，且 $\rho_-(A)\neq\varnothing$，则 $\rho_-(A)$ 是 S 的子半群.

(2)如果 $A\subseteq S$ 是一个左(右，双侧)理想，且 $\rho_-(A)\neq\varnothing$，则 $\rho_-(A)$ 是 S 的一个左(右，双侧)理想.

证 (1)因为 A 是 S 的子半群，且 ρ 是完备的，则有 $\rho_-(A)\rho_-(A)\subseteq\rho_-(AA)\subseteq\rho_-(A)$. 所以 $\rho_-(A)$ 是 S 的一个子半群.

(2)设 A 是 S 的一个左理想，即 $SA\subseteq A$. 注意到 $\rho_-(S)=S$. 则有

$$S\rho_-(A)=\rho_-(S)\rho_-(A)\subseteq\rho_-(SA)\subseteq\rho_-(A),$$

故 $\rho_-(A)$ 是 S 的一个左理想，其他情形可类似证明. □

半群 S 的非空子集 A 称为一个双理想，如果 $ASA\subseteq A$. 一个 S 的子集 A 称为一个上粗双理想，如果 $\rho^-(A)$ 是 S 的一个双理想. 易证下面的定理 5.22 和定理 5.23.

定理 5.22 设 ρ 是 S 上的一个同余关系，如果 $A\subseteq S$ 是 S 的一个双理想，则 A 是 S 的一个上粗双理想.

定理 5.23 设 ρ 是 S 上的一个完备同余关系，如果 $A\subseteq S$ 是 S 的一个双理想，且 $\rho_-(A)\neq\varnothing$，则 $\rho_-(A)$ 是 S 的一个双理想.

定理 5.24 设 ρ 是半群 S 上的一个同余关系，如果 A 是 S 的右理想，B 是 S 的左理想，则 $\rho^-(AB)\subseteq\rho^-(A)\cap\rho^-(B)$，$\rho_-(AB)\subseteq\rho_-(A)\cap\rho_-(B)$.

证 因为 A 是右理想，故 $AB\subseteq AS\subseteq A$；$B$ 是左理想，故 $AB\subseteq SB\subseteq B$. 由 ρ^- 的单调性，$\rho^-(AB)\subseteq\rho^-(A)$，$\rho^-(AB)\subseteq\rho^-(B)$. 故 $\rho^-(AB)\subseteq\rho^-(A)\cap\rho^-(B)$. 同样的道理，有 $\rho_-(AB)\subseteq\rho_-(A)\cap\rho_-(B)$. □

设 α,β 是 S 上的二元关系，我们用 $\alpha\cdot\beta$ 表示关系的复合. 若 α,β 是半群 S 上的同余关系，则众所周知，$\alpha\cdot\beta$ 是 S 上的同余关系当且仅当 $\alpha\cdot\beta=\beta\cdot\alpha$.

定理 5.25 设 ρ,λ 是半群 S 上的同余关系，且 $\rho\lambda=\lambda\rho$. 如果 A 是 S 的子半群，则 $\rho^-(A)\lambda^-(A)\subseteq(\rho\lambda)^-(A)$.

证 设 $c\in\rho^-(A)\lambda^-(A)$，则存在 $a\in\rho^-(A)$，$b\in\lambda^-(A)$，使得 $c=ab$. 故存在 $x,y\in S$ 使得 $x\in[a]_\rho\cap A$，$y\in[b]_\lambda\cap A$. 由 A 是子半群可知 $xy\in A$. $(x,a)\in\rho$，$(y,b)\in\lambda$ 蕴涵 $(xy,ay)\in\rho$ 且 $(ay,ab)\in\lambda$. 故 $(xy,ab)\in\rho\lambda$，

即 $xy\in[ab]_{\rho\lambda}$. 因此有 $xy\in[ab]_{\rho\lambda}\cap A$, 故 $c=ab\in(\rho\lambda)^-(A)$. 所以 $\rho^-(A)\lambda^-(A)\subseteq(\rho\lambda)^-(A)$. □

半群 S 上的一个同余关系 ρ 被称为是一个幂等同余,如果商半群 S/ρ 是一个幂等半群(即每个元都是幂等元). S 的一个子集 P 称为是半素的,如果 $a^2\in P$ 蕴涵 $a\in P$.

定理 5.26 设 ρ 是半群 S 的一个幂等同余, A 是 S 的一个非空子集,则 $\rho^-(A)$ 是半素的.

证 设 $a^2\in\rho^-(A)$. 由于 ρ 是幂等同余,故有 $[a]_\rho\cap A=[a]_\rho[a]_\rho\cap A=[a^2]_\rho\cap A\neq\varnothing$. 这意味着 $[a]_\rho\subseteq\rho^-(A)$. 故 $a\in\rho^-(A)$. 所以 $\rho^-(A)$ 是半素的. □

定理 5.27 设 ρ 是 S 上的幂等同余, A,B 是 S 的非空子集,则 $\rho^-(A)\cap\rho^-(B)\subseteq\rho^-(AB)$, $\rho_-(A)\cap\rho_-(B)\subseteq\rho_-(AB)$.

证 (1)设 $c\in\rho^-(A)\cap\rho^-(B)$, 则 $c\in\rho^-(A)$ 且 $c\in\rho^-(B)$. 故存在元 $a,b\in S$, 使得 $a\in[c]_\rho\cap A$, $b\in[c]_\rho\cap B$. 于是 $ab\in[c]_\rho[c]_\rho=[c]_\rho$. 这意味着 $AB\cap[c]_\rho\neq\varnothing$, 故 $c\in\rho^-(AB)$. 所以 $\rho^-(A)\cap\rho^-(B)\subseteq\rho^-(AB)$.

(2)若 $[c]_\rho\subseteq\rho_-(A)\cap\rho_-(B)$, 则 $[c]_\rho\subseteq A\cap B$. 所以 $[c]_\rho[c]_\rho=[c]_\rho\subseteq AB$, 即 $[c]_\rho\subseteq\rho_-(AB)$. 故 $\rho_-(A)\cap\rho_-(B)\subseteq\rho_-(AB)$. □

由定理 5.26 和定理 5.27,立即有下面的推论.

推论 5.4 设 ρ 是半群 S 的幂等同余. A,B 分别是 S 的右理想和左理想,则 $\rho^-(A)\cap\rho^-(B)=\rho^-(AB)$, $\rho_-(A)\cap\rho_-(B)=\rho_-(AB)$.

设 S,T 是两个半群, f 是 S 到 T 的同态映射,令 $k=\{(a,b)\in S\times S\mid f(a)=f(b)\}$. 则 k 是 S 的一个同余.

定理 5.28 设 S 和 T 是两个半群, f 是 S 到 T 的同态,对 $\forall A\subseteq S$, 有 $f(k^{-1}(A))=f(A)$.

证 因为 $A\subseteq k^{-1}(A)$, 故 $f(A)\subseteq f(k^{-1}(A))$. 又若 $[x]_k\subseteq k^{-1}(A)$, 有 $a\in A$ 使得 $a\in[x]_k$. 根据 k 的定义 $f([x]_k)=\{f(a)\}$. 故 $f(k^{-1}(A))\subseteq f(A)$. 所以 $f(k^{-1}(A))=f(A)$. □

下面我们讨论商半群中的粗糙集.

设 ρ 是半群 S 上的同余关系,显然 S 的子集 A 的上、下近似可以等价地表示为如下形式:

$$\rho_-(A)/\rho=\{X\in S/\rho\mid X\subseteq A\},$$

$$\rho^-(A)/\rho=\{X\in S/\rho\mid X\cap A\neq\varnothing\}.$$

这样 $\rho_-(A)/\rho$, $\rho^-(A)/\rho$ 作为商半群 S/ρ 中的子集有下面的结果.

定理 5.29 设 ρ 是半群 S 上的完备同余关系,

(1)若 $A\subseteq S$ 是 S 的上粗子半群,则 $\rho^-(A)/\rho$ 是 S/ρ 的子半群.

(2)对 $A\subseteq S$,若 $\rho_-(A)\neq\varnothing$ 是 S 的子半群,则 $\rho_-(A)/\rho$ 是 S/ρ 的子半群.

(3)若 $A\subseteq S$ 是 S 的上粗左(右,双侧)理想,则 $\rho^-(A)/\rho$ 是 S/ρ 的左(右,双侧)理想.

(4)对 $A\subseteq S$,若 $\rho_-(A)\neq\varnothing$ 是 S 的左(右,双侧)理想,则 $\rho_-(A)/\rho$ 是 S/ρ 的左(右,双侧)理想.

(5)若 $A\subseteq S$ 是 S 的上粗双理想,则 $\rho^-(A)/\rho$ 是 S/ρ 的双理想.

(6)对 $A\subseteq S$,若 $\rho_-(A)\neq\varnothing$ 是 S 的双理想,则 $\rho^-(A)/\rho$ 是 S/ρ 的双理想.

证 因为 ρ 是完备的,这保证了 S/ρ 的子集的乘积仍然为 S/ρ 的子集.

(1)对 $\forall[a]_\rho,[b]_\rho\in\rho^-(A)/\rho$,由于 $\rho^-(A)$ 是 S 的子半群,$[a]_\rho[b]_\rho=[ab]_\rho\subseteq\rho^-(A)$. 这意味着 $\rho^-(A)/\rho\rho^-(A)/\rho\subseteq\rho^-(A)/\rho$. 故 $\rho^-(A)/\rho$ 是 S/ρ 的子半群.

(2)类似于(1)可证.

(3)对 $\forall[s]_\rho\in S/\rho,[a]_\rho\in\rho^-(A)/\rho$,由于 A 是 S 的上粗左理想,$[s]_\rho[a]_\rho=[sa]_\rho\subseteq\rho^-(A)$. 这意味着 $S/\rho\rho^-(A)/\rho\subseteq\rho^-(A)/\rho$. 故 $\rho^-(A)/\rho$ 是 S/ρ 的左理想.

(4),(5),(6)均可类似地证明. □

第六章　粗糙逻辑与决策

粗糙集理论成功地从数据库中提炼出知识库，进一步利用粗糙集理论给出知识库的推理方法显然是需要进一步研究的课题.本章介绍粗糙逻辑及其在决策中的应用.

§6.1　基于完备信息系统的粗糙逻辑

近似真的概念很早就引起了哲学家和逻辑学家们的注意，而近年来主要是被从事人工智能的研究者所关注.这主要是由于人工智能中专家系统、近似推理等方面的研究需要.本节的主要目的是要为这一问题提供一个理论框架，这里"真"不是确切而是以某种程度被知道.

我们的讨论从粗糙集、近似空间等概念开始，一般地讲我们关于论域的知识并非足够的精确，所以我们不能区别论域中的所有对象，因而也就不能区别论域的所有子集.换句话说，我们只能以某种近似的方式来观察论域的子集.本节所提出的粗糙逻辑有五个逻辑值，即真、假、粗糙真、粗糙假和粗糙不相容.

下面我们先给出几个下面将要用到的定义.

设 $A=(U,R)$ 是一个近似空间，不可定义集 $X\subseteq U$ 可以分为下面的四类：

(1) X 是粗糙可区分的(RDIS)，如果 $\underline{R}X\neq\varnothing,\overline{R}X\neq U$.

(2) X 是内不可区分的(IIND)，如果 $\underline{R}X=\varnothing$.

(3) X 是外不可区分的(EIND)，如果 $\overline{R}X=U$.

(4) X 是完全不可区分的(TIND)，如果 $\underline{R}X=\varnothing,\overline{R}=U$.

显然我们有：X 是可区分的(粗糙可区分的、完全不可区分的)，则 $\sim X$ 也是；X 是内(外)不可区分的，则 $\sim X$ 是外(内)不可区分的.

在粗糙逻辑中我们使用下面的符号：$p,q,r\cdots$ 表示命题变量；T、F 表示真和假；基本的命题联结词是($\sim$)，($\wedge$)及($\vee$)；小写的希腊字母如 φ,ψ 等记粗糙逻辑的公式，其定义按照标准的方式进行.粗糙逻辑的语义通过解释函数($|\ \ |$)按照标准的方式来定义，即对粗糙逻辑的公式，其在论域中的意义如下：

(1) $|T|=U$.

(2) $|F|=\varnothing$.

(3) $|\sim\varphi|=U-|\varphi|$.

(4) $|\varphi\vee\psi|=|\varphi|\cup|\psi|$.

(5) $|\varphi\wedge\psi|=|\varphi|\cap|\psi|$.

我们称 $M=(U,R,|\quad|)$ 为一个模型，这里 (U,R) 是一个近似空间，$|\quad|$ 是一个解释函数.

我们说公式 φ 在 M 中关于 $x\in U$ 是真的，如果 $x\in|\varphi|$，记作 $\models_x\varphi$ 或 $\mathrm{val}_x(\varphi)=T$；否则 φ 在 M 中关于 x 称为是假的.

如果 $|\varphi|=U$，则称公式 φ 在 M 中是真的；否则称为 φ 在 M 中是假的，分别记作 $\models_M(\varphi)$(或 $\mathrm{val}_M(\varphi)=T$)和 $\not\models_M(\varphi)$(或 $\mathrm{val}_M(\varphi)=F$). 当然公式 φ 在 M 中是真的当且仅当 φ 在 M 中关于每一个 $x\in U$ 都是真的，公式 φ 在 M 中是假的当且仅当对某一个 $x\in U$ 是假的.

称公式 φ 在 M 中关于 $x\in U$ 确为真的(记作 $\models_{x,\underline{R}}\varphi$)，如果 $x\in\underline{R}|\varphi|$；称公式 φ 在 M 中关于 $x\in U$ 可能为真的(记作 $\models_{x,\overline{R}}\varphi$)，如果 $x\in\overline{R}|\varphi|$.

事实 6.1 (1) $\models_{x,\underline{R}}\varphi$ 当且仅当 $[x]_R\subseteq|\varphi|$.

(2) $\models_{x,\overline{R}}\varphi$ 当且仅当 $[x]_R\cap|\varphi|\neq\varnothing$.

如果 $|\varphi|$ 是外不可区分的(EIND)，则称 φ 在 M 中粗糙真，记作 $\models_R\varphi$(或 $\mathrm{val}_R(\varphi)=RT$).

显然如果 φ 在 M 中粗糙真，则 φ 在 M 中关于每一个 $x\in U$ 可能真.

如果 φ 在 M 中粗糙真，则 $\sim\varphi$ 称为在 M 中粗糙假记作 $\not\models_R\varphi$ 或 $\mathrm{val}_R(\varphi)=RF$.

事实 6.2 公式 φ 在 M 中粗糙假当且仅当 $|\varphi|$ 是内不可区分的(IIND).

公式 φ 在 M 中是粗糙不相容，如果 φ 在 M 中既是粗糙真又是粗糙假.

事实 6.3 公式 φ 在 M 中粗糙不相容当且仅当 $|\varphi|$ 是完全不可区分的(tind).

在粗糙逻辑中，我们可以按下面的方式讨论真的程度.

设 φ,ψ 都在 M 中粗糙真，我们说在 M 中，φ 比 ψ 更真，记作 $\models_R\varphi>\models_R\psi$，如果 $|\varphi|\supseteq|\psi|$.

设 φ 和 ψ 都在 M 中粗糙假，我们说在 M 中 φ 比 ψ 更假，如果 $|\varphi|\subseteq|\psi|$.

当然，如果 φ 和 ψ 都在 M 中粗糙真，且 φ 比 ψ 更真，则 $\sim\varphi$ 比 $\sim\psi$ 更假.

事实 6.4 如果 φ 和 ψ 都在 M 中粗糙真. 则 $\varphi\vee\psi$ 比 φ 或 ψ 都更真；若 $\varphi\wedge\psi$ 也在 M 中粗糙真，则 φ,ψ 均比 $\varphi\wedge\psi$ 更真.

设 $\varphi_1,\varphi_2,\cdots,\varphi_n,\cdots$ 是一个无限公式序列，使得 φ_{i+1} 比 φ_i 在 M 中更真，则下面的事实是显然的.

事实 6.5 (1)$\overline{R}(\bigcup_{i=1}^{\infty}|\varphi_i|)=\bigcup_{i=1}^{\infty}\overline{R}|\varphi_i|=U$.

(2)$\underline{R}(\bigcap_{i=1}^{\infty}|\sim\varphi_i|)=\bigcap_{i=1}^{\infty}\underline{R}|\sim\varphi_i|=\varnothing$.

事实 6.6 下面的两个条件是等价的：

(1)$\bigcup_{i=1}^{\infty}\underline{R}|\varphi_i|=U$.

(2)$\bigcap_{i=1}^{\infty}\overline{R}|\sim\varphi_i|=\varnothing$.

上述性质的直观意义可以按照下面的方式给予解释：如果 φ_i 在 M 中粗糙真，我们也可以说它在 M 中归纳真，即按照我们当前的知识状态它是真的，随着我们知识的增长，更多的知识允许我们用一个更真的公式 φ_{i+1} 去代替公式 φ_i，最终这个过程将导致一个真的公式 φ，它可以被看作一系列粗糙真公式的极限. 这个极限公式 φ 可以被称为演绎真，因此知识获取的过程就是由部分真(或者说归纳真)到完全真(或者说演绎真)的过程.

§6.2 决策逻辑与决策

在本节中我们将讨论一种具有归纳特征的针对数据分析的逻辑方法，它反映的是带决策属性的信息系统即决策表的完全形式化的性质. 我们首先介绍知识的逻辑表达，决策的逻辑语言、语义，然后介绍决策规则和决策算法，最后讨论相容和不相容算法的简化、最小化以及决策规则的简化问题.

6.2.1 决策逻辑语言

信息系统是我们讨论问题的一个基本概念，与以前不同，在本节中我们将从另一种角度来看信息系统，即它可以被看作是一个关于现实和结果的一些命题的集合，这种理解使得我们可以用逻辑工具来处理信息系统. 通常为方便，信息系统总是被表示为一个二维表，即属性-数值表，我们可以将它看作一种特殊的逻辑模型，这里称为决策逻辑，它将被用来从信息系统中获得蕴涵于其中的知识所导出的结论. 与以前的方法不同，我们将使用逻辑推演中的符号工具，发现知识的依赖性和进行知识简化.

在决策逻辑中有一个基本的概念需要说明，即所谓的决策算法(decision algorithm)，一般而言，算法是指一组能完成某类任务的指令序列，但我们这里所谓的算法却不是这方面的含义，它是指一个由一些决策规则组成的集合. 公式可以是真的或假的，但决策算法作为一组公式的集合却没有真值属性. 决策算法的基本特征是相容性或者称为协调性，即它可以是相容或不相容的，在决策逻辑中，我们所关心的不是证明技巧的研究而是侧重于分析蕴涵于数据中

的事实.

我们定义和讨论的决策逻辑语言由原子公式组成,原子公式是一种属性-属性值对.用命题连接词:与,或,非等通过标准的方法可以构成复合公式,决策逻辑的语言可以形式地归纳定义如下:

(1)语言的符号集

(a) A——属性常量集.

(b) $V=\bigcup_{a\in A}V_a$——属性值常量集.

(c) 命题连接词集$\{\sim,\vee,\wedge,\rightarrow,\equiv\}$,其中的元分别称为非、析取、合取、蕴涵和等价.

(d) 括号(,).

(2) 决策逻辑语言的公式是满足下面条件的极小集

(a) 形如(a,v),或简写为 a_v 的表达式称为原子公式,对$\forall a\in A,v\in V_a$,(a,v)是决策逻辑语言的公式.

(b) 如果 φ,ψ 是决策逻辑语言的公式,则$\sim\varphi,(\varphi\vee\psi),(\varphi\wedge\psi),(\varphi\rightarrow\psi)$和$(\varphi\equiv\psi)$也是公式.

6.2.2 决策逻辑的语义

公式是论域中对象的描述工具,由于论域中的对象可以有相同的描述,故公式同样可以描述子集.特别地,原子公式(a,v)可以被解释为对那些在属性 a 下取值为 v 的对象的描述,复合公式可以按照通常的方式给予解释.为了精确表达这一思想,我们采用模型和可满足性的概念来定义决策逻辑的语义.在决策逻辑中,一个模型是指一个信息系统 $S=(U,A)$.在一个模型 S 中,我们通过适当地解释公式,使得每一个公式变成表达某些对象性质的有意义的语句.这个概念可以通过一个公式被一个对象所满足的概念得到精确的陈述.

一个对象 $x\in U$ 在模型$S=(U,A)$中满足公式 φ,记作 $x\models_S\varphi$ 或$x\in\varphi$,当且仅当下面的条件被满足:

(1) $x\models(a,v)$当且仅当 $a(x)=v$.

(2) $x\models\sim\varphi$ 当且仅当非 $x\models\varphi$.

(3) $x\models\varphi\vee\psi$ 当且仅当$x\models\varphi$ 或$x\models\psi$.

(4) $x\models\varphi\wedge\psi$ 当且仅当$x\models\varphi$ 且$x\models\psi$.

由上面的条件我们有

(5) $x\models\varphi\rightarrow\psi$ 当且仅当$x\models\sim\varphi\vee\psi$.

(6) $x\models\varphi\equiv\psi$ 当且仅当$x\models\varphi\rightarrow\psi$ 且$x\models\psi\rightarrow\varphi$.

如果 φ 是一个公式,定义$|\varphi|_S=\{x\in U|x\models_S\varphi\}$,$|\varphi|_S$ 称为公式 φ 在 S 中的含义.

下面的事实解释了公式的含义.

事实 6.7

(1) $|(a,v)|_S=\{x\in U|a(x)=v\}$.

(2) $|\sim\varphi|_S=-|\varphi|_S$.

(3) $|\varphi\vee\psi|_S=|\varphi|_S\cup|\psi|_S$.

(4) $|\varphi\wedge\psi|_S=|\varphi|_S\cap|\psi|_S$.

(5) $|\varphi\rightarrow\psi|_S=-|\varphi|_S\cup|\psi|_S$.

(6) $|\varphi\equiv\psi|_S=(|\varphi|_S\cap|\psi|_S)\cup(-|\varphi|_S\cap-|\psi|_S)$,

公式的含义就是所有通过公式 φ 表达的对象的集合,或者说公式 φ 是对象集合$|\varphi|_S$ 在决策逻辑语言中的描述.

在决策逻辑中,我们也需要"真"的概念.当且仅当$|\varphi|_S=U$,即论域中的所有对象满足该公式,公式 φ 称为在信息系统 S 中是真的.当且仅当$|\varphi|_S=|\psi|_S$,我们称 φ 与 ψ 在 S 中是等价的.

事实 6.8

(1) $\models_S\varphi$ 当且仅当$|\varphi|_S=U$.

(2) $\models_S\sim\varphi$ 当且仅当$|\varphi|_S=\varnothing$.

(3) $\models_S\varphi\rightarrow\psi$ 当且仅当$|\varphi|_S\subseteq|\psi|_S$.

(4) $\models_S\varphi\equiv\psi$ 当且仅当$|\varphi|_S=|\psi|_S$.

在这里,我们强调:一个公式在一个信息系统中是真的,但在另一个系统中却可能是假的,但有些公式的真值却不依赖于出现在其中的真实属性的值,而是依赖于其本身的结构,这种公式在我们的讨论中具有特殊的作用.

6.2.3 决策逻辑的推演

下面我们准备考虑决策逻辑中的演绎结构,为此我们必须引入推理规则,在我们详细讨论该问题之前,先给出问题解决的一些直观背景.

我们提出的决策逻辑的目的是要表达包含于信息系统中的知识,但同一个语言可以作为许多不同的信息系统的共同语言,从语法角度讲这些语言并没有什么不同,但它们的语义却可以随着模型的不同而不同.为了定义决策逻辑,我们必须证明公式的语义等价性,为此我们需要制定将公式进行不改变语义意义下的进行变换的规则.当然,从理论上讲我们要按照公式含义的定义来证明语义等价性.遗憾的是,这种方法太不实际了,因为它显然是一个无限验证的过程,所以对于形式地证明公式的等价性来讲,适合的公理和推理规则总是必要的.实际上,我们所需要的公理是由经典的命题演算系统的公理和关于信息系统的一些特殊的公理构成的,而推理规则只有假言推理.

在列出所需的特殊公理以前,我们先定义一些辅助概念.

我们将使用如下的缩写形式：$\varphi\wedge\sim\varphi\overset{\wedge}{=}0,\varphi\vee\sim\varphi\overset{\wedge}{=}1$.

显然$\models 1,\models\sim 0$.因此0和1可以分别代表"假"和"真".

$v_i\in v_{a_i},P=\{a_1,a_2,\cdots,a_n\},P\subseteq A$，则称形如$(a_1,v_1)\wedge(a_2,v_2)\wedge\cdots\wedge(a_n,v_n)$的公式为$P$基本公式，或简称为$P$公式，一个$A$基本公式称为一个基本公式.

设$P\subseteq A$，φ是一个P公式且$x\models\varphi$，则φ称为x在S中的P描述.所有在S中可满足的A基本公式的集合称为S中的基本知识，我们用记号$\sum_S(P)$表示所有在S中可满足的P基本公式的析取，有时$\sum_S(P)$可以缩写为$\sum(P)$.如果$P=A$，则$\sum_S(A)$被称为信息系统$S=(U,A)$的特征公式.因此，特征公式代表了一个信息系统的所有知识.

例 6.1 考虑表6.1所示的信息系统

按照上面所引进的概念，$a_1b_0c_2,a_2b_0c_3,a_1b_1c_1,a_2b_1c_3,a_1b_0c_3$就是基本公式(基本知识)，注意这里为了简单，略去了合取符$\wedge$，该系统的特征公式为

$$a_1b_0c_2\vee a_2b_0c_3\vee a_1b_1c_1\vee a_2b_1c_3\vee a_1b_0c_3.$$

表 6.1

U	a	b	c
1	1	0	2
2	2	0	3
3	1	1	1
4	1	1	1
5	2	1	3
6	1	0	3

作为例子，读者容易得到如下几个公式的含义：

$$|a_1\vee b_0c_2|=\{1,3,4,6\},$$

$$|\sim(a_2b_1)|=\{1,2,3,4,6\},$$

$$|b_0\rightarrow c_2|=\{1,3,4,5\},$$

$$|a_2\equiv b_0|=\{2,3,4\}.$$

下面是决策逻辑的特殊公理：

(1) $(a,v)\wedge(a,u)\equiv 0,a\in A,u,v\in V$且$u\neq v$.

(2) $\bigvee\limits_{v\in V_a}(a,v)\equiv 1,a\in A$.

(3) $\sim(a,v)\equiv\bigvee\limits_{u\in V_a,u\neq v}(a,u),a\in A$.

我们有下面的事实.

事实 6.9 对任意 $P\subseteq A$, $\models_S \sum_S(P)\equiv 1$.

我们称公式 φ 可从公式集 Ω 中导出,记作 $\Omega\vdash\varphi$,当且仅当它可以由公理及 Ω 中的公式通过有限次应用推理规则导出.称公式 φ 是决策逻辑的一个定理,记作 $\vdash\varphi$,如果它可以从公理集中导出.一个公式集 Ω 被称为是相容的,当且仅当 $\varphi\wedge\sim\varphi$ 不能从 Ω 中导出.

6.2.4 规范形

设 $P\subseteq A$ 是一个属性子集,φ 是一个公式,称 φ 是 S 中的一个 P 规范形,当且仅当要么 φ 是 0 或 1,要么 φ 是非空的 P 基本公式的析取,这里所谓 φ 是非空的是指 $|\varphi|\neq\varnothing$.一个 A 规范形称为一个规范形.

事实 6.10 设 φ 是一个公式,P 是出现在 φ 中的所有属性构成的属性子集,则存在一个 P 规范形 ψ 使得 $\vdash\varphi\equiv\psi$.

证明类似于经典命题演算中任一公式等价于一个标准析取范式.

例 6.2 在例 6.1 中,我们有

$$a_1\vee b_0c_2\equiv a_1b_0c_2\vee a_1b_1c_1\vee a_1b_0c_3,$$
$$(a_2b_1)\equiv a_1b_0c_2\vee a_2b_0c_3\vee a_1b_1c_1\vee a_1b_0c_3,$$
$$b_0\to c_2\equiv a_1b_0c_2\vee a_1b_1c_1\vee a_2b_1c_3,$$
$$(a_2\equiv b_0)\equiv a_2b_0c_3\vee a_1b_1c_1.$$

6.2.5 决策规则和决策算法

本节我们将定义决策语言的基本概念——决策规则和决策算法.任何一个蕴涵式 $\varphi\to\psi$ 称为一个决策规则.φ,ψ 分别称为 $\varphi\to\psi$ 的前件和后件.若 $\varphi\to\psi$ 在 S 中为真,则称 $\varphi\to\psi$ 在 S 中是相容的,否则,称为在 S 中是不相容的.对决策规则 $\varphi\to\psi$,如果 φ 是一个 P 基本公式,ψ 是一个 Q 基本公式,则称 $\varphi\to\psi$ 是一个 PQ 基本决策规则.若 P,Q 都是明白的,则 PQ 基本决策规则可以简称为基本决策规则.

若 $\varphi_1\to\psi,\varphi_2\to\psi,\cdots,\varphi_n\to\psi$ 都是基本决策规则,则决策规则 $\varphi_1\vee\varphi_2\vee\cdots\vee\varphi_n\to\psi$ 称为由 $\varphi_1\to\psi,\varphi_2\to\psi,\cdots,\varphi_n\to\psi$ 所得的复合决策规则.一个 PQ 规则 $\varphi\to\psi$ 在 S 中是容许的,如果 $\varphi\wedge\psi$ 在 S 中是可满足的.

在本节的剩余内容中,除非特别声明外,我们讨论的规则都是容许的.下面的事实可以被用来检验一个 PQ 规则的相容性.

事实 6.11 一个 PQ 规则 $\varphi\to\psi$ 是相容的,当且仅当出现在该规则前件 φ 的 $P\cup Q$ 规范形中的所有 $P\cup Q$ 基本公式也出现在后件 ψ 的 $P\cup Q$ 规范形

中，否则 PQ 规则 $\varphi\to\psi$ 是不相容的.

例 6.3 如在例 6.1 给定的系统中，规则 $b_0\to c_2$ 不是相容的. 因为 b_0 的 $\{b,c\}$ 规范形为 $b_0c_2\vee b_0c_3$，c_2 的 $\{b,c\}$ 规范形为 b_0c_2. 可见 b_0c_3 没有在规则后件的 $\{b,c\}$ 规范形中出现. 故 $b_0\to c_2$ 是不相容的. 另一方面，容易检验规则 $a_2\to c_3$ 是相容的.

任何一个决策规则的有限集称为一个决策算法，对应地，任何一个基本决策规则的有限集称为一个基本决策算法. 如果在一个基本决策算法中，其所有的决策规则都是 PQ 规则，则称该决策算法是一个 PQ 决策算法，简称为 PQ 算法，记作 (P,Q).

一个 PQ 算法在 S 中是可许的，如果它是所有的可许 PQ 规则所成的集合.

一个 PQ 算法在 S 中是完备的，如果对每一个 $x\in U$，存在算法中的一个 PQ 决策规则 $\varphi\to\psi$ 使得 $x\models_S\varphi\wedge\psi$；否则称该算法是非完备的.

接下来，除非特别声明，我们所考虑的算法都是可许且完备的 PQ 算法.

一个 PQ 算法在 S 中是相容的，如果其中的每一个决策规则都是相容的；否则该算法称为是不相容的.

例 6.4 考虑如下一个信息系统

表 6.2

U	a	b	c	d	e
1	1	0	2	1	1
2	2	1	0	1	0
3	2	1	2	0	2
4	1	2	2	1	1
5	1	2	0	0	2

设 $P=\{a,b,c\}$，$Q=\{d,e\}$ 分别是条件属性和决策属性，则 P,Q 惟一地联系于如下的 PQ 决策算法

$$a_1b_0c_2\to d_1e_1,$$
$$a_2b_1c_0\to d_1e_0,$$
$$a_2b_1c_2\to d_0e_2,$$
$$a_1b_2c_2\to d_1e_1,$$
$$a_1b_2c_0\to d_0e_2.$$

如果设 $R=\{a,b\}$，$T=\{c,d\}$ 分别是条件属性和决策属性，则由表 6.2 决定的 RT 算法为

$$a_1b_0 \to c_2d_1,$$
$$a_2b_1 \to c_0d_1,$$
$$a_2b_1 \to c_2d_0,$$
$$a_1b_2 \to c_2d_1,$$
$$a_1b_2 \to c_0d_0.$$

6.2.6　决策规则的相容性与不可分辨性

为了检查一个决策算法是否是相容的，我们必须检查算法中所有的规则是否是真的. 事实 6.11 为我们提供了一个方法，但这个方法有点过于复杂了，下面的事实 6.12 是一个较简单的方法.

事实 6.12　一个 PQ 决策算法中的规则 $\varphi \to \psi$ 在 S 中是真的(相容的)当且仅当对任意该算法中的规则 $\varphi' \to \psi'$，$\varphi = \varphi'$ 蕴涵 $\psi = \psi'$.

为了检验决策规则 $\varphi \to \psi$ 是否为真，我们要用该规则的前件从算法的其他决策类中分辨出决策类 ψ，这表明"真"这一概念与不可分辨性的概念相联系. 显然，若相同的前件具有不同的后件，则这种规则是不相容的. 我们通过一个例子来说明上面的思想.

例 6.5　表 6.3 给定了一个信息系统.

表 6.3

U	a	b	c	d	e
1	1	0	2	1	1
2	2	1	0	1	0
3	2	1	2	0	2
4	1	2	2	1	1
5	1	2	0	0	2

设 $F = \{a, b, c\}$，$Q = \{d, e\}$ 分别是条件属性和决策属性，问 PQ 算法

$$a_1b_0c_2 \to d_1e_1,$$
$$a_2b_1c_0 \to d_1e_0,$$
$$a_2b_1c_2 \to d_0e_2,$$
$$a_1b_2c_2 \to d_1e_1,$$
$$a_1b_2c_0 \to d_0e_2.$$

是否相容呢?

因为在这个算法中的每一个规则的前件都不相同，所以所有的决策规则都是相容的，因而该算法是相容的.

设 $R = \{a, b\}$，$T = \{c, d\}$ 分别是条件属性和决策属性，则 RT 算法

$$a_1b_0 \to c_2d_1,$$
$$a_2b_1 \to c_0d_1,$$
$$a_2b_1 \to c_2d_0,$$
$$a_1b_2 \to c_2d_1,$$
$$a_1b_2 \to c_0d_0.$$

不是相容的.因为 $a_2b_1 \to c_0d_1$,$a_2b_1 \to c_2d_0$ 这两个规则有相同的前件,但它们的后件却不相同.即我们不能通过 a_2b_1 来区分决策 c_0d_1 和 c_2d_0,所以这两个规则不相容.同样规则 $a_1b_2 \to c_2d_1$ 和 $a_1b_2 \to c_0d_0$ 也是不相容的.故 RT 算法不相容.当然,这一点可以容易地从表 6.4 中看出.

表 6.4

U	a	b	c	d
1	1	0	2	1
4	1	2	2	1
5	1	2	0	0
2	2	1	0	1
3	2	1	2	0

这里为了更清晰,我们调整了对象的次序.

6.2.7 决策算法的相容度

设 $S=(U,A)$是一个模型,$P,Q\subseteq A$,称 Q 完全依赖于P,如果在 S 中存在一个相容的PQ 算法记作$P\Rightarrow_S Q$ 或简写为$P\Rightarrow Q$.

我们还可以定义部分依赖的概念,如果在 S 中存在一个不相容的PQ 算法,则称 Q 部分依赖于P.对于部分依赖就有一个依赖程度的问题,我们用相容度来刻画它.

设(P,Q)是 S 中的一个 PQ 算法,称该算法中,相容的 PQ 规则所成的集合为该算法的正区域,记作 $\mathrm{pos}(P,Q)$,换句话说,一个不相容的算法的正区域是指该算法中相容的部分.称数 $k=|\mathrm{pos}(P,Q)|/|(P,Q)|$为 PQ 算法的相容度.显然,$0\leqslant k\leqslant 1$,当 $k=1$ 时,该算法是相容的;当 $k\neq 1$ 时 Q 部分依赖于P,这时称 Q 对P 的依赖度为k,记作 $P\Rightarrow_k Q$.当给定一个决策算法后,计算其相容度是容易的.

6.2.8 相容算法的简化

当我们依照一个给定的信息系统制定决策时,条件属性和决策属性是预先给定的.按我们前面的讨论我们可以获得一个决策算法.问题是对于决策属

性类是否所有的条件都是必要的呢？这就是算法的简化问题. 这个问题与第一章的属性约简相对应，不过在这里我们采用的方法是逻辑的. 本子节所讨论的都是针对相容算法而言的.

设(P,Q)是一个相容算法，$a \in P$，在(P,Q)算法中，当且仅当算法$(P-\{a\},Q)$相容时，称属性a在(P,Q)中是可省的，否则a是不可省的.

如果P中所有属性都是不可省的，则称算法(P,Q)是独立的. 对于$R \subseteq P$，当算法(R,Q)是独立且相容的时，称(R,Q)为(P,Q)是一个约简. 算法的约简是去掉不必要的条件属性，是对知识表达空间维数进行简化.

算法(P,Q)中所有不可省的属性的集合称为算法(P,Q)的核，记作$\mathrm{core}(P,Q)$.

事实 6.13　$\mathrm{core}(P,Q) = \bigcap \mathrm{red}(P,Q)$. 这里$\mathrm{red}(P,Q)$表示算法$(P,Q)$的所有约简的集合.

例 6.6　考虑表 6.5 所示的信息系统

表 6.5

U	a	b	c	d	e
1	1	0	2	1	1
2	2	1	0	1	0
3	2	1	2	0	2
4	1	2	2	1	1
5	1	2	0	0	2

设$P=\{a,b,c\}$，$Q=\{d,e\}$分别是条件属性集和决策属性集，PQ决策算法如下：

$$a_1b_0c_2 \to d_1e_1,$$
$$a_2b_1c_0 \to d_1e_0,$$
$$a_2b_1c_2 \to d_0e_2,$$
$$a_1b_2c_2 \to d_1e_1,$$
$$a_1b_2c_0 \to d_0e_2.$$

从P中去掉属性a，得到算法$(P-\{a\},Q)$如下：

$$b_0c_2 \to d_1e_1,$$
$$b_1c_0 \to d_1e_0,$$
$$b_1c_2 \to d_0e_2,$$
$$b_2c_2 \to d_1e_1,$$
$$b_2c_0 \to d_0e_2.$$

算法$(P-\{a\},Q)$显然是相容的.

从 P 中去掉 b,得到算法 $(P-\{b\},Q)$ 如下:

$$a_1c_2\to d_1e_1,$$
$$a_2c_0\to d_1e_0,$$
$$a_2c_2\to d_0e_2,$$
$$a_1c_2\to d_1e_1,$$
$$a_1c_0\to d_0e_2.$$

算法 $(P-\{b\},Q)$ 是不相容的.

从 P 中去掉 c,得到算法 $(P-\{c\},Q)$ 如下:

$$a_1b_0\to d_1e_1,$$
$$a_2b_1\to d_1e_0,$$
$$a_2b_1\to d_0e_2,$$
$$a_1b_2\to d_1e_1,$$
$$a_1b_2\to d_0e_2.$$

因为规则 $a_2b_1\to d_1e_0$ 和 $a_2b_1\to d_0e_2$ 不相容,所以 $(P-\{c\},Q)$ 算法不相容. 可见 a,b 是可省的(注意它们不是同时可省的,即 $\{a,b\}$ 不是可省的),c 是不可省的,故 $\{c\}$ 是 $P=\{a,b,c\}$ 的核,算法 (P,Q) 有两个约简 $(\{b,c\},Q)$ 和 $(\{a,c\},Q)$.

6.2.9 不相容算法的约简

对于知识表达中不相容的情况,也可以类似于相容的情况那样进行算法约简,即考察去掉某些属性后其正区域是否发生变化,以判定该属性是否是可省的. 若 (P,Q) 为不相容算法,$a\in P$,当 $\mathrm{pos}(P,Q)=\mathrm{pos}(P-\{a\},Q)$ 时,(P,Q) 算法中属性 a 是可省的,否则称为是不可省的.

当 (P,Q) 算法中所有 $a\in P$ 均是不可省的时,则称 (P,Q) 算法是独立的. 对 $R\subseteq P$,(R,Q) 算法独立且 $\mathrm{pos}(P,Q)=\mathrm{pos}(R,Q)$ 时,称算法 (R,Q) 为 (P,Q) 的一个约简. (P,Q) 的条件属性集的所有约简的交称为 (P,Q) 的核,记作 $\mathrm{core}(P,Q)$.

例 6.7 考虑如下的信息系统(表 6.6)

表 6.6

U	a	b	c	d
1	1	0	2	1
2	1	2	2	1
3	2	1	0	1
4	2	1	2	0
5	1	2	0	0

设 $R=\{a,b\}$，$T=\{c,d\}$分别是条件属性集和决策属性集，则 RT 算法如下：

$$a_1b_0\to c_2d_1,$$
$$a_1b_2\to c_2d_1,$$
$$a_2b_1\to c_0d_1,$$
$$a_2b_1\to c_2d_0,$$
$$a_1b_2\to c_0d_0.$$

在这个算法中，规则 $a_1b_2\to c_2d_1$，$a_1b_2\to c_0d_0$ 不相容，$a_2b_1\to c_0d_1$，$a_2b_1\to c_2d_0$ 也不相容，故 $\text{pos}(R,T)=\{a_1b_0\to c_2d_1\}=\{$能识别决策类 c_2d_1 的算法$\}$，从 R 中去掉属性 a，得算法$(R-\{a\},T)$如下：

$$b_0\to c_2d_1,$$
$$b_2\to c_2d_1,$$
$$b_1\to c_0d_1,$$
$$b_1\to c_2d_0,$$
$$b_2\to c_0d_0.$$

$\text{pos}(R-\{a\},T)=\{$能识别决策类 c_2d_1 的算法$\}$，故 a 是可省的.

容易验证当去掉属性 b 时，算法$(R-\{b\},T)$的相容度为 0，故 b 是不可省的. 故(R,T)只有一个约简，即$(R-\{a\},T)$.

6.2.10 决策规则的约简

决策规则的约简是利用决策逻辑分别消去决策算法中每一个决策规则的不必要条件，它不是整体上的简化属性，而是针对每一个决策规则，去掉表达该规则的冗余属性值，以便进一步简化决策算法.

若 φ 是一个 P 基本公式，$Q\subseteq P$，我们以 φ/Q 表示从 φ 中移去所有 $(a,v_a)(a\in P-Q)$剩余的 Q 基本公式.

设 $\varphi\to\psi$ 是一个PQ 规则，$a\in P$，我们称 a 在该规则中是冗余的，当且仅当：$\models_S\varphi\to\psi$ 蕴涵$\models_S\varphi/(P-\{a\})\to\psi$. 否则，$a$ 在规则 $\varphi\to\psi$ 中称为是必要的. 如果对任一 $a\in P$ 在该PQ 规则中都是必要的，则称该规则是独立的.

P 的一个子集R 称为PQ 规则$\varphi\to\psi$ 的一个约简，如果 $\varphi/R\to\psi$ 是独立的且$\models_S\varphi\to\psi$ 蕴涵$\models_S\varphi/R\to\psi$. 此时 R 称为是既约的.

规则 $\varphi\to\psi$ 的所有必要属性所成的集合，称为该规则的核，记作 $\text{core}(\varphi\to\psi)$.

事实 6.14 $\text{core}(\varphi\to\psi)=\bigcap\text{red}(\varphi\to\psi)$. 这里 $\text{red}(\varphi\to\psi)$表示规则 $\varphi\to\psi$ 的所有约简的集合.

一个决策算法的约简包含两部分,一部分是在整个算法中去掉所有可省略的条件属性,另一部分是对算法中的每一个规则进行约简.这两个部分的顺序没有特别的要求,先进行哪部分都行.下面我们通过例子来说明先进行规则约简再进行整体条件属性约简的方法,另一种方法也大同小异.

例 6.8 考虑由表 6.7 所给定的信息系统

表 6.7

U	a	b	c	d	e
1	1	0	2	1	1
2	2	1	0	1	0
3	2	1	2	0	2
4	1	2	2	1	1
5	1	2	0	0	2

设 $P=\{a,b,c\}$,$Q=\{d,e\}$分别为条件属性和决策属性,PQ 算法如下:

$$a_1b_0c_2\rightarrow d_1e_1,$$

$$a_2b_1c_0\rightarrow d_1e_0,$$

$$a_2b_1c_2\rightarrow d_0e_2,$$

$$a_1b_2c_2\rightarrow d_1e_1,$$

$$a_1b_2c_0\rightarrow d_0e_2.$$

我们先消去算法中每一规则的不必要条件.

对第一条规则 $a_1b_0c_2\rightarrow d_1e_1$;去掉 a_1:$b_0c_2\rightarrow d_1e_1$;去掉 b_0:$a_1c_2\rightarrow d_1e_1$;去掉 c_2:$a_1b_0\rightarrow d_1e_1$;注意到它们在系统中都是真的,故该规则的核是空的,a 或 b 或 c 都是该规则中可省略的属性.注意到 $a_1b_0\rightarrow d_1e_1$ 和 $b_0c_2\rightarrow d_1e_1$ 都不是独立的,它们还可以进一步约简为 $b_0\rightarrow d_1e_1$.故原算法的第一条规则有两个约简,即 $b_0\rightarrow d_1e_1$ 和 $a_1c_2\rightarrow d_1e_1$.

对于 PQ 算法中的第二条规则 $a_2b_1c_0\rightarrow d_1e_0$,去掉 a_2:$b_1c_0\rightarrow d_1e_0$;去掉 b_1:$a_2c_0\rightarrow d_1e_0$;去掉 c_0:$a_2b_1\rightarrow d_1e_0$.因为 $a_2c_0\rightarrow d_1e_0$,$b_1c_0\rightarrow d_1e_0$ 均为真,所以 b 或 a 在该规则中是可省略的.但 $a_2b_1\rightarrow d_1e_0$ 在该系统中是不相容的,故属性 c 是不可省略的.所以该规则有两个约简,即 $b_1c_0\rightarrow d_1e_0$ 和 $a_2c_0\rightarrow d_1e_0$.

类似地,我们可以讨论剩余的规则,得到各规则的核及各规则的约简,它们分别表示在表 6.8 和表 6.9 中.

表 6.8

U	a	b	c	d	e
1	*	*	*	1	1
2	*	*	0	1	0
3	*	*	2	0	2
4	*	*	2	1	1
5	*	*	0	0	2

由表 6.9 可见,该算法的每条规则都有两个简化形式.为使算法简化,我们对每一条规则的两个简化形式取其一构成原算法的一个约简算法.注意到 1′和 4 所表示的规则是一个,所以构成原算法的约简的可能性共有 16 个之多,但无论哪一个都已是最简算法,比如我们可以取其一为

$$b_0 \to d_1e_1,$$
$$b_1c_0 \to d_1e_0,$$
$$a_2c_2 \to d_0e_2,$$
$$b_2c_2 \to d_1e_1,$$
$$a_1c_0 \to d_0e_2.$$

为方便,还可以将该算法合并为

$$b_0 \vee b_2c_2 \to d_1e_1,$$
$$b_1c_0 \to d_1e_0,$$
$$a_2c_2 \vee a_1c_0 \to d_0e_2.$$

表 6.9

U	a	b	c	d	e
1	*	0	*	1	1
1′	1	*	2	1	1
2	2	*	0	1	0
2′	*	1	0	1	0
3	2	*	2	0	2
3′	*	1	2	0	2
4	1	*	2	1	1
4′	*	2	2	1	1
5	1	*	0	0	2
5′	*	2	0	0	2

6.2.11 决策算法的极小化

在本节的最后一部分,我们考虑是否一个决策算法中所有的决策规则都

是必要的呢？确切地说，我们的目标是对每一个决策类删除多余的决策规则．明显地，有时删除一些决策规则并不影响决策的制定，因为被去掉的那些规则的作用可以由留下来的规则来完成．

我们将这一想法形式化：设 $S=(U,A)$ 是一个信息系统，F 是一个基本算法，以 F_ψ 记算法 F 中所有具有后件 ψ 的基本规则所成的集合；以 P_ψ 记 F_ψ 中所有规则的前件所成的集合．

算法 F 中的一个规则 $\varphi\to\psi$ 是多余的，如果 $\models_S \vee P_\psi = \vee(P_\psi-\{\varphi\})$，否则称该规则不是多余的．如果 F_ψ 中的所有规则都不是多余的，则称 F_ψ 是独立的．

F_ψ 的一个子集 F'_ψ 称为是 F_ψ 的一个约简，如果 F'_ψ 是独立的，且 $\models_S \vee P_\psi = \vee P'_\psi$．一个决策规则集 F_ψ 称为是既约的，如果它是自己的一个约简．

一个基本算法 F 称为是极小的，如果其中的每一个规则是既约的且对每一个决策规则 $\varphi\to\psi$，F_ψ 是既约的．

因此，为了简化一个 PQ 算法，我们首先必须进行整体属性约简，再对每一个规则进行规则约简，最后从算法中去掉多余的规则，下面我们通过一个例子来说明．

例 6.9 设有如下一个信息系统

表 6.10

U	a	b	c	d	e
1	1	0	0	1	1
2	1	0	0	0	1
3	0	0	0	0	0
4	1	1	0	1	0
5	1	1	0	2	2
6	2	2	0	2	2
7	2	2	2	2	2

表 6.11

U	a	b	d	e
1	1	0	1	1
2	1	0	0	1
3	0	0	0	0
4	1	1	1	0
5	1	1	2	2
6	2	2	2	2
7	2	2	2	2

设 $P=\{a,b,c,d\}$，$Q=\{e\}$分别是条件属性和决策属性. 显然 PQ 算法是相容的. 我们从 P 中去掉属性 c 后，可得一个$(P-\{c\})Q$ 算法，该算法也是相容的，故条件属性 c 是冗余的，所以我们有一个等价的决策信息系统.

接下来，我们计算各规则的约简. 先计算各规则的核值. 对$(P-\{c\})Q$ 算法中的第一条规则：$a_1b_0d_1 \to e_1$，

去掉 a_1 得 $b_0d_1 \to e_1$ 相容，故 a_1 是多余的.

去掉 b_0 得 $a_1d_1 \to e_1$ 不相容，故 b_0 是核值.

去掉 d_1 得 $a_1b_0 \to e_1$ 相容，故 d_1 是多余的.

表 6.12

U	a	b	d	e
1	*	0	*	1
2	1	*	*	1
3	0	*	*	0
4	*	1	1	0
5	*	*	2	2
6	*	*	*	2
7	*	*	*	2

表 6.13

U	a	b	d	e
1	1	0	*	1
1′	*	0	1	1
2	1	0	*	1
2′	1	*	0	1
3	0	*	*	0
4	*	1	1	0
5	*	*	2	2
6	2	*	*	2
6′	*	2	*	2
6″	*	*	2	2
7	2	*	*	2
7′	*	2	*	2
7″	*	*	2	2

注意这里规则 $b_0d_1 \to e_1$ 和 $a_1b_0 \to e_1$ 已是规则 $a_1b_0d_1 \to e_1$ 的最简形式了.

依次对每一个规则求其核值并获得所有最简形式. 它们被表达在如下的

表 6.12 和表 6.13 中.

最后,我们进行多余规则的删除.

对 $P_{e_1}=\{a_1b_0,b_0d_1,a_1d_0\}$, $P'_{e_1}=\{a_1b_0\}$, $P''_{e_1}=\{b_0d_1,a_1d_0\}$,我们有 $\models_S \vee P_{e_1}=\vee P'_{e_1}$ 及 $\models_s \vee P_{e_1}\equiv\vee P''_{e_1}$ 且 $a_1b_0\to e_1$, $b_0d_1\to e_1$, $a_1d_0\to e_1$ 均是既约规则,故 $\{a_1b_0\to e_1\}$ 或 $\{b_0d_1\vee a_1d_0\to e_1\}$ 即为算法 $\{a_1b_0\to e_1,b_0d_1\to e_1,a_1d_0\to e_1\}$ 的约简.

而对 $P_{e_0}=\{a_0,b_1d_1\}$ 其本身即为既约的.

算法 $\{a_0\to e_0,b_1d_1\to e_0\}$ 的既约算法可写为 $\{a_0\vee b_1d_1\to e_0\}$.

对 $P_{e_2}=\{d_2,a_2,b_2\}$,只有子集 $P'_{e_2}=\{d_2\}$ 满足 $\models_S \vee P_{e_2}\equiv\vee P'_{e_2}$,故算法 $\{d_2\to e_2,a_2\to e_2,b_2\to e_2\}$ 只有一个既约算法即 $\{d_2\to e_2\}$. 故原 PQ 算法一共有两个极小算法:即

$$
\begin{aligned}
&a_1b_0\to e_1,\\
&a_0\vee b_1d_1\to e_0,\\
&d_2\to e_2,
\end{aligned}
$$

或者

$$
\begin{aligned}
&b_0d_1\vee a_1d_0\to e_1,\\
&a_0\vee b_1d_1\to e_0,\\
&d_2\to e_2.
\end{aligned}
$$

§6.3 基于不完备信息系统的模态逻辑

在上两节中,我们详细地讨论了基于完备信息系统的逻辑及其在决策中的应用,但正如我们将要在第十一章中看到的那样,在实际问题中,我们所遇到的并非都是完备信息系统. 本节中,我们将考虑基于不完备信息的知识表示的逻辑基础. 这里所说的不完备信息是指信息系统的信息函数不是完全确定的情况. 通过定义两个模态运算[ ]和〈 〉,我们定义基于不完备信息的模态逻辑,然后讨论该逻辑系统与 Brouwersche 系统的关系. 最后,简要介绍它在决策逻辑中的应用.

在§6.2 中,我们已经有了信息系统的概念,鉴于它与本节中将要介绍的不完备信息系统的区别,我们称之为完备信息系统. 下面我们来说明如何将完备信息系统模型拓展为一个描述不完备信息的模型的不完备信息系统.

设 $S=(U,A,\mathrm{val},f)$ 是一个完备信息系统. 我们知道,这里的 f 是一个 $U\times A$ 到 val 的一个映射,称之为信息函数. 我们将 f 拓展为一个 $U\times A$ 到 2^{val} 的一个映射,这样获得的模型我们称之为一个不完备信息系统. 这样以来,

一个不完备信息系统 $S=(U,A,\mathrm{val},f)$，f 的函数值是一个集合，它就可以描述一个对象在某个属性下取一个或多个不确定属性值的情况. 当然，若对某个 $x\in U, a\in A, f(x,a)=\varnothing$，则表示对象 x 在属性 a 下不取值，在以二维表表达的不完备信息系统中，我们常省略集合符号，空集表示为 * . 如表 6.14.

表 6.14

U	a	b	c
x_1	1	0	1
x_2	1	*	1
x_3	2	1	2
x_4	1,2	0	1

在后面我们将要提出的逻辑的语义中，要用到所谓的不完备信息系统的完备化的概念，设 K 是一个不完备信息系统，我们依据 K 可以构造完备信息系统 K'，即令 K 的论域 U 和属性集 A 及 val 保持不变，对 $\forall x\in U, a\in A$，$f'(x,a)$ 取 $f(x,a)$ 是某一特定的元，当 $f(x,a)=\varnothing$ 时，我们令 $f'(x,a)=e$ 表示一个超常元作为对象 x 在属性 a 下的取值. 这样得到一个完备信息系统 $K'=(U,A,\mathrm{val},f')$，它称为信息系统 K 的一个完备化. 显然，一般地，一个不完备信息系统，通常有多个完备化. 例如表 6.15 即为表 6.14 所描述的不完备信息系统的一个完备化.

表 6.15

U	a	b	c
x_1	1	0	1
x_2	1	*	1
x_3	2	1	2
x_4	1	0	1

若 K' 是不完备信息系统 K 的完备化，则记作 $K'\geqslant K$. 完备化意味着随着新的信息的增加，我们的知识变得更为精确.

我们先定义基于不完备信息系统逻辑(INCRL)的语法. INCRL 的符号集包含如下四类：

(1)命题符号：$p,q,\cdots,p_1,p_2,\cdots$.

(2)命题运算符：$\neg$(非)，$\vee$(析取)，$\wedge$(合取)，$\rightarrow$(蕴涵)，$\equiv$(等价).

(3)模态运算符：[ ](必然)，$\langle\rangle$(可能).

(4)括号:(,).

INCRL 的合式公式按如下的方式归纳地定义:

(1)命题符号是合式公式.

(2)若 A,B 是合式公式,则$(\neg A)$,$(A\wedge B)$,$(A\vee B)$,$(A\rightarrow B)$,$(A\equiv B)$是合式公式.

(3)若 A 是合式公式,则$([\,]A)$,$(\langle\rangle A)$是合式公式.

(4)合式公式仅限于有限次使用(1)~(3)所成的符号串.

在不致引起误解的情况下,合式公式中的括号是可省略的,一般地,我们使用 $A,B,F,G\cdots$来表示合式公式.

在 INCHL 中,一个模型 M 是指一个不完备信息系统 $K=(U,A,\mathrm{val},f)$ 附带一个赋值函数 v_K,这里 v_K 按下面的方式归纳地定义:

对一个命题符号 p,$v_K(p)$是 U 的一个子集,对任意合式公式,

$$v_K(\neg F)=-\,v_K(F),$$
$$v_K(F\wedge G)=v_K(F)\cap v_K(G),$$
$$v_K(F\vee G)=v_K(F)\cup v_K(G),$$
$$v_K(F\rightarrow G)=v_K(\neg F\vee G),$$
$$v_K(F\equiv G)=v_K((F\rightarrow G)\wedge(G\rightarrow F)),$$
$$v_K([\,]F)=\bigcap_{K'\geqslant K}\underline{R(K')}v_K(F),$$
$$v_K(\langle\rangle F)=\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F).$$

在上面的定义中,K'是K 的完备化.$R(K')$表示完备化 K'所诱导的等价关系.

给定一个模型 $M=(K,v_K)$,称公式 F 在模型M 中被对象x 满足当且仅当 $x\in v_K(F)$,记作 $M,x\models F$.

我们定义 $\mathrm{ext}_M F=\{x\in U\mid M,x\models F\}$.

设 Γ 是一个公式集,M 是一个模型,$x\in U$,称 Γ 在M 中被x 满足当且仅当对任意的$F\in\Gamma$ 有 $M,x\models F$,在不致发生混淆的情况下 $x\in\mathrm{ext}_M F$ 可缩写为$x\in F$.

一个公式 F 被称为在M 中是真的(记作$\models_M F$)当且仅当 $\mathrm{ext}_M F=U$.

一个公式 F 被称为有效的(记作$\models F$),如果它在所有的模型中都是真的.

下面我们给出 INCRL 的一些性质.

事实 6.15

(1) $\models[\,]F\rightarrow F$,$\models F\rightarrow\langle\rangle F$.

(2) 如果$\models F\rightarrow G$,则$\models[\,]F\rightarrow[\,]G$ 且$\models\langle\rangle F\rightarrow\langle\rangle G$.

(3) $\models[\,]F\equiv\neg\langle\rangle\neg F, \models\langle\rangle F\equiv\neg[\,]\neg F$.

(4) $\models\langle\rangle(F\wedge G)\rightarrow\langle\rangle F\wedge\langle\rangle G$.

(5) $\models\langle\rangle(F\vee G)\equiv\langle\rangle F\vee\langle\rangle G$.

(6) $\models[\,](F\wedge G)\equiv[\,]F\wedge[\,]G$.

(7) $\models[\,]F\vee[\,]G\rightarrow[\,](F\vee G)$.

(8) $\models F\rightarrow[\,]\langle\rangle F$.

证 设 $M=(K,v_K)$是任一模型.

(1)因为 $v_K([\,]F)=\bigcap_{K'\geqslant K}\underline{R(K')}v_K(F)$,而对 K 的每一个完备化K',有$\underline{R(K')}v_K(F)\subseteq v_K(F)$,故有 $v_K([\,]F)=\bigcap_{K'\geqslant K}\underline{R(K')}v_K(F)\subseteq v_K(F)$. 所以$\models[\,]F\rightarrow F$. 利用$\overline{R(K')}v_K(F)\supseteq v_K(F)$即可证明$\models F\rightarrow\langle\rangle F$.

(2)设 K_i 是K 的任一个完备化. $\models F\rightarrow G$ 意味着 $v_K(F)\subseteq v_K(G)$. 由上下近似的单调性,有$\underline{R(K_i)}v_K(F)\subseteq\underline{R(K_i)}v_K(G)$和$\overline{R(K_i)}v_K(F)\subseteq\overline{R(K_i)}\cdot v_K(G)$. 由 K_i 的任意性,有$\bigcap_{K'\geqslant K}\underline{R(K')}v_K(F)\subseteq\bigcap_{K'\geqslant K}\underline{R(K')}v_K(G)$和$\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F)\subseteq\bigcup_{K'\geqslant K}\overline{R(K')}v_K(G)$,即 $v_K([\,]F)\subseteq v_K([\,]G)$和$v_K(\langle\rangle F)\subseteq v_K(\langle\rangle G)$所以$\models[\,]F\rightarrow[\,]G, \models\langle\rangle F\rightarrow\langle\rangle G$.

(3)由上、下近似的性质,$v_K(\neg\langle\rangle\neg F)=-v_K(\langle\rangle\neg F)=-(\bigcup_{K'\geqslant K}\overline{R(K')}v_K(\neg F))=\bigcap_{K'\geqslant K}\underline{R(K')}v_K(F)=v_K([\,]F)$故$\models[\,]F\equiv\neg\langle\rangle\neg F$. 同理可证$\models\langle\rangle F\equiv\neg[\,]\neg F$.

(4)
$$\begin{aligned}v_K(\langle\rangle(F\wedge G))&=\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F\wedge G)\\&=\bigcup_{K'\geqslant K}\overline{R(K')}(v_K(F)\cap v_K(G))\\&\subseteq(\bigcup_{K'\geqslant K}(\overline{R(K')}v_K(F)))\\&\quad\cap(\bigcup_{K'\geqslant K}(\overline{R(K')}v_K(G))),\end{aligned}$$

所以

$$\models\langle\rangle(F\wedge G)\rightarrow\langle\rangle F\wedge\langle\rangle G.$$

(5)
$$\begin{aligned}v_K(\langle\rangle F\vee G)&=\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F\vee G)\\&=\bigcup_{K'\geqslant K}\overline{R(K')}(v_K(F)\cup v_K(G))\\&=\bigcup_{K'\geqslant K}(\overline{R(K')}v_K(F)\cup\overline{R(K')}v_K(G))\\&=(\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F))\cup(\bigcup_{K'\geqslant K}\overline{R(K')}v_K(G))\\&=v_K(\langle\rangle F)\cup v_K(\langle\rangle G)\\&=v_K(\langle\rangle F\vee\langle\rangle G),\end{aligned}$$

所以

$$\models\langle\rangle(F\vee G)\equiv\langle\rangle F\vee\langle\rangle G.$$

(6)由(3),(5)可得.

(7)由(3),(4)可得.

(8)对 K 的任一完备化K',有$\overline{R(K')}v_K(F)\supseteq v_K(F)$故$\bigcup_{K'\geq K}\overline{R(K')}\cdot v_K(F)\supseteq v_K(F)$.而对 K 的任一完备化K'',由上、下近似的性质

$$\begin{aligned}\underline{R(K'')}(\bigcup_{K'\geq K}\overline{R(K')}v_K(F))&\supseteq\bigcup_{K'\geq K}\underline{R(K'')}\ \overline{R(K')}v_K(F)\\&\supseteq\underline{R(K'')}\ \overline{R(K'')}v_K(F)\\&=\overline{R(K'')}v_K(F)\\&\supseteq v_K(F),\end{aligned}$$

由 K''的任意性,有$\bigcap_{K''\geq K}\underline{R(K'')}(\bigcup_{K'\geq K}\overline{R(K')}v_K(F))\supseteq v_K(F)$.故$\models F\to[\,]\langle\rangle F$.

众所周知,模态逻辑是由 Kripke 模型(W,R,V)刻画的.W 上的二元关系R 的性质刻画了系统的模态算子的性质,从此观点来看,事实 6.15(1)刻画了自反性而(8)刻画了对称性,而刻画传递性的$\langle\rangle\langle\rangle F\to\langle\rangle F$ 在 INCRL 中却不成立.下面的例子可以说明这一点.

例 6.10 设不完备信息系统 K 由表 6.16 给定.

令 $v_K(p)=\{x_3\}$,则 $x_1\in\bigcup_{K'\geq K}\overline{R(K')}\bigcup_{K''\geq K}\overline{R(K'')}v_K(p)=\{x_1,x_2,x_3,x_4\}=U$.但 $x_1\notin\bigcup_{K'\geq K}\overline{R(K')}v_K(p)=v_K(\langle\rangle p)=\{x_2,x_3,x_4\}$.所以$\models\langle\rangle\langle\rangle F\to\langle\rangle F$在 INCRL 中不成立.因此 INCRL 可以被看作是一种 KTB 模态逻辑.

表 6.16

U	a	b
x_1	1	0
x_2	1,2	0
x_3	2,3	0
x_4	1,2	0

下面讨论 INCRL 与 Brouwersche 系统即 KTB 系统的关系.我们将证明 KTB 系统的定理在 INCRL 中都是有效的.KTB 系统由下面的公理和推理规则组成:

(Ax1)经典的命题逻辑的公理.

(Ax2)

(i) $[\,](A\to B)\to([\,]A\to[\,]B)$.

(ii) $[\,]A\to A$.

(iii) $A\to[\,]\langle\rangle A$.

推理规则:

(i)(假言推理) 由 $A,A\to B$,得 B.

(ii)(推广规则) 由 A,得$[\,]A$.

事实 6.16 所有 KTB 的可证明公式在 INCRL 中都是有效的.

证 对(Ax1),即经典命题逻辑的公理 $A\to(B\to A)$,$(A\to(B\to C))\to((A\to B)\to(A\to C))$和$(\neg B\to\neg A)\to(A\to B)$.容易验证它们在 INCRL 中是有效的.

假言推理 A,$A\to B$ 得 B 在 INCRL 中是成立的.因为假定 A,$A\to B$ 有效,即对任意模型 $M=(K,v_K)$,$v_K(A)=U$,$v_K(A\to B)=U$,由 $v_K(A\to B)=-v_K(A)\cup v_K(B)=U$,立即得 $v_K(B)=U$,再由模型 M 的任意性,即知 B 有效,同理可证若 A 有效,则$[\,]A$ 有效.

对(Ax2)中(i)由事实 6.15 中的(1)得$\models[\,](A\to B)\to(A\to B)$.由事实 6.15 中的(2)得$\models(A\to B)\to([\,]A\to[\,]B)$.再由假言推理得$\models[\,](A\to B)\to([\,]A\to[\,]B)$.故(Ax2)的(i)成立.而(Ax2)中的(ii)即事实 6.15 中的(1),(Ax2)中的(iii)即事实 6.15 中的(8).

下面依据 INCRL 的模型(K,v_K)定义 Kripke 模型(W,R,V_K)使得我们能按标准的方式解释 INCRL.

我们把论域 U 当作 W,定义 W 上的二元关系 R 如下:

$(x_i,x_j)\in R\Leftrightarrow$存在 K 的一个完备化K',使得 x_i,x_j 在 $R(K')$下是不可区分的,即$(x_i,x_j)\in R(K')$或者说 x_i,x_j 属于信息系统K'的同一个基本集.

赋值函数 V_K 按标准的方式定义,即

$$V_K(p,x_i)=\begin{cases}1, & 若\ x_i\in v_K(p),\\ 0, & 若\ x_i\notin v_K(p).\end{cases}$$

事实 6.17 上面定义的 R 是自反、对称的.

证 由 R 的定义,事实 6.17. 是显然的. □

事实 6.18 上面定义的 R 不是传递的.

证 我们使用例 6.10 中的模型来说明事实 6.18.显然 K 有如下两个完备化K_1 和 K_2 分别如表 6.17 和表 6.18 所示.

表 6.17

U	a	b
x_1	1	0
x_2	1	0
x_3	3	0
x_4	2	0

表 6.18

U	a	b
x_1	1	0
x_2	2	0
x_3	2	0
x_4	1	0

按 R 的定义$(x_1,x_2)\in R,(x_2,x_3)\in R$.但 x_1,x_3 在 K 的任一个完备化中都是可区别的,即$(x_1,x_3)\notin R$.故 R 不是传递的. □

下面我们通过 R 按标准的方式定义 $V_K(\langle\rangle F,x)$和 $V_K([]F,x)$,即

$$V_K(\langle\rangle F,x) = \max_{\langle x,x_i\rangle\in R}(V_K(F,x_i)),$$

$$V_K([]F,x) = V_K(\neg\langle\rangle\neg F,x).$$

事实 6.19 一个合式公式 F 在 INCRL 中是有效的,当且仅当 F 在上述的 Kripke 模型中是有效的.

证 设 $M=(K,v_K)$是任一 INCRL 模型,当公式 F 不含模态算子时,事实 6.19 是显然的.对含有模态算子的公式 F 我们只需证明 $x\in v_K(\langle\rangle F)\Leftrightarrow V_K(\langle\rangle F,x)=1$.由 $v_K(\langle\rangle F)=\bigcup_{K'\geqslant K}\overline{R(K')}v_K(F)$,可知 $x\in v_K(\langle\rangle F)$.则至少存在一个 K 的完备化K'使得$x\in\overline{R(K')}v_K(F)$.由上、下近似的定义可知,存在 $x_1\in v_K(F)$,即 $V_K(F,x_1)=1$ 使得$(x,x_1)\in R(K')$.由 R 的定义可知,$(x,x_1)\in R$.故 $V_K(\langle\rangle F,x)=\max_{(x,x_i)\in R}(V_K(F,x_i))=1$.

反过来,若 $V_K(\langle\rangle F,x)=1$,则存在某个 $x_i\in W=U$ 使得$(x,x_i)\in R$ 且 $V_K(F,x_i)=1$.由 V_K 和R 的定义可知,$x\in v_K(\langle\rangle F)$. □

我们较详细地讨论过完备信息条件下的决策逻辑和决策,但是在不完备信息条件下,同样存在决策问题.下面我们将 INCRL 的思想基础即不完备信息系统的完备化用于不完备信息条件下的决策逻辑.我们首先推广上节中的特征公式,相容性等概念.

例如 K 是如表 6.19 所示的一个不完备信息系统.

表 6.19

U	a	b	c
x_1	1	0,1	1
x_2	1	*	1
x_3	2	1	2
x_4	1,2	0	1

设 $P=\{a,b\}$, $Q=\{c\}$ 分别是条件属性和决策属性，则称

$$((a,\{1\})\wedge(b,\{0,1\})\to(c,\{1\}))\wedge$$
$$((a,\{1\})\wedge(b,\varnothing)\to(c,\{1\}))\wedge$$
$$((a,\{2\})\wedge(b,\{1\})\to(c,\{2\}))\wedge$$
$$((a,\{1,2,\})\wedge(b,\{0\})\to(c,\{1\}))$$

为系统 K 的特征公式，我们记上面 K 的特征公式为 F，考虑表达式 $\mathscr{P}$，称为 K 的可能性特征公式(PCF)，这里的符号 $\mathscr{P}$ 意为特征公式 F 的可能性应该在 K 的每一个完备化中被考虑. 即 $\mathscr{P}(F)$ 被定义为如下的形式：

设 $F_{K'}$ 是 K 的一个完备化 K' 的特征公式，定义 $\mathscr{P}(F)=\vee_{K'\geqslant K}F_{K'}$.

对上面的例子

$$\begin{aligned}\mathscr{P}(F) &= \vee_{K'\geqslant K}F_{K'}\\ &=[((a,1)\wedge(b,0)\to(c,1))\wedge((a,1)\wedge(b,e)\to(c,1))\\ &\quad\wedge((a,2)\wedge(b,1)\to(c,2))\wedge((a,1)\wedge(b,0)\to(c,1))]\\ &\quad\vee[((a,1)\wedge(b,1)\to(c,1))\wedge((a,1)\wedge(b,e)\to(c,1))\\ &\quad\wedge((a,2)\wedge(b,1)\to(c,2))\wedge((a,1)\wedge(b,0)\to(c,1))]\\ &\quad\vee[((a,1)\wedge(b,0)\to(c,1))\wedge((a,1)\wedge(b,e)\to(c,1))\\ &\quad\wedge((a,2)\wedge(b,1)\to(c,2))\wedge((a,2)\wedge(b,0)\to(c,1))]\\ &\quad\vee\{((a,1)\wedge(b,1)\to(c,1))\wedge((a,1)\wedge(b,e)\to(c,1))\\ &\quad\wedge((a,2)\wedge(b,1)\to(c,2))\wedge((a,2)\wedge(b,0)\to(c,1))].\end{aligned}$$

我们称一个不完备信息系统 K 是可能相容的，如果 K 至少有一个完备化是相容的，K 称为是不可能相容的，如果 K 的每一个完备化都不相容.

事实 6.20　如果系统 K 的可能性特征公式 $\mathscr{P}(F)$ 在 K 中是真的，则系统 K 是可能相容的.

事实 6.20 是显然的，它意味着一个不完备系统 K 的可能相容性能保证我们依据 K 的条件属性相容地选择行为(决策). 当然，如果 $\mathscr{P}(F)$ 是假的，则不可能选择一个相容的行为(决策). 对偶地，我们可以定义一个算子 $\mathscr{A}=\neg\mathscr{P}\neg$，它意味着对应的公式表达了一个不完备信息系统的每一个完备化的相容性. 即设 F 是 K 的特征公式，$\mathscr{A}(F)$ 在 K 中为真表达的若 K 的每一个完备化都相容，否则当 K 至少有一个完备化不相容时 $\mathscr{A}(F)$ 在 K 中为假.

当然，在不完备信息系统情况下，对决策规则同样有相容性问题，以及决策规则及决策算法的约简问题. 但这首先需要对不完备信息系统的约简的进一步深入的研究，我们期待着这方面的结果.

第七章　变精度粗糙集模型

粗糙集理论的中心问题是分类分析. Pawlak 粗糙集模型的一个局限性是它所处理的分类必须是完全正确的或肯定的,因为它是严格按照等价类来分类的,因而它的分类是精确的,亦即"包含"或"不包含",而没有某种程度上的"包含"或"属于". Pawlak 粗糙集模型的另一个局限性是它所处理的对象是已知的且从模型中得到的所有结论仅仅适用于这些对象集. 但在实际应用中,往往需要将一些小规模的对象集中得到的结论应用到大规模的对象集中去.

本章介绍的变精度粗糙集模型(variable precision rough set model)是 Pawlak 粗糙集模型的扩充,它是在基本粗糙集模型的基础上引入了 $\beta(0\leqslant\beta<0.5)$,即允许一定程度的错误分类率存在. 这一方面完善了近似空间的概念,另一方面也有利于用粗糙集理论从认为不相关的数据中发现相关数据. 当然,变精度粗糙集模型的主要任务是解决属性间无函数或不确定关系的数据分类问题. 当 $\beta=0$ 时, Pawlak 粗糙集模型是变精度粗糙集模型的一个特例.

§7.1　多数包含关系

设 X 和 Y 表示有限论域 U 的非空子集. 如果对于每一个 $e\in X$ 有 $e\in Y$,则称 Y 包含 X,记作 $Y\supseteq X$. 令

$$c(X,Y)=\begin{cases}1-|X\cap Y|/|X|, & |X|>0,\\ 0, & |X|=0,\end{cases}\tag{7.1}$$

其中 $|X|$ 表示集合 X 的基数. 称 $c(X,Y)$ 为集合 X 关于集合 Y 的相对错误分类率. 即如果我们将集合 X 中的元素分到集合 Y 中,则做出分类错误的比例为 $c(X,Y)\times 100\%$,真正错分类的元素数目为 $c(X,Y)\times|X|$. 称 $c(X,Y)\times|X|$ 为绝对分类误差.

令 $0\leqslant\beta<0.5$,多数包含关系(majority inclusion relation)定义为

$$Y\overset{\beta}{\supseteq}X\Leftrightarrow c(X,Y)\leqslant\beta.\tag{7.2}$$

"多数"要求隐含着 X 与 Y 中的公共元素的数目大于 X 中元素数目的 50%.

显然, $Y\supseteq X$ 当且仅当 $c(X,Y)=0$.

例 7.1　令

$$X_1=\{x_1,x_2,x_3,x_4\},\qquad X_2=\{x_1,x_2,x_5\},$$

$$X_3 = \{x_4, x_6, x_7\}, \qquad Y = \{x_1, x_2, x_3, x_8\}.$$

按照 β 多数包含关系的定义，下列关系成立：

$$Y \overset{0.25}{\supseteq} X_1, \qquad Y \overset{0.33}{\supseteq} X_2.$$

但对每个 β，下式不成立：

$$Y \overset{\beta}{\supseteq} X_3.$$

从定义立即可以推出，如果 $\beta=0$，则多数包含关系就变成了标准的包含关系.

定理 7.1 如果 $A \cap B = \varnothing$，$B \overset{\beta}{\supseteq} X$，则 $A \overset{\beta}{\supseteq} X$ 不成立.

定理 7.2 如果 $\beta_1 < \beta_2$，则 $Y \overset{\beta_1}{\supseteq} X$ 蕴涵 $Y \overset{\beta_2}{\supseteq} X$.

§7.2 变精度粗糙集模型中的近似集

设 (U,R) 为近似空间，其中论域 U 为非空有限集合，R 为 U 上的等价关系，$U/R = \{E_1, E_2, \cdots E_n\}$ 为 R 的等价类或基本集构成的集合. 对于 $X \subseteq U$，定义 X 的 β 下近似为

$$\underline{R}_\beta X = \bigcup \{E \in U/R \mid X \overset{\beta}{\supseteq} E\}, \tag{7.3}$$

或者

$$\underline{R}_\beta X = \bigcup \{E \in U/R \mid c(E,X) \leqslant \beta\}. \tag{7.4}$$

$\underline{R}_\beta X$ 也称为 β 正区域，记为 $\mathrm{posr}_\beta(X)$.

定义 X 的 β 上近似为

$$\overline{R}_\beta X = \bigcup \{E \in U/R \mid c(E,X) < 1-\beta\}. \tag{7.5}$$

定义 X 的 β 边界域为

$$\mathrm{bnr}_\beta(X) = \bigcup \{E \in U/R \mid \beta < c(E,X) < 1-\beta\}. \tag{7.6}$$

定义 X 的 β 负区域为

$$\mathrm{negr}_\beta(X) = \bigcup \{E \in U/R \mid c(E,X) \geqslant 1-\beta\}. \tag{7.7}$$

X 的 β 正区域(或 X 的 β 下近似)可理解为将 U 中的对象以不大于 β 的分类误差分于 X 的集合. X 的 β 负区域相应理解为将 U 中的对象以不大于 β 的分类误差分于 X 的补集(即 $\sim X$)的集合. 后者的解释可由定理 7.3 给出.

定理 7.3 对于 $\forall X \subseteq U$，下列关系式成立.

$$\mathrm{posr}_\beta(\sim X) = \mathrm{negr}_\beta(X),$$

其中 $\sim X = U - X$.

X 的 β 边界域是由那些以不大于 β 的分类误差既不能分类于 X 又不能分类十 $\sim X$ 的 U 中对象所构成的集合.

如果 $\mathrm{bnr}_\beta(X)=\varnothing$,则

$$\mathrm{posr}_\beta(X)\cup \mathrm{negr}_\beta(X)=U.$$

X 的 β 上近似是由那些以不大于 β 的分类误差不能分类于 $\sim X$ 的 U 中对象所构成的集合.

将上面近似集的定义与 Pawlak 粗糙集模型相比较,人们会发现,如果 $\beta=0$,那么 Pawlak 粗糙集模型就变成了变精度粗糙集模型的特殊情况.这个事实可由下列定理来解释.

定理 7.4 设 $X\subseteq U$,则

(1) $\underline{R}_0X=\underline{R}X$,其中 $\underline{R}X$ 是由 $\underline{R}X=\bigcup\{E\in U/R\mid X\supseteq E\}$ 定义的下近似集.

(2) $\overline{R}_0X=\overline{R}X$,其中 $\overline{R}X$ 是由 $\overline{R}X=\bigcup\{E\in U/R\mid E\cap X\neq\varnothing\}$ 定义的上近似集.

(3) $\mathrm{bnr}_0X=\mathrm{bn}_RX$,其中 bn_RX 是由 $\mathrm{bn}_RX=\overline{R}X-\underline{R}X$ 定义的 X 的边界域.

(4) $\mathrm{negr}_0X=\mathrm{neg}_RX$,其中 neg_RX 是由 $\mathrm{neg}_RX=U-\overline{R}X$ 定义的负区域.

对于 $0\leqslant\beta<0.5$,下列定理成立.

定理 7.5

$$\underline{R}_\beta X\supseteq\underline{R}X,$$
$$\overline{R}X\supseteq\overline{R}_\beta X,$$
$$\mathrm{bn}_RX\supseteq\mathrm{bnr}_\beta X,$$
$$\mathrm{negr}_\beta X\supseteq\mathrm{neg}_RX.$$

易知,随着分类误差 β 的减少,X 的正域与负域将缩小,而边界域将扩大.反之,随着 β 的增大,X 的正域与负将扩大.而边界域将缩小.当 β 趋于极限 0.5 时,即 $\beta\to0.5$,近似集将趋于下列极限.

定理 7.6

$$\underline{R}_\beta X\to\underline{R}_{0.5}X=\bigcup\{E\in U/R\mid c(E,X)<0.5\},$$
$$\overline{R}_\beta X\to\overline{R}_{0.5}X=\bigcup\{E\in U/R\mid c(E,X)\leqslant0.5\},$$
$$\mathrm{bnr}_\beta X\to\mathrm{bnr}_{0.5}X=\bigcup\{E\in U/R\mid c(E,X)=0.5\},$$
$$\mathrm{negr}_\beta X\to\mathrm{negr}_{0.5}X=\bigcup\{E\in U/R\mid c(E,X)>0.5\}.$$

集合 $\mathrm{bn}r_{0.5}X$ 称为 X 的绝对边界,因为它包含在 X 的其他每一个边界域中.下列定理描述了集合 X 在 0.5 精度水平和更高精度水平计算得到的可辨别区域之间的主要关系.

定理 7.7

$$\mathrm{bnr}_{0.5}X=\bigcap_\beta\mathrm{bnr}_\beta X,$$

$$\overline{R}_{0.5}X = \bigcap_{\beta}\overline{R}_{\beta}X,$$
$$\underline{R}_{0.5}X = \bigcup_{\beta}\underline{R}_{\beta}X,$$
$$\mathrm{negr}_{0.5}X = \bigcup_{\beta}\mathrm{negr}_{\beta}X.$$

例 7.2　为了阐述扩充的近似集概念,我们考虑近似空间 $K=(U,R)$,其中 $U=\{x_1,x_2,\cdots,x_{20}\}$,等价关系 R 的等价类由下面给出:

$$E_1 = \{x_1,x_2,x_3,x_4,x_5\},$$
$$E_2 = \{x_6,x_7,x_8\},$$
$$E_3 = \{x_9,x_{10},x_{11},x_{12}\},$$
$$E_4 = \{x_{13},x_{14}\},$$
$$E_5 = \{x_{15},x_{16},x_{17},x_{18}\},$$
$$E_6 = \{x_{19},x_{20}\}.$$

对于两个精度水平:$\beta_1=0$ 和 $\beta_2=0.25$,我们计算集合 $X=\{x_4,x_5,x_8,x_{14},x_{16},x_{17},x_{18},x_{19},x_{20}\}$ 的近似.

如果假设分类误差 $\beta_1=0$,则由定理 7.4 知,集合 X 的 β 近似集等于标准近似集.因此,

$$\underline{R}_0X = E_6,$$
$$\overline{R}_0X = E_1 \cup E_2 \cup E_4 \cup E_5 \cup E_6,$$
$$\mathrm{bnr}_0X = E_1 \cup E_2 \cup E_4 \cup E_5,$$
$$\mathrm{negr}_0X = E_3.$$

如果 $\beta_2 = 0.25$,则有

$$\underline{R}_{0.25}X = E_5 \cup E_6,$$
$$\overline{R}_{0.25}X = E_1 \cup E_2 \cup E_4 \cup E_5 \cup E_6,$$
$$\mathrm{bnr}_{0.25}X = E_1 \cup E_2 \cup E_4,$$
$$\mathrm{negr}_{0.25}X = E_3.$$

为了描述集合 X 被近似空间 $A=(U,R)$ 中的基本集近似分类的程度,定义 β 精度($0\leqslant\beta<0.5$)为

$$\alpha(R,\beta,X) = |\underline{R}_{\beta}X| / |\overline{R}_{\beta}X|. \tag{7.8}$$

β 精度表示集合 X 相对于分类误差 β 的近似描述的不精确性.需要注意的是,随着 β 的增大,集合 X 上近似的基数将变小,下近似的基数将变大.这说明随着分类误差的增大,相对精度将增加.

§7.3　集合的相对可辨别性

集合边界的可辨别概念是相对的,如果允许一个大的分类误差,则在假定

的分类误差限内，集合 X 可能有较大的可辨别性.

如果集合 X 的 β 边界域 $\mathrm{bnr}_\beta(X)=\varnothing$，或者 $\underline{R}_\beta X=\overline{R}_\beta X$ 时，称集合 X 为 β 可辨别，否则称为 β 不可辨别.

对于 β 可辨别集，其相对精度 $\alpha(R,\beta,X)=1$. 集合的可辨别性依赖于 β 值. 一般地，下面性质成立.

定理 7.8 如果 X 在分类水平 $0\leqslant\beta<0.5$ 上是可辨别的，那么 X 在任何水平 $\beta_1>\beta$ 上也是可辨别的.

定理 7.9 如果 $\overline{R}_{0.5}X\neq\underline{R}_{0.5}X$，那么集合 X 在每一个分类误差水平 $0\leqslant\beta<0.5$ 上都是不可辨别的.

定理 7.9 强调，一个具有非空绝对边界的集合总是不可辨别的. 一般来说，我们容易证实下列定理.

定理 7.10 如果 X 在分类误差水平 $0\leqslant\beta<0.5$ 上是不可辨别的，那么 X 在任何水平 $\beta_2<\beta$ 上也是不可辨别的.

如果集合 X 对每个 β 是不可辨别的，则称集合 X 为绝对不可辨别的或绝对粗糙的. 集合 X 是绝对粗糙的当且仅当 $\mathrm{bnr}_{0.5}X\neq\varnothing$. 若粗糙集合不是绝对粗糙的，则称为相对粗糙的(或弱可辨别的). 对每一个相对粗糙集 X，存在一个分类误差水平 β 使得集合 X 在这个水平上是可辨别的.

令 $\mathrm{ndis}(R,X)=\{0\leqslant\beta<0.5 \mid \mathrm{bnr}_\beta(X)\neq\varnothing\}$，$\mathrm{ndis}(R,X)$ 是满足 X 是不可辨别的所有 β 值的全体. 满足 X 为可辨别的分类误差 β 最小值称为可辨别的阈值，这个阈值等于 $\mathrm{ndis}(R,X)$ 最小上界，即

$$\xi(R,X)=\sup\mathrm{ndis}(R,X). \tag{7.9}$$

定理 7.11 $\xi(R,X)=\max(m_1,m_2)$，其中

$$m_1=1-\min\{c(E,X)\mid E\in U/R, 0.5<c(E,X)\},$$

$$m_2=\max\{c(E,X)\mid E\in U/R, c(E,X)<0.5\}.$$

例 7.3 为了说明定理 **7.11**，我们假设基本类为

$$E_1=\{x_1,x_2,x_3,x_4,x_5\},$$
$$E_2=\{x_6,x_7,x_8\},$$
$$E_3=\{x_9,x_{10},x_{11},x_{12}\},$$
$$E_4=\{x_{13},x_{14},x_{15},x_{16}\},$$
$$E_5=\{x_{17},x_{18}\}.$$

令 $X=\{x_4,x_5,x_8,x_{14},x_{15},x_{16},x_{17},x_{18}\}$. X 的分类误差为

$$c(E_1,X)=0.6,$$
$$c(E_2,X)=0.66,$$
$$c(E_3,X)=1.0,$$

$$c(E_4, X) = 0.25,$$
$$c(E_5, X) = 0.0.$$

基于定理 7.11,容易证实满足 X 为可辨别的最小值为 $\xi(R, X) = 0.4$.

§7.4 β 近似的性质

这一节我们给出 β 近似的一些基本性质.

定理 7.12 对于每个 $0 \leqslant \beta < 0.5$,下列关系成立:

(1a) $X \stackrel{\beta}{\supseteq} \underline{R}_\beta X$.

(1b) $\overline{R}_\beta X \supseteq \underline{R}_\beta X$.

(2) $\underline{R}_\beta \varnothing = \overline{R}_\beta \varnothing = \varnothing$; $\underline{R}_\beta U = \overline{R}_\beta U = U$.

(3) $\overline{R}_\beta(X \cup Y) \supseteq \overline{R}_\beta X \cup \overline{R}_\beta Y$.

(4) $\underline{R}_\beta X \cap \underline{R}_\beta Y \supseteq \underline{R}_\beta(X \cap Y)$.

(5) $\underline{R}_\beta(X \cup Y) \supseteq \underline{R}_\beta X \cup \underline{R}_\beta Y$.

(6) $\overline{R}_\beta X \cap \overline{R}_\beta Y \supseteq \overline{R}_\beta(X \cap Y)$.

(7) $\underline{R}_\beta(\sim X) = \sim \overline{R}_\beta X$.

(8) $\overline{R}_\beta(\sim X) = \sim \underline{R}_\beta X$.

证 (1a)为了证明 $X \stackrel{\beta}{\supseteq} \underline{R}_\beta X$,我们只须证明对于任何两个基本集 E_1 和 E_2,如果 $c(E_1, X) \leqslant \beta$ 且 $c(E_2, X) \leqslant \beta$,那么

$$c(E_1 \cup E_2, X) \leqslant \beta.$$

从 $c(X, Y)$ 的定义我们有

(i) $c(E_1, X) = |E_1 \cap (\sim X)| / |E_1| = |E_1 \cap (\sim X)| / (|E_1 \cap (\sim X)| + |E_1 \cap X|) \leqslant \beta$.

(ii) $c(E_2, X) = |E_2 \cap (\sim X)| / |E_2| = |E_2 \cap (\sim X)| / (|E_2 \cap (\sim X)| + |E_2 \cap X|) \leqslant \beta$.

从(i)和(ii)可以推出

$$|E_1 \cap (\sim X)| + |E_2 \cap (\sim X)|$$
$$\leqslant \beta(|E_1 \cap (\sim X)| + |E_1 \cap X| + |E_2 \cap (\sim X)| + |E_2 \cap X|).$$

这意味着

$$c(E_1 \cup E_2, X) = |(E_1 \cup E_2) \cap (\sim X)| / |E_1 \cup E_2| \leqslant \beta.$$

(1b) $\overline{R}_\beta X \supseteq \underline{R}_\beta X$ 可以从上、下近似的定义直接推出.

(2) 因为对于任何基本集 E 有 $c(E, \varnothing) = 1$,因此推出 $\underline{R}_\beta \varnothing = \varnothing$ 且 $\overline{R}_\beta \varnothing = \varnothing$. 类似地,由 $c(E, U) = 0$ 可以推出 $\underline{R}_\beta(U) = U$ 和 $\overline{R}_\beta(U) = U$.

(3) 对于任何 $X,Y\subseteq U$,我们有 $c(E,X\cup Y)\leqslant c(E,X)$,且 $c(E,X\cup Y)\leqslant c(E,Y)$.因此 $\overline{R}_\beta(X\cup Y)\supseteq\overline{R}_\beta X\cup\overline{R}_\beta Y$.

(4) 对于任何 $X,Y\subseteq U$,我们有 $c(E,X\cap Y)\geqslant c(E,X)$且$c(E,X\cap Y)\geqslant c(E,Y)$,因此 $\underline{R}_\beta X\cap\underline{R}_\beta Y\supseteq\underline{R}_\beta(X\cap Y)$.

(5) 证明同(3).

(6) 证明同(4).

(7) 由性质 $c(E,\sim X)=1-c(E,X)$可以直接推出.

(8) 证明同(7). □

标准粗糙集模型与变精度粗糙集模型的一个根本区别是缺少包含 $\overline{R}_\beta X\overset{\beta}{\supseteq}X$.这是由于对任何正数 β,β 负域 $\mathrm{negr}_\beta X$ 一般与 X 相交.

§7.5 属性的近似依赖性

设 $S=(U,A,V,f)$是一个信息系统;$P,Q\subseteq A$ 为条件属性集和决策属性集.$\mathrm{ind}(P)$,$\mathrm{ind}(Q)$表示由 P,Q 决定的不可区分关系,关系 $\mathrm{ind}(P)$的等价类的集合称为条件类,用 U/P 表示,关系 $\mathrm{ind}(Q)$的等价类的集合称为决策类,用 U/Q 表示.

决策属性集 Q 与条件属性集 P 的 β 依赖性定义为

$$\gamma(P,Q,\beta)=|\mathrm{pos}(P,Q,\beta)|/|U|,\tag{7.10}$$

其中

$$\mathrm{pos}(P,Q,\beta)=\bigcup_{Y\in U/Q}\mathrm{ind}\underline{(P)}_\beta Y.\tag{7.11}$$

近似依赖性是粗糙依赖性思想的推广,当 $\beta=0$ 时,它就变成粗糙依赖性.我们知道,粗糙依赖性量度是对执行准确的对象分类整个能力的评价,而近似依赖性量度是对执行具有分类误差 β 的对象分类能力的评价.近似依赖性与粗糙依赖性不同,它不能解释为属性的函数或部分依赖.因为近似依赖性的性质是比函数依赖的性质更弱,例如,传递性不成立.

例 7.4 表 7.1 为一个信息系统.

为了阐述近似依赖的思想,对于不同的容许限 β 值,我们计算属性集 $P=\{a,b,c\}$和 $Q=\{d\}$之间的依赖性.

划分 U/P 有四个类:

$X_1=\{1,2,19,20,21\}$, $X_2=\{3\}$, $X_3=\{4\sim13\}$, $X_4=\{14\sim18\}$.

划分 U/Q 有三个类:

$Y_1=\{1\sim12\}$, $Y_2=\{13\sim17\}$, $Y_3=\{18\sim21\}$.

当 $\beta=0$ 时,有 $Y_1\supseteq X_2$,且

$$\gamma(P,Q,0)=|X_2|/|U|=1/21=0.047.$$

当 $\beta=0.1$ 时，有 $Y_1 \overset{0.1}{\supseteq} X_2$，$Y_1 \overset{0.1}{\supseteq} X_3$，且

$$\gamma(P,Q,0.1)=|X_2 \cup X_3|/|U|=11/21=0.52.$$

当 $\beta=0.2$ 时，有 $Y_1 \overset{0.2}{\supseteq} X_2$，$Y_1 \overset{0.2}{\supseteq} X_3$，$Y_2 \overset{0.2}{\supseteq} X_4$，且

$$\gamma(P,Q,0.2)=|X_2 \cup X_3 \cup X_4|/|U|=16/21=0.76.$$

表 7.1

U	a	b	c	d
1	1	0	0	1
2	1	0	0	1
3	1	1	1	1
4	0	1	1	1
5	0	1	1	1
6	0	1	1	1
7	0	1	1	1
8	0	1	1	1
9	0	1	1	1
10	0	1	1	1
11	0	1	1	1
12	0	1	1	1
13	0	1	1	2
14	1	1	0	2
15	1	1	0	2
16	1	1	0	2
17	1	1	0	2
18	1	1	0	3
19	1	0	0	3
20	1	0	0	3
21	1	0	0	3

§7.6 近似约简

属性约简是粗糙集模型中最重要的概念之一．所谓一个约简是指保持和决策属性 Q 的依赖性相同的最小条件属性子集．通过使用近似依赖的定义，我们来引入近似约简概念．

条件属性集 P 关于决策属性集 Q 的 β 约简或近似约简是 P 的一个子集

red (P,Q,β)，且满足：

(1) $\gamma(P,Q,\beta)=\gamma(\mathrm{red}(P,Q,\beta),Q,\beta)$.

(2) 从 $\mathrm{red}(P,Q,\beta)$中去掉任何一个属性，都将使(1)不成立.

从上述讨论可知，引入 $\beta(0\leqslant\beta<0.5)$参数后，扩充了基本粗糙集理论，更好地体现了数据分析中的数据相关性，从而为获取近似决策规则奠定了基础.

第八章　概率粗糙集模型

概率论是研究自然界，人类社会及技术过程中大量随机现象中的规律性的一门学科．概率，作为随机事件的一种度量，它反映了一种不确定性，因此它在不确定的推理中有重要的应用．

Pawlak 粗糙集模型是基于确定性知识库的，即它的近似空间是完全确定的，因此它忽视了可利用信息库的不确定性．若我们仍旧按 Pawlak 粗糙集模型来处理由随机产生的知识库的数据分析等问题就不能完全反映问题的实质．本章我们讨论概率粗糙集模型．

§8.1　有限论域上概率测度的基本知识

概率从产生之日起就有两种理解，一种理解为信任的程度，它反映了人们的经验和知识，常称为主观概率，另一种理解为(随机)事件在大量重复试验中试验结果出现的相对频率，因此常称为客观概率．即使是主观概率，也能反映或符合某种统计规律性，也具有客观性，因此，我们可以认为概率是对不确定的随机事件的一种客观的反映．

定义 8.1　设 U 是有限论域，集函数 $P:2^U\to[0,1]$ 称为概率测度，若

(1) $P(U)=1$.

(2) 当 $A\cap B=\varnothing$，有 $P(A\cup B)=P(A)+P(B)$.

容易验证，概率测度满足以下性质：

(1) $P(\sim A)=1-P(A)$，$P(\varnothing)=0$.

(2) $P(A\cup B)=P(A)+P(B)-P(A\cap B)$.

(3) $P(A\setminus B)=P(A)-P(A\cap B)$.

(4) $P(\bigcup_{i=1}^{n}A_i)=\sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}}(-1)^{|I|+1}P(\bigcap_{i\in I}A_i)$,

$$P(\bigcap_{i=1}^{n}A_i)=\sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}}(-1)^{|I|+1}P(\bigcup_{i\in I}A_i).$$

由于概率测度具有可加性，因此 U 上的概率测度完全被每一个对象 x_i 上的测度 $p_i=P(x_i)$，$1\leqslant i\leqslant n$，所决定，可表示为

$$X=\begin{Bmatrix}x_1,x_2,\cdots,x_n\\p_1,p_2,\cdots,p_n\end{Bmatrix}.$$

上式称为 U 上的一个概率分布,在信息论中常将 X 称为信息源.

若 P 是 U 上的概率测度,$A,B\subseteq U$ 且 $P(B)>0$,称

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}$$

为在事件 B 发生的条件下事件 A 发生的条件概率.

显然,$P(A\cap B)=P(A\mid B)P(B)$,并且可以用数学归纳法证明当 $P(\bigcap_{i=1}^{k-1}A_i)>0$时有以下乘法公式成立:

$$P(\bigcap_{i=1}^{k}A_i)=P(A_1)P(A_2\mid A_1)P(A_3\mid A_1\cap A_2)\cdots P(A_n\mid\bigcap_{i=1}^{k-1}A_i).$$

由概率可加性可以得到全概率公式:若$\{A_1,A_2,\cdots,A_k\}$是 U 的一个划分且$P(A_i)>0$,则

$$P(A)=\sum_{i=1}^{k}P(A\cap A_i)=\sum_{i=1}^{k}P(A\mid A_i)P(A_i).$$

由乘法公式和全概率公式可以得到 Bayes 公式:若$\{A_1,A_2,\cdots,A_k\}$是 U 的一个划分,则

$$P(A_i\mid A)=\frac{P(A_i)P(A\mid A_i)}{\sum_{j=1}^{k}P(A_j)P(A\mid A_j)}.$$

全概率公式是由原因到结果的计算公式,而 Bayes 公式是已知某种结果发生的情况下,寻求使这个结果发生的原因,因此 Bayes 公式也常称为后验概率公式.

§8.2 信息熵

设有某一随机试验(信息源),其每次试验的结果只能出现 $X_1,X_2,\cdots,X_n$ 中的一个,而且每次试验中每个 X_i 在结果中可能出现的概率 $p_i=P(X_i)$是事先知道的,即试验结果服从分布

$$X=\begin{Bmatrix}X_1,X_2,\cdots,X_n\\ p_1,\ p_2,\cdots,\ p_n\end{Bmatrix},\tag{8.1}$$

有时简记为 $X=(p_1,p_2,\cdots,p_n)$,但是在试验之前 X_i 在试验中是否一定出现是不知道的.

如何来理解和度量信息源 X 的不确定性呢?从直观上来看,若上述概率分布中某个 p_i 为 1,比如 $p_1=1,p_2=p_3=\cdots=p_n=0$,则每次试验后结果 X_1 一定出现,也就是说试验结果是完全确定的,这样,每次试验并没有提供任何信息量,所以在这种情况下信息源的不确定性的量应该是 0. 若 $p_1=p_2=p_3$

$=\cdots=p_n=\dfrac{1}{n}$，则每次试验结果哪一个 X_i 会出现是最不确定的，这时每次试验能提供最大的信息量，所以在这种情况下信息源的不确定性的量也应该是最大的．当然，对于一个给定的信息源，随着试验次数的增加，信息源的信息量会逐渐减少．基于上述理解，我们认为试验中如果提供的信息量越大则信息源的不确定性也越大，试验中如果提供的信息量越少则信息源的不确定性也越小．这样我们就采用 Shannon 的信息熵作为对信息源 X 的不确定性的度量．

定义 8.2　设 U 是论域，$X_1,X_2\cdots,X_n$ 是 U 的一个划分，其上有概率分布(8.1)，则称

$$H(X)=-\sum_{i=1}^{n}p_i\log p_i$$

为信息源 X 的信息熵，其中对数取以 2 为底，而当某个 p_i 为零时，则理解为 $0\cdot\log 0=0$.

$H(X)$具有以下性质：

(1)有界性：$0\leqslant H(X)\leqslant\log n$.

(2)确定性：当 X 是完全确定时，即存在 $X_i\subseteq U$ 使 $P(X_i)=1$，则 $H(X)=0$.

(3)最大性：当 $p_1=p_2=p_3=\cdots=p_n=\dfrac{1}{n}$ 时，$H(X)$达到最大值 $\log n$.

(4)连续性：$I(p_1,p_2,\cdots,p_n)=-\sum_{i=1}^{n}p_i\log p_i$ 是 p_i 的连续函数.

设 $Y=\begin{Bmatrix}Y_1,Y_2,\cdots,Y_m\\ q_1,\ q_2,\cdots,\ q_m\end{Bmatrix}$是另一个信息源，即 $Y_1,Y_2,\cdots Y_m$ 是 U 的另一划分，$P(Y_j)=q_j$，$\sum_{j=1}^{m}q_j=1$，则已知信息源 X 时信息源 Y 的条件熵 $H(Y|X)$定义为

$$H(Y\mid X)=\sum_{i=1}^{n}P(X_i)H(Y\mid X_i),$$

其中 $H(Y\mid X_i)=-\sum_{j=1}^{m}P(Y_j\mid X_i)\log P(Y_j\mid X_i)$ 为事件 X_i 发生时信息源 Y 的条件熵.

定理 8.1　设 Z 是X 和 Y 的联合分布，记为 $Z=X\cdot Y$，则下列关系成立：

(1) $H(Z)=H(X)+H(Y|X)=H(Y)+H(X|Y)$.

(2) $H(X)\geqslant H(X|Y)$，$H(Y)\geqslant H(Y|X)$.

上述不等式说明对信息源 X 的知道使得信息源 Y 的不确定性的量(信息量)变小而不是增大，称

$$I(X;Y) = H(X) - H(X \mid Y)$$

为关于信息源 X 和 Y 的互信息量，它反映了一个信息源从另一个信息源获取的信息量．正如下述定理所描述的那样，互信息量 $I(X;Y)$ 关于 X 和 Y 是对称的．

定理 8.2 关于互信息量，下列公式成立：

$$\begin{aligned} I(X;Y) &= H(X) - H(X \mid Y) \\ &= H(Y) - H(Y \mid X) = I(Y;X) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i \cap Y_j)\log\frac{P(Y_j \mid X_i)}{P(Y_j)} \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i \cap Y_j)\log\frac{P(X_i \mid Y_j)}{P(X_i)}. \end{aligned}$$

§8.3 概率粗糙集模型

Pawlak 粗糙集模型是基于可利用信息的完全性的，因而忽视了可利用信息的不完全性和可能存在的统计信息，这类模型对于不协调的决策表的规则提取往往显得无能为力．本节我们从概率论的观点出发来研究粗糙集理论，为研究不确定信息系统提供了新的粗糙集模型．

定义 8.3 设 U 是有限对象构成的论域，R 是 U 上的等价关系，其构成的等价类为

$$U/R = \{X_1, X_2, \cdots, X_n\}.$$

仍记 x 所在的等价类为 $[x]$，令 P 为定义在 U 的子集类构成的 σ 代数上的概率测度，三元组 $A_p = (U, R, P)$ 称为概率近似空间．U 中的每个子集称为概念，它代表了一个随机事件．$P(X \mid Y)$ 表示事件 Y 发生下 X 出现的条件概率，也可解释为随机选择的对象在概念 Y 的描述下属于 X 的概率．

设 $0 \leqslant \beta < \alpha \leqslant 1$，对于任意 $X \subseteq U$，我们定义 X 关于概率近似空间 $A_p = (U, R, P)$ 依参数 α, β 的概率(I)型下近似 $\underline{\mathrm{PI}}_\alpha(X)$ 和上近似 $\overline{\mathrm{PI}}_\beta(X)$ 如下：

$$\underline{\mathrm{PI}}_\alpha(X) = \{x \in U \mid P(X \mid [x]) \geqslant \alpha\},$$
$$\overline{\mathrm{PI}}_\beta(X) = \{x \in U \mid P(X \mid [x]) > \beta\}.$$

X 关于 A_p 依参数 α, β 的概率(I)型正域，边界和负域分别为

$$\begin{aligned} \mathrm{pos}(X, \alpha, \beta) &= \underline{\mathrm{PI}}_\alpha(X) = \{x \in U \mid P(X \mid [x]) \geqslant \alpha\}, \\ \mathrm{bn}(X, \alpha, \beta) &= \{x \in U \mid \beta < P(X \mid [x]) < \alpha\}, \\ \mathrm{neg}(X, \alpha, \beta) &= U \setminus \overline{\mathrm{PI}}_\beta = \{x \in U \mid P(X \mid [x]) \leqslant \beta\}. \end{aligned}$$

显然，上述定义等价于下列式子

$$\underline{\mathrm{PI}}_\alpha(X) = \mathrm{pos}(X,\alpha,\beta) = \bigcup\{[x] \mid P(X \mid [x]) \geqslant \alpha\},$$
$$\overline{\mathrm{PI}}_\beta(X) = \bigcup\{[x] \mid P(X \mid [x]) > \beta\},$$
$$\mathrm{bn}(X,\alpha,\beta) = \bigcup\{[x] \mid \beta < P(X \mid [x]) < \alpha\},$$
$$\mathrm{neg}(X,\alpha,\beta) = \bigcup\{[x] \mid P(X \mid [x]) \leqslant \beta\}.$$

X 关于 A_p 依参数 α,β 的概率(I)型正域，边界和负域构成了论域 U 的划分．而且显然有

$$\overline{\mathrm{PI}}_\beta(x) = \mathrm{pos}(X,\alpha,\beta) \cup \mathrm{bn}(X,\alpha,\beta).$$

或者

$$\mathrm{bn}(X,\alpha,\beta) = \mathrm{PI}_\beta(X) \setminus \mathrm{pos}(X,\alpha,\beta).$$

当 $\underline{PI}_\alpha(X)=\overline{\mathrm{PI}}_\beta(X)$ 时，或等价地当 $\mathrm{bn}(X,\alpha,\beta)=\varnothing$ 时，称 X 依参数 α,β 关于 A_p 是概率(I)型可定义，否则称 X 依参数 α,β 关于 A_p 是概率(I)型粗糙集．

定理 8.3　设 $0\leqslant\beta<\alpha\leqslant 1$，$X,Y\subseteq U$，则概率(I)型近似算子满足下列对偶性质：

(1) $\underline{\mathrm{PI}}_\alpha(\varnothing)=\overline{\mathrm{PI}}_\alpha(\varnothing)=\varnothing$，$\underline{\mathrm{PI}}_\alpha(U)=\overline{\mathrm{PI}}_\alpha(U)=U$．

(2) $\underline{\mathrm{PI}}_\alpha(X)\subseteq\overline{\mathrm{PI}}_\beta(X)$．

(3) $\underline{\mathrm{PI}}_\alpha(X)=\sim\overline{\mathrm{PI}}_{1-\alpha}(\sim X)$，$\overline{\mathrm{PI}}_\beta(X)=\sim\underline{PI}_{1-\beta}(\sim X)$．

(4) $\overline{\mathrm{PI}}_\beta(X\cup Y)\supseteq\overline{\mathrm{PI}}_\beta(X)\cup\overline{\mathrm{PI}}_\beta(Y)$．
　$\underline{\mathrm{PI}}_\alpha(X\cap Y)\subseteq\underline{\mathrm{PI}}_\alpha(X)\cap\underline{\mathrm{PI}}_\alpha(Y)$．

(5) $\overline{\mathrm{PI}}_\beta(X\cap Y)\subseteq\overline{\mathrm{PI}}_\beta(X)\cap\overline{\mathrm{PI}}_\beta(Y)$．
　$\underline{\mathrm{PI}}_\alpha(X\cup Y)\supseteq\underline{\mathrm{PI}}_\alpha(X)\cup\underline{\mathrm{PI}}_\alpha(Y)$．

(6) 若 $X\subseteq Y$，则

$$\underline{\mathrm{PI}}_\alpha(X)\subseteq\underline{\mathrm{PI}}_\alpha(Y),\overline{\mathrm{PI}}_\beta(X)\subseteq\overline{\mathrm{PI}}_\beta(Y).$$

(7) 若 $\alpha_1\leqslant\alpha_2$，$\beta_1\leqslant\beta_2$ 则

$$\underline{\mathrm{PI}}_{\alpha_2}(X)\subseteq\underline{\mathrm{PI}}_{\alpha_1}(X),\overline{\mathrm{PI}}_{\beta_2}(X)\subseteq\overline{\mathrm{PI}}_{\beta_1}(X).$$

证　由定义直接可得．　□

注 1　当 $\alpha=1,\beta=0$，若取 $P(X\mid[x])=\dfrac{|X\cap[x]|}{|[x]|}$，则

$$\underline{\mathrm{PI}}_\alpha(X) = \underline{\mathrm{PI}}_1(X) = \{x\in U \mid [x]\subseteq X\},$$
$$\overline{\mathrm{PI}}_\beta(X) = \overline{\mathrm{PI}}_0(X) = \{x\in X \mid [x]\cap X\neq\varnothing\}.$$

这时概率(I)型下近似和概率(I)型上近似分别成为 Pawlak 意义下的下近似和上近似，因此本节提出的概率粗糙集模型(I)是 Pawlak 粗糙集模型的推广形式．

注 2　由定理 8.3 的(2)和(7)可知

$$\underline{R}(X)=\{x\in U\mid[x]\subseteq X\}\subseteq\underline{\mathrm{PI}}_\alpha(X)$$
$$\subseteq\overline{\mathrm{PI}}_\beta(X)\subseteq\overline{R}(X)=\{x\in X\mid[x]\cap X\neq\varnothing\}.$$

因此,概率(I)型粗糙集中的边界一般要比相应的 Pawlak 粗糙集的边界小,而正域和负域都比 Pawlak 粗糙集的正域和负域大,其主要原因是概率(I)型粗糙集允许将 Pawlak 粗糙集边界中条件概率大于零小于 1 的某一部分等价类也计人到它的正域里面,还将 Pawlak 粗糙集边界中一部分并入到它的负域中去.从这里可以看出概率粗糙集模型(I)比 Pawlak 粗糙集模型的应用更加广泛.

由定理 8.3 的(7)可以看出概率(I)型粗糙集的正域随着 α 的减少而增大,负域随着 β 的增大而增大,同时边界缩小.但是一般情况下

$$\lim_{\alpha\downarrow r}\underline{\mathrm{PI}}_\alpha(X)=\underline{\mathrm{PI}}_r(X),$$
$$\lim_{\beta\uparrow r}\overline{\mathrm{PI}}_\beta(X)=\overline{\mathrm{PI}}_r(X),$$

却不成立.

例 8.1 设 $U=\{x_i\mid 1\leqslant i\leqslant 20\}$,令 $P(B)=\dfrac{|B|}{|U|}$,$B\subseteq U$,$U/R=\{X_1,X_2,X_3,X_4,X_5,X_6\}$,其中 $X_1=\{x_1,x_2,x_3,x_4,x_5\}$,$X_2=\{x_6,x_7,x_8\}$,$X_3=\{x_9,x_{10},x_{11},x_{12}\}$,$X_4=\{x_{13},x_{14}\}$,$X_5=\{x_{15},s_{16},x_{17},x_{18}\}$,$X_6=\{x_{19},x_{20}\}$.令 $X=\{x_6,x_7,x_8,x_{13},x_{17}\}$,则$\underline{\mathrm{PI}}_{0.5}(X)=X_2\cup X_4$,但 $\lim\limits_{\alpha\downarrow 0.5}\underline{\mathrm{PI}}_\alpha(X)=X_2$,而 $\lim\limits_{\beta\uparrow 0.5}\overline{\mathrm{PI}}_\beta(X)=X_2\cup X_4\neq\overline{\mathrm{PI}}_{0.5}(X)=X_2$.

定理 8.4 设 $0<r<1$,则对于任意 $X\subseteq U$ 有

(1)$\lim\limits_{\alpha\downarrow r}\underline{\mathrm{PI}}_\alpha(X)=\bigcup\limits_{\alpha>r}\underline{\mathrm{PI}}_\alpha(X)=\overline{\mathrm{PI}}_r(X)$.

(2)$\lim\limits_{\beta\uparrow r}\overline{\mathrm{PI}}_\beta(X)=\bigcap\limits_{\beta<r}\overline{\mathrm{PI}}_\beta(X)=\underline{\mathrm{PI}}_r(X)$.

证 (1)当 $\alpha>r$ 时,由定义知

$$\underline{\mathrm{PI}}_\alpha(X)=\{x\in U\mid P(X\mid[x])\geqslant\alpha\}$$
$$\subseteq\{x\in U\mid \mathrm{P}(X\mid[x])>r\}=\overline{\mathrm{PI}}_r(X).$$

又由定理 8.3 知$\underline{\mathrm{PI}}_\alpha(X)$随着 α 的减小而增大,因此

$$\lim_{\alpha\downarrow r}\underline{\mathrm{PI}}_\alpha(X)=\bigcup_{\alpha>r}\underline{\mathrm{PI}}_\alpha(X)\subseteq\overline{\mathrm{PI}}_r(X).$$

若存在 $x_0\in\overline{\mathrm{PI}}_r(X)\setminus\bigcup\limits_{\alpha>r}\underline{\mathrm{PI}}_\alpha(X)$则有 $P(X\mid[x_0])>r$,但是对任意 $\alpha>r$ 有 $x_0\notin\underline{\mathrm{PI}}_\alpha(X)$,即

$$\forall\alpha>r,\text{有 }P(X\mid[x_0])<\alpha,$$

这说明 $P(X\mid[x_0])\leqslant r$ 与$P(X\mid[x_0])>r$ 矛盾!这意味着

$$\bigcup_{\alpha>r}\underline{\mathrm{PI}}_\alpha(X)=\overline{\mathrm{PI}}_r(X).$$

因此,结论(1)成立.

(2)由 $\beta<r$ 可得

$$\begin{aligned}\overline{\mathrm{PI}}_\beta(X) &= \{x\in U \mid P(X\mid[x])>\beta\}\\ &\supseteq \{x\in U\mid P(X\mid[x])\geqslant r\} = \underline{\mathrm{PI}}_r(X).\end{aligned}$$

从而由$\overline{\mathrm{PI}}_\beta(X)$ 关于 β 的单调递减性得

$$\lim_{\beta\uparrow r}\overline{\mathrm{PI}}_\beta(X)=\bigcap_{\beta<r}\overline{\mathrm{PI}}_\beta(X)\supseteq\underline{\mathrm{PI}}_r(X).$$

若存在 $y\in\bigcap\limits_{\beta<r}\overline{\mathrm{PI}}_\beta(X)\setminus\underline{\mathrm{PI}}_r(X)$，则对于 $\forall\beta<r$，$y\in\overline{\mathrm{PI}}_\beta(X)$，但 $y\notin\underline{\mathrm{PI}}_r(X)$，于是

$$P(X\mid[y])>\beta,\qquad\forall\beta<r,\tag{8.2}$$

但是，$P(X|[y])<r$. 由(8.2)知 $P(X|[y])\geqslant r$，这与 $P(X|[y])<r$ 矛盾！故结论(2)成立. □

定理 8.5　设 $0<r<1$，则

$$\begin{aligned}\lim_{\beta\uparrow r,\alpha\downarrow r}\mathrm{bn}(X,\alpha,\beta)&=\bigcap_{\beta<r<\alpha}(\overline{\mathrm{PI}}_\beta(X)\setminus\underline{\mathrm{PI}}_\alpha(X)).\\ &=\underline{\mathrm{PI}}_r(X)\setminus\overline{\mathrm{PI}}_r(X)=\{x\in U\mid P(X\mid[x])=r\}.\end{aligned}$$

证　显然，$\mathrm{bn}(X,\alpha,\beta)\supseteq\underline{\mathrm{PI}}_r(X)\setminus\overline{\mathrm{PI}}_r(X)=\{x\in U|P(X|[x])=r\}$，而当 α 单调下降趋于 r，β 单调上升趋于 r 时，由定理 8.4 知 $\mathrm{bn}(X,\alpha,\beta)$是单调下降的，从而

$$\begin{aligned}\lim_{\beta\uparrow r,\alpha\downarrow r}\mathrm{bn}(X,\alpha,\beta)&-\bigcap_{\beta<r<\alpha}(\overline{\mathrm{PI}}_\beta(X)\setminus\underline{\mathrm{PI}}_\alpha(X))\\ &\supseteq\mathrm{PI}_r(X)\setminus\overline{\mathrm{PI}}_r(X)=\{x\in U\mid P(X\mid[x])=r\}.\end{aligned}$$

若 $\exists x_0\in\bigcap\limits_{\beta<r<\alpha}(\overline{\mathrm{PI}}_\beta(X)\setminus\underline{\mathrm{PI}}_\alpha(X))\setminus\{x\in U|P(X|[x])=r\}$，则对于任意 $\beta<r<\alpha$ 有 $x_0\in\mathrm{bn}(X,\alpha,\beta)=\overline{\mathrm{PI}}_\beta(X)\setminus\underline{\mathrm{PI}}_\alpha(X)$，但 $P(X|[x_0])\neq r$. 从而

$$P(X\mid[x_0])>\beta,\forall\beta<r.\tag{8.3}$$

而且

$$P(X\mid[x_0])<\alpha,\forall\alpha>r.\tag{8.4}$$

由(8.3)得 $P(X|[x_0])\geqslant r$，由(8.4)得 $P(X|[x_0])\leqslant r$，这样我们就得到了 $P(X|[x_0])=r$，这与假设矛盾！定理得证. □

注　从上述定理可以看出，随着 β 逐渐增大(到 r)，α 逐渐下降(到 r)，边界 $\mathrm{bn}(X,\alpha,\beta)$逐渐收缩(成为$\{x\in U|P(X|[x])=r\}$)，$\{x\in U|P(X|[x])=r\}$称为 X 关于概率近似空间依参数 r 的绝对边界，之所以给予这个名称是因为对于任意满足 $\beta<r<\alpha$ 的 α 与 β，边界 $\mathrm{bn}(X,\alpha,\beta)$都包含$\{x\in U|P(X|[x])=r\}$.

尽管例 8.1 表明$\underline{\mathrm{PI}}_\alpha(X)$关于 α 不是右连续的，$\overline{\mathrm{PI}}_\beta(X)$关于 β 不是左连续的，但是下面定理表明他们分别关于 α 和 β 是左连续和右连续的.

定理 8.6 设 $0<r<1$，则对于任意 $X\subseteq U$ 有

(1) $\lim\limits_{\alpha\uparrow r}\underline{\mathrm{PI}}_\alpha(X)=\bigcap\limits_{\alpha<r}\underline{\mathrm{PI}}_\alpha(X)=\underline{\mathrm{PI}}_r(X)$；

(2) $\lim\limits_{\beta\downarrow r}\overline{\mathrm{PI}}_\beta(X)=\bigcup\limits_{\beta>r}\overline{\mathrm{PI}}_\beta(X)=\overline{\mathrm{PI}}_r(X)$.

证 (1)当 $\alpha<r$ 时，由定理 8.3 知

$$\underline{\mathrm{PI}}_\alpha(X)\supseteq\underline{\mathrm{PI}}_r(X).$$

又由定理 8.3 知$\underline{\mathrm{PI}}_\alpha(X)$随着 α 的增大而减少，因此

$$\lim_{\alpha\uparrow r}\underline{\mathrm{PI}}_\alpha(X)=\bigcap_{\alpha<r}\underline{\mathrm{PI}}_\alpha(X)\supseteq\underline{\mathrm{PI}}_r(X).$$

若存在 $x_0\in\bigcap\limits_{\alpha<r}\underline{\mathrm{PI}}_\alpha(X)\setminus\underline{\mathrm{PI}}_r(X)$，则对任意 $\alpha<r$ 有 $x_0\in\underline{\mathrm{PI}}_\alpha(X)$，但是 $x_0\notin\underline{\mathrm{PI}}_r(X)$，即

$$\forall\alpha<r,\text{有 }P(X\mid[x_0])\geqslant\alpha,\text{而 }P(X\mid[x_0])<r.$$

这是矛盾的，因为由 $\forall\alpha<r,P(X|[x_0])\geqslant\alpha$ 可得 $P(X|[x_0])\geqslant r$. 这意味着

$$\bigcap_{\alpha<r}\underline{\mathrm{PI}}_\alpha(X)=\underline{\mathrm{PI}}_r(X).$$

因此，结论(1)成立.

(2)由 $\beta>r$ 和定理 8.3 可得

$$\overline{\mathrm{PI}}_\beta(X)\subseteq\overline{\mathrm{PI}}_r(X).$$

从而由$\overline{\mathrm{PI}}_\beta(X)$关于 β 的单调递减性得

$$\lim_{\beta\downarrow r}\overline{\mathrm{PI}}_\beta(X)=\bigcup_{\beta>r}\overline{\mathrm{PI}}_\beta(X)\subseteq\overline{\mathrm{PI}}_r(X).$$

若存在 $y\in\overline{\mathrm{PI}}_r(X)\setminus\bigcup\limits_{\beta>r}\overline{\mathrm{PI}}_\beta(X)$，即 $y\in\overline{\mathrm{PI}}_r(X)$，但对于 $\forall\beta>r,y\notin\overline{\mathrm{PI}}_\beta(X)$，于是$P(X|[y])>r$，而

$$P(X\mid[y])\leqslant\beta,\forall\beta>r.\tag{8.5}$$

由(8.5)知 $P(X|[y])\leqslant r$，这与 $P(X|[y])>r$ 矛盾！故结论(2)成立. □

若 $0\leqslant\beta<\alpha\leqslant1$，则 X 关于概率近似空间 A_p 依参数 α、β 的粗糙度 $\rho(X,\alpha,\beta)$和近似精度 $\eta(X,\alpha,\beta)$分别定义为

$$\rho(X,\alpha,\beta)=1-\frac{|\underline{\mathrm{PI}}_\alpha(X)|}{|\overline{\mathrm{PI}}_\beta(X)|},$$

$$\eta(X,\alpha,\beta)=1-\rho(X,\alpha,\beta)=\frac{|\underline{\mathrm{PI}}_\alpha(X)|}{|\overline{\mathrm{PI}}_\beta(X)|}.$$

由此可见，X 关于概率近似空间 A_p 依参数 α,β 是可定义的当且仅当其近似精度为 1 而粗糙度为 0.

§8.4 概率粗糙集模型的其他形式

定义 8.4 设 U 是有限论域，R 是 U 上的二元等价关系，P 是定义在 U

的子集类构成的σ代数上的概率测度,仍称三元有序组 $A_p=(U,R,P)$是概率近似空间.对于 $X\subseteq U$,$0<\beta\leqslant\alpha<1$,则 X 关于近似空间$A_p=(U,R,P)$依参数 α,β 的概率(II)型下近似$\underline{\mathrm{PII}}_\alpha(X)$和上近似$\overline{\mathrm{PII}}_\beta(X)$分别定义为

$$\begin{aligned}\underline{\mathrm{PII}}_\alpha(X)&=\{x\in U\mid P(X\mid[x])>\alpha\}\\&=\bigcup\{[x]\mid P(X\mid[x])>\alpha\},\\\overline{\mathrm{PII}}_\beta(X)&=\{x\in U\mid P(X\mid[x])\geqslant\beta\}\\&=\bigcup\{[x]\mid P(X\mid[x])\geqslant\beta\}.\end{aligned}$$

若$\underline{\mathrm{PII}}_\alpha(X)=\overline{\mathrm{PII}}_\beta(X)$,则称 X 关于近似空间 $A_p=(U,R,P)$依参数 α,β 是概率(II)型可定义的,否则称 X 是概率(II)型不可定义的或概率(II)型粗糙集.

对于参数 α,β,概率(II)型粗糙集可分为下列四类:

(1)若$\underline{\mathrm{PII}}_\alpha(X)\neq\varnothing$,$\overline{\mathrm{PII}}_\beta(X)\neq U$,称 X 是部分可定义的.

(2)若$\underline{\mathrm{PII}}_\alpha(X)\neq\varnothing$,$\overline{\mathrm{PII}}_\beta(X)=U$,称 X 是内部可定义的.

(3)若$\underline{\mathrm{PII}}_\alpha(X)=\varnothing$,$\overline{\mathrm{PII}}_\beta(X)\neq U$,称 X 是外部可定义的.

(4)若$\underline{\mathrm{PII}}_\alpha(X)=\varnothing$,$\overline{\mathrm{PII}}_\beta(X)=U$,称 X 是完全不可定义的.

X 关于近似空间$A_p=(U,R,P)$依参数 α,β 的概率(II)型正域,负域和边界分别定义为

$$\begin{aligned}\mathrm{pos}_\alpha(X)&=\underline{\mathrm{PII}}_\alpha(X),\\\mathrm{neg}_\beta(X)&=\sim\overline{\mathrm{PII}}_\beta(X)=\{x\in U\mid P(X\mid[x])<\beta\},\\\mathrm{bn}(X,\alpha,\beta)&=\overline{\mathrm{PII}}_\beta(X)\setminus\underline{\mathrm{PII}}_\alpha(X)\\&=\{x\in U\mid\beta\leqslant P(X\mid[x])\leqslant\alpha\}.\end{aligned}$$

X 关于近似空间$A_p=(U,R,P)$依参数 α,β 的粗糙度$\rho(X,\alpha,\beta)$和近似精度$\eta(X,\alpha,\beta)$分别定义为

$$\rho(X,\alpha,\beta)=1-\frac{|\underline{\mathrm{PII}}_\alpha(X)|}{|\overline{\mathrm{PII}}_\beta(X)|},$$

$$\eta(X,\alpha,\beta)=1-\rho(X,\alpha,\beta)=\frac{|\underline{\mathrm{PII}}_\alpha(X)|}{|\overline{\mathrm{PII}}_\beta(X)|}.$$

由定义可以发现,我们这里定义的下近似和上近似分别就是上一节概率粗糙集模型(I)中的上近似和下近似,即

$$\underline{\mathrm{PII}}_\alpha(X)=\overline{\mathrm{PI}}_\alpha(X),\quad\overline{\mathrm{PII}}_\beta(X)=\underline{\mathrm{PI}}_\beta(X).\tag{8.6}$$

定理 8.7 对于参数 α,β 和$X,Y\subseteq U$,概率(II)型粗糙近似算子满足下列对偶性质:

(1) $\underline{\mathrm{PII}}_\alpha(\varnothing)=\overline{\mathrm{PII}}_\beta(\varnothing)$,$\underline{\mathrm{PII}}_\alpha(U)=\overline{\mathrm{PII}}_\beta(U)=U$.

(2) $\underline{\mathrm{PII}}_\alpha(X)\subseteq\overline{\mathrm{PII}}_\beta(X)$.

(3) $\underline{\mathrm{PII}}_\alpha(X)=\sim\overline{\mathrm{PII}}_{1-\alpha}(\sim X),\overline{\mathrm{PII}}_\beta(X)=\sim\underline{\mathrm{PII}}_{1-\beta}(\sim X)$.

(4) $\overline{\mathrm{PII}}_\beta(X\cup Y)\supseteq\overline{\mathrm{PII}}_\beta(X)\cup\overline{\mathrm{PII}}_\beta(Y)$,

$\underline{\mathrm{PII}}_\alpha(X\cap Y)\subseteq\underline{\mathrm{PII}}_\alpha(X)\cap\underline{\mathrm{PII}}_\alpha(Y)$.

(5) $\overline{\mathrm{PII}}_\beta(X\cap Y)\subseteq\overline{\mathrm{PII}}_\beta(X)\cap\overline{\mathrm{PII}}_\beta(Y)$,

$\underline{\mathrm{PII}}_\alpha(X\cup Y)\supseteq\underline{\mathrm{PII}}_\alpha(X)\cup\underline{\mathrm{PII}}_\alpha(Y)$.

(6) 若 $X\subseteq Y$,则

$$\underline{\mathrm{PII}}_\alpha(X)\subseteq\underline{\mathrm{PII}}_\alpha(Y),\quad \overline{\mathrm{PII}}_\beta(X)\subseteq\overline{\mathrm{PII}}_\beta(Y).$$

(7) 若 $r\leqslant s$,则

$$\underline{\mathrm{PII}}_s(X)\subseteq\underline{\mathrm{PII}}_r(X),\quad \overline{\mathrm{PII}}_s(X)\subseteq\overline{\mathrm{PII}}_r(X).$$

证 由定义直接证或由定理 8.3 及(8.6)式可得. □

注 当取 $\alpha=\beta=\gamma$ 时,这时所得到的粗糙边界就是绝对边界,即

$$\mathrm{bn}(X,\gamma,\gamma)=\overline{\mathrm{PII}}_\gamma(X)\setminus\underline{\mathrm{PII}}_\gamma(X)=\{x\in U\mid P(X\mid[x])=\gamma\}.$$

类似于定理 8.4,定理 8.5 和定理 8.6,并且由这三个定理和(8.6)式可得概率(II)型粗糙集下,上近似和边界关于参数的连续性的定理.

定理 8.8 设 $0<r<1$,则对于任意 $X\subseteq U$ 有

(1) $\lim\limits_{\alpha\downarrow r}\underline{\mathrm{PII}}_\alpha(X)=\bigcup\limits_{\alpha>r}\underline{\mathrm{PII}}_\alpha(X)=\underline{\mathrm{PII}}_r(X)$.

(2) $\lim\limits_{\beta\uparrow r}\overline{\mathrm{PII}}_\beta(X)=\bigcap\limits_{\beta<r}\overline{\mathrm{PII}}_\beta(X)=\overline{\mathrm{PII}}_r(X)$.

定理 8.9 设 $0<r<1$,则

$$\begin{aligned}\lim_{\beta\uparrow r,\alpha\downarrow r}\mathrm{bn}(X,\alpha,\beta)&=\bigcap_{\beta<r<\alpha}(\overline{\mathrm{PII}}_\beta(X)\setminus\underline{\mathrm{PII}}_\alpha(X))\\&=\overline{\mathrm{PII}}_r(X)\setminus\underline{\mathrm{PII}}_r(X)\\&=\{x\in U\mid P(X\mid[x])=r\}.\end{aligned}$$

定理 8.10 设 $0<r<1$,则对于任意 $X\subseteq U$ 有

(1) $\lim\limits_{\alpha\uparrow r}\underline{\mathrm{PII}}_\alpha(X)=\bigcap\limits_{\alpha<r}\underline{\mathrm{PII}}_\alpha(X)=\overline{\mathrm{PII}}_r(X)$.

(2) $\lim\limits_{\beta\downarrow r}\overline{\mathrm{PII}}_\beta(X)=\bigcup\limits_{\beta>r}\overline{\mathrm{PII}}_\beta(X)=\underline{\mathrm{PII}}_r(X)$.

我们这一节提出的概率粗糙集模型(II)与上一节提出的模型都是基于同一个概率近似空间的,虽然表面上差别不大,但在应用上一般是不能互相替代的,它们各自有不同的含义,具体见下面一节.细心的读者可能会发现,类似于前面两类,概率粗糙集模型应该还可以有另外两种不同的类型,我们分别称为概率粗糙集模型(III)和概率粗糙集模型(IV).这里我们将下近似和上近似的概念给出,其他问题留给读者,应用背景可见下一节.

定义 8.5 设 $A_p=(U,R,P)$是概率近似空间,$X\subseteq U$,则 X 关于近似空间 $A_p=(U,R,P)$依参数 $\alpha,\beta(0\leqslant\beta\leqslant\alpha<1)$的概率(III)型下近似$\underline{\mathrm{PIII}}_\alpha(X)$和上近似$\overline{\mathrm{PIII}}_\beta(X)$分别定义为

$$\underline{\mathrm{PIII}}_\alpha(X)=\{x\in U\mid P(X\mid[x])>\alpha\}$$
$$=\bigcup\{[x]\mid P(X\mid[x])>\alpha\},$$
$$\overline{\mathrm{PIII}}_\beta(X)=\{x\in U\mid P(X\mid[x])>\beta\}$$
$$=\bigcup\{[x]\mid P(X\mid[x])>\beta\}.$$

X 关于近似空间 $A_p=(U,R,P)$ 依参数 $\alpha,\beta(0<\beta\leqslant\alpha\leqslant1)$ 的概率(IV)型下近似 $\underline{\mathrm{PIV}}_\alpha(X)$ 和上近似 $\overline{\mathrm{PIV}}_\beta(X)$ 分别定义为

$$\underline{\mathrm{PIV}}_\alpha(X)=\{x\in U\mid P(X\mid[x])\geqslant\alpha\}$$
$$=\bigcup\{[x]\mid P(X\mid[x])\geqslant\alpha\},$$
$$\overline{\mathrm{PIV}}_\beta(X)=\{x\in U\mid P(X\mid[x])\geqslant\beta\}$$
$$=\bigcup\{[x]\mid P(X\mid[x])\geqslant\beta\}.$$

§8.5 Bayes 决策与粗糙近似

这一节将讨论 Bayes 决策过程与上两节给出的粗糙集模型的联系，从中也可以看出概率粗糙集模型的应用背景.

在介绍 Bayes 决策前，我们先来看一个大家都熟悉的背景知识.

例 8.2(医疗诊断问题) 众所周知，一个医生在给一群病人诊病时是根据他掌握的知识或经验(比如说他以前给人看病的临床记录等)以及从病人的临床表现来确定每一位病人是否患了某种特殊的疾病 w，当然具有相同的临床表现 x 的病人应该施以相同的治疗方案，这样的病人集合记为 $[x]$. 如果病人的临床表现提示这个病人肯定患了疾病 w，即确诊，那么医生会毫不犹豫地立即给病人以正确的治疗. 在生活中我们经常听到误诊的事. 在实际情况中经常会碰到这样的局面：病人的临床表现对医生诊断是否患了疾病 w 的证据不充分，这时他不能肯定病人是否一定患有疾病 w，而此时病人正处于紧急状态(如有生命危险)必须立即作出决定. 在这种情况下，医生一方面必须从病人的临床表现 x 中判断有几成把握认为病人患了疾病 w，另一方面还必须考虑到如果病人实际患有疾病 w 而没有给予医治以及没病当病治的后果与代价(或风险).

对于疾病 w 这个概念，用 $\mathrm{pos}(w)$ 表示将得到治疗的病人的集合，用 $\mathrm{neg}(w)$ 表示不用立即治疗的病人的集合，当然对于具有相同的临床表现 x 的病人应该归入同一类，这就是说概念 w 是用 $\mathrm{pos}(w)$ 来近似的. 有时医生简单地把病人集合分成这样两类(如特别紧急情况时)，但很多情况下医生把全体病人分成三类，另一类 $\mathrm{bn}(w)$ 表示在作出是否给予治疗这个决定前需要进一步观察的病人集合. 不管分两类还是三类，在作出决策时总是要考虑风险最小.

设 $K=(U,P,AT,V,\rho)$为一概率知识表示系统,即 U 是论域,P 是 U 的子集全体构成的σ代数上的概率测度,$AT=\{a_1,a_2,\cdots,a_n\}$是有限个属性构成的集合,$V=V_1\times V_2\times\cdots\times V_n$,$V_i$ 是属性a_i 的值域,$\rho:U\to V$ 是信息函数,对于 U 中的每个对象x,$\rho(x)$称为 x 的描述,具有相同描述的对象是不可分辨的,记与 x 具有相同描述的对象全体为$[x]$,这样我们在问题叙述时在不至于引起误解的情况下,描述 $\rho(x)$在一一对应的意义下也常用$[x]$来代替.

设 $\Omega=\{w_1,w_2,\cdots,w_s\}$是具有有限个特征状态的集合,每个 w_i 是 U 的子集,常称为概念,$A=\{r_1,r_2,\cdots,r_m\}$是由 m 个可能决策行为构成的集合,$P(w_j\mid[x])$表示一个对象在描述$[x]$下处于状态 w_j 的概率,一般假定$P(w_j\mid[x])$是已知的. 令 $\lambda(r_i\mid w_j)$表示状态为 w_j 时采用决策 r_i 的风险损失.

假定一个对象的描述为$[x]$,对于这个对象实施了决策 r_i,由于$P(w_j\mid[x])$是在给定描述$[x]$下处于状态 w_j 的概率,因此对象在给定描述$[x]$下采用决策 r_i 的期望损失(常称为条件风险)可由全概率公式得

$$R(r_i\mid[x])=\sum_{j=1}^{s}\lambda(r_i\mid w_j)P(w_j\mid[x]). \tag{8.7}$$

对于给定的描述$[x]$,记 $\tau(x)$为一个决策规则,即 $\tau(x)\in A$,则 τ 是描述空间到A 的一个函数. 令 R 是在给定一个总体决策规则 τ 下的期望总体风险. 由于 $R(\tau(x)\mid[x])$是在描述$[x]$下取决策 $\tau(x)$的条件风险率,因此总体风险

$$R=\sum_{[x]}R(\tau(x)\mid[x])P([x]),$$

其中的和是对整个知识表示系统而言. 显然,如果决策规则 $\tau(x)$使得对于每个$[x]$而言条件风险率 $R(\tau(x)\mid[x])$尽可能地小,那么总体风险就能达到最小值.

Bayes 决策过程简述如下:对于每个对象 $x\in U$,计算出(8.7)式给出的条件风险 $R(r_i\mid[x])$,$i=1,2,\cdots,m$,选取其中一个使条件风险达到最小的决策作为最佳决策行为,如果有两个或两个以上的决策使条件风险达到最小则根据实际情况取其中之一.

为了问题叙述得直观与便于理解,我们用例 8.2 的问题来讨论 Bayes 决策与概率粗糙近似的关系.

设 U 是医生的一群病人(论域),设 w 是某种特殊的疾病(被近似的概念),则 w 把 U 分成两部分,要么患病(记全体为 w),要么没有患病 w(记全体为$\sim w$),这样 $\Omega=\{w,\sim w\}$. 记 pos(w),neg(w)和 bn(w)为 w 的正域

(立即治疗的病人全体),负域(肯定不用治疗的病人全体)和边界(须进一步观察的病人全体).

我们分两种情形来讨论这个例子:

情形 1 全体病人被分成两部分:pos(w)和 neg(w).

情形 2 全体病人被分成三部分:pos(w),neg(w)和 bn(w).

对于情形 1,每一位病人 x 在临床表现(对象的描述)$[x]$下都面临两种可能的决策

(Y) 决策 $r_1: x\in \mathrm{pos}(w)$, 即 $r_1:[x]\to \mathrm{pos}(w)$,

(N) 决策 $r_2: x\in \mathrm{neg}(w)$,即 $r_2:[x]\to \mathrm{neg}(w)$.

这时,$A=\{r_1,r_2\}$. 令 $\lambda(r_i|w)$为病人实际患病(对象实际属于 w)而采取决策 r_i 时的风险,$\lambda(r_i|\sim w)$为病人实际没有患病(对象实际不属于 w)而采取决策 r_i 时的风险,$P(w|[x])$为病人在临床表现$[x]$下患病 w 的概率,$P(\sim w|[x])$为病人在临床表现$[x]$下没有患病 w 的概率. 这样,病人在临床表现$[x]$下采取决策 r_i 的条件风险$R(r_i|[x])$可由全概率公式得

$$R(r_i\mid[x])=\lambda_{i1}P(w\mid[x])+\lambda_{i2}P(\sim w\mid[x]), \tag{8.8}$$

其中 $\lambda_{i1}=\lambda(r_i|w)$,$\lambda_{i2}=\lambda(r_i|\sim w)$,$i=1,2$.

由 Bayes 决策过程可得最小风险规则为

$$\text{(Y)}\quad r_1:[x]\to \mathrm{pos}(w),\text{若 } R(r_1|[x])\leqslant R(r_2|[x]), \tag{8.9}$$

$$\text{(N)}\quad r_2:[x]\to \mathrm{neg}(w),\text{若 } R(r_2|[x])\leqslant R(r_1|[x]). \tag{8.10}$$

由于在实际情况中,对于实际患病的人来说,进行治疗的风险比不进行治疗的风险小;而对于实际没有患病的人来说,进行治疗的风险比不进行治疗的风险大,即 λ_{ij},$i=1,2$,$j=1,2$,满足关系式

$$\lambda_{11}<\lambda_{21},\text{且 }\lambda_{12}>\lambda_{22}. \tag{8.11}$$

将(8.11)分别代人(8.9)和(8.10),并利用 $P(w|[x])+P(\sim w|[x])=1$ 经计算可得最小风险决策规则重新表达为

(Y) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\alpha$,

(N) $r_2:[x]\to \mathrm{neg}(w)$,若 $P(w|[x])\leqslant\alpha$,

其中 $\alpha=\dfrac{\lambda_{12}-\lambda_{22}}{(\lambda_{21}-\lambda_{11})+(\lambda_{12}-\lambda_{22})}$. 显然,由(8.11)知 $0<\alpha<1$.

当决策 $r_1:[x]\to \mathrm{pos}(w)$和 $r_2:[x]\to \mathrm{neg}(w)$的风险相同时,即对象 x 在描述$[x]$下处于状态 w 下的条件概率$P(w|[x])=\alpha$ 时,这时在实际情况中有两种可选方案(但必须选其中之一):

方案 1 将对象列人 pos(w),此时的决策规则为

(Y_1) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\alpha$,

(N_1) $r_2:[x]\to \mathrm{neg}(w)$,若 $P(w|[x])<\alpha$,

于是概念 w 关于参数 α 的正域和负域可表示为

$$\mathrm{pos}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])\geqslant\alpha\},$$
$$\mathrm{neg}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])<\alpha\}.$$

这样我们最终得到了概念 w 关于参数 α 的近似(全体被治疗的病人)为

$$\mathrm{apr}_\alpha(w)=\mathrm{pos}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])\geqslant\alpha\}.$$

即当临床表现为$[x]$的情况下患病 w 的概率大于或等于 α 的那些病人肯定被接受治疗.

方案 2 将对象列人 $\mathrm{neg}(w)$,此时的决策规则为

(Y_2) $r_1:[x]\to\mathrm{pos}(w)$,若 $P(w|[x])>\alpha$,

(N_2) $r_2:[x]\to\mathrm{neg}(w)$,若 $P(w|[x])\leqslant\alpha$,

于是概念 w 关于参数 α 的正域和负域可表示为

$$\mathrm{pos}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])>\alpha\},$$
$$\mathrm{neg}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])\leqslant\alpha\}.$$

这样我们最终得到了概念 w 关于参数 α 的近似(全体被治疗的病人)为

$$\mathrm{apr}_\alpha(w)=\mathrm{pos}_\alpha(w)=\bigcup\{[x]\mid P(w\mid[x])>\alpha\}.$$

即当临床表现为$[x]$的情况下患病 w 的概率大于 α 的那些病人肯定被接受治疗.

对于情形 2,全体病人被分成三部分:$\mathrm{pos}(w)$,$\mathrm{neg}(w)$和 $\mathrm{bn}(w)$. 于是对于每一位病人 x 在临床表现(对象的描述)$[x]$下都面临三种可能的决策(肯定被治疗(Y),肯定不用治疗(N),不确定或怀疑须进一步观察(D))

(Y) 决策 $r_1:x\in\mathrm{pos}(w)$,即 $r_1:[x]\to\mathrm{pos}(w)$,

(N) 决策 $r_2:x\in\mathrm{neg}(w)$,即 $r_2:[x]\to\mathrm{neg}(w)$,

(D) 决策 $r_3:x\in\mathrm{bn}(w)$,即 $r_3:[x]\to\mathrm{bn}(w)$.

这时,$A=\{r_1,r_2,r_3\}$.令 $\lambda(r_i|w)$为病人实际患病(对象实际属于 w)而采取决策 r_i 的风险,$\lambda(r_i|\sim w)$为病人实际没有患病(对象实际不属于 w)而采取决策 r_i 的风险,$P(w|[x])$为病人在临床表现$[x]$下患病的概率,$P(\sim w|[x])$为病人在临床表现$[x]$下没有患病的概率.这样,病人在临床表现$[x]$下采取决策 r_i 的条件风险$R(r_i|[x])$由全概率公式得

$$R(r_i\mid[x])=\lambda_{i1}P(w\mid[x])+\lambda_{i2}P(\sim w\mid[x]),\qquad(8.12)$$

其中 $\lambda_{i1}=\lambda(r_i|w)$,$\lambda_{i2}=\lambda(r_i|\sim w)$,$i=1,2,3$.

由 Bayes 决策过程可得最小风险规则为

(Y) $r_1:[x]\to\mathrm{pos}(w)$,若 $R(r_1|[x])\leqslant R(r_2|[x])$且 $R(r_1|[x])\leqslant R(r_3|[x])$, (8.13)

(N) $r_2:[x]\to\mathrm{neg}(w)$,若 $R(r_2|[x])\leqslant R(r_1|[x])$且 $R(r_2|[x])\leqslant R(r_3|[x])$, (8.14)

(D) $r_3:[x]\to \mathrm{pos}(w)$,若 $R(r_3|[x])\leqslant R(r_1|[x])$且 $R(r_3|[x])\leqslant R(r_2|[x])$. (8.15)

由 $P(w|[x])+P(\sim w|[x])=1$ 可得 $P(\sim w|[x])=1-P(w|[x])$,将它代入到上面三个式子,并利用式子(8.8)将 $R(r_i|[x])$也代入,这时我们发现在风险损失 λ_{ij},($i=1,2,3;j=1,2$)已知时,究竟取哪一条规则只与对象 x 在描述$[x]$下属于概念 w 的条件概率$P(w|[x])$有关.

在实际情况中,如果病人确实患病,那么肯定接受治疗的风险一般不会大于进一步进行观察的风险,而进一步进行观察的风险小于肯定不治疗的风险;如果病人实际上没有病,那么没病当病治的风险大于进一步进行观察的风险,而肯定不治疗的风险不会大于进一步进行观察的风险.这就是说 λ_{ij},($i=1,2,3,j=1,2$)满足关系式

$$\lambda_{11}\leqslant\lambda_{31}<\lambda_{21},\text{且 }\lambda_{12}>\lambda_{32}\geqslant\lambda_{22}. \tag{8.16}$$

将(8.12)分别代入(8.13),(8.14)和(8.15),并利用 $P(w|[x])+P(\sim w|[x])=1$ 经计算可得最小风险决策规则重新表达为

(Y) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\alpha, P(w|[x])\geqslant\gamma$,

(N) $r_2:[x]\to \mathrm{neg}(w)$,若 $P(w|[x])\leqslant\gamma$,且 $P(w|[x])\leqslant\beta$,

(D) $r_3:[x]\to \mathrm{bn}(w)$,若 $\beta\leqslant P(w|[x])\leqslant\alpha$, (8.17)

其中

$$\alpha=\frac{\lambda_{12}-\lambda_{32}}{(\lambda_{31}-\lambda_{11})+(\lambda_{12}-\lambda_{32})}, \tag{8.18}$$

$$\gamma=\frac{\lambda_{12}-\lambda_{22}}{(\lambda_{21}-\lambda_{11})+(\lambda_{12}-\lambda_{22})}, \tag{8.19}$$

$$\beta=\frac{\lambda_{32}-\lambda_{22}}{(\lambda_{21}-\lambda_{31})+(\lambda_{32}-\lambda_{22})}. \tag{8.20}$$

由(8.16)可知 $\alpha\in(0,1]$,$\gamma\in(0,1)$,且 $\beta\in[0,1)$.

由(8.17)可知参数 λ_{ij}必须满足条件$\beta\leqslant\alpha$,否则问题转化为情形 1.这样我们可以看出决策规则完全由从医生或专家输入的 λ_{ij}生成的参数α,β,γ决定.

对于满足条件 $\beta\leqslant\alpha$ 的参数,我们分两种来讨论:(i)$\beta<\alpha$;(ii)$\beta=\alpha$.

(i)若 $\beta<\alpha$,则由于在实数中有关系式

$$\frac{b}{a}<\frac{d}{c}\Rightarrow\frac{b}{a}<\frac{b+d}{a+c}<\frac{d}{c}.$$

并利用(8.18)~(8.20)可知 $\beta<\gamma<\alpha$.这时可能决策规则变为

(Y) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\alpha$,

(N) $r_2:[x]\to \mathrm{neg}(w)$,若且 $P(w|[x])\leqslant\beta$,

(D) $r_3:[x]\to \mathrm{bn}(w)$,若 $\beta\leqslant P(w|[x])\leqslant\alpha$.

下面我们分别将其完全确定.

若 $P(w|[x])=\alpha$ 时采用决策 pos(w),若 $P(w|[x])=\beta$ 时采用决策 neg(w),则我们得到最小风险决策规则为

(Y1) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\alpha$.

(N1) $r_2:[x]\to \mathrm{neg}(w)$,若且 $P(w|[x])\leqslant\beta$.

(D1) $r_3:[x]\to \mathrm{bn}(w)$,若 $\beta<P(w|[x])<\alpha$.

于是得到概念 w 的正域,负域和边界分别为

$$\begin{aligned}\mathrm{pos}(w) &= \bigcup\{[x]\mid P(w\mid[x])\geqslant\alpha\},\\ \mathrm{neg}(w) &= \bigcup\{[x]\mid P(w\mid[x])\leqslant\beta\},\\ \mathrm{bn}(w) &= \bigcup\{[x]\mid \beta<P(w\mid[x])<\alpha\}.\end{aligned}$$

从面得到概念 w 的下近似和上近似分别为

$$\begin{aligned}\underline{\mathrm{apr}}_\alpha(w) &= \mathrm{pos}(w)=\bigcup\{[x]\mid P(w\mid[x])\geqslant\alpha\},\\ \overline{\mathrm{apr}}_\beta(w) &= \mathrm{pos}(w)\cup\mathrm{neg}(w)=\bigcup\{[x]\mid P(w\mid[x])>\beta\}.\end{aligned}$$

这就是我们在§8.3中介绍的概率粗糙集模型(I)中定义的概率(I)型下近似和上近似.这样我们就可以利用概率粗糙集模型(I)来分析类似的具有最小风险的 Bayes 决策问题.

对于对象在描述$[x]$下位于状态 w 时的条件概率为$P(w|[x])=\alpha$ 或 $P(w|[x])=\beta$ 时,除了上面这种情况,另外还有三种确定的最小风险决策规则,方法完全与上面类似,可以验证它们分别对应于§8.4介绍的三类概率粗糙集模型.这样我们可以把(i)的最小风险 Bayes 决策问题转化为概率粗糙集模型下的问题来解决.

(ii)当 $\beta=\alpha$ 时,由(8.18)~(8.20)可得 $\beta=\gamma=\alpha$.这时可能的决策规则为

(Y2) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])\geqslant\gamma$.

(N2) $r_2:[x]\to \mathrm{neg}(w)$,若且 $P(w|[x])\leqslant\gamma$.

(D2) $r_3:[x]\to \mathrm{bn}(w)$,若 $P(w|[x])=\gamma$.

如果在描述$[x]$下对象归入 pos(w)或 bn(w)的风险相同(即$P(w|[x])=\gamma$)时,我们采用决策 bn(w).对于在描述$[x]$下对象归入 neg(w)或 bn(w)的风险相同(即 $P(w|[x])=\gamma$)时,我们采用决策 bn(w).这时可确定最小风险决策规则为

(Y3) $r_1:[x]\to \mathrm{pos}(w)$,若 $P(w|[x])>\gamma$.

(N2) $r_2:[x]\to \mathrm{neg}(w)$,若 $P(w|[x])<\gamma$.

(D3) $r_3:[x]\to \mathrm{bn}(w)$,若 $P(w|[x])=\gamma$.

从而得到概念 w 的正域,负域和边界为

$$\begin{aligned}\mathrm{pos}(w) &= \bigcup\{[x] \mid P(w \mid [x]) > \gamma\},\\ \mathrm{neg}(w) &= \bigcup\{[x] \mid P(w \mid [x]) < \gamma\},\\ \mathrm{bn}(w) &= \bigcup\{[x] \mid P(w \mid [x]) = \gamma\}.\end{aligned}$$

于是得到概念 w 的下近似和上近似为

$$\begin{aligned}\underline{\mathrm{apr}}_\gamma(w) &= \mathrm{pos}(w) = \{[x] \mid P(w \mid [x]) > \gamma\},\\ \overline{\mathrm{apr}}_\gamma(w) &= \mathrm{pos}(w) \cup \mathrm{neg}(w) = \{[x] \mid P(w \mid [x]) \geqslant \gamma\}.\end{aligned}$$

这样我们可以把(ii)的最小风险 Bayes 决策问题转化为概率粗糙集模型(II)的绝对边界情形的问题来解决.

§8.6 粗糙隶属函数与概念的联合

本节首先给出概率粗糙集模型中的粗糙隶属函数的概念,然后讨论求概念 X 和 Y 的联合 $X\cap Y$ 和 $X\cup Y$ 的粗糙近似的问题.

设 $A_p=(U,R,P)$ 是概率近似空间,$X\subseteq U$,则 X 关于 A_p 的粗糙隶属函数定义为

$$\mu_X(x) = P(X \mid [x]), x \in U.$$

粗糙隶属函数值 $\mu_X(x)$ 可以理解为 x 所在的等价类 $[x]$ 中的任何一个对象属于 X 的概率. 利用粗糙隶属函数,X 可以看成一个模糊集,但是一般情况下,粗糙隶属函数不是 Zadeh 意义下的隶属函数,事实上有以下

定理 8.11 粗糙隶属函数满足下列性质:

(1) $\mu_{\sim X}(x)=1-\mu_X(x)$.

(2) $\mu_{X\cup Y}(x)=\mu_X(x)+\mu_Y(x)-\mu_{X\cap Y}(x)$.

(3) $\max\{0,\mu_X(x)+\mu_Y(x)-1\}\leqslant\mu_{X\cap Y}(x)\leqslant\min\{\mu_X(x),\mu_Y(x)\}$.

(4) $\max\{\mu_X(x),\mu_Y(x)\}\leqslant\mu_{X\cup Y}(x)\leqslant\min\{1,\mu_X(x)+\mu_Y(x)\}$.

证 由概率论知识即得. □

在医疗诊治问题中,医生经常会碰到一个病人同时得几种病的情况,设 U 是医生的一群病人,X 和 Y 是两种特殊的疾病,医生(专家)知道在病人的临床表现 $[x]$ 下得每一种病的概率 $P(X|[x])$ 和 $P(Y|[x])$,根据上一节讨论的内容及概率粗糙集方法他能得出每一种病的治疗方案(决策),即可以得到概念 X 和 Y 的正域,负域和边界. 现在的问题是:他能否给出 $X\cap Y$ 和 $X\cup Y$ 的治疗方案? 由于概率粗糙集模型中的粗糙集全体一般不能构成拓扑,即定理 8.3 或定理 8.7 中关系式(4)和(5)一般不是等式,这样 $X\cap Y$ 和 $X\cup Y$ 的治疗方案不能由 X 和 Y 的治疗方案的合成得到. 但是,如果他能知道 $P(X\cap Y|[x])$ 或 $P(X\cup Y|[x])$,那么就能根据概率性质

$$P(X\cup Y\mid[x])=P(X\mid[x])+P(Y\mid[x])-P(X\cap Y\mid[x])$$
求得另一式，这样也能得到 $X\cap Y$ 和 $X\cup Y$ 的近似，然而在实际情况中往往只知道 $P(X|[x])$ 和 $P(Y|[x])$ 却不知道 $P(X\cap Y|[x])$ 或 $P(X\cup Y|[x])$，这样就必须从 $P(X|[x])$ 和 $P(Y|[x])$ 出发找 $P(X\cap Y|[x])$ 或 $P(X\cup Y|[x])$ 的好的估计. 概率统计理论中有多种估计联合分布的方法，我们在这里用最大熵和最小熵原理的方法来估计联合分布.

如果 $P=(p_1,p_2,\cdots,p_n)$ 是一个已知的联合概率分布，那么熵 $H(P)=-\sum_{i=1}^{n}p_i\log p_i$ 就反映了这个分布的不确定性程度. 如果概率分布的信息不全，而只知道分布的某些约束条件，那么我们可以根据最大熵或最小熵的原理来估计联合概率分布. 要注意的是对于同一个问题这两种方法不能同时采用，究竟是用最大熵原理还是用最小熵原理要根据实际情况来决定. 一般的原则是：如果要求想得到的分布是尽可能公平或无偏的(分布的不确定性大)，则采用最大熵原理，如果想得到的分布最不公平或最有偏的(分布的不确定性小)，则采用最小熵原理. 下面我们用最大熵和最小熵原理来估计条件概率的联合分布.

设 $a=P(X|[x]),b=P(Y|[x])$，记

$$a_{11}=P(X\cap Y\mid[x]),a_{12}=P(X\cap(\sim Y)\mid[x]),$$
$$a_{21}=P((\sim X)\cap Y\mid[x]),a_{22}=P((\sim X)\cap(\sim Y)\mid[x]).$$

显然，联合概率分布满足下列约束条件：

$$\begin{aligned}&a_{11}+a_{12}=a,a_{21}+a_{22}=1-a,\\&a_{11}+a_{21}=b,a_{12}+a_{22}=1-b,\end{aligned}\tag{8.21}$$

于是联合概率分布 $P=(a_{11},a_{12},a_{21},a_{22})$ 的熵函数为

$$H(P)=-\sum_{i,j=1}^{2}a_{ij}\log a_{ij}.\tag{8.22}$$

由约束条件(8.21)可知，这里的 a_{ij} 只有一个是自由变量，我们取 a_{11} 为自由变量，则

$$\begin{cases}a_{12}=a-a_{11},\\a_{21}=b-a_{11},\\a_{22}=1-(a+b)+a_{11}.\end{cases}$$

于是代入(8.22)得

$$\begin{aligned}H(P)=&-a_{11}\log a_{11}-(a-a_{11})\log(a-a_{11})\\&-(b-a_{11})\log(b-a_{11})\\&-[1-(a+b)+a_{11}]\log[1-(a+b)+a_{11}].\end{aligned}\tag{8.23}$$

由定理 8.11 中(3) 可得

$$\max\{0,a+b-1\}\leqslant a_{11}\leqslant \min\{a,b\}.$$

对(8.23)两边关于 a_{11} 求导并令导数为零得

$$\frac{\mathrm{d}H(P)}{\mathrm{d}a_{11}}=\log\frac{(a-a_{11})(b-a_{11})}{a_{11}[1-(a+b)+a_{11}]}=0.$$

解得 $a_{11}=ab$,容易验证

$$\max\{0,a+b-1\}\leqslant ab\leqslant \min\{a,b\}.$$

并且当 $a_{11}\in[\max\{0,a+b-1\},ab)$ 时有 $\frac{\mathrm{d}H(P)}{\mathrm{d}a_{11}}>0$,当 $a_{11}\in(ab,\min\{a,b\}]$ 时有 $\frac{\mathrm{d}H(P)}{\mathrm{d}a_{11}}<0$,这样 $H(P)$ 在 $a_{11}=ab$ 处达到最大值,而 $H(P)$ 的最小值只能在区间 $[\max\{0,a+b-1\},\min\{a,b\}]$ 的端点处达到.

如果我们用最大熵原理来估计联合条件概率分布,那么联合条件概率为

$$P(X\cap Y\mid[x])=P(X\mid[x])P(Y\mid[x]),$$
$$P(X\cup Y\mid[x])=P(X\mid[x])+P(Y\mid[x])-P(X\mid[x])P(Y\mid[x]).$$

事实上,这和事件 X 与 Y 的独立性假设是一样的.这时,相应的粗糙隶属函数满足

$$\mu_{X\cap Y}(x)=\mu_X(x)\mu_Y(x),$$
$$\mu_{X\cup Y}(x)=\mu_X(x)+\mu_Y(y)-\mu_X(x)\mu_Y(x).$$

如果我们用最小熵原理来估计联合条件概率分布,那么由于 $H(P)$ 的最小值在区间 $[\max\{0,a+b-1\},\min\{a,b\}]$ 的端点处达到,因此当 $a_{11}=\max\{0,a+b-1\}$ 是 $H(P)$ 的最小值点时有

$$P(X\cap Y\mid[x])=\max\{0,P(X\mid[x])+P(Y\mid[x])-1\},$$
$$P(X\cup Y\mid[x])=P(X\mid[x])+P(Y\mid[x])-P(X\cap Y\mid[x])=\min\{1,P(X\mid[x])+P(Y\mid[x])\}.$$

这时,相应的粗糙隶属函数满足

$$\mu_{X\cap Y}(x)=\max\{0,\mu_X(x)+\mu_Y(x)-1\},$$
$$\mu_{X\cup Y}(x)=\min\{1,\mu_X(x)+\mu_Y(x)-1\}.$$

当 $a_{11}=\min\{a,b\}$ 是 $H(P)$ 的最小值点时有

$$P(X\cap Y\mid[x])=\min\{P(X\mid[x]),P(Y\mid[x])\},\tag{8.24}$$
$$P(X\cup Y\mid[x])=P(X\mid[x])+P(Y\mid[x])-P(X\cap Y\mid[x])=\max\{P(X\mid[x]),P(Y\mid[x])\}.\tag{8.25}$$

这时,相应的粗糙隶属函数满足

$$\mu_{X\cap Y}(x)=\min\{\mu_X(x),\mu_Y(x)\},$$
$$\mu_{X\cup Y}(x)=\max\{\mu_X(x),\mu_Y(x)\}.$$

这与 Zadeh 意义下的隶属函数完全相符.

这样,在实际应用中如果我们知道了 $P(X|[x])$ 和 $P(Y|[x])$,就能根据最大熵或最小熵原理得到 $P(X\cap Y)|[x])$ 和 $P(X\cup Y|[x])$,然后由概率粗糙集模型求出 $X\cap Y$ 与 $X\cup Y$ 的近似,对于多个概念的联合可以重复运用上述办法. 特别值得指出的是当用最小熵原理,而且 $a_{11}=\min\{a,b\}$ 是 $H(P)$ 的最小值点时得到的 $X\cap Y$ 与 $X\cup Y$ 的近似可以通过 X 与 Y 的近似合成,即有下列定理.

定理 8.12 设 $\underline{PA}_\alpha(X)$ 与 $\overline{PA}_\beta(X)$ 是 X 关于概率近似空间依参数 α,β 的概率下近似和概率上近似(四类之一),若 $a_{11}=\min\{a,b\}$ 是(8.23)中熵函数 $H(P)$ 的最小值点,则对于任意 $X,Y\subseteq U$,下面性质成立:

(1) $\overline{PA}_\beta(X\cup Y)=\overline{PA}_\beta(X)\cup\overline{PA}_\beta(Y)$,

$\underline{PA}_\alpha(X\cap Y)=\underline{PA}_\alpha(X)\cap\underline{PA}_\alpha(Y)$.

(2) $\overline{PA}_\beta(X\cap Y)=\overline{PA}_\beta(X)\cap\overline{PA}_\beta(Y)$,

$\underline{PA}_\alpha(X\cup Y)=\underline{PA}_\alpha(X)\cup\underline{PA}_\alpha(Y)$.

证 对于任意的参数 α,由(8.24)和(8.25)可知

$$P(X\mid[x])>\alpha \wedge P(Y\mid[x])>\alpha \Leftrightarrow P(X\cap Y\mid[x])>\alpha,$$

$$P(X\mid[x])>\alpha \vee P(Y\mid[x])>\alpha \Leftrightarrow P(X\cup Y\mid[x])>\alpha.$$

从而由近似定义即证. □

§8.7 知识的不确定性度量

粗糙集理论中知识的不确定性主要是由两个原因引起的:一个原因来自于在给定近似空间的粗糙集的边界,当边界为空集时知识是完全确定的,边界越大知识就越粗糙或越模糊,这种不确定性称为系统的不确定性. 粗糙集理论处理这类知识的不确定性是通过引进粗糙性测度或近似精度来实现的.

粗糙集理论的知识的不确定性引起的另一个原因是直接来自于论域上的二元关系及其产生的知识模块,即近似空间本身. 如果二元等价关系产生的每一个等价类中只有一个元素,那么等价产生的划分不含有任何信息. 划分越粗,每一个知识模块越大,知识库中的知识越粗糙,相对于近似空间的概念和知识就越不确定,这种不确定性称为概念的不确定性. 处理概念不确定性的方法往往用信息熵来刻画,知识的粗糙性与信息熵的关系比较密切,知识的粗糙性实质上是其所含信息多少的更深层次的刻画.

对于概率近似空间 (U,R,P),系统的不确定性用系统的熵 $H(R^*)$ 来表示,即

$$H(R^*) = -\sum_{i=1}^{n} P(X_i)\log P(X_i),$$

其中 $R^* = U/R = \{X_1, X_2, \cdots, X_n\}$ 为 R 在 U 上导出的划分，系统的熵常称为系统的初始熵.

注 若近似空间中没有定义概率测度，则可以将论域上的任何一个等价关系或划分看成是定义在论域上的子集组成的 σ 代数上的一个随机变量，这时在每一个集合上的概率定义为该集合的基数与论域的基数的比值.

若 S 为 U 上另一等价关系，设 $S^* = \{Y_1, Y_2, \cdots, Y_m\}$，则在已知知识 R^* 时知识 S^* 的条件熵为

$$H(S^* \mid R^*) = -\sum_{i=1}^{n} P(X_i)\sum_{j=1}^{m} P(Y_j \mid X_i)\log P(Y_j \mid X_i).$$

条件熵 $H(S^* \mid R^*)$ 提供了知识 S^* 关于知识 R^* 的信息依赖性的一个较为合理的度量.

表 8.1 一个小汽车知识表示系统

U	条件属性 C			决策属性 D	
对象 汽车	a 内部设计	b 机型	c 颜色	d 速度	e 加速性能
1	适中	柴油	银色	中	差
2	拥挤	汽油	白色	高	极好
3	宽敞	柴油	黑色	高	好
4	适中	汽油	黑色	中	极好
5	适中	柴油	银色	低	好
6	宽敞	丙烷	黑色	高	好
7	宽敞	汽油	白色	高	极好
8	拥挤	汽油	白色	低	好

知识 S^* 与知识 R^* 的互信息量为

$$I(R^*; S^*) = H(S^*) - H(S^* \mid R^*) = H(R^*) - H(R^* \mid S^*).$$

互信息量反映知识 S^* 从知识 R^* 上获取的信息量.

例 8.3 表 **8.1** 给出了一个小汽车的知识表示系统.

在这个系统中，条件属性 C 的信息熵为 $H(C^*) = 2.5$，决策属性 D 的信息熵为 $H(D^*) = 2.25$，条件熵 $H(D^* \mid C^*) = 0.499$，互信息量 $I(D^*; C^*) = 1.751$，由于条件属性 C 由三个子属性 a（内部设计），b（机型）和 c（颜色）构成，经计算

$$H(C^* \mid \{a\}^*) = 0.939, I(C^*; \{a\}^*) = 1.561,$$
$$H(C^* \mid \{b\}^*) = 1.094, I(C^*; \{b\}^*) = 1.406,$$
$$H(C^* \mid \{c\}^*) = 0.939, I(C^*; \{c\}^*) = 1.561.$$

从上面三式我们可以发现，由于分别对条件属性中"内部设计"，"机型"和"颜色"的确定，系统中条件属性所提供知识的不确定性也相应地减少了，同时我们也得到了每一个子属性对总体条件属性所提供的信息量(即互信息量). 同样，也可以分析总体决策属性 D 关于子属性"速度"和"加速性能"的条件熵和互信息量.

有些读者可能会问：条件属性对决策属性提供的信息量如何？或者说条件属性和决策属性之间有怎么样的信息依赖性？下面我们就来讨论这个问题.

设 $K = (U, A, V, \rho)$ 为一个信息系统，$R^* = U/R = \{X_1, X_2, \cdots, X_n\}$ 为等价关系 R 在 U 上导出的划分，S 为 U 上另一等价关系，且 $S^* = \{Y_1, Y_2, \cdots, Y_m\}$，则在已知知识 R^* 时知识 S^* 的正则条件熵定义为

$$H_0(S^* \mid R^*) = -\sum_{i=1}^{n} P(X_i) \sum_{j=1}^{m} P(Y_j \mid X_i) \log P(Y_j \mid X_i) / \log m.$$

显然 $0 \leqslant H_0(S^* \mid R^*) \leqslant 1$.

我们用 $H_0(S^* \mid R^*)$ 表示系统中属性 S 基于属性 R 的信息依赖性，常记为

$$R \xrightarrow{H(S^* \mid R^*)} S.$$

特别地，划分 S^* 称为完全信息依赖于划分 R^*，若 $H_0(S^* \mid R^*) = 0$. 划分 S^* 称为完全信息独立于划分 R^*，若 $H_0(S^* \mid R^*) = 1$.

属性集 A 的一个子集 M 称为统计依赖的，若存在 M 的一个真子集 N 使 $H_0(M^* \mid N^*) = H_0(M^* \mid M^*)$，否则称 M 是统计独立的. N 称为 M 的一个统计约简，若 N 是 M 的一个极大统计独立子集，M 的所有统计约简记为 $\mathrm{sred}(M)$. M 的所有统计约简的交集称为 M 的核，记为 $\mathrm{score}(M)$.

系统 K 中的属性子集 $M \subseteq A$ 称为关于属性集 $B \subseteq A$ 是统计依赖的，若存在 M 的一个真子集 L 使 $H_0(B^* \mid M^*) = H_0(B^* \mid L^*)$，否则称 M 关于 B 是统计独立的. N 称为 M 关于 B 的一个相对约简，若 N 是 M 关于 B 的一个极大统计独立子集，M 关于 B 的所有相对约简记为 $\mathrm{sred}_B(M)$，所有这些约简的交集称为 M 关于 B 的相对核，记为 $\mathrm{score}_B(M)$.

属性 a 在 A 中称为统计可省略的，若 $H_0(A^* \mid (A \setminus \{a\})^*) = H_0(A^* \mid A^*)$，否则称 a 在 A 中为统计不可省略的. 属性 a 在 A 中称为关于属性集 D 是统计可省略的，若 $H_0(D^* \mid (A \setminus \{a\})^*) = H_0(D^* \mid A^*)$，否则称

a 在 A 中关于属性集 D 是统计不可省略的.

在实际应用中,找出属性的统计约简在很多概率分类问题中有助于决策规则的简化,我们仍以例 8.3 的小汽车知识表示系统来说明:由于 $H_0(D^* \mid C^*)=0.215$,这说明条件属性 C 与决策属性 D 之间有较强的信息依赖性.又经计算有

$$H_0(D^* \mid (C \setminus \{a\})^*) - H_0(D^* \mid C^*) = 0.041,$$
$$H_0(D^* \mid (C \setminus \{b\})^*) - H_0(D^* \mid C^*) = 0,$$
$$H_0(D^* \mid (C \setminus \{c\})^*) - H_0(D^* \mid C^*) = 0.$$

这说明条件属性集 C 中的属性"机型"或"颜色"关于决策属性是多余的,换言之,从 C 中去掉属性"机型"或"颜色"不会影响决策属性集 D 关于 C 的总体依赖性.例如,我们在条件属性集 C 中去掉属性"颜色",考虑 $C_1 = C \setminus \{c\}$,经计算有

$$H_0(D^* \mid (C_1 \setminus \{a\})^*) - H_0(D^* \mid C_1^*) = 0.364,$$
$$H_0(D^* \mid (C_1 \setminus \{b\})^*) - H_0(D^* \mid C_1^*) = 0.297.$$

这说明条件属性中"内部设计"和"机型"在对"速度"和"加速性能"进行决策时起着重要作用.事实上,条件属性 C 关于决策属性 D 的统计相对约简和核分别为

$$\mathrm{sred}_D(C) = \{内部设计,机型\}, \quad \mathrm{score}_D(C) = \{内部设计\}.$$

度量概念或粗糙集合 X 的不确定性的方法之一是粗糙度或近似精度,近似精度是 X 的下近似的概率与上近似的概率之比,它刻画了所给粗糙集 X 的知识的完全性,粗糙度是 1 减去 X 的近似精度,它反映了所给粗糙集 X 的知识的不完全性,近似精度或粗糙度的计算要求近似区域的每一部分(正域和负域)的概率是已知的,用它们来表示由于边界引起的不确定性是一个很好的度量.

度量粗糙集 X 的不确定性的另一种方式是正则条件熵 $H_0(X^* \mid R^*)$,其中 $X^* = \{X, \sim X\}$ 为由 X 产生的 U 的一个划分,正则条件熵 $H_0(X^* \mid R^*)$ 为概率近似空间中概念 X 的可定义性提供了一种自然测度,它提供了概念 X 被知识 $X^* = \{X_1, X_2, \cdots, X_n\}$ 刻画后反映近似好坏的一个总体测度.

尽管近似精度,粗糙度和正则条件熵能反映概念 X 的不确定性,但是它们并没有提供给我们那些完全属于 X 的下近似区域(正域)里面与不可分辨关系的知识粒度有关的知识的不确定性.

例 8.4 设 $U = \{a_1, a_2, a_3, a_4, b_1, b_2, b_3, c_1, c_2\}$, $X = \{a_1, a_2, a_3, a_4, b_1, c_1\}$, R_1, R_2 和 R_3 为 U 上三个等价关系,其导出的划分分别为

$$R_1^* = \{\{a_1, a_2, a_3, a_4\}, \{b_1, b_2, b_3\}, \{c_1, c_2\}\},$$

$$R_2^* = \{\{a_1,a_2\},\{a_3,a_4\},\{b_1,b_2,b_3\},\{c_1,c_2\}\},$$
$$R_3^* = \{\{a_1\},\{a_2\},\{a_3\},\{a_4\},\{b_1,b_2,b_3\},\{c_1,c_2\}\}.$$

显然，X 关于三个近似空间 $A_i=(U,R_i)$，$i=1,2,3$ 的下近似和上近似都是相同的

$$\underline{\mathrm{apr}}_i X = \{a_1,a_2,a_3,a_4\}, \quad \overline{\mathrm{apr}}_i X = U, \qquad i = 1,2,3.$$

当然集合 X 关于近似空间 A_i 的近似精度是相同的 $\eta_{A_i}(X)=\frac{4}{9}$，$i=1,2,3$；经计算，$X$ 关于 R_i，$i=1,2,3$，的正则条件熵 $H_0(X^*\mid R_i^*)$ 也是相同的，$H_0(X^*\mid R_i^*)=0.304$，$i=1,2,3$. 事实上，与划分 R_1^* 相比，划分 R_2^* 与 R_3^* 被重新划分部分恰好是 X 关于三个近似空间的下近似部分，但是很明显知识 R_1^* 的不确定性大于知识 R_2^* 的不确定性，知识 R_2^* 的不确定性大于知识 R_3^* 的不确定性，这说明光用近似精度，粗糙度和正则条件熵还不能完全反映 X 的不确定性，为此我们引进 X 的粗糙熵来刻画 X 的不确定性.

集合 X 关于概率近似空间(U,R,P)的粗糙熵 $E_r(X)$定义为

$$E_r(X) = -\eta_R(X)\sum_{i=1}^{n} P(X_i)\log p_i,$$

其中$\{X_1,X_2,\cdots,X_n\}$为 R 导出的划分R^*，p_i 为第i 个等价类X_i 中的任何一个对象 x 关于X_i 的条件概率. 特别当 U 中任何集合 Y 的概率定义为 Y 的基数与 U 的基数之比时，即 $P(Y)=\frac{|Y|}{|U|}$，则取 $p_i \doteq \frac{1}{|X_i|}$.

在例 8.4 中，X 关于近似空间(U,R_i,P)的粗糙熵 $E_{r_i}(X)$分别为

$$E_{r_1}(X) = -\frac{5}{9}\left(\frac{4}{9}\log\frac{1}{4} + \frac{3}{9}\log\frac{1}{3} + \frac{2}{9}\log\frac{1}{2}\right) = 0.274,$$

$$E_{r_2}(X) = -\frac{5}{9}\left(\frac{2}{9}\log\frac{1}{2} + \frac{2}{9}\log\frac{1}{2} + \frac{3}{9}\log\frac{1}{3} + \frac{2}{9}\log\frac{1}{2}\right) = 0.20,$$

$$E_{r_3}(X) = -\frac{5}{9}\left(\frac{1}{9}\log 1 + \frac{1}{9}\log 1 + \frac{1}{9}\log 1 + \frac{1}{9}\log 1 + \frac{3}{9}\log\frac{1}{3} + \frac{2}{9}\log\frac{1}{2}\right) = 0.048.$$

从上述结果可以看出，尽管 X 关于每个近似空间的粗糙度和正则条件熵是相同的，但是划分越细粗糙熵就越小.

§8.8 概率粗糙集模型和确定性粗糙集模型的比较

在信息系统中，论域 U 中的概念 X 用知识库(近似空间)的知识来近似刻画时一般有两种类型：

类型(1):知识库的知识是确定的,即论域 U 中所有对象的描述都是已知的.

类型(2):知识库的知识是不确定的,即论域 U 中所有对象的描述只有部分是已知的.

带有决策表的知识表示系统是类型(1)的一种典型应用,因为在这一种决策表中不但所有条件属性和决策属性都是已知的,而且所有对象的描述也是已知的.然而在很多实际应用中论域的对象的描述只有部分是已知的,即知识库中的知识是不确定的,它只能通过试验样本集 $E\subset U$ 所提供的信息来刻画概念 X.从训练样本集中获得广义决策规则的机器学习是类型(2)的一种典型应用.值得注意的是为了使从训练样本集获得的规则适用于整个论域的对象,一般要求样本应在论域中随机抽取,且有足够的数量以保证所作假设的统计可行性,即抽取样本时应符合统计规律性,当然这个工作属于统计学的范畴,粗糙集理论不涉及对这类问题处理的具体内容.从以上分析可知确定性粗糙集模型和概率粗糙集模型的基本差别在于基本知识$\{X_1,X_2,\cdots,X_n\}$用怎样的方法去刻画专家的概念 X.

在规则提取方面,确定性粗糙集模型对于有矛盾的事例会显得无能为力,因为它对概念的刻画仅限于三值决策(即"是","否","不知道"),即根据满足已知概念 X_i 的描述 $\mathrm{des}(X_i)$的研究对象是否满足概念 X 的描述 $\mathrm{des}(X)$来抽取规则的,其具体是

规则 1 若 $X_i\subseteq \mathrm{pos}(X)$,则 $\mathrm{des}(X_i)\rightarrow \mathrm{des}(X)$.

规则 2 若 $X_i\subseteq \mathrm{neg}(X)$,则 $\mathrm{des}(X_i)\rightarrow$非 $\mathrm{des}(X)$.

规则 3 若 $X_i=\mathrm{bn}(X)$,则 $\mathrm{des}(X_i)\rightarrow$不知道.

比如对于例 8.3 给出的决策表,我们要根据属性"内部设计"的值来刻画概念("加速性能"="好"),容易得到概念("加速性能"="好")关于由"内部设计"构成的近似空间$(U,\{a\})$的正域和负域都是空集,故只能获得如下规则:

规则 1 ("内部设计"="适中")→不知道.

规则 2 ("内部设计"="宽敞")→不知道.

规则 3 ("内部设计"="拥挤")→不知道.

从上面例子可知确定性粗糙集模型没有获得和利用近似边界区域所提供的统计信息,然而概率粗糙集模型能充分利用近似边界区域所提供的统计信息,并能给概念 X 一个更完整(带有某种确定度)的刻画,如以统计粗糙集模型(II)的绝对边界情形($\alpha=\beta=0.5$)为例,可以提取带有确定性因子 c 的决策规则如下:

规则 1 若 $P(X\mid X_i)>0.5$,则 $\mathrm{des}(X_i)\xrightarrow{c}\mathrm{des}(X)$.

规则 2 若 $P(X|X_i)<0.5$,则 $\mathrm{des}(X_i)\xrightarrow{c}$ 非 $\mathrm{des}(X)$.

规则 3 若 $P(X|X_i)=0.5$,则 $\mathrm{des}(X_i)\rightarrow$不知道.

其中规则中的确定性因子定义为 $c=\max\{P(X|X_i),1-P(X|X_i)\}$.

显然,能用确定性粗糙集方法作出的决策规则一定能用概率粗糙集方法得到同样的决策规则,反之则不然.当所有用概率粗糙集方法获得的规则的确定性因子 $c<1$ 时,确定性粗糙集方法都判断为"不知道".

如以前述"小汽车的决策表"为例,根据属性"内部设计"的值来刻画概念("加速性能"="好"),如用概率粗糙集方法来处理可获得如下决策规则:

规则 1 ("内部设计"="适中")$\xrightarrow{0.66}$ 非("加速性能"="好").

规则 2 ("内部设计"="宽敞")$\xrightarrow{0.66}$("加速性能"="好").

规则 3 ("内部设计"="拥挤")$\rightarrow$("加速性能"不知道).

由此可见,概率粗糙集方法虽不能说是万能的但至少可以说是确定性粗糙集模型的一个很好的补充与推广.

第九章　模糊粗糙集模型

模糊集理论是美国计算机与控制论专家 L. A. Zadeh 于 1965 提出来的. 目前,以模糊推理为核心的人工智能技术在许多领域取得了明显的成果和经济效益.

在本章前,各种粗糙集模型中所涉及的概念和知识都是清晰的,即所有的集合都是经典集合. 然而,在人们的实际生活中,涉及更多的是模糊概念和模糊知识. 反映在粗糙集模型中主要有两类,一类是知识库的知识是清晰的而被近似的概念是模糊的,另一类是知识库的知识和被近似的概念都是模糊的. 这一章我们讨论有关模糊概念近似的粗糙集模型.

§9.1　模糊集的基本概念

设 U 是由一些确定的可识别的对象构成的集合称为论域,对于 U 中任何一个集合 A,我们可以引进一个特征函数 $A(x)$,即

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

U 中特征函数是从 U 到 $\{0,1\}$ 的一个映射, U 中任何一个特征函数也完全确定了 U 中的一个经典子集合,即

$$A = \{x \in U : A(x) = 1\}.$$

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. 从集合论的角度来看,一个论域中的对象或属于这个集合或不属于这个集合,二者必居其一. 这样,经典集合可以用来描述清晰的概念和知识.

但是在实际问题中,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"就不能完全反映在二值逻辑中. 这里"年轻人"和"非年轻人","老年人"和"非老年人"之间并没有明确的界限,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,我们将特征函数扩充为隶属函数. 所谓论域 U 上的一个隶属函数是指 U 到 $[0,1]$ 的一个映射.

定义 9.1　论域 U 上的一个模糊集合(fuzzy set) A 是由 U 上的一个隶属函数

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

来表示,其中 $A(x)$(有时用 $\mu_A(x)$ 表示)表示元素 x 隶属于模糊集合 A 的程

度.

这样,对于论域 U 的一个对象 x 和 U 上的一个模糊集合 A,我们不能简单地问 x 是"绝对"属于还是不属于 A,而只能问 x 在多大程度上属于 A.隶属度 $A(x)$ 正是 x 属于 A 的程度的数量指标.若 $A(x)=0$,则认为 x 完全不属于 A;若 $A(x)=1$,则认为 x 完全属于 A;若 $0<A(x)<1$,则说 x 依程度 $A(x)$ 属于 A,这时在完全属于和完全不属于 A 的元素之间呈现出一种中间的过渡状态.

一般地,一个模糊集合 A 可以表示为

$$A = \{(x, A(x)) ; x \in U\}.$$

如果论域 U 是有限集合或可数集合,那么可以表示为

$$A = \sum x_i / A(x_i).$$

如果论域 U 是无限不可数集合,那么可以表示为

$$A = \int x / A(x).$$

记 U 上的模糊集合全体为 $\mathscr{F}(U)$.

定义 9.2 设 $A, B \in \mathscr{F}(U)$,若对于任意的 $x \in U$,有 $A(x) \leqslant B(x)$,则称 A 含于 B 或 B 包含 A,记作 $A \subseteq B$.若 $A \subseteq B$,$B \subseteq A$ 同时成立,则称 A 与 B 相等,记作 $A=B$.空集 $\varnothing$ 表示隶属函数恒为 0 的模糊集合,U 表示隶属函数恒为 1 的模糊集合.记 $A \cup B$ 为模糊集合 A 与 B 的并,其隶属函数定义为

$$(A \cup B)(x) = A(x) \vee B(x) = \max\{A(x), B(x)\}.$$

记 $A \cap B$ 为模糊集合 A 与 B 的交,其隶属函数定义为

$$(A \cap B)(x) = A(x) \wedge B(x) = \min\{A(x), B(x)\}.$$

记 A^C 或 $\sim A$ 为 A 的补集,其隶属函数定义为

$$A^C(x) = 1 - A(x).$$

记 $A-B$ 为 A 与 B 的差集,其隶属函数定义为

$$(A - B)(x) = [A \cap (\sim B)](x) = A(x) \wedge (1 - B(x)).$$

模糊集合的交,并,补和包含依次表示了模糊概念的析取,合取,否定和蕴涵,这对于分析实际问题和理论研究有重要的意义.

容易证明模糊集合的运算满足下列性质:

(1)交换律 $A \cup B = B \cup A$, $A \cap B = B \cap A$.

(2)结合律 $(A \cup B) \cup C = A \cup (B \cup C)$,

$(A \cap B) \cap C = A \cap (B \cap C)$.

(3)分配律 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$,

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

(4)吸收律 $A \cup (A \cap B) = A$,

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

(5)对偶律 $(A\cap B)^C=A^C\cup B^C$,

$(A\cup B)^C=A^C\cap B^C$.

(6)对合律 $(A^C)^C=A$.

(7)幂等律 $A\cup A=A\cap A=A$.

(8)两极律 $U\cap A=\varnothing\cup A=A$,

$U\cup A=U,\varnothing\cap A=\varnothing$.

以上运算性质对于经典集合全部成立,但是经典集合中的互补律对于模糊集合一般不再成立.例如取 $U=[0,1]$,$A(x)=x$ 则 $A\cup A^C\neq U$,$A\cap A^C\neq\varnothing$.

模糊集合运算不满足互补律,给研究工作带来了一定困难,而正是由于这一性质,它更能客观地反映实际情况.

对于 $A\in\mathscr{F}(U)$和 $\lambda\in[0,1]$,记

$$A_\lambda=\{x\in U\mid A(x)\geqslant\lambda\},\quad A_\lambda^S=\{x\in U\mid A(x)>\lambda\},$$

A_λ 和A_λ^S 分别称为A 的λ 截集和强λ 截集.当 $\lambda=1$ 时,A_1 称为 A 的核,而集合$\{x\in U|A(x)>0\}$称为 A 的支集记为 $\mathrm{supp}A$.

在处理实际问题时,有时要对模糊概念有个明确的认识与判定,要判断某个对象对模糊集合的明确的归属,这样就要求模糊集合与经典集合之间依一定的法则进行转化,模糊集的截集是解决这个问题的一种比较令人满意的手段.

分解定理 设 $A\in\mathscr{F}(U)$,则 $A=\bigcup\{\lambda A_\lambda|\lambda\in[0,1]\}$,其中 λA_λ 称为λ 与A_λ 的模糊截积,其隶属函数定义为

$$(\lambda A_\lambda)(x)=\lambda\wedge A_\lambda(x).$$

定义 9.3 设 U 是论域,$A\in\mathscr{F}(U)$,则模糊集 A 的基数$|A|$定义如下:

$$|A|=\sum_{x\in U}A(x).$$

§9.2 模糊关系

定义 9.4 设 U,V 是两个论域,若 R 是$U\times V$ 上的一个模糊集,则称R 是从U 到V 的一个模糊关系.特别当 $U=V$ 时,称 R 是U 上的一个模糊关系.

若 U 和 V 都是有限集,则类似于普通关系可以与布尔矩阵建立一一对应关系,模糊关系可以与模糊矩阵建立一一对应关系,这里模糊矩阵是指矩阵中的任何元素都属于$[0,1]$.以后我们只考虑有限论域,并且将模糊关系与模

糊矩阵视为同一.

模糊关系作为定义在 $U\times V$ 上的模糊集,也有交,并,补和模糊截积等运算,这只要将相应的模糊矩阵中的对应元素作 $\wedge,\vee,1-,\lambda\wedge$ 等运算即可,这里不再详细展开了.

定义 9.5 设 U,V,W,是三个论域,R 是从 U 到 V 的一个模糊关系,S 是从 V 到 W 的一个模糊关系,则 R 对 S 的合成是从 U 到 W 的一个模糊关系,记为 $R\circ S$,其隶属函数定义为

$$(R\circ S)(x,z)=\sup\{R(x,y)\wedge S(y,z)\mid y\in V\},(x,z)\in U\times W.$$

在模糊关系 R 中,每对元素(x,y)都对应有一个介于 0 与 1 之间的数 $R(x,y)$,表示 x 对 y 有关系 R 的程度,或称为 x 对 y 的关于关系 R 的相关程度.

定义 9.6 设 U 是论域,R 是 U 上的模糊关系,其隶属函数为 $R(x,y)$,我们称 R 是自反的,若对于任意的 $x\in U$ 有 $R(x,x)=1$;称 R 是对称的,若对于任意的 $x,y\in U$ 有 $R(x,y)=R(y,x)$;称 R 是传递的,若 $R\circ R\subseteq R$;称 R 是等价的,若 R 是自反,对称和传递的模糊关系;称 R 是相似的,若 R 是自反对称的模糊关系.

可以证明一个模糊关系的 λ 截集是一个普通的二元关系,模糊关系与对应普通关系有下列关系:

定理 9.1 设 R 是 U 上的模糊关系,则 R 是自反的当且仅当对于任意的 $\lambda\in[0,1]$,R_λ 是自反的普通关系;R 是对称的当且仅当对于任意的 $\lambda\in[0,1]$,R_λ 是对称的普通关系;R 是传递的当且仅当对于任意的 $\lambda\in[0,1]$,R_λ 是传递的普通关系;R 是等价的当且仅当对于任意的 $\lambda\in[0,1]$,R_λ 是等价的普通关系;R 是相似的当且仅当对于任意的 $\lambda\in[0,1]$,R_λ 是相似的普通关系.

§9.3 模糊粗糙集

在 Pawlak 粗糙集模型中,论域 U 上任意一个经典集合 A 不一定能用知识库(U,R)中的知识来精确地描述,这时就用 A 关于(U,R)的一对上下近似来描述.但在实际生活中,人们涉及到的知识或概念往往是模糊的不确定的,即 A 是 U 上的一个模糊集合,现在的问题是:A 如何用(U,R)中的知识来描述?模糊粗糙集模型就是针对这类问题提出来的.

定义 9.7 设(U,R)是 Pawlak 近似空间,即 R 是论域 U 上的一个等价关系.若 A 是 U 上的一个模糊集合,则 A 关于(U,R)的一对下近似 $\underline{A_R}$ 和上近似 $\overline{A_R}$ 定义为 U 上的一对模糊集合,其隶属函数分别定义为

$$\underline{A}_R(x)=\inf\{A(y)\mid y\in[x]_R\},x\in U,$$
$$\overline{A}_R(x)=\sup\{A(y)\mid y\in[x]_R\},x\in U,$$

其中$[x]_R$为元素x在关系R下的等价类.若$\underline{A}_R=\overline{A}_R$,则称$A$是可定义的,否则称$A$是模糊粗糙集(Fuzzy rough set).称$\underline{A}_R$是A关于(U,R)的正域,称$\sim\overline{A}_R$是A关于(U,R)的负域,称$\overline{A}_R\cap(\sim\underline{A}_R)$为$A$的边界.

以后如果关系R比较明确,我们将下标R省掉.我们知道,在Pawlak近似空间(U,R)中,属于同一等价类中的两个对象是不可分辨的,从上面的定义可以看出,$\underline{A}$和$\overline{A}$中的同一等价类中的隶属函数都是常数,这符合直观的意义.$\underline{A}(x)$可理解为对象x肯定属于模糊集A的隶属程度;$\overline{A}(x)$可理解为x可能属于模糊集A的隶属程度.

可以验证,当A是U上的经典集合时,$\underline{A}$和$\overline{A}$就退化为A在Pawlak意义下关于(U,R)的下近似和上近似,因此定义9.7是Pawlak意义下的推广形式.

定理9.2 由定义9.7给出的下近似$\underline{A}$和上近似$\overline{A}$满足下列(对偶)性质:

(1) $\underline{A}\subseteq A\subseteq\overline{A}$.

(2) $\overline{A\cup B}=\overline{A}\cup\overline{B},\underline{A\cap B}=\underline{A}\cap\underline{B}$.

(3) $\underline{A}\cup\underline{B}\subseteq\underline{A\cup B},\overline{A\cap B}\subseteq\overline{A}\cap\overline{B}$.

(4) $\underline{\sim A}=\sim\overline{A},\overline{\sim A}=\sim\underline{A}$.

(5) $\overline{(\overline{A})}=\underline{(\overline{A})}=\overline{A},\underline{(\underline{A})}=\overline{(\underline{A})}=\underline{A}$.

(6) $\underline{U}=U,\overline{\varnothing}=\varnothing$.

(7) 若$A\subseteq B$,则$\underline{A}\subseteq\underline{B}$,且$\overline{A}\subseteq\overline{B}$.

证 我们只证明一半性质,其余一半完全类似.

(1)显然.

(2)由于

$$\begin{aligned}\overline{A\cup B}(x)&=\sup\{(A\cup B)(y)\mid y\in[x]\}\\&=\sup\{A(y)\vee B(y)\mid y\in[x]\}\\&=\sup\{A(y)\mid y\in[x]\}\vee\sup\{B(y)\mid y\in[x]\}\\&=\overline{A}(x)\vee\overline{B}(x)=(\overline{A}\cup\overline{B})(x),\forall x\in U,\end{aligned}$$

因此$\overline{A\cup B}=\overline{A}\cup\overline{B}$.

$$\begin{aligned}(3)(\underline{A}\cup\underline{B})(x)&=\underline{A}(x)\vee\underline{B}(x)\\&=\inf\{A(y)\mid y\in[x]\}\vee\inf\{B(y)\mid y\in[x]\}\\&\leqslant\inf\{A(y)\vee B(y)\mid y\in[x]\}\\&=\inf\{(A\cup B)(y)\mid y\in[x]\}\end{aligned}$$

$$=(\underline{A}\cup\underline{B})(x),\ \forall x\in U,$$

所以 $\underline{A}\cup\underline{B}\subseteq\underline{A\cup B}$.

(4)由于 $\underline{\sim A}(x)=\inf\{(\sim A)(y)\mid y\in[x]\}$

$$\begin{aligned}&=\inf\{1-A(y)\mid y\in[x]\}\\&=1-\sup\{A(y)\mid y\in[x]\}\\&=1-\overline{A}(x)=(\sim\overline{A})(x),\ \forall x\in U,\end{aligned}$$

所以 $\underline{\sim A}=\sim\overline{A}$.

(5)由于 $y\in[x]\Leftrightarrow[x]=[y]$,因此

$$\begin{aligned}\overline{(\overline{A})}(x)&=\sup\{\overline{A}(y)\mid y\in[x]\}\\&=\sup\{\sup\{A(z)\mid z\in[y]\}\mid y\in[x]\}\\&=\sup\{\sup\{A(z)\mid z\in[x]\}\mid y\in[x]\}\\&=\inf\{\sup\{A(z)\mid z\in[x]\}\mid y\in[x]\}\\&=\underline{(\overline{A})}(x)=\sup\{A(z)\mid z\in[x]\}=\overline{A}(x),\ \forall x\in U,\end{aligned}$$

从而 $\overline{(\overline{A})}=\underline{(\overline{A})}=\overline{A}$.

(6)由于对于任意的 $x\in U$,$U(x)=1$,所以

$$\underline{U}(x)=\inf\{U(y)\mid y\in[x]\}=1,$$

因此 $\underline{U}=U$.

(7)显然. □

定义 9.8 设(U,R)是 Pawlak 近似空间,A,B 是 U 上的模糊集合,称 A 与 B 是模糊粗下相等的,若 $\underline{A}=\underline{B}$,记作 $A\eqsim B$;称 A 与 B 是模糊粗上相等的,若 $\overline{A}=\overline{B}$,记作 $A\simeq B$;称 A 与 B 是模糊粗相等的,若 $\underline{A}=\underline{B}$ 且 $\overline{A}=\overline{B}$,记作 $A\approx B$.

易见,对于 U 上的等价关系 R,$\eqsim$,$\simeq$ 和 $\approx$ 都是 $\mathscr{F}(U)$上的等价关系.以后对于 $\eqsim$,$\simeq$ 和 $\approx$,我们都是指在某个特定的等价关系 R 下的.

定理 9.3 设(U,R)是近似空间,则在 $\mathscr{F}(U)$中下列性质成立:

(1) $A\eqsim B$ 当且仅当 $A\cap B\eqsim A$,且 $A\cap B\eqsim B$.

(2) $A\simeq B$ 当且仅当 $A\cup B\simeq A$,且 $A\cup B\simeq B$.

(3) 若 $A\simeq A'$,且 $B\simeq B'$,则 $A\cup B\simeq A'\cup B'$.

(4) 若 $A\eqsim A'$,且 $B\eqsim B'$,则 $A\cap B\eqsim A'\cap B'$.

(5) 若 $A\subseteq B$,且 $B\simeq\varnothing$,则 $A\simeq\varnothing$.

(6) 若 $A\subseteq B$,且 $A\eqsim U$,则 $B\eqsim U$.

(7) 若 $A\eqsim\varnothing$,或 $B\eqsim\varnothing$,则 $A\cap B\eqsim\varnothing$.

(8) 若 $A\simeq U$ 或 $B\simeq U$,则 $A\cup B\simeq U$.

(9) $A\eqsim U$ 当且仅当 $A=U$.

(10) $A\simeq\varnothing$ 当且仅当 $A=\varnothing$.

证 由定义直接可得. □

定理 9.4 设(U,R)是近似空间,$A\in\mathscr{F}(U)$,则

(1) $\underline{A}=\bigcap\{B\in\mathscr{F}(U)\mid B\eqsim A\}$.

(2) $\overline{A}=\bigcup\{B\in\mathscr{F}(U)\mid B\simeq A\}$.

由定义 9.7 给出的下近似和上近似是一对模糊集合,如果我们将模糊集合 A 用知识库(U,R)中的经典集合来描述,则可以通过模糊集的截集来过渡.

定义 9.9 设(U,R)是 Pawlak 近似空间,A 是 U 上的模糊集合,则 A 关于近似空间(U,R)依参数 $0<\beta\leqslant\alpha\leqslant 1$ 的下近似 $\underline{A}_\alpha$ 和上近似 $\overline{A}_\beta$ 分别定义为

$$\underline{A}_\alpha=\{x\in U\mid \underline{A}(x)\geqslant\alpha\},$$
$$\overline{A}_\beta=\{x\in U\mid \overline{A}(x)\geqslant\beta\}.$$

$\underline{A}_\alpha$ 可以解释为 U 中肯定属于模糊集 A 的隶属程度不小于 α 的那些对象的全体,$\overline{A}_\beta$ 可以解释为 U 中可能属于模糊集 A 的隶属程度不小于 β 的那些对象的全体. 显然

$$\underline{A}_\alpha=\bigcup\{[x]\mid \underline{A}(x)\geqslant\alpha\},$$
$$\overline{A}_\beta=\bigcup\{[x]\mid \overline{A}(x)\geqslant\beta\}.$$

这样,$\underline{A}_\alpha$ 又可以解释为 U 中肯定属于模糊集 A 的隶属程度不小于 α 的那些对象所在等价类的并集,$\overline{A}_\beta$ 可以解释为 U 中可能属于模糊集 A 的隶属程度不小于 β 的那些对象所在的等价类的并集.

注 1 当 A 是经典集合时,对于任意 $\alpha,\beta\in(0,1]$,$\underline{A}_\alpha$ 和 $\overline{A}_\beta$ 分别退化为 A 在 Pawlak 意义下关于(U,R)的下近似 $\underline{R}A$ 和上近似 $\overline{R}A$.

事实上,对于任意 $\alpha\in(0,1]$,由于 A 是经典集合,因此 $A(x)\in\{0,1\}$,于是

$$\begin{aligned}\underline{A}_\alpha&=\{x\in U\mid \underline{A}(x)=1\}\\&=\{x\in U\mid A(y)=1,\forall y\in[x]\}\\&=\{x\in U\mid [x]\subseteq A\}=\underline{R}A.\end{aligned}$$

同理可证,$\overline{A}_\beta=\overline{R}A$.

注 2 若(U,R,P)是概率近似空间,对于任意 $A\in 2^U$,若记

$$A(x)=\mu_A(x)=P(A\mid[x]),x\in U,$$

则 A 可看成是一个模糊集,并且当 $0<\beta\leqslant\alpha\leqslant 1$ 时,有

$$\underline{A}_\alpha=\underline{\mathrm{PIV}}_\alpha(A)=\{x\in U\mid P(A\mid[x])\geqslant\alpha\},$$
$$\overline{A}_\beta=\overline{\mathrm{PIV}}_\beta(A)=\{x\in U\mid P(A\mid[x])\geqslant\beta\}.$$

这时模糊集 A 关于近似空间(U,R)依参数 $0<\beta\leqslant\alpha\leqslant 1$ 的下近似 $\underline{A}_\alpha$ 和上近

似 $\overline{A}_\beta$ 就是经典集 A 关于概率近似空间 (U,R,P) 依参数 $0<\beta\leqslant\alpha\leqslant 1$ 的概率(IV)型下近似 $\underline{\mathrm{PIV}}_\alpha(A)$ 和上近似 $\overline{\mathrm{PIV}}_\beta(A)$，这说明定义 9.9 的近似算子是概率近似算子的一种推广形式，并且由此可见，类似于概率粗糙集模型有四种不同的定义形式，模糊集 A 关于近似空间 (U,R) 依参数 α,β 的下近似 $\underline{A}_\alpha$ 和上近似 $\overline{A}_\beta$ 还有其他三种不同的定义形式，类似的讨论留给读者.

定理 9.5 设 (U,R) 是近似空间，$A,B\in\mathscr{F}(U)$，则对于 $0<\beta\leqslant\alpha\leqslant 1$ 有

(1) $\overline{A\cup B}_\beta=\overline{A}_\beta\cup\overline{B}_\beta$，$\underline{A\cap B}_\alpha=\underline{A}_\alpha\cap\underline{B}_\alpha$.

(2) $\underline{A}_\alpha\cup\underline{B}_\alpha\subseteq\underline{A\cup B}_\alpha$，$\overline{A\cap B}_\beta\subseteq\overline{A}_\beta\cap\overline{B}_\beta$.

(3) 若 $A\subseteq B$，则 $\overline{A}_\beta\subseteq\overline{B}_\beta$，且 $\underline{A}_\alpha\subseteq\underline{B}_\alpha$.

(4) $\underline{A}_\alpha\subseteq\overline{B}_\beta$.

(5) $\overline{\sim A}_\beta=\sim\underline{A}^S_{1-\beta}$，$\underline{\sim A}_\alpha=\sim\overline{A}^S_{1-\alpha}$.

证 由定义和定理 9.2 直接证得. □

定义 9.10 设 (U,R) 是近似空间，$A\in\mathscr{F}(U)$，定义 A 关于 (U,R) 的粗糙度 $\rho_R(A)$ 为

$$\rho_R(A)=1-\frac{|\underline{A}|}{|\overline{A}|}.$$

当 $|\overline{A}|=0$ 时，约定 $\rho_R(A)=0$；称 $\eta_R(A)=\dfrac{|\underline{A}|}{|\overline{A}|}$ 为 A 关于 (U,R) 的近似精度.

显然，$0\leqslant\rho_R(A)\leqslant 1$，$0\leqslant\eta_R(A)\leqslant 1$，若 A 是可定义的，则有 $\rho_R(A)=0$，$\eta_R(A)=1$.

定义 9.11 设 (U,R) 是近似空间，$A\in\mathscr{F}(U)$，对于 $0<\beta\leqslant\alpha\leqslant 1$，定义 A 在近似空间 (U,R) 中关于参数 α,β 的粗糙度 $\rho_A^{\alpha,\beta}$ 如下

$$\rho_A^{\alpha,\beta}=1-\frac{|\underline{A}_\alpha|}{|\overline{A}_\beta|},$$

约定当 $\overline{A}_\beta=\varnothing$ 时，$\rho_A^{\alpha,\beta}=0$.

由定义可以直接得到下列性质.

定理 9.6 (1) $0\leqslant\rho_A^{\alpha,\beta}\leqslant 1$.

(2) 若 β 固定，则 $|\underline{A}_\alpha|$ 随 α 增加而减少，从而 $\rho_A^{\alpha,\beta}$ 随 α 增加而增加；若 α 固定，则 $|\overline{A}_\beta|$ 随 β 增加而减少，从而 $\rho_A^{\alpha,\beta}$ 随 β 增加而减少.

定理 9.7 (1) 若对于 (U,R) 中的任意等价类 $[x]$，都存在 $y\in[x]$ 使得 $A(y)<\alpha$，则 $\underline{A}_\alpha=\varnothing$，从而 $\rho_A^{\alpha,\beta}=1$；

(2) 若模糊集 A 在 (U,R) 的每个等价类中的隶属函数都是常数，则对于任意的 $\alpha\in(0,1]$，有 $\underline{A}_\alpha=\overline{A}_\alpha$，从而 $\rho_A^{\alpha,\alpha}-0$.

证　由定义直接可得. □

定理 9.8　若模糊集 A 的隶属函数恒为常数,即存在 $\delta>0$,使对于任意的 $x\in U$ 有 $A(x)=\delta$,则

(1)当 $0<\beta<\delta<\alpha\leqslant 1$ 时,有 $\rho_A^{\alpha,\beta}=1$.

(2)对于 $0<\beta\leqslant\alpha\leqslant 1$ 的其他情形,有 $\rho_A^{\alpha,\beta}=0$.

证　(1)当 $0<\beta<\delta<\alpha\leqslant 1$ 时,显然 $\underline{A}_\alpha=\varnothing$,$\overline{A}_\beta=U$,从而 $\rho_A^{\alpha,\beta}=1$;

(2)此时有两种情形:情形 1. $\delta<\beta\leqslant\alpha$,这时 $\underline{A}_\alpha=\overline{A}_\beta=\varnothing$,从而根据约定有 $\rho_A^{\alpha,\beta}=0$;情形 2. $\beta\leqslant\alpha\leqslant\delta$,这时 $\underline{A}_\alpha=\overline{A}_\beta=U$,从而根据定义有 $\rho_A^{\alpha,\beta}=0$. □

定理 9.9　设 $A,B\in\mathscr{F}(U)$,且 $A\subseteq B$,$0<\beta\leqslant\alpha\leqslant 1$,则

(1)如果 $\overline{A}_\beta=\overline{B}_\beta$,那么 $\rho_B^{\alpha,\beta}\leqslant\rho_A^{\alpha,\beta}$.

(2)如果 $\underline{A}_\alpha=\underline{B}_\alpha$,那么 $\rho_A^{\alpha,\beta}\leqslant\rho_B^{\alpha,\beta}$.

(3)如果存在 $r>0$,使对于任意 $x\in U$ 有 $A(x)\geqslant r$,那么当 $\beta\geqslant r$ 时,有 $\rho_B^{\alpha,\beta}\leqslant\rho_A^{\alpha,\beta}$.

(4)如果存在 $r>0$,使对于任意 $x\in U$ 有 $A(x)\geqslant r$,那么当 $\alpha\leqslant r$ 时,有 $\rho_B^{\alpha,\beta}=\rho_A^{\alpha,\beta}=0$.

证　由 $A\subseteq B$ 可得 $\underline{A}_\alpha\subseteq\underline{B}_\alpha$ 和 $\overline{A}_\beta\subseteq\overline{B}_\beta$,从而(1)和(2)成立.

(3)由 $\beta\leqslant r$ 可得 $\overline{A}_\beta=\overline{B}_\beta=U$,从而 由(1)知 $\rho_B^{\alpha,\beta}\leqslant\rho_A^{\alpha,\beta}$.

(4)由 $\alpha\leqslant r$ 可得 $\underline{A}_\alpha=\underline{B}_\alpha=U$,且 $\overline{A}_\beta=\overline{B}_\beta=U$,从而 $\rho_B^{\alpha,\beta}=\rho_A^{\alpha,\beta}=0$. □

从上述定理可以知道由 $A\subseteq B$ 不能简单地判别 $\rho_A^{\alpha,\beta}\leqslant\rho_B^{\alpha,\beta}$ 或 $\rho_B^{\alpha,\beta}\leqslant\rho_A^{\alpha,\beta}$.

定理 9.10　设 $A,B\in\mathscr{F}(U)$,若 $A\approx B$,则对于任意 $0<\beta\leqslant\alpha\leqslant 1$ 有 $\rho_A^{\alpha,\beta}=\rho_B^{\alpha,\beta}$.

证　由定义直接可得. □

定理 9.11　设 $A,B\in\mathscr{F}(U)$,$0<\beta\leqslant\alpha\leqslant 1$,则有

$$\rho_{A\cup B}^{\alpha,\beta}|\overline{A}_\beta\cup\overline{B}_\beta|\leqslant\rho_A^{\alpha,\beta}|\overline{A}_\beta|+\rho_B^{\alpha,\beta}|\overline{B}_\beta|-\rho_{A\cap B}^{\alpha,\beta}|\overline{A}_\beta\cap\overline{B}_\beta|. \tag{9.1}$$

证　由定理 9.5 可得

$$\begin{aligned}\rho_{A\cup B}^{\alpha,\beta}&=1-\frac{|\underline{A\cup B}_\alpha|}{|\overline{A\cup B}_\beta|}\\&=1-\frac{|\underline{A\cup B}_\alpha|}{|\overline{A}_\beta\cup\overline{B}_\beta|}\leqslant 1-\frac{|\underline{A}_\alpha\cup\underline{B}_\alpha|}{|\overline{A}_\beta\cup\overline{B}_\beta|},\end{aligned} \tag{9.2}$$

同理可得

$$\rho_{A\cap B}^{\alpha,\beta}=1-\frac{|\underline{A\cap B}_\alpha|}{|\overline{A\cap B}_\beta|}$$

$$= 1 - \frac{|\underline{A}_\alpha \cap \underline{B}_\alpha|}{|\overline{A} \cap \overline{B}_\beta|} \leqslant 1 - \frac{|\underline{A}_\alpha \cap \underline{B}_\alpha|}{|\overline{A}_\beta \cap \overline{B}_\beta|}. \tag{9.3}$$

由于对于任意的有限集 X, Y 有

$$|X \cup Y| = |X| + |Y| - |X \cap Y|. \tag{9.4}$$

因此,利用(9.2),(9.3)和(9.4)有

$$\begin{aligned}
&\rho_{A\cup B}^{\alpha,\beta}|\overline{A}_\beta \cup \overline{B}_\beta| \leqslant |\overline{A}_\beta \cup \overline{B}_\beta| - |\underline{A}_\alpha \cup \underline{B}_\alpha| \\
&= |\overline{A}_\beta| + |\overline{B}_\beta| - |\overline{A}_\beta \cap \overline{B}_\beta| - |\underline{A}_\alpha| - |\underline{B}_\alpha| + |\underline{A}_\alpha \cap \underline{B}_\alpha| \\
&\leqslant |\overline{A}_\beta| + |\overline{B}_\beta| - |\underline{A}_\alpha| - |\underline{B}_\alpha| - \rho_{A\cap B}^{\alpha,\beta}|\overline{A}_\beta \cap \overline{B}_\beta|.
\end{aligned}$$

利用 $\rho_A^{\alpha,\beta}$ 和 $\rho_B^{\alpha,\beta}$ 的定义可知式(9.1)成立. □

设 S 是 U 上另一等价关系,仍记 $\underline{A}_\alpha$ 或 $\underline{A}_\alpha^R$,$\overline{A}_\beta$ 或 $\overline{A}_\beta^R$ 是 A 关于 (U,R) 的近似,记 $\underline{A}_\alpha^S$,$\overline{A}_\beta^S$ 是 A 关于 (U,S) 的近似,记 $\rho_R^{\alpha,\beta}(A)$ 和 $\rho_S^{\alpha,\beta}(A)$ 分别为 A 在 (U,R) 和 (U,S) 中关于 α,β 的粗糙度.

定理 9.12 若 $S \subseteq R$,$A \in \mathscr{F}(U)$,则

(1) $\underline{A}_R \subseteq \underline{A}_S$,$\overline{A}_S \subseteq \overline{A}_R$.

(2) $\underline{A}_\alpha^R \subseteq \underline{A}_\alpha^S$,$\overline{A}_\beta^S \subseteq \overline{A}_\beta^R$.

证 由于 $S \subseteq R$ 当且仅当 $[x]_S \subseteq [x]_R$,$\forall x \in U$,于是

(1)对于 $\forall x \in U$,有

$$\begin{aligned}
\underline{A}_R(x) &= \inf\{A(y) \mid y \in [x]_R\} \\
&\leqslant \inf\{A(y) \mid y \in [x]_S\} = \underline{A}_S(x),
\end{aligned}$$

从而 $\underline{A}_R \subseteq \underline{A}_S$. 同理,$\overline{A}_S \subseteq \overline{A}_R$.

(2)若 $x \in \underline{A}_\alpha$,则 $\underline{A}_R(x) \geqslant \alpha$,从而由(1)得

$$\underline{A}_S(x) \geqslant \underline{A}_R(x) \geqslant \alpha,$$

这样就得到了 $x \in \underline{A}_\alpha^S$,于是 $\underline{A}_\alpha^R \subseteq \underline{A}_\alpha^S$. 同理可得 $\overline{A}_\beta^S \subseteq \overline{A}_\beta^R$. □

定理 9.13 若 $S \subseteq R$,$A \in \mathscr{F}(U)$,则

(1) $\rho_S^{\alpha,\beta}(A) \leqslant \rho_R^{\alpha,\beta}(A)$.

(2) $\rho_S(A) \subseteq \rho_R(A)$.

证 由定义和定理 9.2 即得. □

注 定理 9.13 说明,划分越细,所得近似的粗糙度就越小.

例 9.1 设 $U = \{x_i | 1 \leqslant i \leqslant 8\}$ 为某一组被研究的八名学生,他们被分成三个部分记为

$$U/R = \{\{x_1, x_5\}, \{x_2, x_7, x_8\}, \{x_3, x_4, x_6\}\}.$$

设模糊集合 A 表示模糊概念"个子高",其隶属函数为

$$A = \{x_1/0.5, x_2/0.3, x_3/0.3, x_4/0.6, x_5/0.5,$$

$$x_6/0.8, x_7/1, x_8/0.9\}.$$

则

$$\underline{A}_R = \{x_1/0.5, x_2/0.3, x_3/0.3/x_4/0.3, x_5/0.5, x_6/0.3, x_7/0.3, x_8/0.3\},$$

$$\overline{A}_R = \{x_1/0.5, x_2/1, x_3/0.8, x_4/0.8, x_5/0.5, x_6/0.8, x_7/1, x_8/1\}.$$

近似精度 $\eta_R(A) = \dfrac{|\underline{A}_R|}{|\overline{A}_R|} = 0.431.$

若取 $\alpha < 0.3$，此时每一部分(等价类)的最小隶属度都大于 α，易见 $\rho_R^{\alpha,\beta}(A) = 0$，即模糊概念"个子高"关于阈值 α, β 的粗糙度达到最小值 0.若取 $\alpha > 0.9$，此时每一部分(等价类)的最小隶属度都小于 α，易见 $\rho_R^{\alpha,\beta}(A) = 1$，即模糊概念"个子高"关于阈值 α, β 的粗糙度达到最大值 1.

若我们将这个小组再分得细一些，分成四组，记

$$U/S = \{\{x_1, x_5\}, \{x_2\}, \{x_7, x_8\}, \{x_3, x_4, x_6\}\},$$

则

$$\underline{A}_S = \{x_1/0.5, x_2/0.3, x_3/0.3, x_4/0.3, x_5/0.5, x_6/0.3, x_7/0.9, x_8/0.9\},$$

$$\overline{A}_S = \{x_1/0.5, x_2/0.3, x_3/0.8, x_4/0.8, x_5/0.5, x_6/0.8, x_7/1, x_8/1\}.$$

易见 $\underline{A}_R \subseteq \underline{A}_S \subseteq \overline{A}_S \subseteq \overline{A}_R$.

这时可以验算模糊概念"个子高"关于近似空间 (U, R) 的粗糙度大于关于近似空间 (U, S) 的粗糙度，即划分越细粗糙度越小，从而近似精度越高.

§9.4 基于三角模的模糊粗糙集模型

在 Pawlak 粗糙集模型中知识库中的知识都是清晰的，即近似空间中的集合都是经典集合.但是在实际问题中知识库的知识往往也是模糊的，下面我们就来讨论模糊集合用知识库中的模糊知识来近似的粗糙集模型.

定义 9.12 记 $I = [0,1]$，二元函数 $T: I \times I \to I$，若满足下列条件：

(1)两极律：$T(a,1) = a, a \in I$.

(2)交换律：$T(a,b) = T(b,a), a, b \in I$.

(3)结合律：$T(T(a,b),c) = T(a, T(b,c)), a, b, c \in I$.

(4)单调律：$a \leqslant c, b \leqslant d \Rightarrow T(a,b) \leqslant T(c,d), a, b, c, d \in I$.

则称 T 为 I 上的三角模或 T 模.

定义 9.13 设 T 是 I 上的下半连续三角模，定义 I 上的二元算子 θ_T：$I \times I \to I$ 如下：

$$\theta_T(a,b)=\sup\{c\in I\mid T(a,c)\leqslant b\},\quad a,b\in I,$$

称 θ_T 为 T 的剩余蕴涵(residuation implication of T).

例 9.2 设 $a,b\in I$,记

$$T_1(a,b)=a\wedge b,$$
$$T_2(a,b)=a\cdot b,$$
$$T_3(a,b)=0\vee(a+b-1),$$

则 $T_i,i=1,2,3$,都是 I 上的三角模,容易验证

$$\theta_{T_1}(a,b)=\begin{cases}1, & a\leqslant b,\\ b, & a>b.\end{cases}$$

$$\theta_{T_2}(a,b)=\begin{cases}1, & a\leqslant b,\\ b/a, & a>b.\end{cases}$$

$$\theta_{T_3}(a,b)=1\wedge(1-a+b).$$

在不致引起混淆的情况下,我们将 θ_T 中的下标 T 省去.

对于 $X,Y\in\mathscr{F}(U)$,定义 U 上的模糊集 $T(X,Y)$,其隶属函数为

$$T(X,Y)(x)=T(X(x),Y(x)),\quad x\in U.$$

记 $\theta\langle X,Y\rangle=\bigwedge\limits_{u\in U}\theta(X(u),Y(u))$,于是可以验证

$$\theta\langle X,Y\rangle=\sup\{a\in I\mid T(X,\underline{a})\subseteq Y\},$$

其中 $\underline{a}(u)=a,\forall u\in U$.

定义 U 上的模糊子集 $\Theta(X,Y)$,其隶属函数为

$$\Theta(X,Y)(x)=\theta(X(x),Y(x)),\quad x\in U.$$

定义 U 上的模糊集 1_x,其隶属函数为 $1_x(y)=\begin{cases}1, & y=x,\\ 0, & y\neq x.\end{cases}$ $x,y\in U$. 则可验证

$$\Theta(1_x,\underline{a})=1_{U\setminus\{x\}}\vee\underline{a},$$

其中 1_Z 为集合 $Z\subseteq U$ 的特征函数.

可以证明,关于剩余蕴涵 θ 下述定理成立.

定理 9.14 设 T 是 I 上的下半连续三角模,T 的剩余蕴含 θ 满足下列性质:设 $a,b,c\in I$,则

(θ1) $\theta(a,1)=1,\theta(1,a)=a$.

(θ2) $a\leqslant b\Rightarrow\theta(c,a)\leqslant\theta(c,b)$.

(θ3) $a\leqslant b\Rightarrow\theta(a,c)\geqslant\theta(b,c)$.

(θ4) $T(\theta(a,c),\theta(c,b))\leqslant\theta(a,b)$.

(θ5) $\theta(a\vee b,c)=\theta(a,c)\wedge\theta(b,c)$.

(θ6) $\theta(a,b\wedge c)=\theta(a,b)\wedge\theta(a,c)$.

$(\theta 7)$ $a \leqslant b \Leftrightarrow \theta(a,b)=1$.

$(\theta 8)$ $\theta(a,\theta(b,c))=\theta(b,\theta(a,c))$.

$(\theta 9)$ $\theta(T(a,b),c)=\theta(a,\theta(b,c))$.

$(\theta 10)$ $T(\theta(T(a,b),c),a)\leqslant\theta(b,c)$.

$(\theta 11)$ $\bigwedge\limits_{a\in I}\theta(T(b,\theta(c,a)),a)=\theta(b,c)$.

$(\theta 12)$ $\theta(\theta(a,b),b)\geqslant a$.

$(\theta 13)$ $\bigwedge\limits_{b\in I}\theta(\theta(a,b),b)=a$.

$(\theta 14)$ $T(\theta(a,b),c)\leqslant\theta(a,T(b,c))$.

$(\theta 15)$ $\bigwedge\limits_{b\in I}\theta(\theta(a,b),\theta(c,b))=\theta(c,a)$.

$(\theta 16)$ $\theta(a,b)\leqslant\theta(T(a,c),T(b,c))$.

$(\theta 17)$ $\theta(a,b\vee c)\geqslant\theta(a,b)\vee\theta(a,c)$.

$(\theta 18)$ $a\leqslant\theta(b,T(a,b))$.

定义 9.14　设 T 是 I 上的下半连续三角模，有限论域 U 上二元模糊关系 R 称为 T 相似关系，若 R 是自反，对称和 T 传递(即 $T(R(x,z),R(x,y))\leqslant R(z,y)$, $x,y,z\in U$)的.

对于 U 上的模糊集 A，记 $\|A\|=\sup\limits_{x\in U}A(x)$.

定义 9.15　U 上的模糊子集类 $\Phi=\{A_1,A_2,\cdots,A_k\}\subset\mathscr{F}(U)$ 称为 U 的一个模糊 T 划分，若下列条件满足:

(1)所有 A_i 都是正则的，即 $\|A_i\|=1$.

(2) $\forall x\in U$，存在惟一 $A_i\in\Phi$ 使 $A_i(x)=1$.

(3)若有 $A_i,A_j\in\Phi$ 使 $A_i(x)=A_j(y)=1$，则

$$A_j(x)=A_i(y)=\|T(A_i,A_j)\|.$$

对于 $x\in U$，若 $A_i\in\Phi$ 使 $A_i(x)=1$，则记 $A_i=[x]_\Phi$ 或简记为 $[x]$，称为 x 的模糊 T 相似类.

上述 $[x]_\Phi$ 类似于普通二元等价关系下的 x 所在的等价类. 易见，当 A_i 为经典集时，Φ 构成了 U 上的一个普通划分. 正如论域 U 上的普通二元等价关系与普通划分可以建立一一对应，容易证明二元模糊 T 相似关系与模糊 T 划分也可以建立一一对应，即下述定理成立

定理 9.15　设 Φ 为 U 上一个模糊 T 划分，定义 U 上的一个二元模糊关系 R_Φ，其隶属函数为

$$R_\Phi(x,y)=[x]_\Phi(y)=\|T([x]_\Phi,[y]_\Phi)\|,\quad x,y\in U,$$

则 R_Φ 是 U 上的模糊 T 相似关系. 反之，若 R 是 U 上的模糊 T 相似关系，定义

$$\Phi_R=\{R\langle x\rangle\mid x\in U\},$$

其中 $R\langle x\rangle$ 是由 R 产生的定义在 U 上的一个模糊子集，其隶属函数为

$$R\langle x\rangle(y) = R(x,y), \quad y \in U,$$

则 Φ_R 是 U 上的一个模糊 T 划分.

注 若 U 上的模糊 T 划分中的集合都是经典集，则这时的划分恰好是普通的划分，而对应的模糊 T 相似关系也正是普通的二元等价关系.

定义 9.16 设 R 是有限论域 U 上的二元模糊 T 相似关系，称 (U,R) 是模糊 T 近似空间，对于任意 $X\in\mathcal{F}(U)$，X 关于近似空间 (U,R) 的下近似 $\underline{A}(X)$（或 $\underline{A}X$）和上近似 $\overline{A}(X)$（或 $\overline{A}X$）是定义在 U 上的一对模糊集，其隶属函数定义为

$$\underline{A}(X)(x) = \theta\langle R\langle x\rangle, X\rangle = \bigwedge_{u\in U}\theta(R(u,x),X(u)), x \in U,$$

$$\overline{A}(X)(x) = \| T(R\langle x\rangle, X)\| = \bigvee_{u\in U} T(R(u,x),X(u)), x \in U.$$

称 $(\underline{A}(X),\overline{A}(X))$ 为模糊 T 粗糙集，而 $\underline{A}$ 和 $\overline{A}$ 分别称为模糊 T 下近似算子和模糊 T 上近似算子.

定理 9.16 模糊 T 下近似算子满足下列性质：设 $X,Y\in\mathcal{F}(U)$，$x,y\in U$，$a,b\in I$ 则

(1) $\underline{A}(X)\subseteq X$.

(2) $\underline{A}(\underline{A}(X))=\underline{A}(X)$.

(3) $\underline{A}(X\cap Y)=\underline{A}(X)\cap\underline{A}(Y)$.

(4) $\underline{A}(\Theta(1_x,\underline{a}))(y)=\underline{A}(\Theta(1_y,\underline{a}))(x)=\theta(R(x,y),a)$，特别地 $\underline{A}(\theta(1_x,\underline{a}))(x)=a$.

(5) $\underline{A}\Theta(\underline{a},X))=\Theta(\underline{a},\underline{A}(X))$.

(6) $\underline{A}\Theta(X,\underline{a})\subseteq\Theta(X,\underline{a})\subseteq\Theta(\underline{A}X,\underline{a})$.

(7) $X\subseteq Y\Rightarrow\underline{A}X\subseteq\underline{A}Y$.

(8) $\underline{A}\underline{a}=\underline{a}$.

(9) $\underline{A}((\underline{a})\vee 1_Z)(x)=\bigwedge_{u\notin Z}\theta(R(u,x),a), \forall Z\subseteq U$.

(10) $\underline{A}(\underline{\theta(a,b)}\vee 1_Z)=\Theta(\underline{a},\underline{A}(\underline{b}\vee 1_Z)), \forall Z\subseteq U$.

(11) $\underline{A}(X\cup Y)\supseteq\underline{A}(X)\cup\underline{A}(Y)$.

(12) $\bigwedge_{a\in I}\theta(\underline{A}(\Theta(1_x,\underline{a}))(y),a)=R(x,y)$.

(13) $\underline{A}[x]_R=[x]_R$.

(14) $\inf\{\theta(\underline{A}(X)(x),\underline{A}(X)(y))\mid X\in\mathcal{F}(U)\}=R(x,y)$.

(15) $\inf\{\underline{A}(X)(x)\mid x\in U\}=\inf\{X(x)\mid x\in U\}$.

证 (1)由定理 9.14 性质(θ1)知

$$\underline{A}(X)(x) = \bigwedge_{u\in U}\theta(R(u,x),X(u))$$

$$\leqslant \theta(R(x,x),X(x))$$
$$= R(1,X(x)) = X(x),$$

因此(1)成立.

$$(2)\underline{A}(\underline{A}(X))(x) = \bigwedge_{u\in U}\theta(R(u,x),\underline{A}(X)(u))$$
$$= \bigwedge_{u\in U}\theta(R(u,x),\bigwedge_{v\in U}\theta(R(v,u),X(v)))$$
(由(θ6)) $$= \bigwedge_{u,v\in U}\theta(R(u,x),\theta(R(v,u),X(v)))$$
(由(θ9)) $$= \bigwedge_{u,v\in U}\theta(T(R(v,u),R(u,x)),X(v))$$
(由 R 的 T 传递性和性质(θ3))
$$\geqslant \bigwedge_{v\in U}\theta(R(v,x),X(v)) = \underline{A}(X)(x),$$

因此(2)成立.

$$(3)\underline{A}(X\cap Y)(x) = \bigwedge_{u\in U}\theta(R(u,x),X(u)\wedge Y(u))$$
(由(θ6)) $$= \bigwedge_{u\in U}[\theta(R(u,x),X(u))\wedge\theta(R(u,x),Y(u))]$$
$$= (\bigwedge_{u\in U}\theta(R(u,x),X(u)))$$
$$\wedge(\bigwedge_{u\in U}\theta(R(u,x),Y(u))$$
$$= (\underline{A}(X)(x))\wedge(\underline{A}(Y)(x))$$
$$= [(\underline{A}(X))\cap(\underline{A}(Y))](x),$$

因此(3)成立.

(4)首先,由性质(θ7)和(θ1)得

$$\underline{A}(\Theta(1_x,\underline{a}))(y) = \bigwedge_{u\in U}\theta(R(u,y),\Theta(1_x,\underline{\underline{a}})(u))$$
$$= \bigwedge_{u\in U}\theta(R(u,y),\theta(1_x(u),a))$$
$$= \theta(R(x,y),\theta(1,a))\wedge\bigwedge_{u\neq x}\theta(R(u,y),\theta(0,a))$$
$$= \theta(R(x,y),a)\wedge\bigwedge_{u\neq x}\theta(R(u,y),1)$$
$$= \theta(R(x,y),a)\wedge 1 = \theta(R(x,y),a),$$

从而由上式和 R 的对称性知(4)成立.

(5)由性质(θ8)和(θ6)得

$$\underline{A}(\Theta(\underline{\underline{a}},X))(x) = \bigwedge_{u\in U}\theta(R(u,x),\Theta(\underline{\underline{a}},X)(u))$$
$$= \bigwedge_{u\in U}\theta(R(u,x),\theta(a,X(u)))$$
$$= \bigwedge_{u\in U}\theta(a,\theta(R(u,x),X(u)))$$
$$= \theta(a,\bigwedge_{u\in X}\theta(R(u,x),X(u)))$$
$$= \theta(a,\underline{A}(X)(x)) = \Theta(\underline{\underline{a}},\underline{A}X)(x),$$

因此(5)成立.

(6)由(1)和性质(θ3)即得.

(7)可由(3)简单推得.

(8)$(\underline{A}\underset{=}{a})(x)=\bigwedge_{u\in U}\theta(R(u,x),a)=\theta(\bigvee_{u\in U}R(u,x),a)=\theta(1,a)=a$.

(9)$\underline{A}(\underset{=}{a}\vee 1_Z)(x)=\bigwedge_{u\in U}\theta(R(u,x),a\vee 1_Z(u))$

$$=\bigwedge_{u\in Z}\theta(R(u,x),a\vee 1)\wedge\bigwedge_{u\notin Z}\theta(R(u,x),a)$$

$$=\bigwedge_{u\notin Z}\theta(R(u,x),a).$$

(10)$\underline{A}(\underline{\theta(a,b)}\vee 1_Z)(x)$

$$=\bigwedge_{u\in U}\theta(R(u,x),\theta(a,b)\vee 1_Z(u))$$

$$=\bigwedge_{u\in Z}\theta(R(u,x),1)\wedge\bigwedge_{u\notin Z}\theta(R(u,x),\theta(a,b))$$

$$=1\wedge\bigwedge_{u\notin Z}\theta(R(u,x),\theta(a,b))$$

$$=\theta(a,\bigwedge_{u\notin Z}\theta(R(u,x),b))=\theta(a,\underline{A}(\underset{-}{b}\vee 1_Z)(x)),$$

因此(10)成立.

(11)可由性质(3)简单推得.

(12)由(4)知

$$\bigwedge_{a\in I}\theta(\underline{A}(\Theta(1_x,\underset{=}{a}))(y),a)=\bigwedge_{a\in I}\theta(\theta(R(x,y),a),a),$$

从而由性质(θ13)知(12)成立.

(13)由 R 的 T 相似性和性质(θ18)可得

$$\begin{aligned}\underline{A}[x]_R(y)&=\bigwedge_{u\in U}\theta(R(u,y),[x]_R(u))\\&=\bigwedge_{u\in U}\theta(R(u,y),R(x,u))\\&\geqslant\bigwedge_{u\in U}\theta(R(u,y),T(R(x,y),R(u,y)))\\&\geqslant R(x,y)=[x]_R(y),\end{aligned}\tag{9.5}$$

从而结合上式和(1)可知(13)成立.

(14)首先,由性质(θ15)可得

$$\inf\{\theta(\underline{A}(X)(x),\underline{A}(X)(y))\mid X\in\mathscr{F}(U)\}$$

$$=\inf\{\theta(\bigwedge_{u\in U}\theta(R(u,x),X(u)),\bigwedge_{v\in U}\theta(R(v,y),X(v)))\mid X\in\mathscr{F}(U)\}$$

$$\geqslant\inf\{\bigwedge_{v\in U}\bigvee_{u\in U}\theta(\theta(R(u,x),X(u)),\theta(R(v,y),X(v)))\mid X\in\mathscr{F}(U)\}$$

$$\geqslant\inf\{\bigwedge_{v\in U}\theta(\theta(R(v,x),X(v)),\theta(R(v,y),X(v)))\mid X\in\mathscr{F}(U)\}$$

$$=\bigwedge_{v\in U}\theta(R(v,y),R(v,x)),$$

从而由式(9.5)知

$$\inf\{\theta(\underline{A}(X)(x),\underline{A}(X)(y))\mid X\in\mathscr{F}(U)\}\geqslant R(x,y).\tag{9.6}$$

其次,由于$[x]_R\in\mathscr{F}(U)$,因此利用(13)和性质(θ2)得

$$\begin{aligned}&\inf\{\theta(\underline{A}(X)(x),\underline{A}(X)(y))\mid X\in\mathcal{F}(U)\}\\ \leqslant\ &\theta(\underline{A}[x]_R(x),\underline{A}[x]_R(y))=\theta([x]_R(x),[x]_R(y))\\ =\ &\theta(1,R(x,y))=R(x,y),\end{aligned}\tag{9.7}$$

由(9.6)和(9.7)得(14)成立.

(15)令 $a=\inf\{X(x)\mid x\in U\}$,则 $\underset{=}{a}\subseteq X$,从而由(8)知

$$\underset{=}{a}=\underline{A}\underset{=}{a}\subseteq\underline{A}X\subseteq X,$$

于是

$$a\leqslant\inf\{\underline{A}X(x)\mid x\in U\}\leqslant\inf\{X(x)\mid x\in U\}=a.$$

即(15)成立. □

定理 9.17 模糊 T 上近似算子满足下列性质:设 $X,Y\in\mathcal{F}(U)$,$x,y\in U$,$a,b\in I$,则

(1) $\overline{A}X\supseteq X$.

(2) $\overline{A}(\overline{A}(X))=\overline{A}X$.

(3) $\overline{A}(X\cup Y)=\overline{A}X\cup\overline{A}Y$.

(4) $\overline{A}(1_x)(y)=\overline{A}(1_y)(x)=R(x,y)$.

(5) $\overline{A}(T(\underset{=}{a},X))=T(\underset{=}{a},\overline{A}X)$.

(6) $\overline{A}\Theta(X,\underset{=}{a})\supseteq\Theta(X,\underset{=}{a})\supseteq\Theta(\overline{A}X,\underset{=}{a})$.

(7) $X\subseteq Y\Rightarrow\overline{A}X\subseteq\overline{A}Y$.

(8) $\overline{A}\underset{=}{a}=\underset{=}{a}$.

(9) $\overline{A}(1_x)=\overline{A}[x]_R=[x]_R$.

(10) $\overline{A}(X\cap Y)\subseteq\overline{A}X\cap\overline{A}Y$.

(11) $\overline{A}(X)(x)=\|T(\overline{A}(1_x),X)\|$.

(12) $\overline{A}(1_Z)(z)=\sup\limits_{u\in Z}R(x,u)$,$Z\subseteq U$.

(13) $\|\overline{A}X\|=\|X\|$.

证 (1)由定义和 R 的自反性知

$$\overline{A}(X)(x)\geqslant T(R(x,x),X(x))=X(x),$$

从而(1)成立.

(2)由 T 的下半连续性和结合律知

$$\begin{aligned}\overline{A}(\overline{A}(X))(x)&=\sup_{u\in U}T(R(u,x),\overline{A}X(u))\\&=\bigvee_{u\in U}\bigvee_{v\in U}T(R(u,x),T(R(v,u),X(v)))\\ \text{(由 }R\text{ 的 }T\text{ 传递性)}\quad&\leqslant\bigvee_{v\in U}T(R(v,x),X(v))=\overline{A}X(x),\end{aligned}$$

从而 $\overline{A}(\overline{A}(X))\subseteq\overline{A}X$,于是再结合(1)可知(2)成立.

(3)因为

$$\begin{aligned}\overline{A}(X\cup Y)(x)&=\bigvee_{u\in U}T(R(u,x),X(u)\vee Y(u))\\&=\bigvee_{u\in U}[T(R(u,x),X(u))\vee T(R(u,x),Y(u))]\\&=[\bigvee_{u\in U}T(R(u,x),X(u))]\vee[\bigvee_{u\in U}T(R(u,x),Y(u))]\\&=\overline{A}X(x)\vee\overline{A}Y(x)=(\overline{A}X\cup\overline{A}Y)(x),\end{aligned}$$

因此(3)成立.

(4) $\overline{A}(1_x)(y)=\sup_{u\in U}T(R(u,x),1_x(u))=R(x,y)$

从而由 R 的对称性知(4)成立.

(5)
$$\begin{aligned}\overline{A}(T(\underset{=}{a},X))&=\bigvee_{u\in U}T(R(u,x),T(a,X(u)))\\&=T(a,\bigvee_{u\in U}T(R(u,x),X(u)))\\&=T(a,\overline{A}X(x))=T(\underset{=}{a},\overline{A}X)(x),\end{aligned}$$

因此(5)成立.

(6)由(1)和性质(θ2)即得.

(7)可由(3)简单推得.

(8)由 T 的下半连续性可知

$$\begin{aligned}\overline{A}\underset{=}{a}(x)&=\bigvee_{u\in U}T(R(u,x),a)\\&=T(\bigvee_{u\in U}R(u,x),a)=T(1,a)=a,\end{aligned}$$

从而 $\overline{A}\underset{=}{a}=\underset{=}{a}$成立.

(9)由 R 的自反性和 T 传递性

$$\begin{aligned}\overline{A}[x]_R(y)&=\bigvee_{u\in U}T(R(u,y),[x]_R(u))\\&=\bigvee_{u\in U}T(R(u,y),R(x,u))=R(x,y)=[x]_R(y).\end{aligned}$$

又由(4)知

$$\overline{A}(1_x)(y)=R(x,y)=[x]_R(y),$$

从而可得(9)成立.

(10)可由(3)简单推得.

(11)由(9)知

$$\begin{aligned}\|T(\overline{A}(1_x),X)\|&=\bigvee_{u\in U}T(\overline{A}(1_x)(u),X(u))\\&=\bigvee_{u\in U}T([x]_R(u),X(u))\\&=\bigvee_{u\in U}T(R(x,u),X(u))=\overline{A}(X)(x).\end{aligned}$$

(12)设 $Z\subseteq U$,则

$$\overline{A}(1_Z)(x)=\sup_{u\in U}T(R(u,x),1_Z(u))=\sup_{u\in Z}R(x,u).$$

(13)令 $a=\|X\|$，于是 $X\subseteq \underline{\underline{a}}$，从而由(1)和(8)得

$$X\subseteq \overline{A}X\subseteq \overline{A}\underline{\underline{a}}=\underline{\underline{a}},$$

因此 $a=\|X\|\leqslant\|\overline{A}X\|\leqslant a$，即(13)成立. □

定理 9.18　模糊 T 上近似和下近似满足下列性质：设 $X,Y\in\mathscr{F}(U)$，$x,y\in U$，$a,b\in I$，则

(1) $\overline{A}\underline{A}X=\underline{A}X$.

(2) $\underline{A}\overline{A}X=\overline{A}X$.

(3) $\underline{A}X=\bigwedge_{a\in I}\Theta(\overline{A}\Theta(X,\underline{\underline{a}}),\underline{\underline{a}})$.

(4) $\overline{A}X=\bigwedge_{a\in I}\Theta(\underline{A}\Theta(X,\underline{\underline{a}}),\underline{\underline{a}})$.

(5) $\underline{A}\Theta(X,\underline{a})=\Theta(\overline{A}X,\underline{a})$.

(6) $\underline{A}\Theta(\underline{\underline{a}},\overline{A}X)=\Theta(\underline{\underline{a}},\overline{A}X)$.

(7) $\underline{A}\Theta(\overline{A}X,\underline{\underline{a}})=\Theta(\overline{A}X,\underline{\underline{a}})$.

(8) $\theta\langle X,\underline{A}Y\rangle=\theta\langle\overline{A}X,Y\rangle$.

(9) $\theta\langle\underline{A}X,\underline{A}Y\rangle=\theta\langle\underline{A}X,Y\rangle\geqslant\theta\langle X,Y\rangle$.

(10) $\theta\langle\overline{A}X,\overline{A}Y\rangle=\theta\langle X,\overline{A}Y\rangle\geqslant\theta\langle X,Y\rangle$.

(11) $\underline{A}X(x)=\sup\{a\in I\mid \underline{A}(\Theta(\underline{\underline{a}},X))(x)=1\}$.

(12) $\underline{A}X=X\Leftrightarrow\overline{A}X=X$.

证　(1)
$$\begin{aligned}\overline{A}\underline{A}X(x)&=\bigvee_{u\in U}T(R(u,x),\underline{A}X(u))\\&=\bigvee_{u\in U}T(R(u,x),\bigwedge_{v\in U}\theta(R(v,u),X(v)))\\&\leqslant\bigvee_{u\in U}\bigwedge_{v\in U}T(R(u,x),\theta(R(v,u),X(v)))\end{aligned}$$

(由 R 的 T 传递性，T 的单调律和(θ3))

$$\leqslant\bigvee_{u\in U}\bigwedge_{v\in U}T(R(u,x),\theta(T(R(v,x),R(x,u)),X(v)))$$

(由(θ10))　$$\leqslant\bigwedge_{v\in U}\theta(R(v,x),X(v))=\underline{A}X(x),$$

因此 $\overline{A}\underline{A}X\leqslant\underline{A}X$，从而再结合定理 9.17 的(1)便得结论(1).

(2)
$$\begin{aligned}\underline{A}\overline{A}X(x)&=\bigwedge_{u\in U}\theta(R(u,x),\overline{A}X(u))\\&=\bigwedge_{u\in U}\theta(R(u,x),\bigvee_{v\in U}T(R(v,u),X(v)))\end{aligned}$$

(由(θ17))　$$\geqslant\bigwedge_{u\in U}\bigvee_{v\in U}\theta(R(u,x),T(R(v,u),X(v)))$$

(由 R 的 T 传递性，(θ2)和 T 的结合律)

$$\geqslant\bigwedge_{u\in U}\bigvee_{v\in U}\theta(R(u,x),T(T(R(v,x),R(x,u)),X(v)))$$

(由(θ18))　$$\geqslant\bigvee_{v\in U}T(R(v,x),X(v))=\overline{A}X(x),$$

因此 $\underline{A}\overline{A}X \geqslant \overline{A}X$，从而结合定理 9.16 的性质(1)便知(2)成立.

(3)由于

$$
\begin{aligned}
\bigwedge_{a\in I}\Theta(\overline{A}\Theta(X,a),\underline{a})(x) &= \bigwedge_{a\in I}\theta(\overline{A}\Theta(X,\underline{\underline{a}})(x),a) \\
&= \bigwedge_{a\in I}\theta(\bigvee_{u\in U}T(R(u,x),\theta(X(u),a)),a) \\
(\text{由}(\theta5)) \quad &= \bigwedge_{a\in I}\bigwedge_{u\in U}\theta(T(R(u,x),\theta(X(u),a)),a) \\
(\text{由}(\theta9)) \quad &= \bigwedge_{u\in U}\bigwedge_{a\in I}\theta(R(u,x),\theta(\theta(X(u),a),a)) \\
(\text{由}(\theta6)) \quad &= \bigwedge_{u\in U}\theta(R(u,x),\bigwedge_{a\in I}\theta(\theta(X(u),a),a)) \\
(\text{由}(\theta13)) \quad &= \bigwedge_{u\in U}\theta(R(u,x),X(u)) = \underline{A}X(x),
\end{aligned}
$$

因此(3)成立.

(4)由于

$$
\begin{aligned}
\bigwedge_{a\in I}\Theta(\underline{A}\Theta(X,\underline{\underline{a}}),\underline{a})(x) &= \bigwedge_{a\in I}\theta(\underline{A}\Theta(X,\underline{\underline{a}})(x),a) \\
&= \bigwedge_{a\in I}\theta(\bigwedge_{u\in U}\theta(R(u,x),\theta(X(u),a)),a) \\
(\text{由}(\theta9)) \quad &= \bigwedge_{a\in I}\theta(\bigwedge_{u\in U}\theta(T(R(u,x),X(u)),a),a) \\
(\text{由}(\theta5)) \quad &= \bigwedge_{a\in I}\theta(\theta(\bigvee_{u\in U}T(R(u,x),X(u)),a),a) \\
(\text{由}(\theta13)) \quad &= \bigvee_{u\in U}T(R(u,x),X(u)) = \overline{A}X(x),
\end{aligned}
$$

从而结论(4)成立.

$$
\begin{aligned}
(5)\underline{A}\Theta(X,\underline{a})(x) &= \bigwedge_{u\in U}\theta(R(u,x),\theta(X(u),a)) \\
(\text{由}(\theta9)) \quad &= \bigwedge_{u\in U}\theta(T(R(u,x),X(u)),a) \\
(\text{由}(\theta5)) \quad &= \theta(\bigvee_{u\in U}T(R(u,x),X(u)),a) \\
&= \theta(\overline{A}X(x),a) = \Theta(\overline{A}X,\underline{\underline{a}})(x),
\end{aligned}
$$

从而结论(5)成立.

(6)由定理 9.16 和(2)知

$$\underline{A}\Theta(\underline{\underline{a}},\overline{A}X) = \Theta(\underline{\underline{a}},\underline{A}\overline{A}X) = \Theta(\underline{\underline{a}},\overline{A}X).$$

(7)由(5)知

$$\underline{A}\Theta(\overline{A}X,\underline{a}) = \Theta(\overline{A}(\overline{A}X),\underline{\underline{a}}) = \Theta(\overline{A}X,\underline{a}).$$

$$
\begin{aligned}
(8)\theta\langle X,\underline{A}Y\rangle &= \bigwedge_{u\in U}\theta(X(u),\underline{A}Y(u)) \\
&= \bigwedge_{u\in U}\theta(X(u),\bigwedge_{v\in U}\theta(R(v,u),Y(v))) \\
(\text{由}(\theta6)\text{和}(\theta9)) \quad &= \bigwedge_{u\in U}\bigwedge_{v\in U}\theta(T(R(v,u),X(u)),Y(v)) \\
&= \bigwedge_{v\in U}\theta(\bigvee_{u\in U}T(R(v,u),X(u)),Y(v)) \\
&= \bigwedge_{v\in U}\theta(\overline{A}X(v),Y(v)) = \theta\langle\overline{A}X,Y\rangle.
\end{aligned}
$$

(9)由结论(8),(3)和(θ3)得

$$\theta\langle \underline{A}X,\underline{A}Y\rangle=\theta\langle \overline{A}\underline{A}X,Y\rangle=\theta\langle \underline{A}X,Y\rangle\geqslant\theta\langle X,Y\rangle.$$

(10)由结论(8),(2)和(θ2)得

$$\theta\langle \overline{A}X,\overline{A}Y\rangle=\theta\langle X,\underline{A}\overline{A}Y\rangle\geqslant\theta\langle X,\overline{A}Y\rangle=\theta\langle X,Y\rangle.$$

(11) $\sup\{a\in I\mid \underline{A}\Theta(\underline{a},X)(x)=1\}=\sup\{a\in I\mid\theta(a,\underline{A}X(x))=1\}$

(由(θ7)) $=\sup\{a\in I\mid a\leqslant\underline{A}X(x)\}=\underline{A}X(x)$.

(12)若 $\underline{A}X=X$,则 $\overline{A}X=\overline{A}\underline{A}X=\underline{A}X=X$. 反之,若 $\overline{A}X=X$,则

$$\underline{A}X=\underline{A}\overline{A}X=\overline{A}X=X.$$ □

定理 9.19 设 R 是 U 上的一个经典 T 相似关系(即普通等价关系),则模糊 T 近似算子具有下列表示形式:

(1)对于任意模糊集 $X\in\mathscr{F}(U)$和 $x\in U$

$$\underline{A}X(x)=\inf\{X(u)\mid u\in[x]\},$$
$$\overline{A}X(x)=\sup\{X(u)\mid u\in[x]\}.$$

(2)对于任意经典集 $X\subseteq U$

$$\underline{A}1_X=\{x\in U\mid[x]\subseteq X\},$$
$$\overline{A}1_X=\{x\in U\mid[x]\cap X\neq\varnothing\}.$$

证 (1)对于任意模糊集 $X\in\mathscr{F}(U)$和 $x\in U$,由(θ3)和(θ1)得

$$\begin{aligned}\underline{A}X(x)&=\bigwedge_{u\in U}\theta(R(u,x),X(u))\\&=\bigwedge_{u\in U}\theta(1_{[x]}(u),X(u))\\&=\bigwedge_{u\in[x]}\theta(1,X(u))\\&=\bigwedge_{u\in[x]}X(u)=\inf\{X(u)\mid u\in[x]\}.\end{aligned}$$

同理,由 T 的定义得

$$\begin{aligned}\overline{A}X(x)&=\sup\{T(R(u,x),X(u))\mid u\in U\}\\&=\sup\{T(1_{[x]}(u),X(u))\mid u\in U\}\\&=\sup\{T(1,X(u))\mid u\in[x]\}\\&=\sup\{X(u)\mid u\in[x]\}.\end{aligned}$$

(2)由(1)直接可得. □

注 上述定理表明,前一节的粗糙模糊近似算子和 Pawlak 近似算子都是模糊 T 近似算子的特殊情形.

§9.5 基于包含度的粗糙集模型

我们知道在知识表示系统中,论域中的概念是用知识库中的知识来描述的,在决策表中还可以提取决策规则. 从协调的决策表中可以抽取确定性规

则，而从不协调的决策表中只能抽取不确定性规则或可能性规则，这是因为在不协调的系统中存在着矛盾的事例. 这种从决策表中抽取规则的推理的实质是一种广义的包含关系. 这一节我们就利用包含度的概念来定义模糊粗糙近似算子.

设 U 是有限非空集合，$P(U)$ 表示 U 中经典集合的全体，$\mathscr{F}(U)$ 表示 U 中模糊集合全体.

定义 9.17 设 $\mathscr{F}_0(U)\subseteq\mathscr{F}(U)$，若对于任意 $A,B\in\mathscr{F}_0(U)$ 有数 $D(B/A)$ 对应，且满足：

(1) $0\leqslant D(B/A)\leqslant 1$,

(2) $\forall A,B\in\mathscr{F}_0(U),A\subseteq B\Rightarrow D(B/A)=1$,

(3) 对于 $\forall A,B,C\in\mathscr{F}_0(U),A\subseteq B\subseteq C\Rightarrow D(A/C)\leqslant D(A/B)$,

则称 D 为 $\mathscr{F}_0(U)$ 上的包含度(inclusion degree).

称 D 为 $\mathscr{F}_0(U)$ 上的强包含度，若 D 满足(1)，(2)(3)和以下的(4)：

(4) 对于 $\forall A,B,C\in\mathscr{F}_0(U),A\subseteq B\Rightarrow D(A/C)\leqslant D(B/C)$.

称 D 为 $\mathscr{F}_0(U)$ 上的弱包含度，若 D 满足(1)，(3)和以下的(2)′：

(2)′ 对于 $\forall A,B\in\mathscr{F}_0(U)\cap P(U),A\subseteq B\Rightarrow D(B/A)=1$.

例 9.3 设 U 是有限非空集合，记

$$D(B/A)=\frac{|A\cap B|}{|A|},\quad A,B\in P(U),$$

其中 $|A|$ 表示集合 A 的基数(即元素个数)，约定当 $A=\varnothing$ 时，$D(B/A)=1$，则 D 为 $P(U)$ 上的强包含度.

例 9.4 设 U 是有限非空集合，P 为定义在 $P(U)$ 上的概率测度，记

$$D(B/A)=P(B\mid A)=\frac{P(A\cap B)}{P(A)},\quad A,B\in P(U),$$

则 D 为 $P(U)$ 上的包含度.

例 9.5 设 $\mathscr{F}^*(U)$ 表示 U 上的正则模糊集全体，在 $\mathscr{F}^*(U)$ 上定义

$$\Pi(B/A)=\sup_{x\in U}(A(x)\wedge B(x)),$$

$$N(B/A)=\inf_{x\in U}(A^C(x)\vee B(x)),$$

则 Π 为 $\mathscr{F}^*(U)$ 上的强包含度，N 为 $\mathscr{F}^*(U)$ 上的弱包含度.

例 9.6 设 $A,B\in\mathscr{F}(U)$，记 $n(A,B)=\{x\in\operatorname{supp}B\mid B(x)\leqslant A(x)\}$，定义

$$D(A/B)=\frac{|n(A,B)|}{|\operatorname{supp}B|},\quad B\neq\varnothing,$$

当 $B=\varnothing$ 时，约定 $D(A/B)=1$，则 D 是 $\mathscr{F}(U)$ 上的包含度.

定理 9.20 设 T 是 $I=[0,1]$ 上的三角模，θ 为其对应的剩余蕴涵，则

$D(b/a)=\theta(a,b)$是 I 上的包含度.

证 由定理 9.14 和 D 的定义即知. □

定义 9.18 设 U 是有限非空论域，$A_i \in \mathscr{F}(U)$，$1\leqslant i\leqslant k$，称 $\mathscr{A}=\{A_1, A_2,\cdots,A_k\}$为 U 的一个弱模糊划分，若

$$\bigcup_{1\leqslant i\leqslant k}\mathrm{supp}A_i = U.$$

称 $\mathscr{A}=\{A_1,A_2,\cdots,A_k\}$为 U 的一个正则弱模糊划分，若下列条件满足：

(1) A_i 为 U 上的正则模糊集，$1\leqslant i\leqslant k$.

(2) $\bigcup\limits_{1\leqslant i\leqslant k}\mathrm{supp}A_i = U$.

(3) $i\neq j\Rightarrow \sup\limits_{x\in U}(A_i\cap A_j)(x)<1$.

定义 9.19 设 $\mathscr{A}=\{A_1,A_2,\cdots,A_k\}$为 U 上的弱模糊划分，D 为$\mathscr{F}(U)$上的包含度，称 $\mathrm{A}=(U,\mathscr{A},D)$为模糊包含近似空间，对于任意 $X\in\mathscr{F}(U)$，X 关于 A 依参数 $0\leqslant\beta<\alpha\leqslant 1$ 的下近似 $\underline{A}_\alpha(X)$和上近似 $\overline{A}_\beta(X)$是定义在 U 上的一对模糊子集：

$$\underline{A}_\alpha(X) = \bigcup\{A_i \mid D(X/A_i)\geqslant\alpha\},$$
$$\overline{A}_\beta(X) = \bigcup\{A_i \mid D(X/A_i) > \beta\}.$$

X 关于 A 依参数 α,β 的正域，边界和负域分别为

$$\mathrm{pos}(X,\alpha,\beta)=\underline{A}_\alpha(X) = \bigcup\{A_i \mid D(X/A_i)\geqslant\alpha\},$$
$$\mathrm{bn}(X,\alpha,\beta)=\bigcup\{A_i \mid \beta< D(X/A_i) < \alpha\},$$
$$\mathrm{neg}(X,\alpha,\beta)=\bigcup\{A_i \mid D(X/A_i)\leqslant\beta\}.$$

显然有

$$\overline{A}_\beta(X) = \mathrm{pos}(X,\alpha,\beta)\cup\mathrm{bn}(X,\alpha,\beta).$$

注 1 当 A_i 为经典集，且 $\mathscr{A}=\{A_1,A_2,\cdots,A_k\}$构成 U 的划分时. 如果取包含度

$$D(A/B) = \frac{|A\cap B|}{|B|}, A,B\subseteq U.$$

令 $\alpha = 1,\beta = 0$，则

$$\begin{aligned}\underline{A}_\alpha(X) &= \bigcup\{A_i \mid D(X/A_i)\geqslant 1\}\\ &= \bigcup\{A_i \mid D(X/A_i) = 1\}\\ &= \bigcup\{A_i \mid A_i\subseteq X\},\\ \overline{A}_\beta(X) &= \bigcup\{A_i \mid D(X/A_i) > 0\}\\ &= \bigcup\{A_i \mid X\cap A_i\neq\varnothing\}.\end{aligned}$$

这时下近似 $\underline{A}_\alpha(X)$和上近似 $\overline{A}_\beta(X)$分别退化为 Pawlak 意义下的下近似和上近似.

注 2 当 A_i 为经典集,且 $\mathscr{A}=\{A_1,A_2,\cdots,A_k\}$ 构成 U 的划分时,若 A 为概率近似空间,并且取包含度

$$D(A/B)=P(A\mid B)=\frac{P(A\cap B)}{P(B)},\quad A,B\subseteq U,$$

则这时下近似 $\underline{A}_\alpha(X)$ 和上近似 $\overline{A}_\beta(X)$ 分别退化为概率粗糙集模型(I)意义下的下近似和上近似.

当 $\underline{A}_\alpha(X)=\overline{A}_\beta(X)$ 时,或等价地当 $\mathrm{bn}(X,\alpha,\beta)=\varnothing$ 时,称 X 依参数 α,β 关于 A 是可定义,否则称 X 依参数 α,β 关于 A 是不可定义的或称为包含度模糊粗糙集.

类似于经典集情形,对于参数 α,β,包含度模糊粗糙集 X 可分为下列四类:

(1) 若 $\underline{A}_\alpha(X)\neq\varnothing,\overline{A}_\beta(X)\neq U$,称 X 是部分可定义的.

(2)若 $\underline{A}_\alpha(X)\neq\varnothing,\overline{A}_\beta(X)=U$,称 X 是内部可定义的.

(3)若 $\underline{A}_\alpha(X)=\varnothing,\overline{A}_\beta(X)\neq U$,称 X 是外部可定义的.

(4)若 $\underline{A}_\alpha(X)=\varnothing,\overline{A}_\beta(X)=U$,称 X 是完全不可定义的.

定理 9.21 设 $0\leqslant\beta<\alpha\leqslant1$,$D$ 为 $\mathscr{F}(U)$ 上强包含度,则包含度近似算子满足下列性质:

(1) $\underline{A}_\alpha(X)\subseteq\overline{A}_\beta(X),X\in\mathscr{F}(U)$.

(2) $X,Y\in\mathscr{F}(U),X\subseteq Y\Rightarrow\underline{A}_\alpha(X)\subseteq\underline{A}_\alpha(Y),\overline{A}_\beta(X)\subseteq\overline{A}_\beta(Y)$.

(3) $\overline{A}_\beta(X\cup Y)\supseteq\overline{A}_\beta(X)\cup\overline{A}_\beta(Y),X,Y\in\mathscr{F}(U)$,

$\underline{A}_\alpha(X\cap Y)\subseteq\underline{A}_\alpha(X)\cap\underline{A}_\alpha(Y);X,Y\in\mathscr{F}(U)$.

(4) $\overline{A}_\beta(X\cap Y)\subseteq\overline{A}_\beta(X)\cap\overline{A}_\beta(Y),X,Y\in\mathscr{F}(U)$,

$\underline{A}_\alpha(X\cup Y)\supseteq\underline{A}_\alpha(X)\cup A_\alpha(Y),X,Y\in\mathscr{F}(U)$.

(5) 若 $\alpha_1\leqslant\alpha_2,\beta_1\leqslant\beta_2,X\in\mathscr{F}(U)$,则

$$\underline{A}_{\alpha_2}(X)\subseteq\underline{A}_{\alpha_1}(X),\quad\overline{A}_{\beta_2}(X)\subseteq\overline{A}_{\beta_1}(X).$$

证 (1)由定义可知显然成立.

(2)若 $A_i\subseteq\underline{A}_\alpha(X)$,则由定义知 $D(X/A_i)\geqslant\alpha$,因为 $X\subseteq Y$,且由于 D 是强包含度,从而 $D(Y/A_i)\geqslant D(X/A_i)\geqslant\alpha$,于是 $A_i\subseteq\underline{A}_\alpha(Y)$,因此

$$\underline{A}_\alpha(X)\subseteq A_\alpha(Y).$$

同理可证 $\overline{A}_\beta(X)\subseteq\overline{A}_\beta(Y)$ 成立.

(3)和(4)可由(2)直接推得.

(5)由定义即得. □

§9.6 修正型模糊粗糙集模型

在 Pawlak 粗糙集模型中,集合 $X\subseteq U$ 关于近似空间的下近似$\underline{R}X$ 和上近似 $\overline{R}X$ 满足关系式

$$\underline{R}X \subseteq X \subseteq \overline{R}X.$$

如果将上述三个集合都看成是模糊集,则有

$$\mu_{\underline{R}X}(x) \leqslant \mu_X(x) \leqslant \mu_{\overline{R}X}(x), x \in U. \tag{9.8}$$

易见§9.3 介绍的模糊粗糙集一般不满足上述性质,模糊 T 近似算子虽满足(9.8)但对知识库中的知识要求较高.在实际生活中,模糊概念常常用知识库中模糊知识来描述,且知识库中的所有模糊知识不一定构成论域 U 的弱模糊划分.本节给出修正型粗糙集模型就是针对这类问题.

定义 9.20 设 U 是有限论域,$\mathscr{A}=\{X_1,X_2,\cdots,X_k\}$为 U 中的模糊子集类,且满足 $U=\mathrm{supp}\bigcup\limits_{1\leqslant i\leqslant k}X_i$,$D$ 为 $\mathscr{F}(\mathrm{U})$上的包含度,称$(\mathrm{U},\mathscr{A},D)$为模糊包含近似空间.对于 $\lambda\in(0,1]$和 $Y\in\mathscr{F}(U)$,称 Y 在近似空间$(U,\mathscr{A},D)$中是 λ 可近似的,若

$$D(Y/\bigcap_{X_i\in\mathscr{A}}X_i) \wedge D(\bigcup_{X_i\in\mathscr{A}}X_i/Y) \geqslant \lambda.$$

称 $T_{\mathscr{A}}=\dfrac{|\bigcup\limits_{1\leqslant i\leqslant k}X_i-\bigcap\limits_{1\leqslant i\leqslant k}X_i|}{|U|}$为近似容度(approximation capacity).

定义 9.21 设 $Y\in\mathscr{F}(U)$在近似空间$(U,\mathscr{A},D)$中是 λ 可近似的,记

$$L(Y)=\{\Sigma\subseteq\mathscr{A}\mid D(Y/\bigcap_{X_i\in\Sigma}X_i)\geqslant\lambda\},$$

$$U(Y)=\{\Lambda\subseteq\mathscr{A}\mid D(\bigcup_{X_i\in\Lambda}X_i/Y)\geqslant\lambda\}.$$

取 $\mathscr{A}^*\in L(Y),\mathscr{B}^*\in U(Y)$使

$$|\bigcup_{X_i\in\mathscr{B}^*}X_i-\bigcap_{X_i\in\mathscr{A}^*}X_i|=\min_{\mathscr{A}\in L(Y),\mathscr{B}\in U(Y)}|\bigcup_{X_i\in\mathscr{B}}X_i-\bigcap_{X_i\in\mathscr{A}}X_i|,$$

则 Y 关于近似空间$(U,\mathscr{A},D)$的修正模糊下近似和修正模糊上近似定义为

$$\underline{A}_\lambda(Y)=\bigcap_{X_i\in\mathscr{A}^*}X_i,$$

$$\overline{A}_\lambda(Y)=\bigcup_{X_i\in\mathscr{B}^*}X_i.$$

称$(\underline{A}_\lambda(Y),\overline{A}_\lambda(Y))$为修正模糊粗糙集(modified fuzzy rough set).称 $T_{\mathscr{A}}(Y)=\dfrac{|\overline{A}_\lambda(Y)-\underline{A}_\lambda(Y)|}{|U|}$为 Y 的修正模糊近似容度.一般情况下,$T_{\mathscr{A}}$ 较大,而 $T_{\mathscr{A}}(Y)$却远远小于 $T_{\mathscr{A}}$.

注 1 由定义可以看出，Y 的修正模糊下近似是 $\mathscr{A}$ 中那些至少以程度 λ 包含于 Y 的模糊集的最大的交，而 Y 的修正模糊上近似是 $\mathscr{A}$ 中那些至少以程度 λ 包含 Y 的模糊集的最小的并.

注 2 由于满足定义 9.21 条件的 $\mathscr{A}^*$ 和 $\mathscr{B}^*$ 不一定惟一，因此 Y 的修正模糊粗糙集也不一定惟一.

注 3 当 $\lambda=1$ 时，修正模糊粗糙集满足(9.8)式.

如果我们在信息系统中的属性用模糊集来代替，而对象的属性值用对象的隶属度来代替，则可以得到模糊知识的表达系统，它类似于模糊关系数据库系统. 类似于信息系统的决策表，也存在着模糊知识表达系统的决策表，对应的模糊概念分别称为条件概念和决策概念. 我们也可以得到模糊决策规则：

在确定度 λ 下，对象 $x\in U$ 属于模糊集 Y 的隶属度至少为 $\min\{X_i(x)\mid X_i\in\mathscr{A}^*\}$，但不能超过 $\max\{X_i(x)\mid X_i\in\mathscr{B}^*\}$.

例 9.7 设论域 $U=\{1,2,\cdots,8\}$ 为八个信用卡申请者，每个申请者对应的六个模糊条件概念 $C_1,C_2,\cdots,C_6$ 和三个模糊决策概念 D_1,D_2,D_3 的隶属度都被明确地描述出来. 其中 C_1 = 低银行平衡，C_2 = 中等银行平衡，C_3 = 高银行平衡，C_4 = 月消费低，C_5 = 月消费中等，C_6 = 月消费高；D_1 = 令人满意的申请者，D_2 = 尚可的申请者，D_3 = 不令人满意的申请者. 隶属函数见表 9.1.

表 9.1 一个模糊决策表

U	C_1	C_2	C_3	C_4	C_5	C_6	D_1	D_2	D_3
1	0	0.8	0.3	0.7	0.2	0	0.5	0.7	0
2	1	0	0	0	0.1	0.8	0	0	1
3	0.6	0.3	0	0	1	0	0	0.5	0.7
4	0	0.5	0.6	0	0	1	0.6	0.5	0
5	0	1	0	0	0.7	0.5	0	1	0
6	0	0.1	0.8	1	0	0	0.8	0.4	0
7	0	0	1	0.4	0.6	0	1	0	0
8	1	0	0	0.5	0.5	0	0	0	1

由表可见，$U=\operatorname{supp}\bigcup_{1\leqslant i\leqslant 6}C_i$. 下面我们用修正模糊粗糙集方法求出决策概念 D_1,D_2 和 D_3 关于条件概念 $\mathscr{C}=\{C_i\mid 1\leqslant i\leqslant 6\}$ 的近似.

我们用例 9.6 的包含度来计算，即取

$$D(A/B)=\frac{|n(A,B)|}{|\operatorname{supp}B|},\quad B\neq\varnothing,A,B\in\mathscr{F}(U),$$

当 $B=\varnothing$ 时，约定 $D(A/B)=1$.

显然 $\bigcap_{1\leqslant i\leqslant 6}C_i=\varnothing$，$(\bigcup_{1\leqslant i\leqslant 6}C_i)(x_1)=0.8$，$(\bigcup_{1\leqslant i\leqslant 6}C_i)(x_j)=1,2\leqslant j\leqslant 6$. 因

此,对任意 D_j,有

$$D(D_j/\bigcap_{1\leqslant i\leqslant 6}C_i)=1,\quad D(\bigcup_{1\leqslant i\leqslant 6}C_i/D_j)=1.$$

于是表 9.1 中所有决策概念关于近似空间$(U,\mathscr{C},D)$都是 $\lambda=1$ 可近似的,近似容度 $T_{\mathscr{C}}=\frac{7.8}{8}=0.975$. 包含度 $D(D_j/C_i)$和 $D(C_i/D_j)$见表 9.2 和表 9.3.

表 9.2 包含度 $D(D_j/C_i)$

$\mathscr{C}$	D_1	D_2	D_3
C_1	0	0	1
C_2	0.4	0.8	0.2
C_3	1	0.25	0
C_4	0.25	0.25	0.25
C_5	0.33	0.33	0.33
C_6	0	0.33	0.33

表 9.3 包含度 $D(C_i/D_j)$

$\mathscr{C}$	D_1	D_2	D_3
C_1	0	0.2	0.66
C_2	0.25	0.6	0
C_3	0.75	0.4	0
C_4	0.5	0.4	0
C_5	0	0.2	0.33
C_6	0.25	0.2	0

概念 D_1 的近似 由于 C_3 是 $\mathscr{C}$ 中的惟一的条件概念使 $D(D_1/C_3)=1$,因此 $\underline{A}_\lambda(D_1)=C_3$. 由于 $C_i\in\mathscr{C}$ 中包含 D_1 的最大包含度是 $D(C_3/D_1)=0.75$. 又由于 $D(C_3\cup C_4/D_1)=1$, $D(C_3\cup C_2/D_1)=1$, $D(C_3\cup C_i/D_1)<1$, $i=1,5,6$,且 $|C_3\cup C_4-C_3|=1.1$, $|C_3\cup C_2-C_3|=1.8$,因此 $\overline{A}_\lambda(D_1)=C_3\cup C_4$. 于是对于任意 $x\in U$ 有

$$C_3(x)\leqslant D_1(x)\leqslant C_3(x)\vee C_4(x),\tag{9.9}$$

其中近似容度 $T_{\mathscr{C}}(D_1)=\frac{1.1}{8}=0.1375$.

式(9.9)的含义:我们可以预见,申请者 x 属于模糊集"令人满意的申请者"的隶属度至少要大于等于它属于模糊集"高银行平衡"的隶属度,但不能超过它属于模糊集"高银行平衡"和模糊集"月消费低"的隶属度的最大者.

概念 D_2 的近似 模糊集 C_2 是 $\mathscr{C}$ 中属于 D_2 程度的最大者,且 C_2 与其他模糊集 C_i 的交完全属于 D_2. 这样 $\underline{A}_\lambda(D_2)$ 一定是 $C_2\cap C_i$, $i=1,3,4,5,6$ 中之一. 同时,模糊集 C_2 也是 $\mathscr{C}$ 中包含 D_2 程度的最大者,由于 $D(C_2\cup C_i/D_2)<1$, $i\neq 2$,而 $C_2\cup C_1\cup C_3$, $C_2\cup C_5\cup C_3$, $C_2\cup C_5\cup C_4$, $C_2\cup C_1\cup C_4$ 都完全包含 D_2,于是 $\overline{A}_\lambda(D_2)$必为这四者之一,经计算 $\underline{A}_\lambda(D_2)=C_2\cap C_5$, $\overline{A}_\lambda(D_2)=C_2\cup C_4\cup C_5$. 因此对于任意 $x\in U$ 有

$$\min\{C_2(x),C_5(x)\}\leqslant D_2(x)\leqslant\max\{C_2(x),C_4(x),C_5(x)\}.$$

其中近似容度 $T_{\mathscr{C}}(D_2)=0.5375$.

这样要对对象属于模糊概念"尚可的申请者"的隶属度作出近似,必须知道它属于模糊概念"中等银行平衡","月消费低"和"月消费中等"的隶属度.

概念 D_3 的近似 容易计算 $\underline{A}_\lambda(D_3)=C_1$,$\overline{A}_\lambda(D_3)=C_1\cup C_5$,其中近似容度 $T_{\mathscr{C}}(D_3)=0.2375$.

§9.7 粗糙集与模糊集的比较

模糊集和粗糙集理论在处理不确定性和不精确性问题方面都推广了经典集合论,它们都可以用来描述知识的不精确性和不完全性.然而,它们的侧重面不同,从知识的“粒度”上来看,模糊集主要着眼于知识的模糊性,面粗糙集着眼于知识的粗糙性;从知识描述的方法上来看,模糊集是通过对象关于集合的隶属程度来近似描述的,而粗糙集是通过一个集合关于某个已知的可利用的信息库(即近似空间)的一对上,下近似来描述的;从集合的对象间的关系来看,模糊集强调的是集合边界的病态定义,即边界的不分明性,而粗糙集强调的是集合对象间的不可分辨性;从研究的对象来看,模糊集研究的是属于同一类的不同的对象的隶属关系,重在隶属程度,面粗糙集研究的是不同类中的对象组成的集合之间的关系,重在分类.

Pawlak 粗糙集代数系统$(2^U,\sim,\cap,\cup,\underline{R},\overline{R})$是由经典集合代数系统$(2^U,\sim,\cap,\cup)$加上两个一元集合算子 $\underline{R}$ 和 $\overline{R}$ 而形成的.粗糙集理论中用粗糙隶属函数来刻画知识的模糊性,粗糙集理论中的粗糙隶属函数可以看成特殊的模糊隶属函数.这样,论域 U 中的任何一个经典集 A 都对应一个模糊集 μ_A,而集合 A 的下近似和上近似分别是这个模糊集的核和支撑,即

$$\mathrm{core}(\mu_A)=\{x\mid\mu_A(x)=1\}=\underline{R}A,$$
$$\mathrm{supp}(\mu_A)=\{x\mid\mu_A(x)>0\}=\overline{R}A.$$

由此可见,下近似恰好是 μ_A 的 1 截集而上近似恰好是 μ_A 的强 0 截集.模糊集理论中对象 x 的隶属函数值不依赖于其他对象,一般是由专家直接给出,因此带有很强的主观意志;而粗糙集理论中对象的隶属函数值却依赖于其他对象,它可以从所需处理的数据中直接得到,所以用它来反映知识的模糊性是比较客观的.但粗糙隶属函数一般不是 Zadeh 意义下的隶属函数,它一般不满足 min-max 运算性质.从另外的角度来看,一个粗糙集又可以看成是三值模糊集,即

$$\mu_A(x)=\begin{cases}1, & x\in\mathrm{pos}(A),\\ 1/2, & x\in\mathrm{bn}(A),\\ 0, & x\in\mathrm{neg}(A).\end{cases}$$

当然,1/2 可以用(0,1)中的任何一个数代替,这个函数也不是 Zadeh 意义下的隶属函数.

由于模糊集和粗糙集都有可以用来描述知识的不确定性，各自的特点不同，因此模糊集理论和粗糙集理论有很强的互补性，将这两个理论进行某些整合后，去处理知识的不确定性和不完全性，比它们各自去处理知识的不确定性和不完全性可望显示出更强的功能.

第十章　基于随机集的粗糙集模型

到目前为止，我们给出的粗糙集模型中问题的处理往往都在一个论域里进行，而在实际情况中，我们经常碰到要在不同的论域关系下作出一些决策，另外所涉及的问题中常有数据丢失和信息不完备的情况，因此必须对粗糙集模型进行进一步的扩充．这一章给出了随机集粗糙集模型的一般框架．我们利用随机集去描述粗糙集近似算子，并给出了随机集、粗糙近似算子与证据理论中的信任函数和可能性函数之间的联系．

§10.1　随机集与容度泛函

定义 10.1　所谓随机集(random set)就是取值为闭集合值的随机变量，其严格的数学定义是这样的：设(U,Σ,P)是一个概率空间，$(2^W,\sigma(\beta))$是另一可测空间，其中$\beta\subseteq 2^W$，若映射$F:U\to 2^W$是Σ-$\sigma(\beta)$可测的，即对于任意$A\in\sigma(\beta)$有$\{u\in U|F(u)\in A\}\in\Sigma$，则称$F$是一个随机集．

通常W为一d维欧氏空间R^d或离散空间Z^d，F是随机集当且仅当对于任意$K\in P_k(W)$(其中$P_k(W)$为W的紧子集全体)，有$\{u\in U|F(u)\cap K\neq\varnothing\}\in\Sigma$．随机集$F$的概率分布函数为

$$P_F(A)=P(\{u\in U\mid F(u)\in A\}),\quad A\in\sigma(\beta).$$

由于A是集类，计算分布函数比较复杂，通常考虑F的容度泛函(capacity functional)

$$T_F(K)=P(\{u\in U\mid F(u)\cap K\neq\varnothing\}),K\in P_k(W).$$

容度泛函$T_F(\cdot)$满足下列性质：

(1)$T_F(\varnothing)=0$.

(2)$0\leqslant T_F(K)\leqslant 1,\forall K\in P_k(W)$.

(3)$T_F(\cdot)$在$P_k(W)$上是单调递增的，即对于$K_1,K_2\in P_k(W)$，$K_1\subseteq K_2$蕴涵$T_F(K_1)\subseteq T_F(K_2)$.

(4)$T_F(\cdot)$在$P_k(W)$上是上半连续的，即对于$K,K_n\in P_k(W)$，$n\geqslant 1$，$K_n\downarrow K$蕴涵$T_F(K_n)\downarrow T_F(K)$.

(5)设$K,K_n\in P_k(W)$，若记

$$Q_F^{(0)}(K)=P(\{u\in U\mid F(u)\cap K=\varnothing\})=1-T_F(K),$$

$$Q_F^{(n)}(K;K_1,K_2,\cdots,K_n)$$
$$=Q_F^{(n-1)}(K;K_1,K_2,\cdots,K_{n-1})-Q_F^{(n-1)}(K\cup K_n;K_1,K_2,\cdots,K_{n-1}),n\geqslant 1$$
则有

$$0\leqslant Q_F^{(n)}(K;K_1,K_2,\cdots,K_n)$$
$$=P(\{u\in U\mid F(u)\cap K=\varnothing;F(u)\cap K_i\neq\varnothing,1\leqslant i\leqslant n\})$$
$$\leqslant 1.$$

注 上面$Q_F^{(n)}(K;K_1,K_2,\cdots,K_n)$的直观意义是随机集$F$击中$K_i$,$1\leqslant i\leqslant n$,但没有击中$K$的概率.满足上述性质(3)~(5)的泛函$T_F$常称为Choquet容度或无穷单调序交替容度(alternating capacity of infinite order),随机集的概率分布函数是由F的Choquet容度$T_F(K)$,$K\in P_k(W)$惟一决定的,然而一般容度泛函不构成测度,因为它只满足次可加性,即

$$K_1\cap K_2=\varnothing\Rightarrow T_F(K_1\cup K_2)\leqslant T_F(K_1)+T_F(K_2).$$

容度泛函与信任函数和似然函数有比较密切的联系.

§10.2 信任函数与似然函数

证据理论是由Dempster和Shafer于20世纪60年代初建立的一套新的数学理论,它主要应用于人工智能与专家系统中不确定性问题的处理,它是概率论的进一步扩充,信任函数(belief function)与似然函数(plausibility function)是这个理论的两个最基本也是最重要的概念.

定义10.2 设U是非空有限集,称集函数$B:2^U\to[0,1]$为信任函数,若它满足性质:

(1)$B(\varnothing)=0,B(U)=1$.

(2)对于U中任意子集$X_1,X_2,\cdots,X_n,n\geqslant 1$,有

$$B(\bigcup_{i=1}^{n}X_i)\geqslant\sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}}(-1)^{|I|+1}B(\bigcap_{i\in I}X_i),\tag{10.1}$$

称函数$L:2^U\to[0,1]$为似然函数,若L是B的对偶补,即对于任意$X\subseteq U$有

$$L(X)=1-B(\sim X).$$

定理10.1 似然函数满足下列性质:

(1)$L(\varnothing)=0,L(U)=1$;

(2)对于U中任意子集$X_1,X_2,\cdots,X_n,n\geqslant 1$有

$$L(\bigcap_{i=1}^{n}X_i)\leqslant\sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}}(-1)^{|I|+1}L(\bigcup_{i\in I}X_i).$$

证 由定义和对偶性直接可得. □

通常$B(X)$可以解释为事件X肯定为真的信任程度,而$L(X)$可以解释

为事件 X 为非假的信任程度.

定义 10.3 集函数 $m:2^U\to[0,1]$称为概率分配函数,简称为 mass 函数,如果它满足:

$$(1)m(\varnothing)=0,\qquad (2)\sum_{X\subseteq U}m(X)=1.$$

$m(X)$反映了确切分配到 X 上的信任程度,一般由专家给出,虽然专家可以给一个事件赋以任意大小的信任度,但要求对所有的事件的信任程度的和为 1.

定理 10.2 若 m 是 U 上的一个 mass 函数,则由

$$B(X)=\sum_{D\subseteq X}m(D),X\subseteq U, \tag{10.2}$$

所定义的函数 B 是 U 上的一个信任函数,同时 m 可由 B 的莫比乌斯(Möbius)变换

$$m(X)=\sum_{D\subseteq X}(-1)^{|X\setminus D|}B(D),X\subseteq U, \tag{10.3}$$

求得.反之,若 B 是 U 上的一个信任函数,则由(10.3)定义的函数 m 是 U 上的 mass 函数,此时(10.2)式成立.

证 首先易证 $B(\varnothing)=0,B(U)=1$,且对任意 $X\subseteq U,0\leqslant B(X)\leqslant 1$.对于任意 $X_1,X_2,\cdots,X_n\subseteq U$ 和$D\subseteq U$,记

$$I(D)=\{1,2,\cdots,n\}\cap\{i\mid D\subseteq X_i\}.$$

则易见 $I(D)\neq\varnothing$当且仅当存在 $i\in\{1,2,\cdots,n\}$,使 $D\subseteq X_i$,而 $D\subseteq\bigcap_{i\in I}X_i$ 当且仅当$I\subseteq I(D)$.于是

$$\begin{aligned}
&\sum_{I\subseteq\{1,2,\cdots,n\},I\neq\varnothing}(-1)^{|I|+1}B\Big(\bigcap_{i\in I}X_i\Big)\\
=&\sum_{I\subseteq\{1,2,\cdots,n\},I\neq\varnothing}(-1)^{|I|+1}\Big(\sum_{D\subseteq\bigcap_{i\in I}X_i}m(D)\Big)\\
=&\sum_{I\subseteq\{1,2,\cdots,n\},I\neq\varnothing}(-1)^{|I|+1}\Big(\sum_{I\subseteq I(D),I(D)\neq\varnothing}m(D)\Big)\\
=&\sum_{I\subseteq\{1,2,\cdots,n\},I\neq\varnothing}\ \sum_{I\subseteq I(D),I(D)\neq\varnothing}(-1)^{|I|+1}m(D)\\
=&\sum_{I(D)\neq\varnothing}\ \sum_{I\subseteq I(D),I(D)\neq\varnothing}(-1)^{|I|+1}m(D)\\
=&\sum_{I(D)\neq\varnothing}m(D)\Big(\sum_{I\subseteq I(D),I\neq\varnothing}(-1)^{|I|+1}\Big)\\
=&\sum_{I(D)\neq\varnothing}m(D)\Big(\sum_{I\subseteq I(D)}(-1)^{|I|+1}-\sum_{I=\varnothing}(-1)^{|I|+1}\Big)\\
=&\sum_{I(D)\neq\varnothing}m(D)(-0+1)=\sum_{I(D)\neq\varnothing}m(D)\\
\leqslant&\sum_{D\subseteq\bigcup_{i=1}^{n}X_i}m(D)=B\Big(\bigcup_{i=1}^{n}X_i\Big).
\end{aligned}$$

从而(10.1)成立，所以 B 是信任函数.

同时由于

$$\sum_{D\subseteq X}(-1)^{|X\setminus D|}B(D)=(-1)^{|X|}\sum_{D\subseteq X}(-1)^{|D|}B(D)$$
$$=(-1)^{|X|}\sum_{D\subseteq X}(-1)^{|D|}\sum_{C\subseteq D}m(C)$$
$$=(-1)^{|X|}\sum_{C\subseteq X}m(C)\sum_{C\subseteq D\subseteq X}(-1)^{|D|}.$$

当 $C=X$ 时

$$\sum_{\{D\mid C\subseteq D\subseteq X\}}(-1)^{|D|}=(-1)^{|X|},$$

当 $C\neq X$ 时

$$\sum_{\{D\mid C\subseteq D\subseteq X\}}(-1)^{|C|}=(-1)^{|C|}\cdot(1-1)^{|X\setminus C|}=0,$$

因此(10.3)成立.

反之，若 B 是 U 上的信任函数，且(10.3)式成立，则可证(10.2)式成立.事实上，对于任意 $X\subseteq U$

$$\sum_{D\subseteq X}m(D)=\sum_{D\subseteq X}\sum_{C\subseteq D}(-1)^{|D\setminus C|}B(C)$$
$$=\sum_{C\subseteq X}\sum_{C\subseteq D\subseteq X}(-1)^{|D\setminus C|}B(C)$$
$$=\sum_{C\subseteq X}(-1)^{|C|}B(C)\sum_{C\subseteq D\subseteq X}(-1)^{|D|}$$
$$=\sum_{C=X}(-1)^{|C|}B(C)\sum_{C\subseteq D\subseteq X}(-1)^{|D|}$$
$$+\sum_{C\subset X}(-1)^{|C|}B(C)\sum_{C\subseteq D\subseteq X}(-1)^{|D|}$$
$$=(-1)^{|X|}B(X)(-1)^{|X|}+\sum_{C\subset X}(-1)^{|C|}B(C)\cdot 0$$
$$=B(X).$$

其次

$$m(\varnothing)=\sum_{D\subseteq\varnothing}(-1)^{|\varnothing\setminus D|}B(D)=(-1)^0B(\varnothing)=0,$$

而由(10.2)得

$$\sum_{D\subseteq U}m(D)=B(U)=1. \tag{10.4}$$

最后往证

$$\forall X\in 2^U\setminus\varnothing,0\leqslant m(X)\leqslant 1. \tag{10.5}$$

若 $|X|=1$，则易得 $m(X)\geqslant 0$.

不妨设 $X=\{x_1,x_2,\cdots,x_k\}$，$k>1$，记 $X_i=X\setminus\{x_i\}$，则对于任意 $Y\subset$

X,存在 $I\subseteq\{1,2,\cdots,k\}$,$I\neq\varnothing$,使

$$Y=\bigcap_{i\in I}X_i,\text{且 } X\setminus Y=\{x_i\mid i\in I\}. \tag{10.6}$$

反之,对于任意 $I\subseteq\{1,2,\cdots,k\}$,$I\neq\varnothing$,存在 $Y\subset X$ 使(10.6)成立. 即 X 的真子集与 $\{1,2,\cdots,k\}$ 的非空子集存在一一对应使(10.6)成立. 于是

$$\begin{aligned}
m(X)&=\sum_{Y\subseteq X}(-1)^{|X\setminus Y|}B(Y)\\
&=\sum_{Y=X}(-1)^{|X\setminus Y|}B(Y)+\sum_{Y\subset X}(-1)^{|X\setminus Y|}B(Y)\\
&=(-1)^0B(X)+\sum_{I\subseteq\{1,2,\cdots,k\},I\neq\varnothing}(-1)^{|\{x_i|i\in I\}|}B\Big(\bigcap_{i\in I}X_i\Big)\\
&=B(X)+\sum_{I\subseteq\{1,2,\cdots,k\},I\neq\varnothing}(-1)^{|I|}B\Big(\bigcap_{i\in I}X_i\Big)\\
&=B\Big(\bigcup_{i=1}^{k}X_i\Big)-\sum_{I\subseteq\{1,2,\cdots,k\},I\neq\varnothing}(-1)^{|I|+1}B\Big(\bigcap_{i\in I}X_i\Big)\\
&\geqslant 0.
\end{aligned}$$

这样就证明了对任意 $X\subseteq U$,$m(X)\geqslant 0$. 从而由(10.4)知(10.5)成立. 因此 m 是 mass 函数. □

定理 10.3 若 m 是 U 上的一个 mass 函数,则由

$$L(A)=\sum_{D\cap A\neq\varnothing}m(D),\quad A\subseteq U$$

所定义的函数 L 是 U 上的一个似然函数,并且

$$B(A)\leqslant L(A), \tag{10.7}$$

其中 B 由(10.2)式定义的信任函数.

证 易证 $L(X)=1-B(\sim X)$,由定理 10.2 知 B 是信任函数,从而得 L 是似然函数. (10.7)式成立显然. □

$[B(X),L(X)]$常称为事件 X 的信任区间.

记 $\mathscr{P}$ 为 (U,Σ) 上的概率测度全体,对于 U 上的信任函数 B,B 的核(记为 core B)定义为

$$\mathrm{core}B=\{P\in\mathscr{P}\mid P(X)\geqslant B(X),\forall X\subseteq U\}.$$

由似然函数 L 和 B 的对偶补性,因此与上式等价的有

$$\mathrm{core}B=\{P\in\mathscr{P}\mid P(X)\leqslant L(X),\forall X\subseteq U\}.$$

令 $\mathscr{M}$ 为 B 的莫比乌斯变换 m 的支撑,即

$$\mathscr{M}=\{G\subseteq U\mid m(G)\neq 0\}.$$

概率测度 P 称为 B 的一个配置(allocation),若对任意 $x\in U$

$$P(\{x\})=\sum_{\{G|x\in G\}}\lambda(G,x)m(G), \tag{10.8}$$

其中 $\lambda(G,x)\geqslant 0$,对任意 $G\in\mathscr{M}$ 满足 $\sum_{\{x|x\in G\}}\lambda(G,x)=1$.

定理 10.4　设 B 为 U 上的信任函数，则 B 的配置非空，并且若概率测度 P 是 B 的一个配置，则 $P\in \mathrm{core}B$.

证　设 $U=\{x_1,x_2,\cdots,x_n\}$，m 是 B 的 mass 函数，记

$$\mathscr{M}=\{G\subseteq U\mid m(G)>0\}=\{G_1,G_2,\cdots,G_k\}.$$

显然

$$\sum_{i=1}^{k}m(G_i)=\sum_{G\subseteq U}m(G)=B(U)=1.$$

对于任意的 $A\subseteq U$，A 可以表示为 $A(d_1d_2\cdots d_n)$，其中 $d_i=0$ 或 1，$d_i=1$ 当且仅当 $x_i\in A$. 于是 $m(A)=m(A(d_1d_2\cdots d_n))$，分解 $m(A(d_1d_2\cdots d_n))$ 为

$$m(A(d_1d_2\cdots d_n))=\sum_{i=1}^{n}m(A)\lambda(A,x_i),\tag{10.9}$$

其中 $\lambda(A,x_i)\geqslant 0$，$\sum_{i=1}^{n}\lambda(A,x_i)=1$，并且当 $x_i\notin A$ 时 $\lambda(A,x_i)=0$.

构造集函数 $P:2^U\to[0,1]$ 如下：

$$P(\{x_i\})=\sum_{j=1}^{k}m(G_j)\lambda(G_j,x_i),1\leqslant i\leqslant n,\tag{10.10}$$

$$P(A)=\sum_{x\in A}P(\{x\}),A\in 2^U.$$

显然，$P(A)\geqslant 0$，由于

$$\begin{aligned}\sum_{i=1}^{n}P(\{x_i\})&=\sum_{i=1}^{n}\sum_{j=1}^{k}m(G_j)\lambda(G_j,x_i)\\&=\sum_{j=1}^{k}\sum_{i=1}^{n}m(G_j)\lambda(G_j,x_i)\\&=\sum_{j=1}^{k}m(G_j)\sum_{i=1}^{n}\lambda(G_j,x_i)\\&=\sum_{j=1}^{k}m(G_j)=1.\end{aligned}$$

再由 P 的定义可知 P 是 2^U 上的概率测度. 由式(10.10)知(10.8)成立，即 P 是 B 的一个配置，即 B 的配置非空.

设 P 是 B 的任意一个配置，对于任意 $A\in 2^U$，由于

$$\begin{aligned}P(A)&=\sum_{x\in A}P(\{x\})=\sum_{x\in A}\sum_{j=1}^{k}m(G_j)\lambda(G_j,x)\\&=\sum_{j}^{k}m(G_j)\sum_{x\in A}\lambda(G_j,x)\\&=\sum_{\{j\mid G_j\subseteq A\}}m(G_j)\sum_{x\in A}\lambda(G_j,x)+\sum_{\{j\mid G_j\cap(\sim A)\neq\varnothing\}}m(G_j)\sum_{x\in A}\lambda(G_j,x)\end{aligned}$$

$$= \sum_{\{j|G_j\subseteq A\}} m(G_j) + \sum_{\{j|G_j\cap(\sim A)\neq\varnothing\}} m(G_j)\sum_{x\in A}\lambda(G_j,x)$$

$$= B(A) + \sum_{\{j|G_j\cap(\sim A)\neq\varnothing\}}\sum_{x\in A} m(G_j)\lambda(G_j,x) \geqslant B(A). \quad (10.11)$$

由此可知 $P\in \mathrm{core}B$. □

注 可以证明,B 的核中的每一个概率测度也一定是 B 的配置.

定理 10.5 设 B 是 U 上信任函数,记 $\mathscr{P}_*(X)=\inf\{P(X)\mid P\in \mathrm{core}B\}$, $X\in 2^U$,则 $B=\mathscr{P}_*$,即对于任意 $X\in 2^U$ 有

$$B(X) = \mathscr{P}_*(X). \quad (10.12)$$

证 对于任意 $X\in 2^U$,显然有 $B(X)\leqslant\mathscr{P}_*(X)$,只须证 $\mathscr{P}_*(X)\leqslant B(X)$. 事实上,若有 $G_j\cap(\sim X)\neq\varnothing$,对这样的 G_j 我们在分解式(10.9)中取 λ 使 $\sum_{x\in G_j\cap(\sim X)}\lambda(G_j,x)=1$,则

$$\sum_{\{j|G_j\cap(\sim X)\neq\varnothing\}}\sum_{x\in X} m(G_j)\lambda(G_j,x) = 0.$$

从而由(10.11)知有概率测度 P(P 与 λ 有关)$\in \mathrm{core}B$ 使 $P(X)=B(X)$,于是 $\mathscr{P}_*(X)\leqslant P(X)=B(X)$,因此(10.12)成立. □

例 10.1 设 $U=\{x_1,x_2,x_3\}$, $A_1=\{x_1,x_2\}$, $A_2=\{x_2,x_3\}$, $A_3=U$, mass 函数定义为

$$m(A_1)=0.3, m(A_2)=0.4, m(U)=0.3.$$

信任函数

$$B(A_1)=0.3, B(A_2)=0.4, B(U)=1.$$

分解 $m(A)$ 为

$$m(A_1) = m(A(110) = 0.3\lambda(A_1,x_1)+0.3\lambda(A_1,x_2),$$

其中 $\lambda(A_1,x_1)+\lambda(A_1,x_2)=1, \lambda(A_1,x_i)>0, i=1,2.$

$$m(A_2) = m(A(011)) = 0.4\lambda(A_2,x_2)+0.4\lambda(A_2,x_3),$$

其中 $\lambda(A_2,x_2)+\lambda(A_2,x_3)=1, \lambda(A_2,x_i)>0, i=2,3.$

$$m(U) = m(A(111)) = 0.3\lambda(A_3,x_1)+0.3\lambda(A_3,x_2)+0.3\lambda(A_3,x_3),$$

其中 $\lambda(A_3,x_1)+\lambda(A_3,x_2)+\lambda(A_3,x_3)=1, \lambda(A_3,x_i)>0, i=1,2,3.$ 于是

$$P(\{x_1\}) = 0.3\lambda(A_1,x_1)+0.3\lambda(A_3,x_1),$$
$$P(\{x_2\}) = 0.3\lambda(A_1,x_2)+0.4\lambda(A_2,x_2)+0.3\lambda(A_3,x_2),$$
$$P(\{x_3\}) = 0.4\lambda(A_2,x_3)+0.3\lambda(A_3,x_3).$$

从而

$$P(A_1) = 0.3+0.4\lambda(A_2,x_2)+0.3[\lambda(A_3,x_1)+\lambda(A_3,x_2)] \geqslant B(A_1),$$
$$P(A_2) = 0.4+0.3\lambda(A_1,x_2)+0.3[\lambda(A_3,x_2)+\lambda(A_3,x_3)] \geqslant B(A_2).$$

§10.3 基于随机集的粗糙集模型

在数据处理中人们会经常碰到某些对象或事例描述的数据信息不完全，不精确甚至丢失的情况，面对这种情形人们有几种不同的处理办法，一种简单的处理方法是把这些不精确的模糊的数据或事例从系统或库中删除，使得可利用的信息库成为一个确定的精确的知识库，但是这样做的后果往往会导致大量的信息的丢失，使得所获得的结果失去应有的价值．为了使数据处理的结果尽可能与原系统保持一致，我们可将对象属性的可能取值全部考虑进去，这时一些对象的某些属性取值不再是单点值而是取集合值了，我们可将这种数据库的对象描述函数看成随机集．本节就介绍基于随机集的粗糙集模型．

定义 10.4 设 U 和 W 是两个有限非空集合，$(U,2^U,P)$为概率空间，显然$(2^W,\sigma(2^W))$是一个可测空间，这样任何一个集值函数 $F:U\to 2^W$ 就是一个随机集，称四元有序组 $A=(U,W,F,P)$为随机集近似空间，对于任意 $X\in 2^W$，定义 X 关于A 的下近似$\underline{\mathrm{apr}}_F X$ 和上近似$\overline{\mathrm{apr}}_F X$ 为

$$\underline{\mathrm{apr}}_F X=\{u\in U\mid F(u)\subseteq X\},$$
$$\overline{\mathrm{apr}}_F X=\{u\in U\mid F(u)\cap X\neq\varnothing\}.$$

当$\underline{\mathrm{apr}}_F X=\overline{\mathrm{apr}}_F X$ 时，称 X 关于近似空间A 是可定义的，否则称 X 关于近似空间A 是不可定义的或粗糙的．

注：当 $U=W$ 时，如果我们定义二元关系 $R=\{(u,v)\mid v\in F(u),u\in U\}$，并且可以将 $F(u)$看成 u 的邻域，这时得到的模型就退化为一般关系下的广义粗糙集模型．又由于一般 U 和 W 是两个不同的论域，因此随机集粗糙集模型又与第八章的概率粗糙集模型是不相同的．以下若 F 比较明确我们将下标F 省去．

定理 10.6 设 $F:U\to 2^W$ 是随机集，则近似算子$\underline{\mathrm{apr}}$和$\overline{\mathrm{apr}}$满足下列性质：

(1)$\underline{\mathrm{apr}}X=\sim(\overline{\mathrm{apr}}(\sim X))$，

$\overline{\mathrm{apr}}X=\sim(\underline{\mathrm{apr}}(\sim X))$，$X\subseteq W$．

(2)$\underline{\mathrm{apr}}W=U$，$\overline{\mathrm{apr}}\varnothing=\varnothing$．

(3) $\underline{\mathrm{apr}}(X\cap Y)=\underline{\mathrm{apr}}X\cap\underline{\mathrm{apr}}Y$，

$\overline{\mathrm{apr}}(X\cup Y)=\overline{\mathrm{apr}}X\cup\overline{\mathrm{apr}}Y$，$X,Y\subseteq W$．

(4) $X\subseteq Y\subseteq W\Rightarrow\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}Y$，$\overline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}Y$．

(5) $\underline{\mathrm{apr}}(X\cup Y)\supseteq\underline{\mathrm{apr}}X\cup\underline{\mathrm{apr}}Y$，

$\overline{\mathrm{apr}}(X\cap Y)\subseteq\overline{\mathrm{apr}}X\cap\overline{\mathrm{apr}}Y$，$X,Y\subseteq W$．

证 由定义直接可得． □

定理 10.7 若 $F:U\to 2^W$ 是随机集，则以下等价：

(1) 对于任意 $u\in U$，有 $F(u)\neq\varnothing$.

(2) 对任意 $X\subseteq W$，$\underline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}X$.

(3) $\underline{\mathrm{apr}}\varnothing=\varnothing$.

(4) $\overline{\mathrm{apr}}\,W=U$.

证 类似于定理 3.3 的证明. □

设 $F:U\to 2^U$ 是随机集，$x\in U$，若 $x\in F(x)$，则称 x 是 F 的不动点.

定理 10.8 设 $F:U\to 2^U$ 是随机集，则以下等价：

(1) 任意的 $x\in U$ 是 F 的不动点.

(2) $\underline{\mathrm{apr}}X\subseteq X$，$\forall X\subseteq U$.

(3) $X\subseteq\overline{\mathrm{apr}}X$，$\forall X\subseteq U$.

证 类似于定理 3.5 的证明. □

显然，下近似$\underline{\mathrm{apr}}_F$ 和上近似$\overline{\mathrm{apr}}_F$ 是 2^W 到 2^U 的两个集算子，我们称 $(U,W,\cap,\cup,\sim,\underline{\mathrm{apr}},\overline{\mathrm{apr}})$为随机集粗糙集代数系统.

由定理 10.6 的性质(3)我们可以看出$\overline{\mathrm{apr}}X=\bigcup\limits_{x\in X}\overline{\mathrm{apr}}\{x\}$，类似于一般关系下粗糙集模型，取上近似分布函数 $h(x)=\overline{\mathrm{apr}}\{x\}$，$x\in W$，则不难验证 $h(x)=\{u\in U\mid x\in F(u)\}$，而 $F(u)=\{y\in W\mid u\in h(y)\}$，当然$\overline{\mathrm{apr}}X=\bigcup\limits_{x\in X}h(x)$.

对于随机集 $F:U\to 2^W$，定义关系划分函数 $j:2^W\to 2^U$

$$j(X)=\{u\in U\mid F(u)=X\},\qquad X\subseteq W.$$

若 $j(X)\neq\varnothing$，则称 X 为焦点集，记全体焦点集为 $\mathscr{J}$，即 $\mathscr{J}=\{X\in 2^W\mid j(X)\neq\varnothing\}$.

定理 10.9 关系划分函数 j 满足下列性质：

(1) $\bigcup\limits_{A\subseteq W}j(A)=U$.

(2) $A\neq B,A,B\subseteq W\Rightarrow j(A)\cap j(B)=\varnothing$.

证 由定义直接可得. □

定理 10.10 设 $F:U\to 2^W$ 为随机集，$j:2^W\to 2^U$ 为关系划分函数，则近似算子与关系划分函数有以下关系：

(1) $\underline{\mathrm{apr}}X=\bigcup\limits_{B\subseteq X}j(B)$，$X\subseteq W$.

(2) $\overline{\mathrm{apr}}X=\bigcup\limits_{B\cap X\neq\varnothing}j(B)$，$X\subseteq W$.

(3) $j(X)=\underline{\mathrm{apr}}X\setminus\bigcup\limits_{B\subset X}\underline{\mathrm{apr}}B$.

证 类似于定理 3.18 和定理 3.19. □

对于随机集 $F:U\to 2^W$ 和近似空间 (U,W,F,P),记

$$m(X)=P(j(X)),\quad X\in 2^W,$$

$$B(X)=P(\underline{\mathrm{apr}}X),\quad X\in 2^W,\tag{10.13}$$

$$L(X)=P(\overline{\mathrm{apr}}X),\quad X\in 2^W.\tag{10.14}$$

引理 10.1 集函数 m,B 和 L 有以下关系:

(1) $B(X)=\sum\limits_{A\subseteq X}m(A),\quad X\in 2^W$.

(2) $L(X)=\sum\limits_{A\cap X\neq\varnothing}m(A),\quad X\in 2^W$.

(3) $m(X)=\sum\limits_{A\subseteq X}(-1)^{|X\setminus A|}B(A),\quad X\in 2^W$.

证 (1) 由定理 10.10 得

$$\begin{aligned}B(X)&=P(\underline{\mathrm{apr}}X)=P(\bigcup_{A\subseteq X}j(A))\\&=\sum_{A\subseteq X}P(j(A))=\sum_{A\subseteq X}m(A).\end{aligned}$$

(2)对于任意 $X\in 2^W$,由定理 10.9,定理 10.10 和定理 10.6 知

$$\begin{aligned}&\sum_{A\cap X\neq\varnothing}P(j(A))+\sum_{A\cap X=\varnothing}P(j(A))\\&=\sum_{A\subseteq W}P(j(A))=P(\bigcup_{A\subseteq W}j(A))\\&=P(\underline{\mathrm{apr}}W)=P(U)=1.\end{aligned}$$

从而由近似算子的对偶性和(1)可得

$$\begin{aligned}L(X)&=P(\overline{\mathrm{apr}}X)=P(\sim\underline{\mathrm{apr}}(\sim X))\\&=1-P(\underline{\mathrm{apr}}(\sim X))=1-\sum_{A\subseteq\sim X}P(j(A))\\&=1-\sum_{A\cap X=\varnothing}P(j(A))\\&=\sum_{A\cap X\neq\varnothing}P(j(A))\\&=\sum_{A\cap X\neq\varnothing}m(A).\end{aligned}$$

(3) $$\begin{aligned}&\sum_{A\subseteq X}(-1)^{|X\setminus A|}B(A)=(-1)^{|X|}\sum_{A\subseteq X}(-1)^{|A|}B(A)\\&=(-1)^{|X|}\sum_{A\subseteq X}(-1)^{|A|}\sum_{B\subseteq A}m(B)\\&=(-1)^{|X|}\sum_{A\subseteq X}\sum_{B\subseteq A}(-1)^{|A|}m(B)\\&=(-1)^{|X|}\sum_{B\subseteq X}\sum_{B\subseteq A\subseteq X}(-1)^{|A|}m(B)\\&=(-1)^{|X|}\sum_{B\subseteq X}m(B)\sum_{B\subseteq A\subseteq X}(-1)^{|A|}\end{aligned}$$

$$= (-1)^{|X|}\sum_{B=X} m(B)\sum_{B\subseteq A\subseteq X}(-1)^{|A|} + (-1)^{|X|}\sum_{B\subset X} m(B)\sum_{B\subseteq A\subseteq X}(-1)^{|A|}$$

$$= (-1)^{|X|}\sum_{B=X} m(B)(-1)^{|X|} + (-1)^{|X|}\sum_{B\subset X} m(B)\cdot 0 = m(X). \quad \square$$

定理 10.11 若随机集 $F:U\to 2^W$ 对于任意的 $u\in U$ 有 $F(u)\neq\varnothing$，则 m 是 mass 函数.

证 首先，由于对于任意的 $u\in U$ 有 $F(u)\neq\varnothing$，由定理 10.7 知 $\underline{\mathrm{apr}}\varnothing=\varnothing$，从而由定理 10.10 得

$$m(\varnothing) = P(j(\varnothing)) = P(\varnothing) = 0. \tag{10.15}$$

其次，由定理 10.9 得

$$\sum_{A\subseteq W} m(A) = \sum_{A\subseteq W} P(j(A)) = P(\bigcup_{A\subseteq W} j(A)) = P(U) = 1. \tag{10.16}$$

由(10.15)和(10.16)知 m 是 mass 函数. $\square$

定理 10.12 若随机集 $F:U\to 2^W$ 对于任意的 $u\in U$ 有 $F(u)\neq\varnothing$，则由式(10.13)和(10.14)定义的集函数 B 和 L 分别是 W 上的信任函数和似然函数.

证 由引理 10.1，定理 10.11 和定理 10.2 即得. 以下我们给出 B 是信任函数的一个直接证明.

首先，由定理 10.7 知

$$B(\varnothing) = P(\underline{\mathrm{apr}}\varnothing) = P(\varnothing) = 0. \tag{10.17}$$

其次，由定理 10.6 得

$$B(W) = P(\underline{\mathrm{apr}}W) = P(U) = 1. \tag{10.18}$$

最后，设 $A_1,A_2,\cdots,A_n\subseteq W$，$n\geqslant 1$，则由定理 10.6 和概率测度的性质可得

$$\begin{aligned} B(\bigcup_{i=1}^n A_i) &= P(\underline{\mathrm{apr}}\bigcup_{i=1}^n A_i) \geqslant P(\bigcup_{i=1}^n \underline{\mathrm{apr}}A_i) \\ &= \sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}} (-1)^{|I|+1} P(\bigcap_{i\in I}\underline{\mathrm{apr}}A_i) \\ &= \sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}} (-1)^{|I|+1} P(\underline{\mathrm{apr}}\bigcap_{i\in I}A_i) \\ &= \sum_{\varnothing\neq I\subseteq\{1,2,\cdots,n\}} (-1)^{|I|+1} B(\bigcap_{i\in I}A_i). \end{aligned} \tag{10.19}$$

由式(10.17)，(10.18)和(10.19)知 B 是信任函数. $\square$

下述定理表明论域上的任何一个信任函数本质上就是某一个随机集导出的下近似的概率测度.

定理 10.13 设 B 是 W 上的一个信任函数，则存在有限论域 U，2^U 上的概率测度 P 和随机集 $F:U\to 2^W$ 使得对于任意的 $u\in U$，$F(u)\neq\varnothing$，且

$$P(\underline{\mathrm{apr}}_F X) = B(X), X\in 2^W. \tag{10.20}$$

证 设 m 为 B 的 mass 函数，m 的焦点集为

$$\mathscr{M} = \{X \in 2^W \mid m(X) > 0\}.$$

设 $\mathscr{M} = \{X_1, X_2, \cdots, X_k\}$，由 B 是信任函数知

$$\sum_{i=1}^{k} m(X_i) = \sum_{A \subseteq W} m(A) = 1. \tag{10.21}$$

任取具有 k 个元素的集合 $U = \{u_1, u_2, \cdots, u_k\}$，定义集函数 $P: 2^U \to [0,1]$ 如下：

$$P(\{u_i\}) = m(X_i), \quad i = 1,2,\cdots,k,$$

$$P(A) = \sum_{u \in A} P(u), \quad A \in 2^U.$$

由式(10.21)知

$$\sum_{i=1}^{k} P(u_i) = \sum_{u \in U} P(u) = \sum_{X \in \mathscr{M}} m(X) = 1,$$

因此易见 P 是定义在 2^U 上的概率测度.

定义随机集 $F: U \to 2^W$ 如下

$$F(u_i) = X_i, \quad i = 1,2,\cdots,k.$$

显然，当 $A = X_i$ 时有 $j(A) = \{u_i\}$，而当 $A \notin \mathscr{M}$ 时有 $j(A) = \varnothing$，这样由 P 和 F 的构造可见对于任意 $A \in 2^W$ 有 $P(j(A)) = m(A)$，从而由引理 10.1 和定理 10.12 知以 m 为 mass 函数的信任函数就是下近似的概率，即式(10.20)成立. □

定义 10.5 设 $F: U \to 2^W$ 为随机集，函数 $f: U \to W$ 称为 F 的一个选择，若对于任意的 $u \in U$，有 $f(u) \in F(u)$.

F 的每一个选择 f 可看作是从 $(U, 2^U)$ 到 $(W, 2^W)$ 的一个随机变量，记 F 的选择全体为 S_F，则显然 $F(u) = \{f(u) \mid f \in S_F\}$. 在不完备信息系统中，$F$ 的每一个选择 f 可以解释为与已知的可利用的不精确的知识库相容的一种特殊的不确定性(或可能)描述.

比如 $S = (U, AT, V, F)$ 为一不完备信息系统，U 为论域，AT 为属性集，$V = \bigcup_{a \in AT} V_a$，$V_a$ 为属性 a 的值域，$F = \{F_a \mid a \in AT\}$，$F_a: U \to V_a$ 为对象描述函数，由于 S 是不完备信息系统，因此 F_a 一般是集值函数，F_a 的任何一个选择 f_a 是单值函数，表示不完备信息系统中与属性 a 对应的知识库相容的一种不确定性的描述，若记 $v_a = \{f_a(u) \mid u \in U\}$，$f = \{f_a \mid a \in AT\}$，$v = \bigcup_{a \in AT} v_a$，则系统 $s = (U, AT, v, f)$ 是与原系统 S 相容的一个不精确描述，实际上它就是原系统 S 的一个可能的完备化系统，当 f 取遍 F 的所有可能选择时，我们就得到了原系统的所有完备化系统.

定理 10.14 设 $F: U \to 2^W$ 为随机集，对于任意的 $f \in S_F$，记它在 $(W, 2^W)$ 上导出一个概率测度为 P_f，即

$$P_f(X)=P(\{u\in U\mid f(u)\in X\}),X\in 2^W.$$

若对于任意 $u\in U$ 有 $F(u)\neq\varnothing$，则对于任意的 $f\in S_F$，P_f 是 B 的配置，其中 B 是 F 导出的信任函数.

证 对于任意 $X\in 2^W$，显然有

$$\{u\in U\mid f(u)\in X\}\subseteq\{u\in U\mid F(u)\cap X\neq\varnothing\}.$$

因此 $P_f(X)\leqslant P(\overline{\mathrm{apr}}X)=L(X)$，从而 $P_f\in\mathrm{core}B$，于是 P_f 是 B 的配置. □

在实际问题中，每个人的评价可以看成是一个概率测度，但由于问题的不确定性和专家知识的局限性，每个专家对同一个问题的评价也是不相同的，比如对 k 个专家来说，就形成了一组概率测度. 一种方法就是将这些概率测度加权平均形成一个新的概率测度作为对问题的评价，这是一种简单的方法，但是对专家的知识造成的损失太大；另一种方法是通过这组概率得到概率上界和概率下界形成一个概率区间.

定义 10.6 设 $\mathscr{P}$ 是论域 U 上的一族概率测度，记

$$\mathscr{P}_*(X)=\inf\{P(X)\mid P\in\mathscr{P}\},X\in 2^U,$$

$$\mathscr{P}^*(X)=\sup\{P(X)\mid P\in\mathscr{P}\},X\in 2^U.$$

$\mathscr{P}_*$ 称为 $\mathscr{P}$ 的概率下界，$\mathscr{P}^*$ 称为 $\mathscr{P}$ 的概率上界.

定理 10.15 设 $\mathscr{P}$ 是论域 U 上的一族概率测度，则以下性质成立：

(1) $\mathscr{P}_*(\varnothing)=\mathscr{P}^*(\varnothing)=0,\mathscr{P}_*(U)=\mathscr{P}^*(U)=1$.

(2) $0\leqslant\mathscr{P}_*(X)\leqslant\mathscr{P}^*(X)\leqslant 1,X\in 2^U$.

(3) $\mathscr{P}^*(X)=1-\mathscr{P}_*(\sim X),X\in 2^U$.

(4) 对于任意 $X,Y\in 2^U$，若 $X\cap Y=\varnothing$，则

$$\begin{aligned}\mathscr{P}_*(X)+\mathscr{P}_*(Y)&\leqslant\mathscr{P}_*(X\cup Y)\leqslant\mathscr{P}_*(X)+\mathscr{P}^*(Y)\\&\leqslant\mathscr{P}^*(X\cup Y)\leqslant\mathscr{P}^*(X)+\mathscr{P}^*(Y).\end{aligned}$$

证 (1)与(2)显然.

(3)由于

$$\begin{aligned}&\mathscr{P}_*(\sim X)+\mathscr{P}^*(X)\\&=\inf\{1-P(X)\mid P\in\mathscr{P}\}+\sup\{P(X)\mid P\in\mathscr{P}\}\\&=1-\sup\{P(X)\mid P\in\mathscr{P}\}+\sup\{P(X)\mid P\in\mathscr{P}\}=1,\end{aligned}$$

因此(3)成立.

(4)设 $X\cap Y=\varnothing$，存在 $\mathscr{P}$ 中的概率测度序列 $\{P_i\mid i\geqslant 1\}$，使

$$\lim_{i\to\infty}P_i(X\cup Y)=\mathscr{P}_*(X\cup Y).$$

于是

$$\begin{aligned}\lim_{i\to\infty}P_i(X\cup Y)&=\lim_{i\to\infty}(P_i(X)+P_i(Y))\\&\geqslant\liminf_{i\to\infty}P_i(X)+\liminf_{i\to\infty}P_i(Y)\end{aligned}$$

$$\geq \mathscr{P}_*(X) + \mathscr{P}_*(Y).$$

第一个不等式得证,其他不等式的证明类似. □

对于随机集 $F: U \to 2^W$,记 $\mathscr{P}_F = \{P_f \mid f \in S_F\}$,若对于任意 $u \in U$ 有 $F(u) \neq \varnothing$,则可以证明

$$B(X) = \mathscr{P}_{F*}(X), X \in 2^W.$$

$$L(X) = \mathscr{P}_F^*(X), X \in 2^W.$$

由此我们可以得到求由 F 导出的信任函数和似然函数的另一种方法,即先求出 F 的所有可测选择 S_F 以及每一选择 $f \in S_F$ 的概率分布 P_f,然后分别求出 $\mathscr{P}_F = \{P_f \mid f \in S_F\}$ 的概率下界和概率上界函数得到所求的信任函数和似然函数.

例 10.2(一个不完备信息系统)　设论域(一组汽车)$U = \{1,2,3,4,5\}$,其上有均匀分布的概率测度 P, $P(i) = \frac{1}{5}$,属性集 $AT = \{a, b\}$, a = 价格, b = 内部设计, $V_a = \{\{H\}, \{L\}, \{*\}\}$,其中 H = "高", L = "低",这里 $*$ = "H, L",表示属性值或者是"高"或者是"低"但究竟是"高"还是"低"不能确定.

表 10.1　一个不完备信息系统

U	a	b
1	高	宽敞
2	低	宽敞
3	*	拥挤
4	高	宽敞
5	低	宽敞

表 10.2　系统的完备化

U	a	b
1	高	宽敞
2	低	宽敞
3	高	拥挤
4	高	宽敞
5	低	宽敞

表 10.3　系统的完备化

U	a	b
1	高	宽敞
2	低	宽敞
3	低	拥挤
4	高	宽敞
5	低	宽敞

关于属性 a,对象属性描述函数 F_a 是一个随机集:

$F_a(1) = F_a(4) = \{H\}$,　$F_a(2) = F_a(5) = \{L\}$,　$F_a(3) = \{H, L\}$.

F_a 导出的概率分配函数:

$$m(\{H\}) = m(\{L\}) = \frac{2}{5}, \quad m(\{H, L\}) = \frac{1}{5}.$$

概率分配函数的支撑:

$$\mathscr{M} = \{\{H\}, \{L\}, \{H, L\}\}.$$

F_a 导出的信任函数:

$$B(\{H\}) = P(\underline{\mathrm{apr}}\{H\}) = \frac{2}{5}, \quad B(\{L\}) = P(\underline{\mathrm{apr}}\{L\}) = \frac{2}{5},$$

$$B(\{H, L\}) = P(\underline{\mathrm{apr}}\{H, L\}) = 1.$$

F_a 导出的似然函数:

$$L(\{H\}) = P(\overline{\mathrm{apr}}\{H\}) = \frac{3}{5}, \quad L(\{L\}) = P(\overline{\mathrm{apr}}\{L\}) = \frac{3}{5},$$

$$L(\{H,L\}) = P(\overline{\mathrm{apr}}\{H,L\}) - 1.$$

这时信任函数 B 的配置的一般形式为:

$$\begin{aligned} P(\{H\}) &= \lambda(\{H\},H)m(\{H\}) + \lambda(\{H,L\},H)m(\{H,L\}) \\ &= \frac{2}{5} + \lambda(\{H,L\},H)\frac{1}{5}; \end{aligned}$$

$$\begin{aligned} P(\{L\}) &= \lambda(\{L\},L\}m(\{L\}) + \lambda(\{H,L\},L)m(\{H,L\}) \\ &= \frac{2}{5} + \lambda(\{H,L\},L)\frac{1}{5}, \end{aligned}$$

其中 $\lambda(\{H,L\},H)+\lambda(\{H,L\},L)=1, \lambda(\{H,L\},H)\geqslant 0, \lambda(\{H,L\},L)\geqslant 0$.

F_a 共有两个可测选择 f_1 和 f_2:

$$f_1(1) = f_1(3) = f_1(4) = H, \quad f_1(2) = f_1(5) = L,$$

$$f_2(1) = f_2(4) = H, \quad f_2(2) = f_2(3) = f_2(5) = L.$$

它们相应地产生原系统(表 10.1)的两个完备化系统(表 10.2 和表 10.3). f_1 和 f_2 分别产生$(\{H,L\},2^{\{H,L\}})$上的两个概率测度 P_{f_1} 和 P_{f_2}:

$$P_{f_1}(\{H\}) = P(\{1,3,4\}) = \frac{3}{5}, \quad P_{f_1}(\{L\}) = P(\{2,5\}) = \frac{2}{5}.$$

$$P_{f_2}(\{H\}) = P(\{1,4\}) = \frac{2}{5}, \quad P_{f_2}(\{L\}) = P(\{2,3,5\}) = \frac{3}{5}.$$

可见概率测度 P_{f_1} 和 P_{f_2} 都是信任函数 B 的配置,而且在配置 P_{f_1} 的表达式中,

$$\lambda(\{H,L\},H) = 1, \quad \lambda(\{H,L\},L) = 0.$$

在配置 P_{f_2} 的表达式中,

$$\lambda(\{H,L\},H) = 0, \quad \lambda(\{H,L\},L) = 1.$$

显然 $\mathscr{P}_{F_a} = \{P_{f_1}, P_{f_2}\}$ 是 core B 的一个真子集,但可以验证对于任意 $X \in 2^{\{H,L\}}$ 有

$$\inf\{P_{f_1}(X), P_{f_2}(X)\} = P(\underline{\mathrm{apr}}X) = B(X),$$

$$\sup\{P_{f_1}(X), P_{f_2}(X)\} = P(\overline{\mathrm{apr}}X) = L(X).$$

即信任函数 B 恰好是 $\mathscr{P}_{F_a}$ 的概率下界,L 恰好是 $\mathscr{P}_{F_a}$ 的概率上界.

§10.4 近似算子与可能性测度

Zadah 在模糊集基础上建立了可能性理论,Dubois 和 Prade 进一步发展了该理论.这一节讨论基于随机集的粗糙近似与可能性测度的关系.

定义 10.7　集函数 $N:2^U\to[0,1]$称为必然性测度(necessity measure),若满足

(1) $N(\varnothing)=0, N(U)=1$.

(2) $N(X\cap Y)=N(X)\wedge N(Y), \forall X,Y\in 2^U$.

集函数 $\Pi:2^U\to[0,1]$称为可能性测度(possibility measure),若满足

(1) $\Pi(\varnothing)=0, \Pi(U)=1$.

(2) $\Pi(X\cup Y)=\Pi(X)\vee\Pi(Y), \forall X,Y\in 2^U$.

定理 10.16　(1)若 Π 是 U 上的可能性测度,则

$$N(X)=1-\Pi(\sim X),\quad X\in 2^U.$$

是 U 上的必然性测度.

(2)若 N 是 U 上的必然性测度,则

$$\Pi(X)=1-N(\sim X),\quad X\in 2^U.$$

是 U 上的可能性测度.

证　(1)易证 $N(\varnothing)=0, N(U)=1$ 成立.因为

$$\begin{aligned}N(X\cap Y)&=1-\Pi(\sim(X\cap Y))\\&=1-\Pi((\sim X)\cup(\sim Y))\\&=1-\Pi(\sim X)\vee\Pi(\sim Y)\\&=(1-\Pi(\sim X))\wedge(1-\Pi(\sim Y))\\&=N(X)\wedge N(Y),\end{aligned}$$

所以 N 是必然性测度.

(2)类似可证.　□

定理 10.17　有限集 U 上的每个可能性测度 Π 可以用 U 上的正则模糊集惟一地表示,这个正则模糊集称为 Π 的可能性分布.

证　设 Π 是 U 上的可能性测度,记

$$A(x)=\Pi(\{x\}).$$

对于任意 $X\subseteq U$,根据可能性测度的定义有

$$\Pi(X)=\bigvee_{x\in X}\Pi(\{x\})=\bigvee_{x\in X}A(x),$$

由于 U 是有限集,且 $\Pi(U)=\max\limits_{x\in U}A(x)=1$,因此 A 是 U 上的正则模糊集.惟一性显然.　□

定理 10.18

(1) U 上的每个可能性测度 Π 必为 U 上的似然函数.

(2) U 上的每个必然性测度 N 必为 U 上的信任函数.

证　(1)设 $X_1,X_2,X_3\in 2^U$,不妨设

$$\Pi(X_1)\leqslant\Pi(X_2)\leqslant\Pi(X_3).$$

于是

$$\begin{aligned}
&\sum_{i=1}^{3}\Pi(X_i)-\sum_{i<j}\Pi(X_i\cup X_j)+\Pi(X_1\cup X_2\cup X_3)\\
&=\sum_{i=1}^{3}\Pi(X_i)-\sum_{i<j}\Pi(X_i\cup X_j)+\Pi(X_3)\\
&=\Pi(X_1)+\Pi(X_2)+\Pi(X_3)-\Pi(X_2)\\
&\quad-\Pi(X_3)-\Pi(X_3)+\Pi(X_3)\\
&=\Pi(X_1)\geqslant\Pi(X_1\cap X_2\cap X_3).
\end{aligned}$$

一般情况类似可证.

(2)设 N 是 U 上的必然性测度,则由定理 10.16 可知

$$\Pi(X)=1-N(\sim X)$$

是 U 上的可能性测度.由(1)知 Π 是 U 上的似然函数,于是由对偶性知

$$N(X)=1-\Pi(\sim X)$$

是 U 上的信任函数. □

定义 10.8 设 m 是 U 上的 mass 函数,若 m 的焦点集 $\mathscr{M}=\{X\mid m(X)>0\}$ 满足集合套性质,即

$$X_1\subset X_2\subset\cdots\subset X_k,$$

$X_i\in\mathscr{M}$,则称 mass 函数 m 是协调的.

定理 10.19

(1)一个协调的 mass 函数构成的信任函数 B 是必然性测度.

(2)一个协调的 mass 函数构成的似然函数 L 是可能性测度.

证 (1)设 m 是 U 上的 mass 函数,m 的焦点集 $\mathscr{M}=\{X\mid m(X)>0\}$ 满足集合套性质

$$X_1\subset X_2\subset\cdots\subset X_k.$$

对于任意的 $X,Y\in 2^U$,只须证

$$B(X\cap Y)=B(X)\wedge B(Y). \tag{10.22}$$

记

$$i_1=\max\{i\mid X_i\subseteq X\},\quad i_2=\max\{i\mid X_i\subseteq Y\},$$

则 $X_i\subseteq X\cap Y$ 当且仅当 $i\leqslant\min\{i_1,i_2\}$,于是

$$\begin{aligned}
B(X\cap Y)&=\sum_{i=1}^{\min\{i_1,i_2\}}m(X_i)\\
&=\min\left\{\sum_{i=1}^{i_1}m(X_i),\sum_{i=1}^{i_2}m(X_i)\right\}\\
&=B(X)\wedge B(Y).
\end{aligned}$$

即(10.22)成立,所以 B 是必然性测度.

(2)证法与(1)类似. □

定义 10.9　设 $F:U\to 2^W$ 是随机集,称 F 满足集合套性质,若对于任意 $x,y\in U$,$F(x)\subseteq F(y)$ 或 $F(y)\subseteq F(x)$ 成立.

设 j 是随机集 F 的关系划分函数,$\mathscr{J}$ 为 j 的焦点集全体,记

$$\mathscr{J}=\{X\in 2^W\mid j(X)\neq\varnothing\}=\{X_1,X_2,\cdots,X_k\}.$$

对于任意的 $X,Y\in\mathscr{J}$ 由定理 10.9 知,或者 $X\subset Y$,或者 $Y\subset X$.不失一般性可设

$$X_1\subset X_2\subset\cdots\subset X_k. \tag{10.23}$$

定理 10.20　设随机集 $F:U\to 2^W$ 满足集合套性质,则对于任意 $X,Y\in 2^W$,近似算子满足下列性质:

(1) $\underline{\mathrm{apr}}X\subseteq\underline{\mathrm{apr}}Y$ 或者 $\underline{\mathrm{apr}}Y\subseteq\underline{\mathrm{apr}}X$.

(2) $\overline{\mathrm{apr}}X\subseteq\overline{\mathrm{apr}}Y$ 或者 $\overline{\mathrm{apr}}Y\subseteq\overline{\mathrm{apr}}X$.

(3) $\underline{\mathrm{apr}}(X\cap Y)=\underline{\mathrm{apr}}X$ 或者 $\underline{\mathrm{apr}}(X\cap Y)=\underline{\mathrm{apr}}Y$.

(4) $\overline{\mathrm{apr}}(X\cup Y)=\overline{\mathrm{apr}}X$ 或者 $\overline{\mathrm{apr}}(X\cup Y)=\overline{\mathrm{apr}}Y$.

证　(1)由于 F 满足集合套性质,因此 F 的关系划分函数 j 的焦点集 $\mathscr{J}$ 具有式(10.23)的形式.记

$$\tau(X)=\max\{i\mid X_i\subseteq X\},\quad \tau(Y)=\max\{i\mid X_i\subseteq Y\},$$

则

$$\tau(X)\leqslant\tau(Y)\text{ 或者 }\tau(Y)\leqslant\tau(X). \tag{10.24}$$

又由定理 10.10 知

$$\underline{\mathrm{apr}}X=\bigcup_{1\leqslant i\leqslant\tau(X)}j(X_i),\quad \underline{\mathrm{apr}}Y=\bigcup_{1\leqslant i\leqslant\tau(Y)}j(X_i),$$

从而由(10.24)知性质(1)成立.

(2)证法与(1)类似.

(3)结合(1)和定理 10.6 的性质(3)即得.

(4)结合(2)和定理 10.6 的性质(3)即得. □

定理 10.21　设随机集 $F:U\to 2^W\setminus\{\varnothing\}$满足集合套性质,记

$$N(X)=P(\underline{\mathrm{apr}}X),\Pi(X)=P(\overline{\mathrm{apr}}X),X\in 2^W. \tag{10.25}$$

则

(1)N 是 W 上的必然性测度;

(2)Π 是 W 上的可能性测度.

证　(1)首先由定理 10.7 知

$$N(\varnothing)=P(\underline{\mathrm{apr}}\varnothing)=P(\varnothing)=0, \tag{10.26}$$

$$N(W)=P(\underline{\mathrm{apr}}W)=P(U)=1. \tag{10.27}$$

其次,对于任意 $X,Y\in 2^W$ 由定理 10.20 可知

$$N(X \cap Y) = N(X) \text{ 或者 } N(X \cap Y) = N(Y),$$

从而

$$N(X \cap Y) \geqslant N(X) \wedge N(Y). \tag{10.28}$$

又由定理 10.6 知

$$\begin{aligned} N(X \cap Y) &= P(\underline{\mathrm{apr}}(X \cap Y)) \\ &\leqslant P(\underline{\mathrm{apr}}X \cap \underline{\mathrm{apr}}Y) \\ &\leqslant P(\underline{\mathrm{apr}}X) \wedge P(\underline{\mathrm{apr}}Y) \\ &= N(X) \wedge N(Y). \end{aligned} \tag{10.29}$$

由(10.28)和(10.29)得

$$N(X \cap Y) = N(X) \wedge N(Y). \tag{10.30}$$

由(10.26),(10.27)和(10.30)知 N 是 W 上的必然性测度.

(2)证法与(1)类似. □

定理 10.22 设随机集 $F: U \to 2^W \setminus \{\varnothing\}$ 满足集合套性质,h 为近似空间 (U, W, F, P) 的上近似分布,Π 为(10.25)式定义的可能性测度,记 $\pi(x) = P(h(x))$,则 π 是 Π 的可能性分布.

证 由定理 10.21 对于任意 $X \in 2^W$

$$\begin{aligned} \Pi(X) &= \Pi\Big(\bigcup_{x \in X} \{x\}\Big) = \bigvee_{x \in X} \Pi(\{x\}) \\ &= \bigvee_{x \in X} P(\overline{\mathrm{apr}}\{x\}) = \bigvee_{x \in X} \pi(x). \end{aligned}$$

由定理 10.17 的证明知 π 是 Π 的可能性分布. □

第十一章 不完备信息系统的粗糙集方法

在许多情况下,面临的信息系统是不完备的.主要问题之一是属性的缺省值.在现实生活中,绝大多数信息系统都是不完备的.本章主要介绍不完备信息系统下的粗糙集模型,知识约简和决策规则提取.

§11.1 不完备信息系统

$S=(U,AT)$是信息系统,其中 U 是对象的非空有限集合,AT 是属性的非空有限集合,对于每个 $a\in AT$ 有 $a:U\to V_a$,其中 V_a 称为 a 的值域.

每个属性子集 $A\subseteq AT$ 决定了一个不可区分关系 $\mathrm{ind}(A)$:

$$\mathrm{ind}(A)=\{(x,y)\in U\times U\mid \forall a\in A, a(x)=a(y)\}.$$

关系 $\mathrm{ind}(A)(A\subseteq AT)$构成了 U 的划分,用 $U/\mathrm{ind}(A)$来表示.

对于一个对象,一些属性值可能是缺省的.为了表明这种情况,通常给定一个区分值(即空值 null value)给这些属性.

如果至少有一个属性 $a\in AT$ 使得 V_a 含有空值,则称 S 为一个不完备信息系统,否则它是完备的,我们用 $*$ 表示空值.

令 $A\subseteq AT$,我们定义相似关系如下:

$$SIM(A)=\{(x,y)\in U\times U\mid \forall a\in A, a(x)=a(y) \text{ or } a(x)=* \text{ or } a(y)=*\}.$$

定理 11.1 $SIM(A)$是一个相容关系:

$$SIM(A)=\bigcap_{a\in A}SIM(\{a\}).$$

令 $S_A(x)$表示对象集$\{y\in U\mid (x,y)\in SIM(A)\}$.对于 A 而言,$S_A(x)$是与 x 可能不可区分的对象的最大集合.

令 $D_A(x)$表示对象集$\{y\in U\mid (x,y)\notin SIM(A)\}$.对于 A 而言,$D_A(x)$是与 x 可能可区分的对象的最大集合.

对任意 $x\in U$,$S_A(x)\cap D_A(x)=\varnothing$且 $S_A(x)\cup D_A(x)=U$.

令 $U/SIM(A)=\{S_A(x)\mid x\in U\}$表示分类.$U/SIM(A)$中的任何元素称为相容类.$U/SIM(A)$中的相容类一般不构成 U 的划分,它们构成 U 的覆盖,$\bigcup U/SIM(A)=U$.

例 11.1 表 11.1 给出了几个小汽车的描述,我们按照选择的属性子集对它们分类.

从表 11.1 我们有 $U=\{1,2,3,4,5,6\}$，$AT=\{P,M,S,X\}$，其中 P,M,S,X 表示 Price, Mileage, Size, Max-Speed.

注 $U/SIM(AT)=\{S_{AT}(1),S_{AT}(2),S_{AT}(3),S_{AT}(4),S_{AT}(5),S_{AT}(6)\}$，其中 $S_{AT}(1)=\{1\}$，$S_{AT}(2)=\{2,6\}$，$S_{AT}(3)=\{3\}$，$S_{AT}(4)=\{4,5\}$，$S_{AT}(5)=\{4,5,6\}$，$S_{AT}(6)=\{2,5,6\}$.

容易验证：$U/SIM(\{P,S,X\})=U/SIM(AT)$，而 $U/SIM(\{S,X\})\neq U/SIM(AT)$（$U/SIM(\{S,X\})=\{S_A(1),S_A(2),S_A(3),S_A(4),S_A(5),S_A(6)\}$，其中 $A=\{S,X\}$ 且 $S_A(1)=S_A(2)=\{1,2,6\}$，$S_A(3)=\{3\}$，$S_A(4)=S_A(5)=\{4,5,6\}$，$S_A(6)=\{1,2,4,5,6\}$）.

表 11.1

Car	Price	Milenge	Size	Max-Speed
1	High	High	Full	Low
2	Low	*	Full	Low
3	*	*	Compact	High
4	High	*	Full	High
5	*	*	Full	High
6	Low	High	Full	*

在例 11.1 中，属性集 AT 和 $\{P,S,X\}$ 对小汽车的分类是相同的，而 AT 和 $\{S,X\}$ 对小汽车的分类是不同的. 通常，我们感兴趣的是 AT 的约简.

形式上，一个集合 $A\subseteq AT$ 是信息系统的一个约简，若 $SIM(A)=SIM(AT)$ 且 $\forall B\subset A$，$SIM(B)\neq SIM(AT)$.

对于例 11.1 给定的信息系统，$\{P,S,X\}$ 是一个约简.

对于 $x(x\in U)$，一个集合 $A\subseteq AT$ 是信息系统关于 x 的一个约简，若 $S_A(x)=S_{AT}(x)$ 且 $\forall B\subset A$，$S_B(x)\neq S_{AT}(x)$.

§11.2 近似集

令 $X\subseteq U$ 和 $A\subseteq AT$. $\underline{A}X$ 是 X 的下近似当且仅当

$$\underline{A}X=\{x\in U\mid S_A(x)\subseteq X\}=\{x\in X\mid S_A(x)\subseteq X\}.$$

$\overline{A}X$ 是 X 的上近似当且仅当

$$\overline{A}X=\{x\in U\mid S_A(x)\cap X\neq\varnothing\}=\bigcup\{S_A(x)\mid x\in X\}.$$

与完备信息系统相似,$\underline{A}X$ 是肯定属于 X 的对象的集合,而 $\overline{A}X$ 是可能属于 X 的对象的集合.

定理 11.2 对于 $\forall A,B\subseteq AT$, $\forall X\subseteq U$,有

(1) $\underline{A}X\subseteq X\subseteq\overline{A}X$.

(2) $A\subset B\Rightarrow\underline{A}X\subseteq\underline{B}X$.

(3) $A\subset B\Rightarrow\overline{A}X\supseteq\overline{B}X$.

证 直接验证即得. □

§11.3 决策表,决策规则和知识约简

(不完备)决策表(DT)是一个(不完备)信息系统 $DT=(U,AT\cup\{d\})$,其中 $d(d\notin AT$ 且 $*\notin V_d)$称为决策属性,AT 中的元素称为条件属性.

定义函数$\partial_A:U\rightarrow P(V_d)$,$A\subseteq AT$ 为

$$\partial_A(x)=\{i\mid i=d(y),y\in S_A(x)\}.$$

称∂_A 为 DT 中的广义决策函数,其中 $P(V_d)$表示 V_d 的幂集.

如果对任意 $x\in U$ 有$|\partial_{AT}(x)|=1$,则 DT 是协调的(确定的);否则它是不协调的(不确定的).易证下面定理成立:

定理 11.3 关系 $\mathrm{ind}(\partial_A)$,$A\subseteq AT$,构成了 U 的一个划分.

定理 11.4 对于不完备 DT

$$X\in U/\mathrm{ind}(\partial_A)\Rightarrow\underline{A}X=X=\overline{A}X$$

不成立(尽管对于完备 DT 成立).

任何决策表可以看作为如下形式的(广义)决策规则集:

$$\wedge(c,v)\rightarrow\vee(d,w),$$

其中 $c\in AT$,$v\in V_c$,$w\in V_d$. $\wedge(c,v)$称为规则的条件部分,$\vee(d,w)$称为规则的决策部分.

令 X 是具有性质$\wedge(c,v)(c\in AT,v\in V_c)$的对象集,$Y$ 是具有性质$\vee(d,w)(w\in V_d)$的对象集.

在 DT 中,决策规则$\wedge(c,v)\rightarrow\vee(d,w)$为真当且仅当 $\overline{C}X\subseteq Y$,其中 C 是出现在规则的条件部分的所有属性构成的集合.

在 DT 中,决策规则 $r:\wedge(c,v)\rightarrow\vee(d,w)(c\in AT,v\in V_c,w\in V_d)$是最优的当且仅当该规则为真且由出现在 r 中的合取与析取的真子集构成的任何规则均为假.

例 11.2 通过增加决策属性 $d=$ Acceleration 扩充表 11.1 为表 11.2,然后确定决策类 $U/\mathrm{ind}(d)$和广义决策类 $U/\mathrm{ind}(\partial_{AT})$,并且对子每个决策类计

算其下近似和上近似,最后写出为真的决策规则.

表 11.2

Car	Price	Mileage	Size	Max-Speed	d
1	High	High	Full	Low	Good
2	Low	*	Full	Low	Good
3	*	*	Compact	High	Poor
4	High	*	Full	High	Good
5	*	*	Full	High	Excel
6	Low	High	Full	*	Good

从表 11.2 我们有 $U/\mathrm{ind}(d)=\{X_{\text{good}},X_{\text{poor}},X_{\text{excel}}\}$,其中 $X_{\text{good}}=\{1,2,4,6\}$,$X_{\text{poor}}=\{3\}$,$X_{\text{excel}}=\{5\}$.

因此

$$\underline{AT}X_{\text{good}}=\{1,2\};\qquad \overline{AT}X_{\text{good}}=\{1,2,4,5,6\};$$
$$\underline{AT}X_{\text{poor}}=\{3\};\qquad \overline{AT}X_{\text{poor}}=\{3\};$$
$$\underline{AT}X_{\text{excel.}}=\varnothing;\qquad \overline{AT}X_{\text{excel}}=\{4,5,6\}.$$

在表 11.3 中,我们给出了广义决策值.

表 11.3

Car	∂_{AT}
1	{Good}
2	{Good}
3	{Poor}
4	{Good, Excel}
5	{Good, Excel}
6	{Good, Excel}

$$U/\mathrm{ind}(\partial_{AT})=\{X_{\{\text{good}\}},X_{\{\text{poor}\}},X_{\{\text{good,excel}\}}\},$$

其中 $X_{\{\text{good}\}}=\{1,2\}$,$X_{\{\text{poor}\}}=\{3\}$,$X_{\{\text{good,excel}\}}=\{4,5,6\}$. 因此

$$\underline{AT}X_{\{\text{good}\}}=\{1\};\qquad \overline{AT}X_{\{\text{good}\}}=\{1,2,6\};$$
$$\underline{AT}X_{\{\text{poor}\}}=\{3\};\qquad \overline{AT}X_{\{\text{poor}\}}=\{3\};$$
$$\underline{AT}X_{\{\text{good,excel}\}}=\{4,5\};\qquad \overline{AT}X_{\{\text{good,excel}\}}=\{2,4,5,6\}.$$

对于 DT，我们列出为真的决策规则：

r_1:$(P,\text{high})\wedge(M,\text{high})\wedge(S,\text{full})\wedge(X,\text{low})\rightarrow(d,\text{good})$;

r_2:$(P,\text{low})\wedge(M,*)\wedge(S,\text{full})\wedge(X,\text{low})\rightarrow(d,\text{good})$;

r_3:$(P,*)\wedge(M,*)\wedge(S,\text{compact})\wedge(X,\text{high})\rightarrow(d,\text{poor})$;

r_4:$(P,\text{high})\wedge(M,*)\wedge(S,\text{full})\wedge(X,\text{high})\rightarrow(d,\text{good})\vee(d,\text{excel})$;

r_5:$(P,*)\wedge(M,*)\wedge(S,\text{full})\wedge(X,\text{high})\rightarrow(d,\text{good})\vee(d,\text{excel})$;

r_6:$(P,\text{low})\wedge(M,\text{high})\wedge(S,\text{full})\wedge(X,*)\rightarrow(d,\text{good})\vee(d,\text{excel})$.

从广义决策函数、规则为真和最优决策规则的定义可知，对于 $x(x\in U)$，最优规则的决策部分等于$(d,w_1)\vee(d,w_2)\vee\cdots\vee(d,w_n)$，其中$\{w_1,w_2,\cdots,w_n\}=\partial_{AT}(x)$. 这样，寻找最优规则的问题就限制到了条件属性的约简问题.

形式上，集合 $A\subseteq AT$ 是 DT 的一个约简(相对约简)，若

$$\partial_A=\partial_{AT}\ \text{且}\ \forall B\subset A,\partial_B\neq\partial_{AT}.$$

为了确定具有最小条件数的决策规则，在 DT 中，对于一个对象我们可以使用约简的概念.

在 DT 中，对于 $x(x\in U)$，集合 $A\subseteq AT$ 是 DT 的一个约简，若

$$\partial_A(x)=\partial_{AT}(x)\ \text{且}\ \forall B\subset A,\partial_B(x)\neq\partial_{AT}(x).$$

定理 11.5　设 A 是一个相对约简，则对于不完备 DT

$$X\in U/\mathrm{ind}(\partial_{AT})\Rightarrow \underline{A}X=\underline{AT}X,$$

$$X\in U/\mathrm{ind}(\partial_{AT})\Rightarrow \overline{A}X=\overline{AT}X$$

不成立.

例 11.3　利用表 11.2 中 DT 的相对约简来说明定理 11.5.

容易验证，$A=\{\text{Size, Max-Speed}\}$是表 11.2 所描述的 DT 的一个约简. 对于属性集 A，我们给出 $U/\mathrm{ind}(\partial_{AT})$中的类的下近似和上近似：

$$\underline{A}X_{\{\text{good}\}}=\varnothing,\qquad \overline{A}X_{\{\text{good}\}}=\{1,2,6\},$$

$$\underline{A}X_{\{\text{poor}\}}=\{3\},\qquad \overline{A}X_{\{\text{poor}\}}=\{3\},$$

$$\underline{A}X_{\{\text{good,excel}\}}=\{4,5\},\qquad \overline{A}X_{\{\text{good,excel}\}}=\{1,2,4,5,6\},$$

其中 $X_{\{\text{good}\}}$，$X_{\{\text{poor}\}}$和 $X_{\{\text{good,excel}\}}$与在例 11.2 中的意义相同.

上述近似集与例 11.2 中计算的近似集相比较，有下面关系式：

$$\underline{A}X_{\{\text{good}\}}\subset\underline{AT}X_{\{\text{good}\}},$$

$$\overline{A}X_{\{\text{good,excel}\}}\supset\overline{AT}X_{\{\text{good,excel}\}}.$$

从表 11.2 容易看出：$\{Max\text{-}Speed\}$是对象 1 和 2 的一个约简，Size 是对象

3～6 的一个约简，这些约简允许我们得到下面三个最优决策规则：

$$r'_1:(X,\text{low})\to(d,\text{good}),$$
$$r'_2:(S,\text{compact})\to(d,\text{poor}),$$
$$r'_3:(S,\text{full})\to(d,\text{good})\vee(d,\text{excel}).$$

注：当 DT 中所有或一些缺省值由任意值代替时，上述决策规则将保持为真.

§11.4 区分函数与约简的计算

这一节考虑利用区分函数计算不完备信息系统和不完备决策表的约简.

设 $\alpha_A(x,y)$ 是满足 $(x,y)\notin SIM(\{a\})$ 的属性 $a\in A$ 的集合. 因此，如果 $(x,y)\in SIM(A)$，则 $\alpha_A(x,y)=\varnothing$. 令 $\sum\alpha_A(x,y)$ 是一个布尔表达式. 如果 $\alpha_A(x,y)=\varnothing$，则 $\sum\alpha_A(x,y)=1$；否则，$\sum\alpha_A(x,y)$ 是包含在 $\alpha_A(x,y)$ 中的属性所对应变量的析取.

Δ 是不完备信息系统的区分函数，若

$$\Delta=\prod_{(x,y)\in U\times U}\sum\alpha_{AT}(x,y).$$

$\Delta(x)$ 是不完备信息系统中对象 x 的区分函数，若

$$\Delta(x)=\prod_{y\in U}\sum\alpha_{AT}(x,y).$$

Δ^* 是不完备决策表 DT 的区分函数，若

$$\Delta^*=\prod_{(x,y)\in U\times\{z\in U\mid d(z)\notin\partial_{A_I}(x)\}}\sum\alpha_{A_I}(x,y).$$

$\Delta^*(x)$ 是不完备决策表 DT 中对象 x 的区分函数，若

$$\Delta^*(x)=\prod_{y\in\{z\in U\mid d(z)\notin\partial_{AT}(x)\}}\sum\alpha_{AT}(x,y).$$

例 11.4 利用区分函数 Δ 计算表 11.1 给出的不完备信息系统的所有约简.

表 11.4 给出了 $\alpha_{AT}(x,y)$ 的值，其中 $x,y\in U$. 因此，我们有

$$\Delta=P(S\vee X)X(P\vee X)S=PSX,$$
$$\Delta(1)=P(S\vee X)X=PX,$$
$$\Delta(2)=P(S\vee X)(P\vee X)X=PX,$$
$$\Delta(3)=(S\vee X)S=S,$$
$$\Delta(4)=X(P\vee X)SP=PSX,$$
$$\Delta(5)=SX,$$
$$\Delta(6)=PS.$$

表 11.4

$x \setminus y$	1	2	3	4	5	6
1		P	S,X	X	X	P
2	P		S,X	P,X	X	
3	S,X	S,X		S	S	S
4	X	P,X	S			P
5	X	X	S			
6	P		S	P		

所以,{Price, Size, Max-Speed}是所述不完备信息系统的一个约简,{Price, Max-Speed}是对象 1 和 2 的一个约简等.

例 11.5　利用区分函数 Δ^* 计算表 11.2 给出的不完备决策表 DT 的所有约简.

表 11.5 给出了 $\alpha_{AT}(x,y)$的值,其中 $x \in U, y \in \{z \in U \mid d(z) \notin \partial_{AT}(x)\}$.

表 11.5

$x \setminus y$	1	2	3	4	5	6
1			S,X		X	
2			S,X		X	
3	S,X	S,X		S	S	S
4			S			
5			S			
6			S			

因此,我们有

$$\begin{aligned}
\Delta^* &= (S \vee X)XS = SX, \\
\Delta^*(1) &= (S \vee X)X = X, \\
\Delta^*(2) &= (S \vee X)X = X, \\
\Delta^*(3) &= (S \vee X)X = S, \\
\Delta^*(4) &= S, \\
\Delta^*(5) &= S; \\
\Delta^*(6) &= S.
\end{aligned}$$

所以,{Size, Max-Speed}是所述不完备决策表 DT 的一个约简(相对约简),{Max-Speed}是对象 1 和 2 的一个相对约简等.

参 考 文 献

[1] Ahn B S, Chao S S, Kim C Y. The integrated methodology of rough set theory and artificial neural network for business failure prediction. Expert Systems with Applications, 18 (2000): 65~74

[2] Banerjee M. and Pal S K. Roughness of a fuzzy set. Information Sciences , 93(1996): 235~246

[3] Banerjee M, Mitra S, and Pal S K. Rough fuzzy MLP: knowledge encoding and classification. IEEE Transactions on Neural Networks, 9(1998): 1203~1216

[4] Beaubouef T, Petry F and Arora G. Information-theoretic measures of uncertainty for rough sets and rough relational databases. Information Sciences, 109(1998): 185~195

[5] Beaubouef T, Petry F and Buckles B. Extension of the relational database and its algebra with rough set techniques. Computational Intelligence, 11(1995): 233~245

[6] Bell D A and Guan J W. Computation methods for rough classification and discovery. Journal of the American Society for Information Science, 49(1998): 403~414

[7] Bjorvand A T. Mining time series using rough sets – a case study. Lecture Notes in Artificial Intelligence, 1263(1997): 351~358

[8] Bodjanova S. Approximation of fuzzy concepts in decision making. Fuzzy Sets and Systems, 85(1997): 23~29

[9] Bonikowski Z, Bryniarski E and Wybraniec-Skardowska. Extensions and intentions in the rough set theory. Journal of Information Sciences, 107(1998): 149~167

[10] Bouguettaya A and Le Viet Q. Data clustering analysis in a multidimensional space. Information Sciences, 112(1998): 267~295

[11] Cattaneo G, Giuntini R and Pilla R. $BZMV^{dM}$algebras and stonian MV-algebras (applications to fuzzy sets and rough approximations). Fuzzy Set and Systems, 108(2000): 201~222

[12] Chakrabarty K, Biswas R and Nanda S. Fuzziness in rough sets. Fuzzy Sets and Systems, 110(2000): 247~251

[13] Chan C C. A rough set approach to attribute generalization in data mining. Journal of Information Sciences, 107(1998): 169~176

[14] Chanas S and Kuchta D. Further remarks on the relation between rough and fuzzy sets. Fuzzy Sets and Systems, 47(1992): 391~394

[15] Chateauneuf A and Jaffray A Y. Some characterizations of lower probabilities and other monotone capacities through the use of Mobius inversion. Math. Soc. Sci., 17(1989): 263~283

[16] Chmielewski M R and Grzymala-Busse J W. Global discretization of continuous attributes as preprocessing for machine learning. International Journal of Approximate Reasoning, 15(1996): 319~331

[17] Chow C K and Liu C N. Approximating discrete probability distributions with dependence trees. IEEE Transactions on Information Theory, IT-14(1968): 462~467

[18] Coker D. Fuzzy rough sets are intuitionistic L-fuzzy sets. Fuzzy Sets and Systems, 96 (1998): 381~383

[19] Demri S. A class of decidable information logics. Theoretical Computer Science, 195 (1998): 33~60

[20] Deogun J S, Choubey S K, Raghaven V V and Sever H. Feature selection and effective classifiers. Journal of the American Society for Information Science, 49(1998): 423~434

[21] Dimitras A I, Slowinski R, Susmaga R and Zopounidis C. Business failure prediction using rough sets. European Journal of Operational Research, 114(1999): 263~280

[22] Dubois D. and Prade H. Rough fuzzy sets and fuzzy rough sets. International Journal of General Systems, 17(1990): 191~209

[23] Dubois D and Prade H. Twofold fuzzy sets and rough sets - some issues in knowledge representation. Fuzzy Sets and Systems, 23(1987): 3~18

[24] Dubois D and Prade H. Fuzzy sets in approximate reasoning part 1: inference with possibility distributions. Fuzzy Sets and Systems 100 Supplement(1999): 73~132

[25] Duntsch I. Rough relation algebras. Fundamenta Informaticae, 21(1994): 321~331

[26] Duntsch I and Gediga G. Statistical evaluation of rough set dependency analysis. Int. J. Human-Computer Studies, 46(1997): 589~604

[27] Duntsch I and Gediga G. Simple data filtering in rough set systems. International Journal of Approximate Reasoning, 18(1998): 93~106

[28] Duntsch I. A logic for rough sets. Theoretical Computer Sciences, 179(1997): 427~436

[29] Duntsch I and Gediga G. Uncertainty measures of rough set prediction. Artificial Intelligence, 106(1998): 109~137

[30] Dutta B and Ray D. A concept of egalitarianism under participation constraints. Econometrica, 57(1989): 615~635

[31] Gehrke M and Walker E. On the structure of rough sets. Bulletin of the Polish Academy of Sciences: Computer Science, 40(1992): 235~245

[32] Gorzalczany M B and Piasta Z. Neuro-fuzzy approach versus rough-set inspired methodology for intelligent decision support. Information Sciences, 120(1999): 45~68

[33] Goutsis J et al.. Nguyen. Random Sets: Theory and applications, The IMA Volumes in Mathematics and its applications 39. New York: Springer-Verlag, 1997

[34] Greco S, Matarazzo B and Slowinski. Rough approximation of a preference relation by

dominance relations. European Journal of Operational Research, 117(1999): 63~83

[35] Green J, Horne N, Orlowska E and Siemens P. A rough set model of information retrieval. Fundamenta Informaticae, 28(1996): 273~296

[36] Griffin G and Chen Z. Rough set extension of Tcl for data mining. Knowledge-Based Systems, 11(1998): 249~253

[37] Grzymala-Busse E M and Grzymala-Busse J W. The usefulness of a machine learning approach to knowledge acquisition. Computational Intelligence, 11(1995): 268~279

[38] Guan J W and Bell D A. Rough computational methods for information systems. Artificial Intelligence, 105(1998): 77~103

[39] Harmanec D, Klir J and Resconi G. On modal logic interpretation of Dempster-Shafer theory of evidence. International Journal of Intelligent Systems, 9(1994): 941~951

[40] Hammer R L, Peled U N and Sorensen S. Pseudo-Boolean functions and game theory: Core elements and Shapley value. Cahier du CERO, 19(1977): 1~2

[41] Hashemi R L, Le Blanc L A and Rucks C T. A hybrid intelligent system for predicting bank holding structures. European Journal of Operational Research, 109(1998): 390~402

[42] Hellendoorn H. After the fuzzy wave reached Europe. European Journal of Operational Research, 99(1997): 58~71

[43] Herment M and Orlowska E. Handling information logics in a graphical proof editor. Computational Intelligence, 11(1995): 297~322

[44] Hisdal E. Are grades of membership probabilities? Fuzzy Sets and Systems, 25(1988): 325~348

[45] Xiaohua Hu and Cercone N. Learning in relational databases: a rough sets approach. Computational Intelligence, 11(1995): 323~355

[46] Iwinski T B. Algebraic approach to rough sets. Bulletin of the Polish Academy of Sciences, Mathematics, 35(1987): 673~683

[47] Jackson A G, Pawlak Z and LeClair S R. Rough sets applied to discovery of materials knowledge. Journal of Alloys and Compounds, 279(1998): 14~21

[48] Jagielska I. Using rough sets for knowledge discovery in the development of a decision support system for issuing smog alerts. Lecture Notes in Artificial Intelligence, 1394 (1998): 388~389

[49] Jagielska I, Matthews C and Whitfort T. An investigation into the application of neural networks, fuzzy logic, genetic algorithms, and rough sets to automated knowledge acquisition for classification problems. Neurocomputing, 24(1999): 37~54

[50] Kan Z. On rough sets and inference analysis. Lecture Notes in Computer Sciences, 1396 (1998): 256~265

[51] Kennes R. Computational aspects of the Mobius transformation graphs. IEEE Transactions on System, Man, and Cybernetics, 22(1992): 201~223

[52] Kerre E E. Fuzzy sets and approximate reasoning(English Edition). 西安交通大学出版社, 1999

[53] Komorowski J and Aleksander O. Modelling prognostic power of cardiac tests using rough sets. Artificial Intelligence in Medicine, 15(1999): 167～191

[54] Komorowski J, Polkowski L and Skowron A. Rough sets for data mining and knowledge discovery. Lecture Notes in Artificial Intelligence, 1263(1997): 393

[55] Kortelainen J. On the evaluation of compatibility with gradual rules in information systems: a topological approach. Control Cybernetics, 28(1999): 121～131

[56] Kortelainen J. On relationship between modified sets, topological spaces and rough sets. Fuzzy Sets and Systems, 61(1994): 91～95

[57] De Korvin A and McKeegan C. Knowledge acquisition using rough sets when membership values are fuzzy sets. Journal of Intelligent and Fuzzy Systems, 6(1998): 237～244

[58] Kovalerchuk B. Linguistic context spaces: necessary frames for correct approximate reasoning. International Journal of General Systems, 25(1996): 61～80

[59] Kryszkiewicz M. Rough set approach to incomplete information systems. Information Sciences, 112(1998): 39～49

[60] Kuncheva L I. Fuzzy rough sets: Application to feature selection. Fuzzy Sets and Systems, 51(1992): 147～153

[61] Kuroki N. Rough ideals in semigroups. Information Sciences, 100(1997): 139～163

[62] Deyu Li(李德玉), Yichen Ma(马逸尘). Invariant characters of information systems under some homomorphisms, Information Sciences, (to appear)

[63] Deyu Li(李德玉), Qi Chen(陈琪), Yichen Ma(马逸尘). Invariant characters of information systems under some homomorphisms I. In: Proceedings of the Third World Congress on Intelligent Control and Automation. 2000, 2731～2735

[64] Deyu Li(李德玉), Yichen Ma(马逸尘). On rough group. (to appear)

[65] Jiye Liang(梁吉业), Zongben Xu(徐宗本), Duoqian Miao(苗夺谦), Reduction of knowledge in incomplete information systems, The Sixteenth World Computer Congress, 2000

[66] Jiye Liang(梁吉业), Zongben Xu(徐宗本), Uncertainty measures of roughness of knowledge and rough sets in incomplete information systems, The Third World Congress on Intelligent Control and Automation, 2000, 2526～2529

[67] Jiye Liang(梁吉业), Zongben Xu(徐宗本), Jiang Wang(王江). Uncertainty measures of roughness of knowledge and signifcance of attributes in rough set data analysis. The Third Asian Control Conference. 2000

[68] Lin T Y. A rough logic formalism for fuzzy controllers: A hard and soft computing view. International Journal of Approximate Reasoning, 15(1996): 395～414

[69] Lin T Y. Introduction to the special issue on rough sets. International Journal of Approximate Reasoning, 15(1996): 287～289

[70] Lingras P J and Yao Y Y. Data mining using extensions of the rough set model. Journal of the American Society for Information Science, 49(1998): 415～422

[71] Lingras P. Comparison of neofuzzy and rough neural networks. Information Sciences, 110(1998): 207～215

[72] Lipski W. On semantic issues connected with incomplete information databases. ACM Transactions on Database Systems, 4(1979): 262～296

[73] Lotfi A, Andersen H C and Tsoi A C. Interpretation preservation of adaptive fuzzy inference systems. International Journal of Approximate Reasoning, 15(1996): 367～378

[74] Maler R P S. Representing rules as random sets, I: statistical corelations between rules. Information Sciences, 88(1996): 47～68

[75] McSherry D. Knowledge discovery by inspection. Decision Support Systems, 21(1997): 43～47

[76] Meyerowitz A, Richman F and Walker E. Calculating maximumentropy probability densities for belief functions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2(1994): 377～389

[77] Miyamoto S. Application of rough sets to information retrieval. Journal of the American Society for Information Science, 49(1998): 195～205

[78] Moradi H, Grzymala-Busse J W and Roberts A. Entropy of English text: experiments with humans and a machine learning system based on rough sets. Information Sciences, 104(1998): 31～47

[79] Novotny M and Pawlak Z. On a problem concerning dependence spaces. Fundamenta Informaticae, 16(1992): 275～287

[80] Morsi N N and Yakout M M. Axiomatics for fuzzy rough sets. Fuzzy Sets and Systems, 100(1998): 327～342

[81] Mrozek A. Rough sets and dependency analysis among attributes in computer implementations of experts inference models. International Journal of Man-Machine Studies, 30(1989): 457～473

[82] Nakamura A and Gao J M. A logic for fuzzy data analysis. Fuzzy Sets and Systems, 39(1991): 127～132

[83] Nakamura A. A rough logic based on incomplete information and its application. International Journal of Approximate Reasoning, 15(1996): 367～378

[84] Nanda S and Majumdar S. Fuzzy rough sets. Fuzzy Sets and Systems, 45(1992): 157～160

[85] Nguyen H T. On random sets and belief functions. Journal of Mathematical Analysis and Applications, 65(1978): 531～542

[86] Nguyen H T and Wlker E A. A History and introduction to the algebra of conditional events and probability logic. IEEE Transaction Systems, Man, and Cybernetics, 24(1994): 1671～1675

[87] Nurmi H, Kacprzyk J and Fedrizzi M. Probabilistic, fatty and rough concepts in social choice. European Journal of Operational Research, 95 (1996): 264～277

[88] Pal S K. Uncertainty management in space station autonomous research: pattern recognition perspective. Information Sciences, 72(1993): 1～63

[89] Pawlak Z. Rough sets. International Journal of Computer and Information Sciences, 11 (1982): 341～356

[90] Pawlak Z. Rough sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, Boston, 1991

[91] Pawlak Z. Rough classification. Int. J. Man-Machine Studies, 20(1984): 469～483

[92] Pawlak Z. Rough sets and fuzzy sets. Fuzzy Sets and Systems, 17(1985): 99～102

[93] Pawlak Z. Rough sets and decision tables. Lecture Notes in Computer Science, 208 (1985): 187～196

[94] Pawlak Z. On decision tables. Bulletin of the Polish Academy of Sciences: Technical Sciences, 34(1986): 563～589

[95] Pawlak Z et al.. Rough Sets. Communication of the ACM, 38(1995): 89～95

[96] Pawlak Z. Rough set theory and its applications to data analysis. Cybernetics and Systems: An International Journal, 29(1998): 661～688

[97] Pawlak Z. An inquiry into anatomy of conflicts. Journal of Information Sciences, 109 (1998): 65～78

[98] Pawlak Z. Vagueness and uncertainty: A rough set perspective. Computational Intelligence, 11(1995): 227～232

[99] Pawlak Z, Wong S K M and Ziarko W. Rough sets: probabilistic versus deterministic approach. International Journal of Man-Machine Studies, 29(1988): 81～95

[100] Pawlak Z. Rough set approach to Knowledge-based decision support. European Journal of Operational Research, 99(1997): 48～57

[101] Pedrycz W. Shadowed sets: representing and processing fuzzy sets. IEEE Transactions on Systems, Man, and Cybernetics-part B: Cybernetics, 28(1998): 103～109

[102] Plonka L and Mrozek A. Rule-based stabilization of the inverted pendulum. Computational Intelligence, 11(1995): 348～356

[103] Polkowski L and Skowron A. Rough mereology: a new paradigm for approximate reasoning. International Journal of Approximate Reasoning, 15(1996): 333～365

[104] Pomerol J C. Artificial intelligence and human decision making. European Journal of Operational Research, 99(1997): 3～25

[105] Pomykala J and Pomykala J A. The Stone algebra of rough sets. Bulletin of the Polish Academy of Sciences: Mathematics, 36(1988): 495～512

[106] Quafafou M. α-RST: a generalization of rough set theory. Information Sciences, 124 (2000): 301～316

[107] Richards D and Compton P. An alternative verification and validation technique for an

alternative knowledge representation and acquisition technique. Knowledge-Based Systems, 12(1999): 55~73

[108] Ruspini E H. Approximate reasoning: past, present, future. Information Sciences, 57-58(1991): 297~317

[109] Salonen H and Nurmi H. A note on rough sets and common knowledge events. European Journal of Operational Research, 112(1999): 692~695

[110] Shafer G. A mathematical theory of evidence. Princeton University Press, Princeton, 1976

[111] Shan N and Ziarko W. Data-based acquisition and incremental modification of classification rules. Computational Intelligence, 11(1995): 357~370

[112] Shapley L. Cores of convex games. International Journal of Game Theory, 1(1971): 11~26

[113] En-Wei Shi(施恩伟). Some Properties of the indisernibility relation in rough sets. Chinese Science Bulletin, 35(1990): 338~341

[114] Shore J E and Johnson R W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. IEEE Transactions on Information Theory, IT-26(1980): 26~37

[115] Sivakumar K and Goutsias J. Binary random fields, random closed sets and morphological sampling. IEEE Transactions on Image Processing, 5(1996): 899~912

[116] Skowron A. The rough sets theory and evidence theory. Fundamenta Informatae XIII (1990): 245~262

[117] Skowron A. Extracting laws from decision tables: A rough set approach. Computational Intelligence, 11(1995): 371~388

[118] Slowinski R. Intelligent decision support——handbook of applications and advances of rough sets theory. Dordrecht: Kluwer Academic publishers, 1992

[119] Slowinski R and Stefanowski J. Rough classification in incomplete information systems. Math. Comput. Modelling, 12(1989): 1347~1357

[120] Slowinski R, Zopounidis C and Dimitras A I. Prediction of company acquisition in Greece by means of the rough set approach. European Journal of Operational Research, 100(1997): 1~15

[121] Stepaniuk J. Attribute discovery and rough sets. Lecture Notes in Artificial Intelligence, 1263(1997): 145~155

[122] Suraj Z. Reconstruction of cooperative information systems under cost constraints: A rough set approach. Information Sciences, 111(1998): 273~291

[123] Tanaka T, Ishibuchi H and Matsuda N. Fuzzy expert system based on rough sets and its application to medical diagnosis. International Journal of General Systems, 21(1992): 83~97

[124] Tentush I. On rough mereological Cech topologies. Bulletin of the Polish Academy of

Sciences: Computer Science, 43(1995): 75~85

[125] Tsumoto S. Automated extraction of medical expert system rules from clinical databases based on rough set theory. Information Sciences, 112(1998): 67~84

[126] Tsumoto S and Tanaka H. Primerose: Probabilistic rule induction method based on rough sets and resampling methods. Computational Intelligence, 11(1995): 389~405

[127] Tsumoto S and Tanaka H. Algebraic specification of empirical inductive learning methods based on rough sets and matroid theory. Lecture Notes in Computer Science, 958 (1995): 224~243

[128] Vakarelov D. A modal logic for similarity relations in Pawlak knowledge representation systems. Fundamenta Informaticae, 15(1991): 61~79

[129] Walczak B et. al.. Rough sets theory. Chemometrics and Intelligent Laboratory Systems, 47(1999): 1~16

[130] Webster L, Chen J G, Tan S S, C. Watson and A. De Korwin. Validation of authentic reasoning expert systems. Information Sciences, 117(1999): 19~46

[131] Wong S K M, Wang L S and Yao Y Y. On modeling uncertainty with interval structures. Computational Intelligence, 11(1995): 406~426

[132] Wong S K M and Ziarko W. Comparison of the probabilistic approximate classification and the fuzzy set model. Fuzzy Sets and Systems, 21(1987): 357~362

[133] Wong S K, Ziarko W and Ye R L. Comparison of rough-set and statistical methods in inductive learning. International Journal of Man-Machine Studies, 24(1986): 53~72

[134] Weizhi Wu(吴伟志) and Wenxiu Zhang(张文修). Neighborhood operator systems and approximations. (to appear)

[135] Wygralak M. Rough sets and fuzzy sets——some remarks on interrelations. Fuzzy Sets and Systems, 29(1989): 241~243

[136] Yahia M E, Mahmod R, Sulaiman N and Ahmad F. Rough neural expert systems. Expert Systems with Applications, 18(2000): 87~99

[137] Yao Y Y. A comparative study of fuzzy sets and rough sets. Information Sciences, 109 (1998): 227~242

[138] Yao Y Y. Measuring retrieval effectiveness based on user preference of documents. Journal of the American Society for Information Science, 46(1995): 133~145

[139] Yao Y Y. Relational interpretations of neighborhood operators and rough set approximation. Information Sciences, 111(1998): 239~259

[140] Yao Y Y. Two views of the theory of rough sets in finite universes. International Journal of Approximate Reasoning, 15(1996): 291~317

[141] Yao Y Y and Lin T Y. Generalization of rough sets using modal logic. Intelligent Automation and Soft computing, 2(1996): 103~120

[142] Yao Y Y and Lingras P. Interpretations of belief functions in the theory of rough sets. Information Sciences 104(1998): 81~106

[143] Yao Y Y. A decision theoretic framework for approximating concepts. International Journal of Man-Machine Studies, 37(1992): 793~809

[144] Yao Y Y. Constructive and algebraic methods of the theory of rough sets. Information Sciences, 109(1998): 21~47

[145] Yao Y Y and Wong S K M. Representation, propagation and combination of uncertain information. International Journal of General Systems, 23(1994): 59~83

[146] Yao Y Y, Wong S K M and Wang L S. A non-numeric approach to uncertain reasoning. International Journal of General Systems, 23(1995): 343~359

[147] Yasdi R. Combining rough sets learning-method and neural earning-method to deal with uncertain and imprecise information. Neurocomputing, 7(1995): 61~84

[148] Zadeh A. Fuzzy sets. Information and Control, 8(1965): 338~353

[149] Zhong N, Dong J Z and Ohsuga S. Data mining based on the generalization distribution table and rough sets. Lecture Notes in Artificial Intelligence, 1394(1998): 360~373

[150] Ziarko W. Variable precision rough set model. Journal of Computer and System Sciences, 46(1993): 39~59

[151] Ziarko W(Ed.). Rough sets, fuzzy sets and knowledge discovery. London: Spinger-Verlag, 1994

[152] Ziarko W. Introduction to the special issue on rough sets and knowledge discovery. Computational Intelligence, 11(1995): 223~226

[153] Ziarko W and Shan N. A Method for Computing all maximally general rules in attribute-value systems. Computational Intelligence, 12(1996): 223~234

[154] Ziarko W and Shan N. Discovering attribute relationships, dependencies and rules by using rough sets. In: Proceedings of the 28th Annual Hawaii International Conference on System Sciences, 1995. 293-299

[155] Hoaglin D C, Mosteller F and Tukry J W(陈忠琏,郭德媛译). 探索性数据分析,北京:中国统计出版社,1998

[156] Zhang Wenxiu (张文修), Le Huiling (乐惠玲). The structure of the norm system on Fuzzy sets. Math. Anal. and Appl., 127(1987): 559~568

[157] 张文修,杨金丽.合情推理与计算机思维.西安:西安交通大学出版社,1988

[158] 张文修,王国俊,刘旺金,方锦暄.模糊数学引论.西安:西安交通大学出版社,1991

[159] 张文修,汪振鹏,高勇.集值随机过程.北京:科学出版社,1996

[160] 张文修,梁怡.不确定性推理原理.西安:西安交通大学出版社,1996

[161] 张文修,梁广锡.模糊控制与系统.西安:西安交通大学出版社,1997

[162] 张文修,梁怡.遗传算法的数学基础.西安:西安交通大学出版社,2000

[163] 王珏,苗夺谦,周育健.关于 Rough Set 理论与应用的综述.模式识别与人工智能,9(1996):337~344

[164] 苗夺谦,王珏.粗糙集理论中知识粗糙性与信息熵关系的讨论.模式识别与人工智能,11(1998):34~40

[165] 刘清,黄兆华,刘少辉,姚力文.带 Rough 算子的决策规则及数据挖掘中的软计算.计算机研究与发展,36(1999):800～804

[166] 何新贵.模糊知识处理的理论与技术.北京:国防工业出版社,1998

[167] 曾黄麟.粗集理论及其应用.重庆:重庆大学出版社,1998

[168] 蔡自兴,徐光祐.人工智能及其应用.北京:清华大学出版社,1997

[169] 汪培庄,李洪兴.知识表示的数学理论.天津:天津科学技术出版社,1994

[170] 梁吉业,曲开社,徐宗本.信息系统的属性约简.系统工程的理论与实践(待发表)

[171] 吴伟志,张文修.基于随机集的粗糙集模型(Ⅰ).西安交通大学学报,34(12)(2000):15～19

[172] 张文修,徐萍,陈雁.内涵空间与外延空间.工程数学学报,8(3) (1991):22～28

[173] 陈雁,徐萍,张文修.关系数据库的证据生成与合成.计算机学报,17(1994):729～735

[174] 张文修.合情推理与创新思维的信息模型.西安交通大学学报,25(4) (1991):113～118

[175] 张文修.概率命题的合情推理.西安交通大学学报,26(2)(1992):1～6

[176] 张文修,徐萍,陈雁.随机关系数据库的证据生成与合成.西安交通大学学报,27(4)(1993):107～110

[177] 梁广锡,张文修.包含度及其在专家系统中的应用.工程数学学报,11(4) (1994):13～24

[178] 张梅,张朋柱,万百五,韩崇昭.决策问题模型化及其求解学习机制.管理工程学报,10(1) (1996):22～29

[179] 张文修,徐宗本,梁怡,梁广锡.包含度理论.模糊系统与数学,10(4) (1996):1～9

[180] 张文修.包含度理论——不确定性研究的方法学,中国科学基金,10(3) (1996):254～259

记 号 表

$\neg$	41	$\underline{A}_R$	162	$\mathscr{E}(X)$	78
$\models$	99	$\overline{A}_R$	162	$\mathrm{ext}_M(F)$	117
$\not\models$	99	$\mathscr{A}(F)$	122	$\mathscr{H}(U)$	159
$[\ ]$	116	Atom($\boldsymbol{A}$)	77	$h(x)$	53
$\langle\ \rangle$	116	$a' \rightarrow b$	79	HA	57
$\mid\quad\mid$	98	$acc_a(W)$	30	$H(R^*)$	152
$\sim$	4	$\underline{A}(X)(x)$	171	$H(S^* \mid R^*)$	152
$\eqsim$	11	$\overline{A}(X)(x)$	171	$H(X)$	134
$\simeq$	11,87	$\alpha(x,y)$	22	$H(Y \mid X)$	134
$\approx$	11,87	$\alpha^*(x,y)$	24	$H_0(S^* \mid R^*)$	153
$\equiv$	92	$\alpha_R(X)$	8	ind(**P**)	1
$\dot{\subset}$	11	$\alpha_R(\mathscr{F})$	10	$I(X;Y)$	135
$\subsetneq$	11	$\alpha(R,\beta,X)$	126	$j(X)$	53
$\ddot{\subseteq}$	11	$\Vert A \Vert$	170	$R(r_i \mid [x])$	143
Δ	22	$B(X)$	188	$\mathscr{J}(X)$	78
$\Delta(x)$	211	$\mathrm{bn}_R(X)$	4	LA	57
Δ^*	24	$\mathrm{bn}(X,\alpha,\beta)$	135	$L(X)$	188
$\Delta^*(x)$	211	$\boldsymbol{B}(U)$	77	$\lim\limits_{\alpha\downarrow r}$	137
$\underline{\in}_R$	7	$C(\boldsymbol{L})$	87	$\lim\limits_{\beta\uparrow r}$	137
$\overline{\in}_R$	7	$c(X,Y)$	123	$m(X)$	189
∂_A	208	$C^{\cup}(K)$	91	$n(A,B)$	179
$\underline{a}$	169	$C^{\cap}(K)$	91	$N(X)$	202
A_λ	160	core($\boldsymbol{P}$)	13	$\mathrm{neg}_R(X)$	4
A_λ^S	160	$\mathrm{core}_{\boldsymbol{P}}(\boldsymbol{Q})$	15	$n(X)$	41
$\underline{\mathrm{apr}}X$	43	$\mathscr{C}(X)$	78	$\mathrm{neg}(X,\alpha,\beta)$	136
$\overline{\mathrm{apr}}X$	43	$\mathrm{des}(x)$	21	$\mathrm{ndis}(R,X)$	127
$\underline{A}_\alpha$	164	$\mathrm{D}(A/B)$	179	$N_\theta(A)$	79
$\overline{A}_\beta$	164	$\mathrm{E}_r(X)$	155	$P(A \setminus B)$	132