第42卷 第6期
2015年6月
计 算 机 科 学
Computer Science
Vol. 42 No. 6
June 2015

格观念下的知识约简

马 丽[1,2] 米据生[1]

(河北师范大学数学与信息科学学院 石家庄 050024)[1] (石家庄经济学院信息工程学院 石家庄 050031)[2]

摘 要 经典的信息系统可看作是一种特殊的格结构。从格的角度给出了知识约简和协调集的概念,借助粗糙集的思想定义了一对新的上下近似算子,并在此基础上通过定义可辨识集,给出了约简和相对约简的判定定理。这种表示更深层次地揭示了知识的本质,将知识约简的已有相关结果做了进一步拓广。

关键词 格,分划,知识约简,相对约简

中图法分类号 TP182 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.6.018

Knowledge Reduction under View of Lattice

MA Li[1,2] MI Ju-sheng[1]

(College of Mathematics and Information Science, Hebei Normal University, Shijiazhuang 050024, China)[1]

(College of Information and Engineering, Shijiazhuang University of Economics, Shijiazhuang 050031, China)[2]

Abstract Classic information systems can be viewed as a special case of lattice structure. We proposed some new concepts such as knowledge reduction and consistent set based on rough set theory under the view of lattice. By means of defining lower approximate and upper approximate operators, we gave two specific reductions. Then, we presented the judgement theorems for consistent sets and proved them. These representations deeply reveal the essence of knowledge, and some relevant results of knowledge reduction are obtained.

Keywords Lattice, Partition, Knowledge reduction, Relative reduction

1 引言

计算机网络的发展以及为人们提供的信息量的不断增长,使得人们对信息分析工具的要求越来越高,人们希望从数据中获取其潜在的知识。粗糙集(Rough Set)理论是波兰数学家 Z. Pawlak 教授于 1982 年提出的一种能够定量分析处理不精确、不一致、不完整数据的数学工具[1]。粗糙集理论思想独特,最初的原型来源于较简单的信息模型,它的基本思想是通过关系数据库分类和归纳形成概念和规则的近似来实现知识发现。近 20 年间,粗糙集理论已成为一种重要的智能信息处理技术[2,3],该理论已经在人工智能学界的知识发现(规则提取、数据挖掘、机器学习)等方面得到广泛应用,知识发现的各种不同方法应运而生。

知识约简是粗糙集理论重要的研究内容之一,也是知识获取的关键步骤。所谓知识约简是指在保持知识库的分类能力不变的条件下,删除其中不必要的知识。通过删除冗余知识,可以大大提高信息系统潜在知识的清晰度。知识约简是进行系统简化和维护的前提,是知识发现研究的核心问题之一。目前,各种形式的知识约简和规则提取的研究取得了许多成果[7-10]。另外由于系统本身的不确定因素的广泛存在,导致约简所蕴含的知识易于发现,简化知识的表示方式。约简的过程,实际上是在保持某种不变的条件下,求得最小的知识集的过程。

经典的信息系统可看作是一种特殊的格结构。本文从格的角度出发,首先通过定义格上的分划,借助粗糙集的思想引入一对近似算子,并讨论它们的性质;其次在此基础上给出协调集和约简的定义,进而给出约简的判定条件;最后进一步给出相对协调集和相对约简的定义和判定定理。将格的已有相关结果做进一步推广,为知识约简提供了一种新的思路。

2 基本概念

定义 1 设$(L,\leq)$是偏序集,如果$\forall x,y\in L$,$\{x,y\}$都有最小上界和最大下界,则称 L 关于偏序$\leq$作成一个格。

由于最小上界和最大下界的唯一性,可把求$\{x,y\}$的最小上界和最大下界看成 x 与 y 的二元运算 $\vee$ 和 $\wedge$,即 $x\vee y$ 和 $x\wedge y$ 分别表示 x 与 y 的最小上界和最大下界,因此格可以表示为$(L,\wedge,\vee)$。

定义 2 设$(L,\wedge,\vee)$是格,若对于$\forall x,y,z\in L$,有 $x\wedge(y\vee z)=(x\wedge y)\vee(x\wedge z)$,$x\vee(y\wedge z)=(x\vee y)\wedge(x\vee z)$,则称 L 为分配格。

定义 3 设 L 是格,若存在 $a\in L$,使得$\forall x\in L$,有 $a\leq x$,则称 a 为 L 的全下界;若存在 $b\in L$,使得$\forall x\in L$,有 $x\leq b$,则

到稿日期:2014-04-06 返修日期:2014-05-27 本文受国家自然科学基金(61170107,61300153,61300121),河北省高校创新团队领军人才培育计划项目(LJRC022)资助。

马 丽(1977—),女,博士生,副教授,主要研究方向为粗糙集与概念格、近似推理等,E-mail:sumasoft@163.com;**米据生**(1966—),男,教授,博士生导师,主要研究方向为粗糙集、概念格、人工智能等,E-mail:mijsh@263.net。

称 b 为 L 的全上界。若 L 存在全下界和全上界，则称 L 为有界格。若将全下界和全上界分别记为 0 和 1，则可将有界格 L 记为$(L,\wedge,\vee,0,1)$。

定义 4 设 L 为有界格，若对 $\forall x\in L$，都 $\exists y\in L$，使得 $x\wedge y=0$ 和 $x\vee y=1$ 成立，则称 y 是 x 的补元。此时称 y 与 x 是互补的。若有界格 L 中，每个元素都至少有一个补元，则称 L 为有补格。

定义 5 设 L 是格，若对于 L 的任意子集 S，$\wedge S$ 和 $\vee S$ 都存在，则称 L 是完备格。

定义 6 设 L 是完备格，称 P 为 L 的分划，若满足：

(1) $\forall x\in P, x\neq 0$；

(2) $x,y\in P$ 且 $x\neq y$，则 $x\wedge y=0$；

(3) $\bigvee_{x\in P} x=1$。

引理 1 对 L 的两个分划 P_1 和 P_2，定义

$P_1\leqslant P_2\Leftrightarrow\forall x\in P_1,\exists y\in P_2$，有 $x\leqslant y$。

$P_1\preceq P_2\Leftrightarrow\forall y\in P_2,\exists x\in P_1$，有 $x\leqslant y$。

则关系$\leqslant$和$\preceq$满足下列性质：

(1)若 $P_1\leqslant P_2$，对于 $x_1,x_2\in P_1$，$y_1,y_2\in P_2$，且 $x_1\leqslant y_1$，$x_2\leqslant y_2$，则当 $y_1\neq y_2$ 时有 $x_1\neq x_2$。

(2)若 $P_1\preceq P_2$，对于 $x_1,x_2\in P_1$，$y_1,y_2\in P_2$，且 $x_1\leqslant y_1$，$x_2\leqslant y_2$，则当 $x_1\neq x_2$ 时有 $y_1\neq y_2$。

引理 2 设 L 是满足第一分配律的格，P_1 和 P_2 是 L 的两个分划，定义

$P_1\wedge P_2=\{x\wedge y\mid x\in P_1,y\in P_2,x\wedge y\neq 0\}$

则 $P_1\wedge P_2$ 也为 L 的分划。

证明：(1)由定义可知 $\forall z\in P_1\wedge P_2, z\neq 0$。

(2)若 $z_1,z_2\in P_1\wedge P_2$，不妨设 $z_1=x_1\wedge y_1$，$z_2=x_2\wedge y_2$，其中 $x_1,x_2\in P_1$，$y_1,y_2\in P_2$，由定义知 $x_1\wedge y_1\neq 0$ 且 $x_2\wedge y_2\neq 0$。若 $z_1\neq z_2$，也就是有 $x_1\wedge y_1\neq x_2\wedge y_2$，则有 $x_1\neq x_2$ 或 $y_1\neq y_2$。由于 P_1 和 P_2 是 L 的两个分划，因此有 $x_1\wedge x_2=0$ 或 $y_1\wedge y_2=0$。

因此有 $z_1\wedge z_2=(x_1\wedge y_1)\wedge(x_2\wedge y_2)=0$。

(3) $\forall z\in P_1\wedge P_2$，不妨设 $z=x\wedge y(x\in P_1,y\in P_2)$，则 $\bigvee_{z\in P_1\wedge P_2} z=\bigvee_{x\in P_1\wedge y\in P_2}(x\wedge y)=(\bigvee_{x\in P_1}x)\wedge(\bigvee_{y\in P_2}y)=1$。

故由定义可知，$P_1\wedge P_2$ 也为 L 的分划。

引理 3 设 L 是满足第一分配律的格，P_1 和 P_2 是 L 的两个分划，则 $P_1\leqslant P_2\Leftrightarrow P_1\wedge P_2=P_1$。

证明：($\Rightarrow$) 若 $P_1\leqslant P_2$，则 $\forall x\in P_1$，$\exists y\in P_2$，使得 $x\leqslant y$，于是 $x\wedge y=x$。因此 $x\in P_1\wedge P_2$。

$\forall z\in P_1\wedge P_2$，不妨设 $z=x\wedge y(x\in P_1,y\in P_2)$，由于$P_1\leqslant P_2$，因此 $\exists y'\in P_2$，使得 $x\leqslant y'$。

于是 $0\neq z=x\wedge y\leqslant y'\wedge y$，即 $0\neq y'\wedge y$，故 $y'=y$。所以 $z=x\wedge y'=x\in P_1$。因此有 $P_1\wedge P_2=P_1$。

($\Leftarrow$) $\forall z\in P_1$，则 $z\in P_1\wedge P_2$，故 $\exists x\in P_1,y\in P_2$，使得 $z=x\wedge y$，即有 $z\wedge x=z\neq 0$。

因为 P_1 是 L 的分划，故 $z=x$，即有 $x\leqslant y$。因此 $P_1\leqslant P_2$。

由上面引理很容易得到，若 $P_1\leqslant P_2$ 且 $P_2\leqslant P_1$，则 $P_1=P_2$。

定义 7 设 $\mathscr{R}$ 是格 L 的一簇分划，若记$\overline{\mathscr{R}}=\wedge\{R:R\in\mathscr{R}\}$，则$\overline{\mathscr{R}}$是 L 的分划，称为由 $\mathscr{R}$ 生成的分划。

定理 1 设 L 是满足第一分配律的格，P_1 和 P_2 是 L 的两个分划，若 $P_1\leqslant P_2$，则 $\forall y\in P_2$，有 $y=\vee\{x\in P_1:x\leqslant y\}$。

证明：$\forall y\in P_2$，

$y=y\wedge 1=y\wedge(\vee\{x:x\in P_1\})$

$=\vee\{x\wedge y:x\in P_1\}$

$=\vee\{x\wedge y:x\in P_1,x\wedge y\neq 0\}$

$=\vee\{x\wedge y:x\in P_1,x\leqslant y\}$

$=\vee\{x:x\in P_1,x\leqslant y\}$

定义 8 设 L 是格，P 为 L 的分划，$\forall y\in L$，记$\underline{apr}P(y)=\vee\{x\in P:x\leqslant y\}$，$\overline{apr}P(y)=\vee\{x\in P:x\wedge y\neq 0\}$。

分别称$\underline{apr}P(y)$和$\overline{apr}P(y)$为 y 关于 P 的下近似和上近似。

引理 4 下近似和上近似具有如下性质：

$\forall x,y\in L$，

(1) $\underline{apr}P(y)\leqslant y\leqslant\overline{apr}P(y)$；

(2)若 L 是有补格，则$\sim\underline{apr}P(\sim y)=\overline{apr}P(y)$；

(3) $\underline{apr}P(x\wedge y)=\underline{apr}P(x)\wedge\underline{apr}P(y)$，

$\overline{apr}P(x\vee y)=\overline{apr}P(x)\vee\overline{apr}P(y)$。

证明：(1) $\forall y\in L$，

$y\wedge\overline{apr}P(y)=\vee\{x\wedge y:x\in P,x\wedge y\neq 0\}$

$=\vee\{x\wedge y:x\in P\}$

$=y\wedge(\vee\{x:x\in P\})$

$=y\wedge 1=y$

因此，$y\leqslant\overline{apr}P(y)$。又 $\underline{apr}P(y)\leqslant y$ 显然成立，故有$\underline{apr}P(y)\leqslant y\leqslant\overline{apr}P(y)$。

(2)先证 $\forall y\in L$，$x_1,x_2\in P$，若 $x_1\leqslant\sim y$，$x_2\wedge y\neq 0$，则 $x_1\wedge x_2=0$。

若不然，因为 P 为 L 的分划，则 $x_1=x_2$。

由 $x_1\leqslant\sim y$ 得 $x_1\wedge y\leqslant(\sim y)\wedge y=0$，故有 $x_1\wedge y=0$，即 $x_2\wedge y=0$ 与条件 $x_2\wedge y\neq 0$ 矛盾。

$\underline{apr}P(\sim y)\wedge\overline{apr}P(y)$

$=(\vee\{x\in P:x\leqslant\sim y\})\wedge(\vee\{x\in P:x\wedge y\neq 0\})$

$=\vee\{x_1\wedge x_2:x_1,x_2\in P,x_1\leqslant\sim y,x_2\wedge y\neq 0\}$

$=0$

类似地，

$\underline{apr}P(\sim y)\vee\overline{apr}P(y)$

$=(\vee\{x\in P:x\leqslant\sim y\})\vee(\vee\{x\in P:x\wedge y\neq 0\})$

$=\vee\{x_1\vee x_2:x_1,x_2\in P,x_1\leqslant\sim y,x_2\wedge y\neq 0\}$

$=1$

故有$\sim\underline{apr}P(\sim y)=\overline{apr}P(y)$。

(3) $\underline{apr}P(x\wedge y)$

$=\vee\{z\in P:z\leqslant x\wedge y\}$

$=\vee\{z\in P:z\leqslant x,z\leqslant y\}$

$=(\vee\{z\in P:z\leqslant x\})\wedge(\vee\{z\in P:z\leqslant y\})$

$=\underline{apr}P(x)\wedge\underline{apr}P(y)$

另一等式 $\overline{apr}P(x\vee y)=\overline{apr}P(x)\vee\overline{apr}P(y)$ 由以上结果和(2)立即可得。

3 知识约简

定义 9 设 L 是完备格，$\mathscr{R}$ 是 L 的一簇分划，$\mathscr{S}\subseteq\mathscr{R}$，若$\overline{\mathscr{S}}=\overline{\mathscr{R}}$，则称 $\mathscr{S}$ 与 $\mathscr{R}$ 是协调的。若 $\mathscr{S}$ 与 $\mathscr{R}$ 是协调的，而对 $\forall P\in\mathscr{S}$，$\mathscr{S}\setminus\{P\}$ 与 $\mathscr{R}$ 是不协调的，则称 $\mathscr{S}$ 是 $\mathscr{R}$ 的约简。

$\forall x,y\in\overline{\mathscr{R}}$,定义

$D(x,y)=\{R\in\mathscr{R}:\exists x',y'\in R,x'\wedge y'=0\Rightarrow x\leqslant x',y\leqslant y'\}$

称 $D(x,y)$ 为 x 与 y 的可辨识集。

定理 2 设 L 是完备格,$\mathscr{R}$ 是 L 的一簇分划,$\mathscr{S}\subseteq\mathscr{R}$,则 $\overline{\mathscr{S}}=\overline{\mathscr{R}}\Leftrightarrow\mathscr{S}\cap D(x,y)\neq\varnothing,\forall x,y\in\overline{\mathscr{R}},x\neq y$。

证明:(⇒)设 $\overline{\mathscr{S}}=\overline{\mathscr{R}}$,$\forall x,y\in\overline{\mathscr{R}},x\neq y$,有 $x,y\in\overline{\mathscr{S}}$。于是对于 $\forall P\in\mathscr{S}$,$\exists x_P,y_p\in P$,使得有 $x=\wedge x_P,y=\wedge y_p$ 成立。但是 $x\neq y$,必存在 $P\in\mathscr{S}$,使得 $x_P\neq y_p$,即 $x_P\wedge y_p=0$。显然 $x\leqslant x_P,y\leqslant y_p$,因此 $P\in\mathscr{S}\cap D(x,y)$,即 $\mathscr{S}\cap D(x,y)\neq\varnothing$。

(⇐) $\forall x,y\in\overline{\mathscr{R}},x\neq y$,由于 $\mathscr{S}\cap D(x,y)\neq\varnothing$,不妨设 $P\in\mathscr{S}\cap D(x,y)$,则必有 $P\in\mathscr{S}$,使得 $\exists x',y'\in P,x'\wedge y'=0,x\leqslant x',y\leqslant y'$。因为 $\overline{\mathscr{R}}\leqslant\overline{\mathscr{S}}$,故 $\exists x_1,y_1\in\overline{\mathscr{S}}$,使得 $x\leqslant x_1,y\leqslant y_1$。又因为 $\overline{\mathscr{S}}\leqslant P$,且 P 是 L 的分划,故必有 $x_1\leqslant x',y_1\leqslant y'$。因此 $x_1\wedge y_1\leqslant x'\wedge y'=0$。即说明 $\forall x,y\in\overline{\mathscr{R}},x\neq y,\exists x_1,y_1\in\overline{\mathscr{S}},x_1\wedge y_1=0$,使得 $x\leqslant x_1,y\leqslant y_1$。故必有 $\overline{\mathscr{S}}=\overline{\mathscr{R}}$。

4 相对约简

设 L 是完备格,$\mathscr{R}$ 和 $\mathscr{S}$ 是 L 的两簇分划,$\mathscr{R}\cap\mathscr{S}=\varnothing$。若 $\mathscr{R}\leqslant\mathscr{S}$,则称系统 $(L,\mathscr{R},\mathscr{S})$ 是协调的。

定义 10 设系统 $(L,\mathscr{R},\mathscr{S})$ 是协调的,$\mathscr{R}'\subseteq\mathscr{R}$,若 $\overline{\mathscr{R}}'\leqslant\overline{\mathscr{S}}$,则称 $\mathscr{R}'$ 关于 $\mathscr{S}$ 是相对协调的。若 $\mathscr{R}'$ 关于 $\mathscr{S}$ 是相对协调的,而对于 $\forall P\in\mathscr{R}'$,$\mathscr{R}'\backslash\{P\}$ 关于 $\mathscr{S}$ 不是相对协调的,则称 $\mathscr{R}'$ 是 $\mathscr{R}$ 的相对约简。

对 $\forall x,y\in\overline{\mathscr{R}}$,定义

$$\hat{D}(x,y)=\begin{cases}\mathscr{R}, & \exists d\in\overline{\mathscr{S}},\text{使得 } x\leqslant d,y\leqslant d\\ \{P\in\mathscr{R}, & \exists x',y'\in P,x'\wedge y'=0,\text{使得 } x\leqslant x',y\leqslant y'\}\end{cases}$$

否称 $\hat{D}(x,y)$ 为 x 与 y 的相对可辨识集。

定理 3 设 L 是完备格,若系统 $(L,\mathscr{R},\mathscr{S})$ 是协调的,$\mathscr{R}'\subseteq\mathscr{R}$,则

$$\overline{\mathscr{R}'}\leqslant\overline{\mathscr{S}}\Leftrightarrow\mathscr{R}'\cap\hat{D}(x,y)\neq\varnothing,\forall x,y\in\overline{\mathscr{R}},x\neq y$$

证明:(⇒)设 $\overline{\mathscr{R}}'\leqslant\overline{\mathscr{S}}$。$\forall x,y\in\overline{\mathscr{R}},x\neq y$,若 $\exists d\in\overline{\mathscr{S}}$,使得 $x\leqslant d,y\leqslant d$,则 $\hat{D}(x,y)=\mathscr{R}$。由于 $\mathscr{R}'\subseteq\mathscr{R}$,因此有 $\mathscr{R}'\cap\hat{D}(x,y)\neq\varnothing$。

若不存在 $d\in\overline{\mathscr{S}}$,使得 $x\leqslant d,y\leqslant d$,由于 $\overline{\mathscr{R}}\leqslant\overline{\mathscr{S}}$,因此 $\exists d_1,d_2\in\overline{\mathscr{S}},d_1\wedge d_2=0$,使得 $x\leqslant d_1,y\leqslant d_2$。

又 $\overline{\mathscr{R}}\leqslant\overline{\mathscr{R}'}$,故 $\exists x',y'\in\overline{\mathscr{R}}'$,使得 $x\leqslant x',y\leqslant y'$。而 $\overline{\mathscr{R}}'\leqslant\overline{\mathscr{S}}$,故 $\exists d_1',d_2'\in\overline{\mathscr{S}}$,使得 $x'\leqslant d_1',y'\leqslant d_2'$。

所以有 $x\leqslant x'\leqslant d_1'$ 且 $x\leqslant d_1$,故 $d_1'=d_1$。同样,$d_2'=d_2$。因此,$x\leqslant x'\leqslant d_1,y\leqslant y'\leqslant d_2$。

因为 $d_1\wedge d_2=0$,所以有 $x'\neq y'$。

由于 $x',y'\in\overline{\mathscr{R}'}$,因此可知 $\exists P\in\mathscr{R}'$ 及 $\exists x_1,y_1\in P,x_1\wedge y_1=0$,使 $x\leqslant x_1,y\leqslant y_1$。即有 $\mathscr{R}'\cap\hat{D}(x,y)\neq\varnothing$。

(⇐) $\forall x\in\overline{\mathscr{R}}'$,由于 $\overline{\mathscr{R}}\leqslant\overline{\mathscr{R}}'$,由定理 1 可知 $x=\vee\{y\in\overline{\mathscr{R}}:y\leqslant x\}$。

$\mathscr{R}'\cap\hat{D}(x,y)\neq\varnothing$,于是 $\exists x_1\in\overline{\mathscr{R}}$,使得 $x_1\leqslant x$。由协调性知,$\exists d\in\overline{\mathscr{S}}$,使得 $x_1\leqslant d$。

若 $y_1\in\overline{\mathscr{R}}$,使得 $y_1\leqslant d$ 不成立,则有 $d'\in\overline{\mathscr{S}}$。由条件 $\mathscr{R}'\cap\hat{D}(x,y)\neq\varnothing$ 知 $\exists P\in\mathscr{R}'$ 及 $\exists x',y'\in P,x'\wedge y'=0$,使得 $x_1\leqslant x',y_1\leqslant y'$。

由于 $\overline{\mathscr{R}}\leqslant\overline{\mathscr{R}'}$,$\exists x_0,y_0\in\overline{\mathscr{R}'}$,使得 $x_1\leqslant x_0,y_1\leqslant y_0$。因此有 $x_0\leqslant x',y_0\leqslant y'$,从而 $x_0\wedge y_0=0$。

而 $x_1\leqslant x,x_1\leqslant x_0$,故有 $x=x_0$。这样即有 $x\wedge y_0=0$。

因此 y_1 不满足 $y_1\leqslant x$。否则 $y_1\leqslant x,y_1\leqslant y_0$ 时有 $y_1\leqslant x\wedge y_0=0$ 矛盾。

这说明若 $y_1\in\overline{\mathscr{R}},y_1\leqslant x$,则必有 $y_1\leqslant d$。因此 $x=\vee\{y\in\overline{\mathscr{R}}:y\leqslant x\}\leqslant d$,即有 $\overline{\mathscr{R}}'\leqslant\overline{\mathscr{S}}$。证毕。

5 算例

令 $B=\{a,b,c\}$,则 $\langle P(B),\supseteq\rangle$ 为完备格。其哈斯图如图 1 所示。

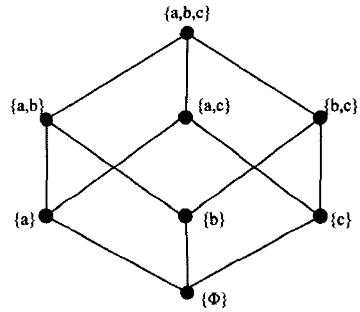

图 1

根据定义,共有 5 个分划,分别如下:

$P_1=\{\{a\},\{b\},\{c\}\},P_2=\{\{a\},\{b,c\}\},P_3=\{\{c\},\{a,b\}\},P_4=\{\{b\},\{a,c\}\},P_5=\{\{a,b,c\}\}=\{B\}$。

有 $P_1\leqslant P_2,P_1\leqslant P_3,P_1\leqslant P_4,P_1\leqslant P_5,P_2\leqslant P_5,P_3\leqslant P_5,P_4\leqslant P_5$,且 $P_1\wedge P_2=P_1\wedge P_3=P_1\wedge P_4=P_1\wedge P_5=P_1,P_2\wedge P_5=P_2,P_3\wedge P_5=P_3,P_4\wedge P_5=P_4,P_2\wedge P_3=P_2\wedge P_4=P_3\wedge P_4=P_1$。

分别以元素 $\{a\},\{b\},\{c\},\{a,b\}$ 为例,考虑关于分划 P_1,P_2,P_3,P_4,P_5 的上下近似。

$\underline{apr}P_1(\{a\})=\{a\},\overline{apr}P_1(\{a\})=\{a\}$;

$\underline{apr}P_1(\{b\})=\{b\},\overline{apr}P_1(\{b\})=\{b\}$;

$\underline{apr}P_1(\{c\})=\{c\},\overline{apr}P_1(\{c\})=\{c\}$;

$\underline{apr}P_1(\{a,b\})=\{a,b\},\overline{apr}P_1(\{a,b\})=\{a,b\}$。

$\underline{apr}P_2(\{a\})=\{a\},\overline{apr}P_2(\{a\})=\{a\}$;

$\underline{apr}P_2(\{b\})=\varnothing,\overline{apr}P_2(\{b\})=\{b,c\}$;

$\underline{apr}P_2(\{c\})=\varnothing,\overline{apr}P_2(\{c\})=\{b,c\}$;

$\underline{apr}P_2(\{a,b\})=\{a\},\overline{apr}P_2(\{a,b\})=B$。

$\underline{apr}P_3(\{a\})=\varnothing,\overline{apr}P_3(\{a\})=\{a,b\}$;

$\underline{apr}P_3(\{b\})=\varnothing,\overline{apr}P_3(\{b\})=\{a,b\}$;

$\underline{apr}P_3(\{c\})=\{c\},\overline{apr}P_3(\{c\})=\{c\}$;

$\underline{apr}P_3(\{a,b\})=\{a,b\},\overline{apr}P_3(\{a,b\})=\{a,b\}$。

$\underline{apr}P_4(\{a\})=\varnothing,\overline{apr}P_4(\{a\})=\{a,c\}$;

$\underline{apr}P_4(\{b\})=\{b\},\overline{apr}P_4(\{b\})=\{b\}$;

$\underline{apr}P_4(\{c\})=\varnothing,\overline{apr}P_4(\{c\})=\{a,c\}$;

(下转第 100 页)

similarity measure[J]. KSCE journal of civil Engineering, 2014, 18(2): 521-530

[4] 刘文军,赵利萍. 粗糙集的相似度量[J]. 数学理论与应用, 2012, 32, (3): 35-42
Liu Wen-jun, Zhao Li-ping. The similarity degree of rough set [J]. Mathematical Theory and Applications, 2012, 32(3): 35-42

[5] 史占红,连玉平. 基于包含度的粗糙集间的相似性度量[J]. 数学教学研究, 2008(2): 53-54
Shi Zhan-hong, Lian Yu-ping. Similarity measure between rough sets based on inclusion degree[J]. Mathematics teaching research, 2008(2): 53-54

[6] 张文修,吴伟志,梁吉业,等. 粗糙集理论与方法[M]. 北京:科学出版社, 2001
Zhang Wen-xiu, Wu Wei-zhi, Liang Ji-ye, et al. Rough set theory and method[M]. Beijing: Science Press, 2001

[7] 张文修,梁怡,徐萍. 基于包含度的不确定推理[M]. 北京: 清华大学出版社, 2007
Zhang Wen-xiu, Liang Yi, Xu Ping. Uncertainty reasoning based on inclusion degree[M]. Beijing: Tsinghua University press, 2007

[8] 徐久成,沈均毅,王国胤. Rough 集之间的相似度量[J]. 计算机科学, 2003, 30(10): 55-60
Xu Jiu-cheng, Shen Jun-yi, Wang Guo-yin. Measure of similarity between rough sets[J]. Computer science, 2003, 30(10): 55-60

[9] Wang G Y, Guan L H, Hu F. Rough set extensions in incomplete information systems[J]. Frontiers of Electrial and Engineering in China, 2008, 3(4): 399-405

[10] Wang G Y, Peters J F, Skowron A, et al. Rough sets and knowledge technology[C]// Proc. of the RSKT 2006. LNCS 4062, Berlin: Springer-Verlag, 2006: 1-32

[11] Wang G Y, Li T R, Grzymala B J, et al. Rough sets and knowledge technology[C]// Proc. of the RSKT 2008. LNCS 5009, Berlin: Springer-Verlag, 2008: 1-18

[12] 王国胤,姚一豫,于洪. 粗糙集理论与应用研究综述[J]. 计算机学报, 2009, 32(7): 1229-1246
Wang Guo-yin, Yao Yi-yu, Yu Hong. A Survey on Rough Set Theory and Applications[J]. Chinese Journal of Computers, 2009, 32(7): 1229-1246

(上接第 81 页)

$\underline{apr}P_4(\{a,b\})=\{b\}, \overline{apr}P_4(\{a,b\})=B$;

$\underline{apr}P_5(\{a\})=\emptyset, \overline{apr}P_5(\{a\})=B$;

$\underline{apr}P_5(\{b\})=\emptyset, \overline{apr}P_5(\{b\})=B$;

$\underline{apr}P_5(\{c\})=\emptyset, \overline{apr}P_5(\{c\})=B$;

$\underline{apr}P_5(\{a,b\})=\emptyset, \overline{apr}P_5(\{a,b\})=B$。

易验证引理 4 的 3 条性质均满足。

若分划族 $\mathscr{R}=\{P_1,P_2,P_3,P_4,P_5\}$,则$\overline{\mathscr{R}}=P_1$ 称为由 $\mathscr{R}$ 生成的分划。则 $\mathscr{S}_1=\{P_1\}$, $\mathscr{S}_2=\{P_1,P_2\}$, $\mathscr{S}_3=\{P_1,P_3\}$, $\mathscr{S}_4=\{P_1,P_4\}$, $\mathscr{S}_5=\{P_1,P_5\}$, $\mathscr{S}_6=\{P_2,P_3\}$, $\mathscr{S}_7=\{P_2,P_4\}$, $\mathscr{S}_8=\{P_3,P_4\}$ 均为 $\mathscr{R}$ 的协调集,其中 $\mathscr{S}_1=\{P_1\}$, $\mathscr{S}_6=\{P_2,P_3\}$, $\mathscr{S}_7=\{P_2,P_4\}$, $\mathscr{S}_8=\{P_3,P_4\}$ 均为 $\mathscr{R}$ 的约简。

而 $\mathscr{S}_9=\{P_2,P_5\}$, $\mathscr{S}_{10}=\{P_3,P_5\}$, $\mathscr{S}_{11}=\{P_4,P_5\}$ 均不为 $\mathscr{R}$ 的协调集。

再比如系统$(L,\mathscr{R},\mathscr{S}_1)$,$(L,\mathscr{R},\mathscr{S}_6)$和$(L,\mathscr{S}_7,\mathscr{S}_8)$均不协调,而$(L,\mathscr{S}_1,\mathscr{S}_6)$,$(L,\mathscr{S}_1,\mathscr{S}_7)$,$(L,\mathscr{S}_1,\mathscr{S}_8)$,$(L,\mathscr{S}_3,\mathscr{S}_7)$均协调,$\mathscr{S}_1$ 为 $\mathscr{S}_6,\mathscr{S}_7,\mathscr{S}_8$ 的相对约简,$\mathscr{S}_3$ 不为 $\mathscr{S}_7$ 的相对约简。

结束语 粗糙集理论(RST)自从 1982 年被波兰数学家 Z. Pawlak 提出用来研究不完整数据、不精确知识的表达和学习以来,已被广泛应用于不同的领域进行数据分析。粗糙集的数据表现方式是信息系统,本文把经典的信息系统看作是一种特殊的格结构,通过在格上定义分划后,借助粗糙集思想给出上下近似算子及可辨识集,进而分析得到约简和相对约简的判定条件。这种表示从更深层次上揭示了知识的本质,将知识约简的已有相关结果做了进一步拓广。本文的讨论还是一个尝试,相关研究还有待进一步深入。

参 考 文 献

[1] Pawlak Z. Rough set[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341-356

[2] Bryniarski E. Formal description of rough sets [C]// Rough Sets, Fuzzy Sets and Knowledge Discovery. Springer-Verlag, 1994: 208-216

[3] Pawlak Z. Rough sets: Theoretical Aspects of Reasoning about Data [M]. Boston: Kluwer Academic Publishers, 1991

[4] Ganter B, Wille R. Formal Concept Analysis: Mathematical Foundations [M]. Berlin: Springer, 1999

[5] Kryszkiewicz M. Rules in incomplete information systems[J]. Information Sciences, 1999, 113: 271-292

[6] Yao Y Y. Concept lattices in rough set theory[C]// Dick S, Kurgan L, Pedrycz W, et al. , eds. Proceedings of 2004 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS 2004). IEEE, June 2004: 796-801

[7] 王珏,王任,苗夺谦,等. 基于 Rough set 理论的"数据浓缩"[J]. 计算机学报, 1998, 21(5): 393-400
Wang J, Wang R, Miao D Q, et al. Data enriching based on rough set theory [J]. Chinese Journal of Computers, 1998, 21(5): 393-400

[8] 张文修,梁怡,吴伟志,等. 信息系统与知识发现[M]. 北京:科学出版社, 2003
Zhang W X, Liang Y, Wu W Z, et al. Information system and knowledge discovery [M]. Beijing: Science Press, 2003

[9] 胡可云,陆玉昌,石纯一. 概念格及其应用进展[J]. 清华大学学报:自然科学版, 2000, 40(9): 77-81
Hu K Y, Lu Y C, Shi C Y. Advances in concept lattice and its application [J]. Journal of Tsinghua University, 2000, 40(9): 77-81

[10] 张文修,魏玲,祁建军. 概念格的属性约简理论与方法[J]. 中国科学(E), 2005, 35(6): 628-639
Zhang W X, Wei L, Qi J J. Attribute reduction theory and approach to concept lattice [J]. Science in China(E), 2005, 35(6): 628-639

[11] 刘清. 邻域值信息表上的邻域逻辑及其数据推理[J]. 计算机学报, 2001, 24(4): 405-410
Liu Q. Neighborhood logic and its data reasoning on neighborhood-valued information table[J]. Chinese Journal of Computers, 2001, 24(4): 405- 410

[12] 宋笑雪,解争龙,张文修. 集值决策信息系统的知识约简与规则提取[J]. 计算机科学, 2007, 34(4): 182-184
Song X X, Xie Z L, Zhang W X. Knowledge reduction and rule extraction in set-valued decision information system[J]. Computer Science, 2007, 34(4): 182-184