第 46 卷 第 2 期
2019 年 2 月
计 算 机 科 学
COMPUTER SCIENCE
Vol. 46 No. 2
Feb. 2019

一种新的基于粒度重要度的三支决策模型

薛占熬　韩丹杰　吕敏杰　赵丽平

（河南师范大学计算机与信息工程学院　河南 新乡 453007）

（“智慧商务与物联网技术”河南省工程实验室　河南 新乡 453007）

摘　要　粒度重要度是多粒度粗糙集中的一项重要研究内容。针对现有粒度重要度只考虑单个粒度对决策的直接影响而忽略了其他粒度对决策综合影响的问题，结合多粒度粗糙集近似质量的概念，通过研究粒度重要度的构造方法，提出了一种新的多粒度间的粒度重要度的计算方法，并给出了基于该方法的粒度约简算法。同时，为减少冗余决策信息，将约简集与三支决策理论相结合，构建了基于粒度重要度的三支决策模型，给出了决策规则。最后通过实例证明，新的粒度约简算法可以获得具有更高区分度的数据，且缩小了延迟域范围，使最终决策更合理。

关键词　多粒度粗糙集，粒度重要度，粒度约简，三支决策

中图法分类号　TP181　　**文献标识码**　A　　**DOI**　10.11896/j.issn.1002-137X.2019.02.036

New Three-way Decisions Model Based on Granularity Importance Degree

XUE Zhan-ao　HAN Dan-jie　LV Min-jie　ZHAO Li-ping

(College of Computer and Information Engineering, Henan Normal University, Xinxiang, Henan 453007, China)

(Engineering Lab of Henan Province for Intelligence Business & Internet of Things, Xinxiang, Henan 453007, China)

Abstract　Granularity importance degree is an important content of multi-granularity rough set. Now, the construction methods of granularity importance degree only consider the direct influence of single granularity on decision-making, ignoring the combined effect of other granularity on decision-making. Combining the concept of multi-granularity approximated quality, this paper studied the construction method of granularity importance degree, proposed a new granularity importance degree calcuation method among multiple granularities, and gave a granular structure reduction algorithm based on this method. Meanwhile, in order to reduce the redundant decision information, through combining reduction set and three-way decisions, this paper constructed the three-way decisions model based on granularity importance degree, and gave the decision rules in detail. Finally, the results of an example show that the proposed granular degree reduction algorithm can obtain the data with larger discrimination, and narrows the range of delay domain, making the final decision more reasonable.

Keywords　Multi-granularity rough set, Granularity importance degree, Granular structure reduction, Three-way decisions

1　引言

1982 年，Pawlak[1] 提出了粗糙集理论。该理论能够很好地处理不精确和不确定性问题，因此引起了学者们的广泛关注与研究，从而出现了一系列相关理论和扩展模型[2-6]。在粗糙集理论中，一个属性集被视为一个粒度空间，而论域被划分成知识粒度或元素的集合。从粒计算[7] 的角度来看，经典粗糙集是单粒度的，但在实际应用中多粒度往往起着更重要的作用。随后，Qian 等[8-9] 提出并阐述了多粒度粗糙集模型，该模型中的粒度约简问题[10] 是一个重要的研究内容。桑妍丽等[11] 通过引入近似分布的概念，以近似分布质量作为多粒度决策粗糙集模型中的不确定性度量方式，提出了一种基于多粒度决策粗糙集的粒度约简算法，但该算法没有考虑粒度间的相互影响。Dai 等[12] 针对多粒度粗糙集的属性约简问题，提出了一种结合粒子群算法的多知识快速约简方法。Yang 等[13] 考虑多粒度结构的代价问题，并基于近似质量的定义，提出了代价敏感多粒度决策系统的约简算法。Jing 等[14] 提出了基于知识粒度与粒度视图的增量式属性约简方法。Yang 等[15] 通过可辨识矩阵的算法来计算覆盖粗糙集的所有粒度约简，并提出了一种基于覆盖粗糙集的粒度约简算法。

到稿日期：2017-12-13　返修日期：2018-03-18　本文受国家自然科学基金计划项目（61772176），河南省科技攻关项目（182102210078，182102210362），河南省科技创新人才项目（184100510003），新乡市科技攻关计划项目（CXGG17002）资助。

薛占熬（1963－），男，博士，教授，主要研究方向为人工智能基础理论、粗糙集理论和三支决策理论，E-mail：xuezhanao@163.com（通信作者）；**韩丹杰**（1992－），女，硕士生，主要研究方向为粗糙集理论、三支决策理论和多粒度；**吕敏杰**（1993－），女，硕士生，主要研究方向为粗糙集理论和直觉模糊集理论；**赵丽平**（1993－），女，硕士生，主要研究方向为粗糙集理论、三支决策理论。

Li 等[16]从规则、规则的支持和确定性因素之间的差异和关系的角度，研究了多粒度粗糙集与概念格之间的关系，并对获取的规则进行了算法复杂度分析。

为了更好地给出粗糙集模型中正域、负域和边界域的语义解释，Yao[17-18]提出了三支决策的概念。随后，许多学者从多个方面对三支决策理论进行了深入的研究。Liang 等[19]考虑直觉模糊点算子对损失函数的影响，重新构建了直觉模糊决策粗糙集的基本模型，分析了三支决策的新推导，将三支决策的应用范围扩展到新的直觉模糊环境中。Qian 等[20]在考虑满足概率正域单调性的条件下，提出了动态粒度下序贯三支决策的属性约简，但是没有考虑代价结构。薛占熬等[21]将三支决策规则与条件属性结合，通过考虑负域对决策的影响，重构了属性权重的计算方法。Hu 等[22-23]系统回答了三支决策理论中的度量问题、条件问题、评价函数问题，总结了三支决策空间及在此空间上的各类三支决策，最后给出了多粒度三支决策空间的转化方法。史进玲等[24]针对决策信息表讨论了知识的粒化，以粒度的划分数量和粒化决策权重为主要指标，寻求最优的粒度空间，提出了基于风险最小化的多粒度三支决策模型，为实际问题的决策和风险处理提供了一个新的研究视角。Li 等[25]从认知层面揭示了三支概念解决决策问题的本质思想，提出了一种用多粒度描述三支概念的公理化方法，且采用集合近似的思想来模拟从给定线索中学习认知概念的认知过程。Chen 等[26]将序贯三支决策应用到动态多尺度决策表的最优尺度选择问题中，提出了基于序贯三支决策的动态多尺度决策表的最优尺度选择方法，该方法适用于在添加新对象的情况下更新多尺度决策表的最优尺度问题。通过研究发现，将粒度约简与三支决策结合的文献较少，但对多粒度粗糙集进行约简可以简化粒度空间的决策规则，因此将约简集和三支决策结合具有一定的创新和研究价值。

基于改进的粗糙集属性重要度[27]计算方法，综合考虑多粒度粗糙集的内部粒度重要度和外部粒度重要度的直接影响与间接影响，本文提出新的粒度重要度计算方法。在此基础上，给出一种新的粒度约简算法，并将约简集与三支决策理论结合，构建新的基于粒度重要度的三支决策模型，推导出决策规则；最后根据实例分析并验证了该模型的有效性。此研究为多粒度粗糙集与三支决策理论的结合研究提供了一种新方法。

2　基础知识

2.1　多粒度粗糙集

定义 1（粗糙集）[1]　设 U 表示论域，为一非空有限集合，$R\subseteq U\times U$ 为论域 U 上的等价关系，$apr=(U,R)$ 表示近似空间。等价关系 R 可形成论域 U 上的一个划分，记为 U/R。设 X 为论域 U 上的子集，$\forall X\subseteq U$，X 的下近似集$\underline{R}(X)$与上近似集 $\overline{R}(X)$分别定义为：

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

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

其中，$[x]_R$ 表示对象 x 在等价关系 R 下的等价类。

当$\underline{R}(X)\neq\overline{R}(X)$时，称 X 是在 U 中关于 R 的粗糙集；当$\underline{R}(X)=\overline{R}(X)$时，称 X 是在 U 中关于 R 的精确集。

定义 2（悲观多粒度粗糙集）[8]　设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 中的 m 个属性子集。$\forall X\subseteq U$，X 关于 A 的悲观多粒度粗糙集的下近似和上近似分别定义为：

$$\underline{\sum_{i=1}^{m}A_i}^{I}(X)=\{x\in U\mid[x]_{A_1}\subseteq X\wedge[x]_{A_2}\subseteq X\wedge[x]_{A_3}\subseteq X\wedge\cdots\wedge[x]_{A_m}\subseteq X\}$$

$$\overline{\sum_{i=1}^{m}A_i}^{I}(X)=\sim(\underline{\sum_{i=1}^{m}A_i}^{I}(\sim X))$$

其中，$[x]_{A_i}$（$1\leqslant i\leqslant m$）是 x 关于 A_i 的等价类，$\sim X$ 是 X 的补集，$\wedge$ 是合取连接词，以下相同，则$(\underline{\sum_{i=1}^{m}A_i}^{I}(X),\overline{\sum_{i=1}^{m}A_i}^{I}(X))$称为关于 X 的悲观多粒度粗糙集。

定义 3（乐观多粒度粗糙集）[8]　设信息系统 $IS=\langle U,AT,V,f\rangle$，$A=\{A_1,A_2,\cdots,A_m\}$是 AT 中的 m 个属性子集。$\forall X\subseteq U$，X 关于 A 的乐观多粒度粗糙集的下近似和上近似分别定义为：

$$\underline{\sum_{i=1}^{m}A_i}^{O}(X)=\{x\in U\mid[x]_{A_1}\subseteq X\vee[x]_{A_2}\subseteq X\vee[x]_{A_3}\subseteq X\cdots\vee[x]_{A_m}\subseteq X\}$$

$$\overline{\sum_{i=1}^{m}A_i}^{O}(X)=\sim(\underline{\sum_{i=1}^{m}A_i}^{O}(\sim X))$$

其中，$\vee$ 是析取连接词，以下相同，则$(\underline{\sum_{i=1}^{m}A_i}^{O}(X),\overline{\sum_{i=1}^{m}A_i}^{O}(X))$表示关于 X 的乐观多粒度粗糙集。

2.2　三支决策

三支决策理论源于决策粗糙集，是传统二支决策的拓展形式，它更好地解释了决策粗糙集的语义，且比二支决策理论更加精确地表达了决策过程的不确定性。Yao[17-18]详细介绍了三支决策理论，简要概述如下。

设 $U=\{x_1,x_2,\cdots,x_n\}$为论域，$\xi=\{\omega_P,\omega_N,\omega_B\}$表示将一个对象 x 划分到接受域 $POS(X)$、拒绝域 $NEG(X)$和边界域 $BND(X)$的决策动作。用代价函数 $\lambda_{PP},\lambda_{NP},\lambda_{BP}$ 分别表示对象 $x\in U$ 时的 3 种决策代价；用代价函数 $\lambda_{PN},\lambda_{NN},\lambda_{BN}$ 分别表示对象 $x\notin U$ 时的 3 种决策代价，如表 1 所列。

表 1　代价矩阵

Table 1　Cost matrix

决策动作	决策函数	
	X	$\sim X$
ω_P	λ_{PP}	λ_{PN}
ω_B	λ_{BP}	λ_{BN}
ω_N	λ_{NP}	λ_{NN}

根据贝叶斯决策风险最小原则，可得到如下所示的三支决策规则。

(1)接受决策规则(P)：若 $P(X\mid[x])\geqslant\alpha$，则 $x\in POS(X)$。

(2)拒绝决策规则(N)：若 $P(X\mid[x])\leqslant\beta$，则 $x\in NEG(X)$。

(3)延迟决策规则(B)：若 $\beta<P(X\mid[x])<\alpha$，则 $x\in BND(X)$。

其中，α，β和γ的计算公式分别如下：

$$\alpha=\frac{\lambda_{PN}-\lambda_{BN}}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})}$$

$$\beta=\frac{\lambda_{BN}-\lambda_{NN}}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{BP})}$$

$$\gamma=\frac{\lambda_{PN}-\lambda_{NN}}{(\lambda_{PN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})}$$

3 基于改进的粒度重要度的粒度约简算法

3.1 粒度重要性度量

定义4(近似质量)[11,13] 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A=\{A_1,A_2,\cdots,A_m\}$是C的m个属性子集族，决策属性D导出的划分为$U/D=\{X_1,X_2,\cdots,X_s\}$。粒度集$A$表示$U/D$的近似质量，其定义为：

$$\gamma(A,D)=\frac{|\cup\{\underline{\sum_{i=1}^{m}A_i}^{\Delta}(X_t)\,|\,1\leqslant t\leqslant s\}|}{|U|}$$

其中，$\Delta\in\{I,O\}$表示悲观和乐观多粒度粗糙集两种情况，以下相同。

定义5(粒度重要性)[10] 设决策信息系统$DIS=(U,C\cup D,V,f)$，$\varnothing\subset A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族，$X\in U/D$。

(1)若$\underline{\sum_{i=1,A_i\in A}^{m}A_i}^{\Delta}(X)\neq\underline{\sum_{i=1,A_i\in A-A'}^{m}A_i}^{\Delta}(X)$，则称$A'$在$A$中关于$X$是重要的；

(2)若$\underline{\sum_{i=1,A_i\in A}^{m}A_i}^{\Delta}(X)=\underline{\sum_{i=1,A_i\in A-A'}^{m}A_i}^{\Delta}(X)$，则称$A'$在$A$中关于$X$是不重要的。

定义6(内部重要度)[11,13] 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族。$\forall A_i\in A'$，定义在粒度集$A'$上，$A_i$关于$D$的内部重要度为：

$$sig_{\rm in}^{\Delta}(A_i,A',D)=|\gamma(A',D)-\gamma(A'-\{A_i\},D)|$$

定义7(外部重要度)[11,13] 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族，$\forall A_i\in A-A'$，定义在粒度集$A'$上，$A_i$关于$D$的外部重要度为：

$$sig_{\rm out}^{\Delta}(A_i,A',D)=|\gamma(A_i\cup A',D)-\gamma(A',D)|$$

定理1 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族。

(1)对于$\forall A_i\in A'$，其在属性子集族A'的基础上相对于D的粒度重要度为：

$$sig_{\rm in}^{\Delta}(A_i,A',D)=\frac{1}{m-1}\sum_{1\leqslant j\leqslant m,j\neq i}|sig_{\rm in}^{\Delta}(\{A_j,A_i\},A',D)-sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)|$$

(2)对于$\forall A_i\in A-A'$，其在属性子集族A的基础上相对于D的粒度重要度为：

$$sig_{\rm out}^{\Delta}(A_i,A',D)=\frac{1}{m-1}\sum_{1\leqslant j\leqslant m,j\neq i}|sig_{\rm out}^{\Delta}(\{A_j,A_i\},A_i\cup A',D)-sig_{\rm out}^{\Delta}(A_j,A',D)|$$

证明：

(1)由定义5可得：

$$\begin{aligned}&sig_{\rm in}^{\Delta}(A_i,A',D)\\&=|\gamma(A',D)-\gamma(A'-\{A_i\},D)|\\&=\frac{m-1}{m-1}|\gamma(A',D)-\gamma(A'-\{A_i\},D)|+\sum_{j\neq i}|\gamma(A'-\{A_j,A_i\},D)-\gamma(A'-\{A_j,A_i\},D)|\\&=\frac{1}{m-1}\sum_{j\neq i}(|\gamma(A',D)-\gamma(A'-\{A_j,A_i\},D)-(\gamma(A'-\{A_i\},D)-\gamma(A'-\{A_j,A_i\},D)|)\\&=\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm in}^{\Delta}(\{A_j,A_i\},A',D)-sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)|\end{aligned}$$

(2)由定义6可得：

$$\begin{aligned}&sig_{\rm out}^{\Delta}(A_i,A',D)=|\gamma(\{A_i\}\cup A',D)-\gamma(A',D)|\\&=\frac{m-1}{m-1}|\gamma(\{A_i\}\cup A',D)-\gamma(A',D)|-\sum_{j\neq i}|\gamma(A'-\{A_j\},D)-\gamma(A'-\{A_j\},D)|\\&=\frac{1}{m-1}\sum_{j\neq i}(|\gamma(\{A_i\}\cup A',D)-\gamma(A'-\{A_j\},D)|-|(\gamma(A'-\{A_j\},D)-\gamma(A',D)|)\\&=\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm out}^{\Delta}(\{A_j,A_i\},\{A_i\}\cup A',D)-sig_{\rm out}^{\Delta}(A_j,A',D)|\end{aligned}$$

定义6和定义7仅给出了单个粒度对整个粒度集的直接影响，并没有考虑其余粒度对决策的间接影响；定义8和定义9将综合分析计算粒度之间的相互依赖关系，并给出一种新的粒度重要度计算方法。

定义8 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族，则对于$\forall A_i\in A'$，其在属性子集族$A$的基础上相对于$D$的内部重要度为：

$$sig_{\rm in}'^{,\Delta}(A_i,A',D)=sig_{\rm in}^{\Delta}(A_i,A',D)+\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{\rm in}^{\Delta}(A_j,A',D)|$$

定义9 设决策信息系统$DIS=(U,C\cup D,V,f)$，$A'\subseteq A=\{A_1,A_2,\cdots,A_m\}$是$C$的$m$个属性子集族，则对于$\forall A_i\in A-A'$，其在属性子集族$A$的基础上相对于$D$的外部重要度为：

$$sig_{\rm out}'^{,\Delta}(A_i,A',D)=sig_{\rm out}^{\Delta}(A_i,A',D)+\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm out}^{\Delta}(A_j,A',D)-sig_{\rm out}^{\Delta}(A_j,\{A_i\}\cup A',D)|$$

定义8中，$sig_{\rm in}^{\Delta}(A_i,A',D)$表示$A_i$对决策的直接影响，$\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{\rm in}^{\Delta}(A_j,A',D)|$表示$A_i$对决策的间接影响。

其中，当$sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{\rm in}^{\Delta}(A_j,A',D)|>0$时，说明$A_j$的粒度重要度因在属性子集$A'-\{A_i\}$中增加了$A_i$而有所增大，故应将$A_j$的粒度重要度加入到$A_i$的粒度重要度中。因此，当共有$m$个属性子集时，应在$\{A_i\}$的粒度重要度上加上$\frac{1}{m-1}\sum_{j\neq i}|sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{\rm in}^{\Delta}(A_j,A',D)|$。

若当$j\neq i$时，$|sig_{\rm in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{\rm in}^{\Delta}(A_j,A',D)|=0$，则表明粒度$A_i$与其他粒度之间没有相互作用，此时

$sig_{in}^{\prime,\Delta}(A_i,A',D)=sig_{in}^{\Delta}(A_i,A',D)$。定理 2 和定理 3 将进一步说明改进的粒度重要度的合理性。

以上是对定义 8 的分析，定义 9 的分析类似，这里不再赘述。

定理 2　改进的内部重要度可以等价表示为：

$$sig_{in}^{\prime,\Delta}(A_i,A',D)=\frac{1}{m-1}\sum_{j\neq i}sig_{in}^{\Delta}(A_i,A'-\{A_j\},D)$$

证明：

$$\begin{aligned}&sig_{in}^{\prime,\Delta}(A_i,A',D)\\&=sig_{in}^{\Delta}(A_i,A',D)+\frac{1}{m-1}\sum_{j\neq i}|sig_{in}^{\Delta}(A_j,A'-\{A_i\},D)-sig_{in}^{\Delta}(A_j,A',D)|\\&=\frac{m-1}{m-1}|\gamma(A',D)-\gamma(A'-\{A_i\},D)|+\frac{1}{m-1}\sum_{j\neq i}\|\gamma(A'-\{A_i\},D)-\gamma(A'-\{A_j,A_i\},D)|-|\gamma(A',D)-\gamma(A'-\{A_j\},D)\|\\&=\frac{1}{m-1}\sum_{j\neq i}|\gamma(A'-\{A_j\},D)-\gamma(A'-\{A_j,A_i\},D)|\\&=\frac{1}{m-1}\sum_{j\neq i}sig_{in}^{\Delta}(A_i,A'-\{A_j\},D)\end{aligned}$$

定理 3　改进的外部重要度可以等价表示为：

$$sig_{out}^{\prime,\Delta}(A_i,A',D)=\frac{1}{m-1}\sum_{j\neq i}sig_{out}^{\Delta}(A_i,\{A_j\}\cup A',D)$$

证明：

$$\begin{aligned}&sig_{out}^{\prime,\Delta}(A_i,A',D)\\&=sig_{out}^{\Delta}(A_i,A',D)+\frac{1}{m-1}\sum_{j\neq i}|(sig_{out}^{\Delta}(A_j,A',D)-sig_{out}^{\Delta}(A_j,\{A_i\}\cup A',D))|\\&=\frac{m-1}{m-1}|\gamma(\{A_i\}\cup A',D)-\gamma(A',D)|+\frac{1}{m-1}\sum_{j\neq i}\|\gamma(A',D)-\gamma(\{A_j\}\cup A',D)|-|\gamma(\{A_i\}\cup A',D)-\gamma(\{A_i,A_j\}\cup A',D)\|\\&=\frac{1}{m-1}\sum_{j\neq i}|\gamma(\{A_i,A_j\}\cup A',D)-\gamma(\{A_i\}\cup A',D)|\\&=\frac{1}{m-1}\sum_{j\neq i}sig_{out}^{\Delta}(A_i,\{A_j\}\cup A',D)\end{aligned}$$

定理 2 表明，当 $j\neq i$，$sig_{in}^{\Delta}(A_i,A'-\{A_j\},D)=0$（$sig_{out}^{\Delta}(A_i,\{A_j\}\cup A',D)=0$）时，$sig_{out}^{\prime,\Delta}(A_i,A',D)=0$（$sig_{out}^{\prime,\Delta}(A_i,A',D)=0$），且各个粒度重要度是在 A' 去掉 A_j 的基础上计算的，这使得各个粒度重要度更容易区分。

3.2　粒度约简算法

由文献[11,13]可以得出悲观多粒度与乐观多粒度的下近似分别为 L^{I} 和 L^{O}。根据定理 2 和定理 3 得出基于改进的粒度重要度的粒度约简算法，具体如算法 1 所示。

算法 1　基于改进的粒度重要度的粒度约简算法

设信息系统 $\mathrm{DIS}=(\mathrm{U},\mathrm{C}\cup\mathrm{D},\mathrm{V},\mathrm{f})$，$\mathrm{A}'\subseteq\mathrm{A}=\{\mathrm{A}_1,\mathrm{A}_2,\cdots,\mathrm{A}_m\}$ 是 C 的 m 个属性子集族，$\forall\mathrm{A}_i\in\mathrm{A}'$，$\mathrm{U/D}=\{\mathrm{X}_1,\mathrm{X}_2,\cdots,\mathrm{X}_s\}$。

输入：决策信息系统 $\mathrm{DIS}=(\mathrm{U},\mathrm{C}\cup\mathrm{D},\mathrm{V},\mathrm{f})$

输出：该决策信息系统的一个粒度约简 Red^{Δ}

Step 1　设置 $\mathrm{Red}^{\Delta}\leftarrow\emptyset$，$1\leqslant \mathrm{k}\leqslant \mathrm{m}$。

Step 2　计算 U/D，计算悲观多粒度与乐观多粒度的下近似 L^{I} 和 L^{O}。

Step 3　对于 $\forall\mathrm{A}_i\in\mathrm{A}$，根据定义 2 计算 $\mathrm{sig}_{in}^{\prime,\Delta}(\mathrm{A}_i,\mathrm{A}',\mathrm{D})$。

Step 4　如果 $\mathrm{sig}_{in}^{\prime,\Delta}(\mathrm{A}_i,\mathrm{A}',\mathrm{D})>0$，则 $\mathrm{Red}^{\Delta}=\mathrm{Red}^{\Delta}\cup\mathrm{A}_i$。

Step 5　对于 $\forall\mathrm{A}_i\in\mathrm{A}-\mathrm{Red}^{\Delta}$：

Step 5.1　如果 $\gamma(\mathrm{Red}^{\Delta},\mathrm{D})=\gamma(\mathrm{A},\mathrm{D})$，转 Step 7；

Step 5.2　否则，根据定义 9，计算 $\mathrm{sig}_{out}^{\prime,\Delta}(\mathrm{A}_i,\mathrm{A}',\mathrm{D})$。

Step 6　如果 $\mathrm{sig}_{out}^{\prime,\Delta}(\mathrm{A}_k,\mathrm{A}',\mathrm{D})=\max\{\mathrm{sig}_{out}^{\prime,\Delta}(\mathrm{A}_k,\mathrm{A}',\mathrm{D})\}$，则 $\mathrm{Red}^{\Delta}=\mathrm{Red}^{\Delta}\cup\mathrm{A}_k$，转 Step 5；

Step 7　对于 $\forall\mathrm{A}_i\in\mathrm{Red}^{\Delta}$，如果 $\gamma(\mathrm{Red}^{\Delta}-\mathrm{A}_i,\mathrm{D})=\gamma(\mathrm{A},\mathrm{D})$，则 $\mathrm{Red}^{\Delta}=\mathrm{Red}^{\Delta}-\mathrm{A}_i$。

Step 8　输出粒度约简集 Red^{Δ}。

根据文献[11,13]，可得下近似的时间复杂度为 $O(|C|\cdot|U|^2)$，则在算法 1 中计算由 m 表示的 $X_t\in U/D=\{X_1,X_2,\cdots,X_s\}$ 的下近似的时间复杂度为 $O((s+1)\cdot|C|\cdot|U|^2)$。因此，算法 1 的时间复杂度为 $O(\sum_{i=1}^{m}(|C|-i+1)\cdot(s+1)\cdot|C|^2\cdot|U|^2)$。

4　基于粒度重要度的三支决策模型

在进行决策时，考虑到冗余信息会增加决策的难度，本节先将冗余信息删除，得到粒度约简，然后对约简集进行决策。

设 $DIS=(U,C\cup D,V,f)$ 是一个完备信息系统，其中 $A=\{A_1,A_2,\cdots,A_m\}$ 是 C 的 m 个属性子集族，令 A 的约简集合 $Red^{\Delta}=\{A_1,A_2,\cdots,A_r\}$，其中 $1\leqslant r\leqslant m$，下面对约简集 Red^{Δ} 进行决策判断。

采取决策动作 $\omega_P,\omega_N,\omega_B$ 的期望损失函数分别表示为：

$$R'(\omega_P|[x]_{A_r})=\sum_{i=1}^{r}\lambda_{PP}\cdot P(X|[x]_{A_r})+\sum_{i=1}^{r}\lambda_{PN}\cdot(1-P(X|[x]_{A_r}))$$

$$R'(\omega_N|[x]_{A_r})=\sum_{i=1}^{r}\lambda_{NP}\cdot P(X|[x]_{A_r})+\sum_{i=1}^{r}\lambda_{NN}\cdot(1-P(X|[x]_{A_r}))$$

$$R'(\omega_B|[x]_{A_r})=\sum_{i=1}^{r}\lambda_{BP}\cdot P(X|[x]_{A_r})+\sum_{i=1}^{r}\lambda_{BN}\cdot(1-P(X|[x]_{A_r}))$$

根据贝叶斯决策规则，决策者通常会选择风险最小的决策作为最佳方案，因此可得到新的基于粒度重要度的三支决策模型的决策规则。

（1）接受决策规则（P′）：若 $R'(\omega_P|[x]_{A_r})\leqslant R'(\omega_B|[x]_{A_r})$ 且 $R'(\omega_P|[x]_{A_r})\leqslant R'(\omega_N|[x]_{A_r})$，则 $x\in POS(X)$；

（2）拒绝决策规则（N′）：若 $R'(\omega_N|[x]_{A_r})\leqslant R'(\omega_P|[x]_{A_r})$ 且 $R'(\omega_N|[x]_{A_r})\leqslant R'(\omega_B|[x]_{A_r})$，则 $x\in NEG(X)$；

（3）延迟决策规则（B′）：若 $R'(\omega_B|[x]_{A_r})\leqslant R'(\omega_N|[x]_{A_r})$ 且 $R'(\omega_B|[x]_{A_r})\leqslant R'(\omega_P|[x]_{A_r})$，则 $x\in BND(X)$。

简化之后，P′，N′，B′规则可以改写为：

（1）接受决策规则（P_1'）：若 $R'(\omega_P|[x]_{A_r})\geqslant\alpha$ 且 $R'(\omega_P|[x]_{A_r})\geqslant\gamma$，则 $x\in POS(X)$；

（2）拒绝决策规则（N_1'）：若 $R'(\omega_N|[x]_{A_r})\leqslant\beta$ 且 $R'(\omega_N|[x]_{A_r})<\gamma$，则 $x\in NEG(X)$；

（3）延迟决策规则（B_1'）：若 $R'(\omega_B|[x]_{A_r})>\beta$ 且 $R'(\omega_B|[x]_{A_r})<\alpha$，则 $x\in BND(X)$。

由于不等式 $0\leqslant\beta<\alpha\leqslant1$ 成立，因此（P_1'），（N_1'），（B_1'）规则可以重写为：

（1）接受决策规则（P_2'）：若 $P'(X|[x]_{A_r})\geqslant\alpha$，则 $x\in POS(X)$；

(2)拒绝决策规则(N_2'):若 $P'(X|[x]_{A_r})\leqslant\beta$,则 $x\in NEG(X)$;

(3)延迟决策规则(B_2'):若 $\beta<P'(X|[x]_{A_r})<\alpha$,则 $x\in BND(X)$。

定义 8 和定义 9 同时考虑单个粒度对决策的直接影响和间接影响,改进了粒度约简算法,使得最后得到的约简集更加合理,最后对约简集进行规则提取。一种新的基于粒度重要度的三支决策模型的构造步骤如下。

Step 1 根据定理 2,计算每个粒度的内部重要度 $sig_{in}'^{,\Delta}(A_i,A',D)$;

Step 2 根据算法 1,计算约简集 Red^{Δ};

Step 3 根据 2.2 节的公式计算边界阈值(α,β);

Step 4 计算 $P(X|[x]_{A_r})$;

Step 5 分别对悲观多粒度和乐观多粒度的约简集进行规则提取。

5 实例分析

本文以文献[11]给出的某企业的 10 个在线投资方案为例进行粒度重要度的计算,通过对获得的各粒度重要度进行对比分析,对其进行粒度约简,找到最小评估站点集,最后利用三支决策规则进行投资方案的选择。

Step 1 由文献[11]可知,粒度集 $A=\{A_1,A_2,A_3,A_4,A_5\}$,$A'\subseteq A$,$U/D=\{\{x_1,x_2,x_4,x_6,x_8\},\{x_3,x_5,x_7,x_9,x_{10}\}\}$,即 $Y_1=\{x_1,x_2,x_4,x_6,x_8\}$,$Y_2=\{x_3,x_5,x_7,x_9,x_{10}\}$。

根据本文定理 2 可得悲观多粒度下每个站点的内部重要度,如表 2 所列。

表 2 悲观多粒度粗糙集的内部重要度

Table 2 Internal importance degree of pessimistic multi-granularity rough sets

	A_1	A_2	A_3	A_4	A_5
$sig_{in}^{I}(A_i,A',D)$	0	0.050	0	0	0
$sig_{in}'^{,I}(A_i,A',D)$	0	0.025	0	0.025	0.025

Step 2 根据算法 1,将内部重要度大于 0 的站点$\{A_2,A_4,A_5\}$加入到粒度约简集中作为起点粒度集,得:

$r^{I}(\{A_2,A_4\},D)=r^{I}(\{A_4,A_5\},D)=r^{I}(A,D)=0.2$

因此,不用计算外部重要度,即可得到悲观多粒度的约简集为 $R'^{,I}=\{A_2,A_4\}$或$\{A_4,A_5\}$。

Step 3 同理,可计算乐观多粒度下的约简集。

根据定理 2 计算乐观多粒度下每个站点的内部重要度,如表 3 所列。

表 3 乐观多粒度粗糙集的内部重要度

Table 3 Internal importance of optimistic multi-granularity rough sets

	A_1	A_2	A_3	A_4	A_5
$sig_{in}^{O}(A_i,A',D)$	0	0.150	0.050	0	0.050
$sig_{in}'^{,O}(A_i,A',D)$	0.025	0.375	0.225	0	0

根据算法 1,得到起点粒度集为$\{A_1,A_2,A_3\}$。

计算 $r^{O}(\{A_2,A_3\},D)=r^{O}(A,D)=1$,因此乐观多粒度的约简集为 $R'^{,O}=\{A_2,A_3\}$。

从表 2 中可以看出,根据定义 6 计算内部重要度时,仅有一个粒度 A_2 的重要度不为 0,而当粒度重要度为 0 时,各个粒度之间没有区分度,难以计算它们的粒度约简或粒度权重。文献[11]中,在得到粒度 A_2 的粒度重要度之后,A_2 的近似质量与整个粒度集的近似质量不同,因此还需再次计算外部重要度。在计算外部重要度时,仅考虑单个粒度对粒度 A_2 的影响,不能充分体现粒度 A_2 对整体决策的影响。

本文在计算粒度内部重要度和外部重要度时,均考虑去掉某个或者某些粒度对决策的影响。根据定理 2,计算悲观多粒度的内部重要度,得到粒度 A_2,A_4 和 A_5 的内部重要度不为 0,并且这 3 个粒度的近似分类质量与整个粒度集的近似分类质量相同,因此不需要计算外部重要度,减少了部分计算量。同理,本实例不需要计算乐观多粒度的外部重要度,使得计算更加简便。

Step 4 根据三支决策理论,对得到的约简集进行规则提取。

根据文献[11]中的损失函数表,以及 2.2 节中的相关公式,可以得到边界阈值为:

$$\alpha=\frac{8-2}{(8-2)+(2-0)}=0.75$$

$$\beta=\frac{2-0}{(2-0)+(6-2)}=0.33$$

因此,边界阈值为(0.33,0.75)。

Step 5 根据(P_2'),(N_2'),(B_2')规则可得:

$A_1:\beta<P(X|[x]_{A_1})=0.5<\alpha$

$A_2:P(X|[x]_{A_2})=0.75=\alpha$

$A_3:P(X|[x]_{A_3})=0.75=\alpha$

$A_4:\beta<P(X|[x]_{A_4})=0.5<\alpha$

$A_5:P(X|[x]_{A_5})=0.25<\beta$

通过实例说明,在计算各个粒度的重要度时,既要考虑对决策的直接影响,又要考虑粒度间的间接影响,以此得到悲观和乐观多粒度粗糙集的粒度约简集,提取了约简集的规则,简化了原始多粒度空间的决策过程。本实例在悲观多粒度的意义下,站点 2 的评估通过,站点 4 需要进一步考察再做决定,而站点 1、站点 3 的和站点 5 的评估可以忽略,或者对站点 4 需要进一步考察再做决定,站点 5 的评估不通过,而站点 1、站点 2 和站点 3 的评估可以忽略;在乐观多粒度的意义下,站点 2 和站点 3 的评估通过,站点 1、站点 4 和站点 5 的评估可以忽略。

结束语 本文在多粒度粗糙集的粒度内部重要度与外部重要度的基础上,进一步考虑了单个粒度对决策的直接影响与间接影响,从而提出新的粒度重要度计算方法,并在此基础上给出一种改进的粒度约简算法,使最终获得的约简集更加合理。通过将约简集与三支决策理论结合,进一步对粒度选择进行决策分析。最后通过对比实例分析证明了本文算法的有效性。下一步将考虑多个粒度对决策的直接影响与间接影响。

参考文献

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

[2] YAO Y Y. Probabilistic Rough Set Approximations[J]. International Journal of Approximate Reasoning,2008,49(2):255-271.

[3] YAO Y Y,WONG S K M. A Decision Theoretic Framework for Approximating Concepts[J]. International Journal of Man-machine Studies,1992,37(6):793-809.

[4] SLEZAK D,ZIARKO W. The Investigation of the Bayesian Rough Set Model[J]. International Journal of Approximate Reasoning,2005,40(1-2):81-91.

[5] ZIARKO W. Variable Precision Rough Set Model[J]. Journal of Computer and System Sciences,1993,46(1):39-59.

[6] YAO Y Y,LIN T Y. Generalization of Rough Sets Using Modal Logics[J]. Intelligent Automation & Soft Computing,1996,2(2):103-119.

[7] LIN T Y. Granular Computing on Binary Relations II:Rough Set Representations and Belief Functions[J]. Rough Sets in Knowledge Discovery,1998(1):122-140.

[8] QIAN Y H,LIANG J Y,YAO Y Y,et al. MGRS:A Multi-granulation Rough Set[J]. Information Sciences,2010,180(6):949-970.

[9] QIAN Y H,LIANG J Y,DANG C Y. Incomplete Multi-granulation Rough Set[J]. IEEE Transactions on Systems,Man,and Cybernetics-Part A:Systems and Humans,2010,40(2):420-431.

[10] SANG Y L,QIAN Y H. A Granular Space Reduction Approach to Pessimistic Multi-Granulation Rough Sets[J]. Pattern Recognition and Artificial Intellgence,2012,25(3):361-366.(in Chinese)
桑妍丽,钱宇华.一种悲观多粒度粗糙集中的粒度约简算法[J].模式识别与人工智能,2012,25(3):361-366.

[11] SANG Y L,QIAN Y H. Granular Structure Reduction Approach to Multigranulation Decision-theoretic Rough Sets[J]. Computer Science,2017,44(5):199-205.(in Chinese)
桑妍丽,钱宇华.多粒度决策粗糙集中的粒度约简方法[J].计算机科学,2017,44(5):199-205.

[12] DAI G Y,WANG Z M,YANG C,et al. A Multi-granularity Rough Set Algorithm for Attribute Reduction through Particles Particle Swarm Optimization[C]// Proceedings of International Computer Engineering Conference. New York:IEEE Press,2016:303-307.

[13] YANG X B,QI Y S,SONG X N,et al. Test Cost Sensitive Multigranulation Rough Set:Model and Minimal Cost Selection[J]. Information Sciences,2013,250(11):184-199.

[14] JING Y G,LI T R,HFUJITA,et al. An Incremental Attribute Reduction Approach Based on Knowledge Granularity with a Multi-granulation View[J]. Information Sciences,2017,411:23-38.

[15] YANG T,LI Z W,YANG X Q. A Granular Reduction Algorithm Based on Covering Rough Sets[J]. Journal of Applied Mathematics,2012,2012:1-13.

[16] LI J H,REN Y,MEI C L,et al. A Comparative Study of Multigranulation Rough Sets and Concept Lattices via Rule Acquisition[J]. Knowledge-Based Systems,2015,91:152-164.

[17] YAO Y Y. The Superiority of Three-way Decisions in Probabilistic Rough Set Models[J]. Information Sciences,2011,181(6):1080-1096.

[18] YAO Y Y. Three-Way Decision:An Interpretation of Rules in Rough Set Theory[C]// International Conference on Rough Sets and Knowledge Technology. Berlin:Springer,2009:642-649.

[19] LIANG D C,XU Z S,LIU D. Three-way Decisions with Intuitionistic Fuzzy Decision-theoretic Rough Sets Based on Point Operators[J]. Information Sciences,2017,375:183-201.

[20] QIAN J,DANG C Y,YUE X D,et al. Attribute Reduction for Sequential Three-way Decisions under Dynamic Granulation[J]. International Journal of Approximate Reasoning,2017,85:196-216.

[21] XUE Z A,ZHU T L,XUE T Y,et al. Methodology of Attribute Weights Acquisition Based on Three-way Decision Theory[J]. Computer Science,2015,42(8):265-268.(in Chinese)
薛占熬,朱泰隆,薛天宇,等.基于三支决策理论的条件属性权重构造方法[J].计算机科学,2015,42(8):265-268.

[22] HU B Q. Three-way Decisions Space and Three-way Decisions[J]. Information Sciences,2014,281:21-52.

[23] HU B Q,WONG H,YIU K F C. On Two Novel Types of Three-way Decisions in Three-way Decision Spaces[M]. New York:Elsevier Science Inc.,2017.

[24] SHI J L,ZHANG Q Y,DU G Y. Multi-granularity Three-way Decision Model Based on Minimum Risks[J]. Journal of Henan Normal University (Natural Science Edition),2017,45(2):101-107.(in Chinese)
史进玲,张全友,杜根远.基于风险最小化的多粒度三支决策模型[J].河南师范大学学报(自然科学版),2017,45(2):101-107.

[25] LI J H,HUANG C C,QI J J,et al. Three-way Cognitive Concept Learning via Multi-granularity[J]. Information Sciences,2017,378(1):244-263.

[26] CHEN H,LI J H,MIN F,et al. Optimal Scale Selection in Dynamic Multi-scale Decision Tables Based on Sequential Three-way Decisions[J]. Information Sciences,2017,415:213-232.

[27] XIAO J S,SUN L M. Improved Attribute Significance Degree Based on Rough Set[J]. Computer Engineering and Applications,2017,53(3):174-176.(in Chinese)
肖劲森,孙立民.改进的粗糙集属性重要度[J].计算机工程与应用,2017,53(3):174-176.