决策粗糙集理论研究现状与展望

于 洪[1),3)] 王国胤[1),2)] 姚一豫[3)]

[1)](重庆邮电大学计算智能重庆市重点实验室 重庆 400065)

[2)](中国科学院重庆绿色智能技术研究院电子信息技术研究所 重庆 400714)

[3)](里贾纳大学计算机科学系 里贾纳 S4S 0A2 加拿大)

摘 要 经典Pawlak粗糙集理论中的核心概念上、下近似集是通过集合相交非空和包含来定义的.由于缺乏对错误的容忍能力,其实际应用受到了限制.20世纪90年代初,Yao等人结合贝叶斯决策理论提出了决策粗糙集模型.近年来,该模型逐渐得到重视,并在不确定性信息处理方面得到了广泛应用.该文首先就为什么要提出决策粗糙集模型、该模型具有什么特点以及该模型中需要解决的几个问题进行了详细讨论.然后,总结了国内外关于决策粗糙集模型的研究现状和进展,详细分析了存在的挑战性问题,并深入探讨了未来的研究方向.

关键词 粗糙集;决策粗糙集;三支决策;数据分析;不确定性;智能信息处理

中图法分类号 TP18 **DOI号** 10.11897/SP.J.1016.2015.01628

Current Research and Future Perspectives on Decision-Theoretic Rough Sets

YU Hong[1),3)] WANG Guo-Yin[1),2)] YAO Yi-Yu[3)]

[1)](*Chongqing Key Laboratory of Computational Intelligence, Chongqing University of Posts and Telecommunications, Chongqing* 400065)

[2)](*Institute of Electronic Information Technology, Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing* 400714)

[3)](*Department of Computer Science, University of Regina, Regina, Saskatchewan, S4S 0A2 Canada*)

Abstract As the central concepts in rough set theory, the classical Pawlak lower and upper approximations are defined based on qualitative set-inclusion and non-empty overlapping relations, respectively. Consequently, the theory suffers from an intolerance of errors, which greatly restricts its real-world applications. To overcome this limitation, Yao and colleagues proposed a decision-theoretic rough sets (DTRS) model in early 1990s' by introducing the Bayesian decision theory into rough sets. In recent years, the model has attracted much attention and has been applied in uncertain information processing. This paper aims at (1) presenting a survey of the motivations for introducing the DTRS model, the main features of the model, and the problems to be studied in the model, (2) reviewing the fundamental results, state-of-art research, and challenges, and (3) pointing out future perspectives and potential research topics.

Keywords rough sets; decision-theoretic rough sets; three-way decisions; data analysis; uncertain; intelligent information processing

1 引 言

粗糙集(Rough Sets,也称Rough集、粗集)理论是Pawlak[1]于1982年提出的一种处理不精确、不一致、不完整信息与知识的数学工具.粗糙集理论作为一种数据分析处理理论,在机器学习、知识发现、数据挖掘、决策支持与分析、信息安全、物联网、

收稿日期:2013-07-18;最终修改稿收到日期:2014-11-12.本课题得到国家自然科学基金(61379114,61272060)、重庆市自然科学基金重点项目(cstc2013jjB40003)资助.于 洪,女,1972年生,博士,教授,中国计算机学会(CCF)会员,主要研究领域为三支决策、三支聚类、粗糙集、区间集、智能信息处理和Web智能和数据挖掘等.E-mail:yuhong@cqupt.edu.cn.**王国胤**,男,1970年生,博士,教授,中国计算机学会(CCF)会员,主要研究领域为粗糙集、粒计算、机器学习、数据挖掘、知识技术和认知计算等.E-mail:wanggy@ieee.org.**姚一豫**,男,1962年生,博士,教授,主要研究领域为三支决策、粗糙集、区间集、粒计算、信息检索、Web智能和数据挖掘等.

云计算、生物信息处理等领域得到了广泛且成功地应用[2-11].

粗糙集的基本思想是用可定义集合来刻画不可定义集合，从而给出一个概念的上近似集和下近似集定义.经典粗糙集的近似是基于概念之间的定性关系(即包含或相交不空)定义的，并不考虑概念相交的程度，因而不适用于处理很多实际问题.为了解决 Pawlak 粗糙集模型过于严格、缺乏容错能力的问题，人们提出了各种概率型粗糙集扩展模型.

1990 年，Yao 等人[12]提出了决策粗糙集模型(Decision-Theoretic Rough Sets，DTRS)，拓广了 Pawlak 等人[13]的 0.5-概率粗糙集模型.决策粗糙集模型的主要出发点是用条件概率定义概念的相交程度，并用两个阈值定义概率上下近似集.1993 年，Ziarko[14]提出了变精度粗糙集模型(Variable Precision Rough Sets，VPRS)，从集合包含度的视角给出了决策粗糙集模型的一个特例(即两个阈值之和为 1).随后，Pawlak 和 Skowron 相继提出了粗糙隶属函数概念[15]、参数化粗糙集模型[16-17]；2005 年，Ślęzak[18]提出了贝叶斯粗糙集模型(Bayesian Rough Sets，BRS)；2008 年，Herbert 和 Yao[19]提出了博弈粗糙集模型(Game-Theoretic Rough Sets，GTRS).这些工作增进了对粗糙集理论的研究，并且扩大了粗糙集理论的应用领域[20-29].

现实世界中更多的是不确定性信息，如何从这些不精确、不一致、不完整的信息中得到我们需要的知识，是广大学者一直关注的问题[30].决策粗糙集模型结合概率论展开研究，给出了粗糙集理论的定量描述，以及基于贝叶斯决策论的一个语义模型，同时也给出了一个实际、有效的解释和计算阈值的方法，为我们研究不确定知识提供了一个新的思路.近年来，在国内外粗糙集学术会议和有关期刊上关于决策粗糙集的研究成果日渐增多[6,8-10].例如，国际认知信息学系列会议(ICCI)在 2010 年、国际粗糙集与知识技术系列会议(RSKT)自 2009 年以来都成功举办了以决策粗糙集为主题的专题讨论；中国 Rough 集与软计算、Web 智能及粒计算联合学术会议(CRSSC-CWI-CGrC)自 2010 年以来，每年都举办了以决策粗糙集为主题的分组讨论.此外，《International Journal of Approximate Reasoning》与《Fundamenta Informaticae》等国际学术期刊出版了以决策粗糙集为主题的专辑.决策粗糙集正在成为当前的研究热点.

本文首先简要介绍了 Pawlak 粗糙集和决策粗糙集的一些基本知识，并给出了概率粗糙集模型理论研究的 3 个基本问题.然后，围绕这些基本问题，解释了决策粗糙集模型的贡献，并综述了该模型在这些问题上的已有解决方案.最后，介绍国内外决策粗糙集模型的研究与应用现状，以及需要重点研究的主要问题.我们将这个理论模型目前的研究状况介绍给信息科学工作者，希望进一步推动并促进该领域的研究工作.

2　经典 Pawlak 粗糙集模型

粗糙集主要研究的问题是集合的近似及相关的数据分析和推理方法与算法[1,31].粗糙集理论的重要贡献是给出了一种基于等价关系的数据分析方法，并给出了一个非常精确、严格的数学描述.粗糙集理论首次形式化地描述了对象不可分辨性、属性冗余性及属性约简等重要概念.

作为一种数据分析方法，粗糙集主要以数据表为工具研究属性之间的依赖关系，从而获得有用的分类知识.一个数据表定义为一个有穷对象集和属性集的二元组，即 $S=(U,At)$.一个属性子集定义一个对象集上的等价关系，记为 E，其等价类是基本的可定义子集.通过等价类，我们可以描述或近似描述 U 的任何一个子集.设子集 $X\subseteq U$ 表示一个概念所包含的对象集，即该概念的外延，它不一定可以准确地用 E 的等价类来描述，也就是说 X 不一定是一组等价类的并集.因此，用一对上近似和下近似来刻画 X：

$$\overline{apr}(X)=\{x\in U\mid [x]\cap X\neq\varnothing\},$$
$$\underline{apr}(X)=\{x\in U\mid [x]\subseteq X\} \tag{1}$$

给定任何一个子集 $X\subseteq U$，基于它的上、下近似，得到 U 的一个划分：

$$\begin{aligned}&\mathrm{POS}(X)=\underline{apr}(X)=\{x\in U\mid [x]\subseteq X)\},\\&\mathrm{NEG}(X)=U-\overline{apr}(X)=\{x\in U\mid [x]\cap X=\varnothing\},\\&\mathrm{BND}(X)=\overline{apr}(X)-\underline{apr}(X)\\&\quad=\{x\in U\mid [x]\cap X\neq\varnothing\wedge\neg([x]\subseteq X)\}\end{aligned} \tag{2}$$

这 3 个子集分别称为 X 的正域 $\mathrm{POS}(X)$、负域 $\mathrm{NEG}(X)$和边界域 $\mathrm{BND}(X)$.

上近似、下近似从定性的角度考虑了两种情况，即可能性和必然性.上近似解释为如果存在一个 x 的等价对象在集合 X 中，那么这个对象可能属于 X；下近似解释为如果一个对象 x 的所有等价对象都在集合 X 中，那么它必然属于 X.

3 概率粗糙集模型

Pawlak 粗糙集可以被视为一种定性的近似,下近似由集合包含定义而上近似由集合相交非空定义.该定义不允许任何不确定性,这种优点同时也带来它的局限性.

在 Pawlak 粗糙集中,由于正域是建立在代数包含关系基础上的,因此难以体现概念表示的容错性,这正是经典粗糙集模型的局限所在.针对 Pawlak 粗糙集模型缺乏容错能力的问题,我们需要考虑 Pawlak 粗糙集的另一种表示,即将概率近似空间引入到粗糙集的研究中,获得定量粗糙集模型.

Wong 和 Ziarko[32-33] 于 1987 年将概率近似空间引入到粗糙集的研究中.令 $\Pr(X|[x])$ 表示任何一个对象在属于 $[x]$ 的条件下属于 X 的条件概率.那么,可以获得下面的等价条件:

$$\Pr(X|[x])=1\Leftrightarrow[x]\subseteq X,$$
$$\Pr(X|[x])=0\Leftrightarrow[x]\cap X=\varnothing,$$
$$0<\Pr(X|[x])<1\Leftrightarrow[x]\cap X\neq\varnothing\wedge\neg([x]\subseteq X).$$

这样,就得到了 Pawlak 三个域的另一种表示:

$$\begin{aligned}&\mathrm{POS}(X)=\{x\in U|\Pr(X|[x])=1\},\\&\mathrm{NEG}(X)=\{x\in U|\Pr(X|[x])=0\},\\&\mathrm{BND}(X)=\{x\in U|0<\Pr(X|[x])<1\}\end{aligned}\quad(3)$$

显然,定性粗糙集中的 3 个域仅仅使用了概率的两个极端值,即 0 和 1.这种表示为定量粗糙集给出了一个很好的启示.如果我们将 0 和 1 用其他的值来表示,那么就可以获得一种定量粗糙集模型.

在 1990 年,Yao 等人[12] 提出了决策粗糙集模型(Decision-theoretic Rough Sets Model,DTRS Model).该模型用一对概率阈值来替换上面所提到的 0 和 1.设 $0\leq\beta<\alpha\leq1$,则 (α,β)-概率正、负和边界域可定义如下:

$$\begin{aligned}&\mathrm{POS}_{(\alpha,\beta)}(X)=\{x\in U|\Pr(X|[x])\geq\alpha\},\\&\mathrm{NEG}_{(\alpha,\beta)}(X)=\{x\in U|\Pr(X|[x])\leq\beta\},\\&\mathrm{BND}_{(\alpha,\beta)}(X)=\{x\in U|\beta<\Pr(X|[x])<\alpha\}\end{aligned}\quad(4)$$

当阈值 (α,β) 取值为 $(1,0)$ 时,我们就获得了 Pawlak 粗糙集.因此,从形式上看,(α,β)-正、负和边界域拓广了 Pawlak 粗糙集.对于构建新的模型来讲,这还远远不够,我们需要探讨和解释该模型所用到的基本概念、基本量和语义解释.

关于概率粗糙集模型,至少有以下 3 个问题需要解决[34]:

(1) 阈值 α 和 β 的解释与计算;

(2) 条件概率 $\Pr(X|[x])$ 的估计;

(3) 概率正、负及边界域的解释与应用.

决策粗糙集模型的研究贡献在于它不仅给出了概率正、负和边界域这个结果,更重要的是给出了解决这 3 个问题的合理方案,比如:基于贝叶斯决策论可以通过决策风险最小化获得阈值的计算和解释[12];通过朴素贝叶斯模型估计条件概率[35];概率 3 个区域可以看做是三支决策理论的应用[36-37].因此,决策粗糙集是一个有坚实理论基础同时又实用的模型[38-40].

4 决策粗糙集理论研究的 3 个问题

在本节中,我们将围绕上一节的 3 个问题介绍决策粗糙集的已有研究结果.

4.1 阈值的解释与计算

与 Pawlak 正、负域不同,概率正、负域包含错误分类.正域的错误分类率是 $1-\Pr(X|[x])\leq1-\alpha$,负域的错误分类率是 $\Pr(X|[x])\leq\beta$.这为 α 和 β 给出了一种基于错误分类率的解释.该解释有其直观易懂的优点.但是,这并没有给出一种指导思想和一套有效的方法来解释和获得这两个阈值.

在 1985 年的科技报告中,Wong 和 Ziarko[33] 提出了 0.5-概率粗糙集模型,该模型随后在 Pawlak 等人[13] 的文章中有更进一步的介绍.这个模型的主要理论依据是多数规则(majority rule).它用一个 0.5 概率阈值来定义概率正、负和边界域:

$$\begin{aligned}&\mathrm{POS}_{0.5}(X)=\{x\in U|\Pr(X|[x])>0.5\},\\&\mathrm{NEG}_{0.5}(X)=\{x\in U|\Pr(X|[x])<0.5\},\\&\mathrm{BND}_{0.5}(X)=\{x\in U|0<\Pr(X|[x])=0.5\}\end{aligned}\quad(5)$$

阈值 0.5 定量地刻画了多数规则,当等价类 $[x]$ 中超过一半的元素属于 X 时,我们可以将 x 放到 X 的正域中;当超过一半的元素不属于 X 时,我们可以将 x 放到 X 的负域中;当刚好一半的元素属于 X 时,我们可以将 x 放到 X 的边界域中.但这种多数规则并不能解释一般的 (α,β) 阈值.

关于一般的 (α,β) 阈值的确定,决策粗糙集简单地使用了贝叶斯决策理论[12].对于一个子集 $X\subseteq U$,可以构造一个状态集合 $\Omega=\{X,X^c\}$,其中 X 和 X^c 互补.

对应于粗糙集中的正域、边界域和负域,我们就可以构造一个决策动作集 $Action=\{a_P,a_B,a_N\}$,

其中，a_P，a_B和a_N分别代表将一个对象分类到正域、边界域和负域的决策动作，即$x\in \mathrm{POS}(X)$，$x\in \mathrm{BND}(X)$，$x\in \mathrm{NEG}(X)$. 不同的决策动作会导致不同的分类后果，可能的 6 种损失函数见表 1. 其中，第 1 列函数表示一个对象属于集合 X 时，采取动作a_P，a_B和a_N带来的损失函数记为λ_{PP}，λ_{BP}和λ_{NP}；第 2 列函数表示一个对象不属于集合 X 时，采取动作a_P，a_B和a_N带来的损失函数记为λ_{PN}，λ_{BN}和λ_{NN}.

表 1　损失函数

	X(正例)	X^c(负例)
a_P	$\lambda_{PP}=\lambda(a_P\mid X)$	$\lambda_{PN}=\lambda(a_P\mid X^c)$
a_B	$\lambda_{BP}=\lambda(a_B\mid X)$	$\lambda_{BN}=\lambda(a_B\mid X^c)$
a_N	$\lambda_{NP}=\lambda(a_N\mid X)$	$\lambda_{NN}=\lambda(a_N\mid X^c)$

损失函数的意义取决于具体的应用. 通常，它可以用其他更直观的概念定义和解释，比如，钱、时间、人力等资源的度量，或者是不同后果的危险程度的度量. 一方面，在建立一个数学模型时，不需要考虑一个具体的解释，因为这会限制模型的一般性. 另一方面，非常需要建立这种抽象概念和应用中的具体概念的关系，因为这才能保证在实际应用中对抽象概念赋予有意义的解释.

对于$[x]$中的对象，采取不同的动作所产生的损失表示如下：

$$R(a_P\mid[x])=\lambda_{PP}\Pr(X\mid[x])+\lambda_{PN}\Pr(X^c\mid[x]),$$
$$R(a_B\mid[x])=\lambda_{BP}\Pr(X\mid[x])+\lambda_{BN}\Pr(X^c\mid[x]),$$
$$R(a_N\mid[x])=\lambda_{NP}\Pr(X\mid[x])+\lambda_{NN}\Pr(X^c\mid[x]) \quad (6)$$

贝叶斯决策论给出了以下最小风险决策规则：

(P) 如果 $R(a_P\mid[x])\le R(a_B\mid[x])$且 $R(a_P\mid[x])\le R(a_N\mid[x])$，则选择 $x\in \mathrm{POS}(X)$，

(B) 如果 $R(a_B\mid[x])\le R(a_P\mid[x])$且 $R(a_B\mid[x])\le R(a_N\mid[x])$，则选择 $x\in \mathrm{BND}(X)$，

(N) 如果 $R(a_N\mid[x])\le R(a_P\mid[x])$且 $R(a_N\mid[x])\le R(a_B\mid[x])$，则选择 $x\in \mathrm{NEG}(X)$.

当两个动作有同样的风险时，需要引入决胜规则(tie-breaking rule)，这样每一个对象只划分到唯一的域中.

因为 $\Pr(X\mid[x])+\Pr(X^c\mid[x])=1$，我们可以只使用概率$\Pr(X\mid[x])$和损失函数$\lambda$来简化这些规则. 一般来说，将一个属于 X 的对象 x 划分到正域 $\mathrm{POS}(X)$的损耗小于或者等于将其划分到边界域 $\mathrm{BND}(X)$的损耗；而且，上述两种损耗应该小于将这个对象划分到负域 $\mathrm{NEG}(X)$的损耗. 反过来，如果将一个不属于 X 的对象 x 划分到负域 $\mathrm{NEG}(X)$的损耗应该小于或等于将其划分到边界域 $\mathrm{BND}(X)$的损耗；而且，这两种损耗应该小于将这个对象划分到正域 $\mathrm{POS}(X)$的损耗. 即有关系：$\lambda_{PP}\le\lambda_{BP}<\lambda_{NP}$，$\lambda_{NN}\le\lambda_{BN}<\lambda_{PN}$. 基于这两个条件，从规则(P)～(N)可以获得以下 3 个阈值：

$$\alpha=\frac{(\lambda_{PN}-\lambda_{BN})}{(\lambda_{PN}-\lambda_{BN})+(\lambda_{BP}-\lambda_{PP})},$$
$$\gamma=\frac{(\lambda_{PN}-\lambda_{NN})}{(\lambda_{PN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{PP})},$$
$$\beta=\frac{(\lambda_{BN}-\lambda_{NN})}{(\lambda_{BN}-\lambda_{NN})+(\lambda_{NP}-\lambda_{BP})} \quad (7)$$

可以证明$\alpha\in(0,1]$，$\gamma\in(0,1)$，$\beta\in[0,1)$. 基于上述 3 个阈值，规则(P)～(N)可以重新表达为：

(P) 如果 $\Pr(X\mid[x])\ge\alpha$ 且 $\Pr(X\mid[x])\ge\gamma$，则选择 $x\in \mathrm{POS}(X)$，

(B) 如果 $\Pr(X\mid[x])\le\alpha$ 且 $\Pr(X[x])\ge\beta$，则选择 $x\in \mathrm{BND}(X)$，

(N) 如果 $\Pr(X\mid[x])\le\beta$ 且 $\Pr(X\mid[x])\le\gamma$，则选择 $x\in \mathrm{NEG}(X)$.

如果进一步假设损失函数满足$\frac{\lambda_{NP}-\lambda_{BP}}{\lambda_{NP}-\lambda_{NN}}>\frac{\lambda_{BP}-\lambda_{PP}}{\lambda_{PN}-\lambda_{BN}}$，可以证明$\alpha>\gamma>\beta$. 那么不再需要阈值$\gamma$，规则(P)～(N)进一步简化为($\mathrm{P}_1$)～($\mathrm{N}_1$)：

(P_1) 如果 $\Pr(X\mid[x])\ge\alpha$，则选择 $x\in \mathrm{POS}(X)$；

(B_1) 如果 $\beta<\Pr(X\mid[x])<\alpha$，则选择 $x\in \mathrm{BND}(X)$；

(N_1) 如果 $\Pr(X\mid[x])\le\beta$，则选择 $x\in \mathrm{NEG}(X)$.

这样，决策粗糙集模型给出了解释和计算阈值α、β的理论和方法.

4.2　条件概率的估计

Yao 和 Zhou[35] 提出了朴素贝叶斯粗糙集模型并给出了估计条件概率 $\Pr(X\mid[x])$的一种简单方法.

因为条件概率 $\Pr(X\mid[x])$很难从可观察到的数值中估计，通常用 Bayes 公式进行变换，即

$$\Pr(X\mid[x])=\frac{\Pr([x]\mid X)\Pr(X)}{\Pr([x])},$$

其中，$\Pr(X\mid[x])$称为后验概率，$\Pr(X)$称为先验概率，$\Pr([x]\mid X)$称为似然函数. 在文献[35]中，作者通过引入赔率(odds)变换、对数(logit)变换巧妙地将条件概率 $\Pr(X\mid[x])$的估计转换为似然比 $\Pr([x]\mid X)/\Pr([x]\mid X^c)$的估计，即

$$\begin{aligned}\mathrm{logit}(\Pr(X\mid[x]))&=\log(O(\Pr(X\mid[x])))\\&=\log\frac{\Pr([x]\mid X)}{\Pr([x]\mid X^c)}+\log\frac{\Pr(X)}{\Pr(X^c)}\end{aligned} \quad (8)$$

假设$[x]$由一个数据表的属性子集$A\subseteq At$定义，即$[x]=[x]_A=\bigcap_{a\in A}[x]_{\{a\}}$. 为了估计$\Pr(X|[x])$和$\Pr([x]|X^c)$，朴素贝叶斯粗糙集做了下面的概率独立假设：

$$\begin{aligned}\Pr([x]\mid X)&=\Pr([x]_A\mid X)\\&=\Pr(\bigcap_{a\in A}[x]_{\{a\}}\mid X)\\&=\prod_{a\in A}\Pr([x]_{\{a\}}\mid X),\\\Pr([x]\mid X^c)&=\Pr([x]_A\mid X^c)\\&=\Pr(\bigcap_{a\in A}[x]_{\{a\}}\mid X^c)\\&=\prod_{a\in A}\Pr([x]_{\{a\}}\mid X^c).\end{aligned}$$

将它们代入logit式(8)，得到

$$\operatorname{logit}(\Pr(X\mid[x]_A))=\sum_{a\in A}\log\frac{\Pr([x]_{\{a\}}\mid X)}{\Pr([x]_{\{a\}}\mid X^c)}+\log\frac{\Pr(X)}{\Pr(X^c)}.$$

那么，基于信息表中的数据，概率$\Pr([x]_{\{a\}}|X)$和$\Pr([x]_{\{a\}}|X^c)$可以通过如下公式估计：

$$\begin{aligned}\Pr([x]_{\{a\}}|X)&=\frac{|[x]_{\{a\}}\bigcap X|}{|X|},\\\Pr([x]_{\{a\}}|X^c)&=\frac{|[x]_{\{a\}}\bigcap X^c|}{|X^c|}\end{aligned}\tag{9}$$

其中$|\cdot|$表示集合的势. 通常，$[x]_{\{a\}}$的对象个数会多一点，相对而言这些估计可能更可靠一点.

就阈值来说，可以建立以下的关系：

$$\begin{aligned}&\Pr(X\mid[x])\geq\alpha\\\Leftrightarrow&\frac{\Pr(X\mid[x])}{\Pr(X^c\mid[x])}\geq\frac{\alpha}{1-\alpha}\\\Leftrightarrow&\frac{\Pr([x]\mid X)}{\Pr([x]\mid X^c)}\frac{\Pr(X)}{\Pr(X^c)}\geq\frac{\alpha}{1-\alpha}\\\Leftrightarrow&\log\frac{\Pr([x]\mid X)}{\Pr([x]\mid X^c)}+\log\frac{\Pr(X)}{\Pr(X^c)}\geq\log\frac{\alpha}{1-\alpha}.\end{aligned}$$

在概率独立假设下有

$$\Pr(X|[x])\geq\alpha\Leftrightarrow\sum_{a\in A}\log\frac{\Pr([x]_{\{a\}}|X)}{\Pr([x]_{\{a\}}|X^c)}\geq\log\frac{\Pr(X^c)}{\Pr(X)}+\log\frac{\alpha}{1-\alpha},$$

令

$$\alpha'=\log\frac{\Pr(X^c)}{\Pr(X)}+\log\frac{\alpha}{1-\alpha}\tag{10}$$

同样对于β，可以建立类似的关系，即可令

$$\beta'=\log\frac{\Pr(X^c)}{\Pr(X)}+\log\frac{\beta}{1-\beta}\tag{11}$$

则可以获得以下概率正、负和边界域的定义：

$$\begin{aligned}\mathrm{POS}_{(\alpha,\beta)}(X)&=\left\{x\in U\middle|\sum_{a\in A}\log\frac{\Pr([x]_{\{a\}}|X)}{\Pr([x]_{\{a\}}|X^c)}\geq\alpha'\right\},\\\mathrm{NEG}_{(\alpha,\beta)}(X)&=\left\{x\in U\middle|\sum_{a\in A}\log\frac{\Pr([x]_{\{a\}}|X)}{\Pr([x]_{\{a\}}|X^c)}\leq\beta'\right\},\\\mathrm{BND}_{(\alpha,\beta)}(X)&=\left\{x\in U|\beta'<\sum_{a\in A}\log\frac{\Pr([x]_{\{a\}}|X)}{\Pr([x]_{\{a\}}|X^c)}<\alpha'\right\}\end{aligned}\tag{12}$$

式(12)中，α'和β'可由α、β、$\Pr(X)$和$\Pr(X^c)$导出，其中概率可通过式(9)计算得到.

4.3 三支决策

三支决策提出的最初目的是为了给粗糙集3个域更合理的语义解释. 最新研究表明，粗糙集可以看作是三支决策的一个特例. 三支决策给出了更广的理论、方法和工具[36,38-40].

三支决策是现实生活中常用的策略之一. 比如，医生通常采用三支决策，即根据初步诊断对病人实施治疗、不治疗或进一步观察，这3类措施可能有错，因而会导致不同的风险. 杂志编辑也常采用三支决策处理稿件，即根据审稿人意见决定接收、拒绝或进一步审查. 投资管理者也采用投资、不投资或进一步观察的三支决策. 这样的例子还有很多.

从很大程度上讲，粗糙集的正域、负域和边界域为三支决策提供了理论基础. 具体地说，正域所对应的规则，简称正规则，表示接收；负域对应的规则，简称负规则，表示拒绝；边界域对应的规则，简称边界规则，对应不做决定或者推迟决定. 不论是接收还是拒绝都可能带有错误，即接收错误或拒绝错误，这与统计学中对一个假设的接收和拒绝相似.

首先来看看基于经典Pawlak粗糙集的三支决策. 设$\mathrm{Des}(x)=\bigwedge_{a\in A}(a=I_a(x))$是合取可定义概念$[x]_A\subseteq U$的描述，即内涵. 从Pawlak粗糙集的正、负和边界域可推导出如下三支决策规则：

$\mathrm{Des}(x)\rightarrow$接收，$x\in X$，$[x]_A\subseteq\mathrm{POS}(X)$，

$\mathrm{Des}(x)\rightarrow$既不接收也不拒绝，

$x\in X$，$[x]_A\subseteq\mathrm{BND}(X)$，

$\mathrm{Des}(x)\rightarrow$拒绝，$x\in X$，$[x]_A\subseteq\mathrm{NEG}(X)$.

对正规则而言，由于$[x]_A\subseteq\mathrm{POS}(X)$，它的置信度(confidence) $\delta=\Pr(X|[x]_A)=1$，同时错误率为$1-\delta=0$. 对负规则而言，由于它的置信度$\delta'=\Pr(X^c|[x]_A)=1$，因而错误率为$1-\delta'=0$. 对于边界规则，它的置信度和错误率介于0和1之间. 因为正、负规则不产生错误，所以基于Pawlak粗糙集的三支决策可以理解为定性决策规则.

在决策粗糙集模型中，正域、边界域和负域由一对阈值(α,β)决定，其中$0\leq\beta<\alpha\leq1$. 因此，它们能够容忍

一定程度的错误，其所对应的定量三支决策规则是：

$\mathrm{Des}(x)\rightarrow$ 接收，$x\in X$，$[x]_A\subseteq \mathrm{POS}_{(\alpha,\beta)}(X)$，

$\mathrm{Des}(x)\rightarrow$ 既不接收也不拒绝，

$x\in X$，$[x]_A\subseteq \mathrm{BND}_{(\alpha,\beta)}(X)$，

$\mathrm{Des}(x)\rightarrow$ 拒绝，$x\in X$，$[x]_A\subseteq \mathrm{NEG}_{(\alpha,\beta)}(X)$.

根据决策粗糙集中正域的定义，正规则的置信度满足 $\delta=\Pr(X|[x]_A)\geq\alpha$，其错误率满足 $1-\delta\leq 1-\alpha$. 负规则的置信度满足 $\delta'=\Pr(X^c|[x]_A)=1-\Pr(X|[x]_A)\geq 1-\beta$，错误率满足 $1-\delta'\leq\beta$.

当阈值 $\alpha=1$ 且 $\beta=0$ 时，经典粗糙集模型可以看作是决策粗糙集模型的一个特例.

定性和定量三支决策的主要区别是定性决策不允许有错误，而定量决策能够容忍一定的错误. 在实际应用中，很难做到完全没有错误的决策，因而定量的接收和拒绝显得更有意义. 换句话说，当支持证据很强时，就可以采取接收策略；当否定证据很强时，就可以采取拒绝策略；当证据不足时，可以延迟决策，并寻求新的证据. 决策粗糙集中的阈值 (α,β) 反映了所需证据的强弱. 这样，决策粗糙集可以和统计学中假设验证相联系，也为实际应用中的三支决策提供了一个理论模型.

三支决策是复杂问题的有效求解方法与实践. 网站 http://www2.cs.uregina.ca/~twd/为三支决策研究者提供了一个信息共享和交流的平台.

5　决策粗糙集模型的研究现状

如前所述，基于经典 Pawlak 粗糙集的知识获取是一种定性的研究，不容许有错误. 而实际应用中，很难达到没有错误的决策，因而基于决策粗糙集模型定量的研究更有意义.

知识获取过程中，人类经验尤其是专家知识对于规则的获取显得非常重要；换句话说，我们在进行知识获取的过程中考虑用户的主观意思/意愿是合理的. 另外一方面，不依赖经验知识，尤其是许多应用领域很难得到经验知识，利用数据驱动的方式进行知识获取的研究一直是机器学习追求的目标，也就是说，我们希望客观地获取知识. 其实，无论是用主观的方式还是客观的方式获取知识，二者相互之间并不矛盾；他们从不同的角度提出了解决问题的方法，很多时候，二者也很好地进行了结合. 所以，接下来我们将从这两方面来介绍基于决策粗糙集模型的拓展研究以及应用.

5.1　结合三支决策语义的研究

经典 Pawlak 代数粗糙集模型中的单个决策类（目标概念）可以采用正域、负域和边界域进行描述. 考虑总体决策类的正域、负域和边界域：因为总体决策类中包含的单个决策类的负域一定可以划分到其补决策类的正域或边界域中，所以总体决策类负域实际上是空集. 这样，论域就是总体决策类正域和边界域的并集，而总体决策类的负域恒为空集. 因此，事实上产生的决策规则是二支决策. 即由总体决策类正域导出的确定性规则和总体决策类边界域导出的可能性规则. 此时负域没有实际语义，被当作是冗余的. 这种不完备粗糙结构反映了 Pawlak 代数粗糙集模型对总体决策类刻画具有不完整性. 另外，在经典的最小风险 Bayes 决策理论中，决策行为仅包含分类为正例和分类为负例，即必须对分类做出确定的选择，其对应的决策形式也为二支决策，没有考虑在数据信息不足背景下的中间决策方式.

经典 Pawlak 代数粗糙集模型和 Bayes 最小风险决策理论中的二支决策语义在决策粗糙集模型中得到了拓广. 在决策粗糙集模型中论域 U 可以通过阈值 α 和 β 划分为 3 个区域，即总体决策类正域、总体决策类边界域和总体决策类负域. 从语义上看，这 3 个区域可以对应 3 种规则类型及三支决策语义，即正域决策、边界域决策和负域决策. 这种语义的完备性使得决策粗糙集在应用时的描述更具完整性.

因此，许多学者结合三支决策从主观语义的方向展开了基于决策粗糙集模型的研究.

(1) 决策模型的研究

Li 和 Zhou[41] 结合决策者不同的风险偏好提出了一种新的基于决策粗糙集的决策模型，该模型从三支风险的视角对决策模型进行了研究；Zhou 和 Li[42] 还研究了基于决策粗糙集的多层次决策规则提取方法.

决策粗糙集模型中的两个状态 X、X^c 分别表示某事件属于 X 和不属于 X，这是一个两分类问题. 然而在实际应用中，决策者可能面临的状态集是多分类的情形，Liu 等人[43-44] 研究了多分类的三支决策粗糙集模型. 此外，Liu 等人[45] 还介绍了基于判别分析的三支决策方法等.

Yang 和 Yao[46] 提出了多用户决策粗糙集模型. 当不同用户（比如专家或者管理者等）根据各自的标准给出不同的决策集时，对于如何制定综合的、有共识的决策标准，该文给出了一种很好的解决方

案.梁吉业等人在文献[40]提出了一种新的基于粗糙集的多属性群决策方法，该方法分析了初步决策集合的结构,给出专家评价相似度的定义和性质,通过计算专家评价之间的相似度确定专家的客观权重.

在许多不确定决策问题中其决策对象往往涉及两个不同但又相互关联的论域,如在一个医疗诊断系统中,其同时涉及症状与疾病这两个彼此相关的不同集合.因此,讨论多论域上的决策问题很有必要.Ma 和 Sun[47-48] 系统地讨论了双论域上概率粗糙集的基本理论,为模型中的参数给出了一种较少依赖决策者主观偏好和直觉经验的选择方式,进一步地给出了双论域上的决策粗糙集模型.

(2)属性约简

属性约简是粗糙集研究中的重要研究内容之一[49].通常说来,一个属性约简是保持某种目标性质不变的独立属性子集,不同的约简目标则约简类型不同.在这些约简中,正域约简讨论是最多的.在 Pawlak 代数粗糙集模型中,因为正域相对于条件属性集具有单调性,所以约简只需保证条件属性相对决策属性的依赖度保持不变即可.然而,在概率粗糙集模型中,正域不再具有相对于条件属性的单调性,仅保持依赖度不变不能作为概率粗糙集约简的判定依据,还需要依靠其他的属性集度量标准.因此,如何给出适合概率型粗糙集模型约简的一般定义以及分析其相关性质是决策粗糙集约简理论研究的主要问题之一[49-50].

Li 等人[50]针对决策粗糙集正域的非单调性特点给出了一种新的 α-正域约简的定义.在该定义中,约简前后的正域不要求严格相等,而只要求正域保持非减特性.新定义中对正域约束条件的改变与决策粗糙集正域的非单调性是相吻合的,在此基础上,作者给出了求 α-正域约简的启发式搜索算法.Ma 等人[51]提出了基于决策粗糙集的多类属性约简模型;Chebrolu 和 Sanjeevi[52]结合遗传算法研究了决策粗糙集模型下的属性约简问题;Liu 和 Min 等人[53-54]研究了测试代价最优情况下的正域属性约简问题.

5.2 结合数据驱动的研究

众多的知识获取算法中,都涉及到一些参数或阈值的设定.事实上,如果没有相关领域知识和经验,这在多数情况下是不可行的,如外太空探测或者是网络入侵检测等领域的先验知识很难获得.因此,如何客观地、数据驱动地结合决策粗糙集理论来进行知识获取也是一个重要的研究方向.

(1)属性约简

Jia 等人[55-56]指出目前决策粗糙集模型下定义的属性约简都要求约简前后正区域保持不变或者非负区域不变等,而属性约简所带来的区域变化的好坏却无法判断,只能人为地偏向于保持或增大正区域或非负区域,这在理论性和可解释性上存在一定的困难.因此,作者给出了决策粗糙集模型下基于决策风险最小化的属性约简定义,要求在约简后的属性集合上作出最小的决策风险.相比传统的基于正域保持的属性约简定义,该定义对于正域、边界域和负域同等对待、不加入主观偏好信息,使得约简的结果更具客观性和可解释性,同时将决策风险形式化为一个最优化问题,为求解最小风险属性约简问题提供了一种理论方法.

关于考虑决策风险的属性约简的工作才刚刚开始,如何完善相关理论以及设计有关算法并应用将是未来的主要工作.

(2)基于 DTRS 的聚类研究

决策粗糙集模型是一个典型的概率粗糙集模型,它具有处理不确定信息和模糊信息的能力.Lingras 等人[57]提出了基于决策粗糙集模型的模糊聚类评估方法.Yu 等人[58-60]研究了基于 DTRS 的自动聚类方法.

Yu 等人在文献[58]中考虑了多种损失函数,提出了一种基于决策粗糙集的聚类模式代价评估方法.该方法结合决策粗糙集模型中的风险函数对聚类过程进行评估,以此来帮助判断哪些子类需要合并帮助找到聚类算法的终止点.同时,该方法利用三支决策思想可以处理不同粒度的交叉重叠聚类,并给出了一个自适应的粒度范围;当然,用户也可以根据需求来调整粒度大小.如何结合粒计算方法找到一个合适的粒度参数来提高算法的时间效率是未来的工作目标.

文献[59-60]在前述工作的基础上,扩展决策粗糙集模型,提出了新的聚类有效性评估函数,给出了自动地确定聚类数的解决方案,并给出了一种快速的启发式算法.Yu 等人[61]结合区间集的描述形式,提出了三支决策聚类的思想.

结合决策粗糙集研究软聚类方法是未来的一个重要研究方向.

(3)代价损失函数和决策阈值的确定

基于决策粗糙集的应用研究中,如何确定代价损失函数 λ_{ij} 是一个关键性问题,多数情况下是人工给定.

Herbert 和 Yao 等人在文献[62-66]中结合博弈论，将最优代价损失函数的确定方法与分类量度之间的博弈竞争问题联系起来，通过寻找两者代价支付矩阵中的博弈平衡点来达到优化各决策域大小的目标. 结合博弈论来研究代价损失函数和决策阈值的确定方法为概率阈值的计算方法提供了一条新的途径.

贾修一等人在文献[67-68]中通过研究三支决策粗糙集模型中的风险损失和建立模型需要的阈值参数之间的关系，提出了一个最优化问题，说明解决该优化问题即可求得所需参数，并给出了一种自适应求阈值参数的方法. 利用学习到的阈值建立的三支决策粗糙集模型能够取得更好的分类性能. Deng 等人[69]结合信息熵刻画三个域的不确定性，从而寻找具有最优的一对(α,β)值，提供了一种自动计算α,β阈值的方法. Liu 等人[70]结合回归分析和决策粗糙集提出了一种新的分类方法，利用回归分析方法计算出 DTRS 中的阈值. Liang 等人[71]提出了三角模糊决策粗糙集 TFDTRS(Triangular Fuzzy Decision-Theoretic Rough Sets)，给出了一种确定决策粗糙集模型中风险函数的方法.

如何寻求新的更有效的确定 λ_{ij} 函数或者 α,β 的方法、或者设计高效的启发式算法是未来的一个研究方向.

5.3　其他方面的研究成果

现有的研究工作大都基于构造性方法(constructive approaches)，我们也可以基于公理性方法(axiomatic approaches)来研究概率粗糙集，Li 和 Yang[72]讨论了概率粗糙集模型中上下近似运算的性质. Qian 等人[73]在 DTRS 的基础上结合粒计算研究了多粒度决策粗糙集，并给出了乐观性多粒度决策粗糙集和悲观性多粒度决策粗糙集模型，讨论了多粒度决策粗糙集模型与其他多粒度粗糙集之间的关系.

接下来，本节给出一些具体的基于决策粗糙集模型的应用研究例子.

(1) 医疗网络支持系统

Yao 等人[65,74]构建了基于 DTRS 的医疗网络支持系统(Web-based medical Decision Support Systems，WDMSS)模型. 在这个模型中，结合博弈粗糙集对诊断中的不确定性进行了讨论，并增加了用于风险分析的决策模块. 通过对患者的症状数据与诊断错误代价的整体评估，从三支决策的角度建议患者基于最小风险进行决策，即立即治疗、或者免于治疗或者进一步评估诊断.

(2) 电子邮件信息过滤系统

Zhao 等人[75]采用决策粗糙集方法提出了一种新的电子邮件信息过滤系统. 该系统采用决策粗糙集的正域、负域和边界域分别描述用户感兴趣的邮件信息、不感兴趣的邮件信息与中性邮件信息，评估函数定义为邮件误判的代价函数对错误分类的风险代价，并依据最小化风险原则给出最优邮件分类规则，实验表明，该系统与传统的二支邮件过滤方法相比具有更优越的性能. 此外，Zhou 等人[76]和 Jia 等人[77]分别从决策粗糙集三支决策的思想对邮件信息过滤系统做了进一步的分析与拓展.

(3) 基于 DTRS 模型的文本分类方法

Li 和 Miao 等人在文献[78]中提出了一种基于决策粗糙集理论并以实例为中心的文本层次分类模型. 为了获得所有可能的分类路径并减少错误在层次分类路径中的传递，构建了层次决策粗糙集模型，其包括两个关键阶段，即层次粗糙代价/收益决策模型和分类路由算法. 给出了一般性损失函数定义以确定层次决策粗糙集模型中的风险函数参数值，通过该定义可较好地度量一个实例分配到一个子类中的损益. 最后基于 SVM 提出了一种新的分类路径选择方法，从而保持分类过程中选择最优分类路径. 实验结果分析表明所提模型优于平面分类和标准层次分类模型.

如何将基于 DTRS 的文本分类思想用于其他多类分类问题、如何提高分类精度、设计合理的一般性分类路由算法是未来研究方向之一.

(4) 石油开采、政策决策等

Liu 等人在文献[43]中研究了 DTRS 方法在石油开采决策问题中的应用. 该文将一个决策能获得的期望总收益视为该决策的收益函数 θ 与成本函数 φ 之差，依次计算各决策的期望总收益，并选择期望总收益最高的决策为最终决策. 文献[43]采用该方法对 8 类不同石油矿产资源收益与成本数据进行分析，根据这 8 类资源属于富矿的概率得到了相应最优的开采策略.

随着政策环境的风云变幻、科学技术的快速发展以及人们不断增长的需求，在政策制定过程中我们应该推行风险管理，当政府在做政策决策时，风险可通过所发生的损失或者成本来定量. 因此，Liu 等人在文献[79]中提出了基于决策粗糙集的三支决策政策分析方法.

如何将基于决策粗糙集的三支决策思想应用于

更多的决策领域将是未来的工作.

(5) E-Learning 学习系统

Ayad 和 Liu[80]结合贝叶斯研究了概率粗糙集在 E-Learning 学习系统中的应用. 为了处理多类分类问题,该系统首先提出了一种改进的贝叶斯粗糙集模型;然后,针对学生信息表数据中各门学科中期成绩、学习习惯以及与最终成绩的关系,运用改进的 Bayes 粗糙集方法进行数据分析,从而找出影响最终成绩好坏的主要因素. 这种基于扩展概率粗糙集模型的方法为提高 E-Learning 系统的学习效果提供了较好的决策支持.

6 展望与总结

作为一种有效的数据分析方法,粗糙集在智能信息处理等研究领域发挥了非常重要的作用. 通过考虑容错能力和三支决策的决策风险,将贝叶斯决策过程引入到粗糙集中,从而形成了决策粗糙集. 决策粗糙集是经典 Pawlak 粗糙集理论的概率拓展模型,是粗糙集理论的重要组成部分. 自从 20 世纪 90 年代提出以来,该模型逐渐得到研究人员的重视,并在不确定性信息处理方面得到广泛应用. 本文综述分析了国内外关于决策粗糙集模型的研究内容. 这些分析将为决策粗糙集将来的研究工作提供理论与应用借鉴.

总结全文,我们认为基于决策粗糙集的研究中还有以下几方面的关键问题:

(1) 决策粗糙集理论模型上的扩展工作. 进一步完善决策粗糙集模型的理论模型,结合其他不确定性信息处理方法进行理论上的拓展,为应用研究奠定基础. 例如,研究者可以结合模糊集、证据理论、粒计算、形式概念分析、知识空间等理论来研究决策粗糙集模型.

(2) 基于决策粗糙集模型的属性约简研究以及应用. 属性约简有着非常广泛的应用价值,比如进行特征提取、数据降维等. 研究属性约简非常有意义,一方面,我们可以完善不确定性情况下属性约简相关理论;另一方面,我们也可以进行相关的算法研究.

(3) 如何解决决策粗糙集模型中的阈值设置问题. 结合数据驱动的自主式学习方法、博弈论、信息论、贝叶斯理论等都可能成为解决此问题的方法.

(4) 决策粗糙集模型与其他数据挖掘方法的结合问题. 比如,已有一些基于决策粗糙集进行聚类和分类研究的成果,未来也可以考虑和数据流挖掘相结合.

(5) 决策规则的获取问题. 从三支决策的角度对已有的二支决策结果进行了扩展,但如何高效地获取三支决策规则以及其中存在的问题还有待进一步地探索.

(6) 如何面对大数据的挑战. 大数据具有数据量大、数据与时俱增、数据类型繁多、数据结构复杂、要求处理迅速等特点. 因此,如何结合决策粗糙集来应对大数据时代到来的挑战是学者们应该思考的问题. 是否可以提出一种结合概率论的决策粗糙集模型的不完备性处理方法;考虑结合代价损失函数解决非结构化数据转化为合理结构化数据的模型问题;如何设计高效算法应对快速变化的动态流数据. 这些都是未来需要解决的问题.

(7) 开展三支决策研究. 为了给粗糙集的 3 个域更合理的语义解释,学者们提出了三支决策;但随着研究的进一步深入,我们发现粗糙集可以看作是三支决策的一个特例. 因此,深入开展三支决策研究将非常有意义. 例如,我们在获得边界域后,如何使用边界域还有待进一步研究.

(8) 扩大其应用领域. 如何将决策粗糙集模型应用到具有不确定性的应用领域是未来的一个研究方向. 比如,我们可以考虑将其应用到网络支持系统、社交网络服务、Web 服务等具有不确定性特征的应用中.

参 考 文 献

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

[2] Chen Yuan-Yuan, Zhang Ji-Long, Li Xiao, et al. Research on concentration of multi-component pollution gas based on SVM with kernel optimized by rough set. Spectroscopy and Spectral Analysis, 2010, 30(12): 3384-3387(in Chinese)
(陈媛媛, 张记龙, 李晓等. 基于粗糙集核优化的支持向量机在多组分污染气体定量分析中的研究与应用. 光谱学与光谱分析, 2010, 30(12): 3384-3387)

[3] Li Xiao-Yong, Gui Xiao-Lin, Mao Qian, Leng Dong-Qi. Adaptive dynamic trust measurement and prediction model based on behavior monitoring. Chinese Journal of Computers, 2009, 32(4): 664-674(in Chinese)
(李小勇, 桂小林, 毛倩, 冷东起. 基于行为监控的自适应动态信任度测模型. 计算机学报, 2009, 32(4): 664-674)

[4] Sakai H, Chakraborty M K, Hassanien A E, et al. Rough sets, fuzzy sets, data mining and granular computing// Proceedings of the 12th International Conference (RSFDGrC 2009). Berlin: Springer, 2009

[5] Wang Guo-Yin, Yao Yi-Yu, Yu Hong. A survey on rough set theory and applications. Chinese Journal of Computers, 2009, 32(7): 1229-1246(in Chinese)
(王国胤，姚一豫，于洪. 粗糙集理论与应用研究综述. 计算机学报，2009，32(7)：1229-1246)

[6] Yu J, Greco S, Lingras P, et al. Rough sets and knowledge technology//Proceedings of the 5th International Conference (RSKT 2010). Berlin: Springer, 2010

[7] Qian Jin, Miao Duo-Qian, Zhang Ze-Hua. Knowledge reduction algorithms in cloud computing. Chinese Journal of Computers, 2011, 34(12): 2332-2343(in Chinese)
(钱进，苗夺谦，张泽华. 云计算环境下知识约简算法. 计算机学报，2011，34(12)：2332-2343)

[8] Yao J T, Ramanna S, Wang G Y, Suraj Z. Rough sets and knowledge technology//Proceedings of the 6th International Conference(RSKT 2011). Berlin: Springer, 2011

[9] Li T R, Nguyen H S, Wang G Y, et al. Rough sets and knowledge technology//Proceedings of the 7th International Conference(RSKT 2012). Berlin: Springer, 2012

[10] Yao J T, Yang Y, Slowinski R, et al. Rough Sets and Current Trends in Computing//Proceedings of the 8th International Conference(RSCTC 2012). Berlin: Springer, 2012

[11] An Jian, Gui Xiao-Lin, Zhang Wen-Dong, et al. Social relation cognitive model of mobile nodes in the Internet of Things. Chinese Journal of Computers, 2012, 35(6): 1164-1174(in Chinese)
(安健，桂小林，张文东等. 物联网移动感知中的社会关系认知模型. 计算机学报，2012，35(6)：1164-1174)

[12] Yao Y Y, Wong S K M, Lingras P. A decision-theoretic rough set model//Ras Z W, Zernankova M, Emrich M L eds. Methodologies for Intelligent Systems, 5. New York: North-Holland, 1990: 17-24

[13] Pawlak Z, Wong S K M, Ziarko W. Rough sets: Probabilistic versus deterministic approach. International Journal of Man-Machine Studies, 1988, 29(1): 81-95

[14] Ziarko W. Variable precision rough set model. Journal of Computer and System Sciences, 1993, 46(1): 39-59

[15] Pawlak Z, Skowron A. Rough membership functions//Yager R R, Fedrizzi M, Kacprzyk J eds. Advances in the Dempster-Shafer Theory of Evidence. New York: Wiley, 1994: 251-271

[16] Pawlak Z, Skowron A. Rough sets: Some extensions. Information Sciences, 2007, 177(1): 28-40

[17] Skowron A, Stepaniuk J. Tolerance approximation spaces. Fundamenta Informaticae, 1996, 27(2-3): 245-253

[18] Ślęzak D. Rough sets and bayes factor//Peters J F, Skowron A eds. Transactions on Rough Sets Ⅲ. Berlin: Springer, 2005: 202-229

[19] Herbert J P, Yao J T. Game-theoretic risk analysis in decision-theoretic rough sets//Proceedings of the 3rd International Conference on Rough Sets and Knowledge Technology. Chengdu, China, 2008: 132-139

[20] Wei W, Liang J Y, Qian Y H. A comparative study of rough sets for hybrid data. Information Sciences, 2012, 190: 1-16

[21] Hu Q H, Yu D R, Guo M Z. Fuzzy preference based rough sets. Information Sciences, 2010, 180(10): 2003-2022

[22] Qian Yu-Hua, Liang Ji-Ye, Wang Feng. A positive-approximation based accelerated algorithm to feature selection from incomplete decision tables. Chinese Journal of Computers, 2011, 34(3): 435-442(in Chinese)
(钱宇华，梁吉业，王锋. 面向非完备决策表的正向近似特征选择加速算法. 计算机学报，2011，34(3)：435-442)

[23] Chen Hao, Yang Jun-An, Zhuang Zhen-Quan. The core of attributes and minimal attributes reduction in variable precision rough set. Chinese Journal of Computers, 2012, 35(5): 1011-1017(in Chinese)
(陈昊，杨俊安，庄镇泉. 变精度粗糙集的属性核和最小属性约简算法. 计算机学报，2012，35(5)：1011-1017)

[24] Mi J S, Wu W Z, Zhang W X. Approaches to knowledge reduction based on variable precision rough set model. Information Sciences, 2004, 159(3-4): 255-272

[25] Xu Yi, Li Long-Shu. Variable precision rough set model based on (α,λ) connection degree tolerance relation. Acta Automatica Sinica, 2011, 37(3): 303-308(in Chinese)
(徐怡，李龙澍. 基于(α,λ)联系度容差关系的变精度粗糙集模型. 自动化学报，2011，37(3)：303-308)

[26] Ningler M, Stockmanns G, Schneider G, et al. Adapted variable precision rough set approach for EEG analysis. Artificial Intelligence in Medicine, 2009, 47(3): 239-261

[27] Ma W, Sun B. Probabilistic rough set over two universes and rough entropy. International Journal of Approximate Reasoning, 2012, 53(4): 608-619

[28] Yu Ying, Maio Duo-Qian, Liu Cai-Hui, Wang Lei. An improved KNN algorithm based on variable precision rough sets. Pattern Recognition and Artificial Intelligence, 2012, 25(4): 617-623(in Chinese)
(余鹰，苗夺谦，刘财辉，王磊. 基于变精度粗糙集的 KNN 分类改进算法. 模式识别与人工智能，2012，25(4)：617-623)

[29] Yanto I T R, Vitasari P, Herawan T, Deris M M. Applying variable precision rough set model for clustering student suffering study's anxiety. Expert Systems with Applications, 2012, 39(1): 452-459

[30] Wang G Y. Rough set based uncertainty knowledge expressing and processing//Proceedings of the Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. Moscow, Russia, 2011: 11-18

[31] Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning About Data. Boston: Kluwer Academic Publishers Press, 1991

[32] Wong S K M, Ziarko W. Comparison of the probabilistic approximate classification and the fuzzy set model. Fuzzy Sets and Systems, 1987, 21(3): 357-362

[33] Wong S K M, Ziarko W. A Probabilistic model of approximate classification and decision rules with uncertainty in inductive learning. Regina: Department of Computer Science, University of Regina, Technical Report CS-85-23, 1985

[34] Li Hua-Xiong, Zhou Xian-Zhong, Li Tian-Rui, et al. Decision-Theoretic Rough Sets Theory and Its Applications. Beijing: Science Press, 2011(in Chinese)
(李华雄，周献中，李天瑞等. 决策粗糙集理论及其研究进展. 北京：科学出版社，2011)

[35] Yao Y Y, Zhou B. Naive Bayesian rough sets//Proceedings of the 5th International Conference on Rough Sets and Knowledge Technology. Beijing, China, 2010: 719-726

[36] Yao Y Y. The superiority of three-way decisions in probabilistic rough set models. Information Sciences, 2011, 181(6): 1080-1096

[37] Yao Y Y. Three-way decisions with probabilistic rough sets. Information Sciences, 2010, 180(3): 341-353

[38] Yao Y Y. An outline of a theory of three-way decisions// Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing. Chengdu, China, 2012: 1-17

[39] Liu Dun, Li Tian-Rui, Miao Duo-Qian, et al. Three-Way Decisions and Granular Computing. Beijing: Science Press, 2013(in Chinese)
(刘盾, 李天瑞, 苗夺谦等(编著). 三支决策与粒计算. 北京: 科学出版社, 2013)

[40] Jia Xiu-Yi, Shang Lin, Zhou Xian-Zhong, et al. Three-Way Decisions Theory and Its Applications. Nanjing: Nanjing University Press, 2012(in Chinese)
(贾修一, 商琳, 周献中等. 三支决策理论与应用. 南京: 南京大学出版社, 2012)

[41] Li H X, Zhou X Z. Risk decision making based on decision-theoretic rough set: A three-way view decision model. International Journal of Computational Intelligence Systems, 2011, 4(1): 1-11

[42] Zhou X Z, Li H X. A multi-view decision model based on decision-theoretic rough set//Proceedings of the 4th International Conference on Rough Sets and Knowledge Technology. Gold Coast, Australia, 2009: 650-657

[43] Liu D, Yao Y Y, Li T R. Three-way investment decisions with decision-theoretic rough sets. International Journal of Computational Intelligence Systems, 2011, 4(1): 66-74

[44] Liu D, Li T R, Li H X. A multiple-category classification approach with decision-theoretic rough sets. Fundamenta Informaticae, 2012, 115(2-3): 173-188

[45] Liu D, Li T R, Liang D C. A new discriminant analysis approach under decision-theoretic rough sets//Proceedings of the 6th International Conferences on Rough Sets and Knowledge Technology. Banff, Canada, 2011: 473-482

[46] Yang X P, Yao J T. Modelling multi-agent three-way decisions with decision theoretic rough sets. Fundamenta Informaticae, 2012, 115(2-3): 157-171

[47] Ma W M, Sun B Z. On relationship between probabilistic rough set and Bayesian risk decision over two universes. International Journal of General Systems, 2012, 41(3): 225-245

[48] Ma W M, Sun B Z. Probabilistic rough set over two universes and rough entropy. International Journal of Approximate Reasoning, 2012, 53(4): 608-619

[49] Yao Y Y, Zhao Y. Attribute reduction in decision theoretic rough set models. Information Sciences, 2008, 178(17): 3356-3373

[50] Li H X, Zhou X Z, Zhao J B, Liu D. Attribute reduction in decision-theoretic rough set model: A further investigation// Proceedings of the 6th International Conference on Rough Sets and Knowledge Technology. Banff, Canada, 2011: 466-475

[51] Ma X A, Wang G Y, Yu H, Li Tian-Rui. Decision region distribution preservation reduction in decision-theoretic rough set model. Information Sciences, 2014, 278: 614-640

[52] Chebrolu S, Sanjeevi S G. Attribute reduction in decision-theoretic rough set models using genetic algorithm//Proceedings of the 2nd International Conference on the Swarm, Evolutionary, and Memetic Computing. Visakhapatnam, India, 2011: 307-314

[53] Liu J B, Min F, Liao S J, Zhu W. Minimal test cost feature selection with positive region constraint//Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing. Chengdu, China, 2012: 259-266

[54] Min F, Hu Q H, Zhu W. Feature selection with test cost constraint. International Journal of Approximate Reasoning, 2014, 55(1): 167-179

[55] Jia X Y, Liao W H, Tang Z M, Shang L. Minimum cost attribute reduction in decision-theoretic rough set models. Information Sciences, 2013(219): 151-167

[56] Jia Xiu-Yi, Shang Lin, Chen Jia-Jun. Attribute reduction based on three-way decision//Chinese Association for Artificial Intelligence: Progress of Artificial Intelligence in China (2009). Beijing: Beijing University of Posts and Telecommunicaions Press, 2009: 193-198(in Chinese)
(贾修一, 商琳, 陈家骏. 基于三值决策的属性约简//中国人工智能学会: 中国人工智能进展(2009). 北京: 北京邮电大学出版社, 2009: 193-198)

[57] Lingras P, Chen M, Miao D. Rough cluster quality index based on decision theory. IEEE Transactions on Knowledge and Data Engineering, 2009, 21(7): 1014-1026

[58] Yu H, Chu S S, Yang D C. Autonomous knowledge-oriented clustering using decision-theoretic rough set theory. Fundamenta Informaticae, 2012, 115(2-3): 141-156

[59] Yu H, Liu Z G, Wang G Y. Automatically determining the number of clusters using decision-theoretic rough set// Proceedings of the 6th International Conference on Rough Sets and Knowledge Technology. Banff, Canada, 2011: 504-513

[60] Yu H, Liu Z G, Wang G Y. An automatic method to determine the number of clusters using decision-theoretic rough set. International Journal of Approximate Reasoning, 2014, 55(1): 101-115

[61] Yu H, Wang Y, Jiao P. A three-way decisions approach to density-based overlapping clustering//Peters J F, et al, eds. Transactions on Rough Sets XVIII. LNCS 8449. Berlin: Springer, 2014: 92-109

[62] Herbert J P, Yao J T. Game-theoretic risk analysis in decision- theoretic rough sets//Proceedings of the 3rd International Conference on Rough Sets and Knowledge Technology. Chengdu, China, 2008: 132-139

[63] Herbert J P, Yao J T. Game-theoretic rough sets. Fundamenta Informaticae, 2011, 108(3-4): 267-286

[64] Herbert J P, Yao J T. Analysis of data-driven parameters in game-theoretic rough sets//Proceedings of the 6th International Conference on Rough Sets and Knowledge Technology. Banff, Canada, 2011: 447-456

[65] Zhang Y, Yao J T. Determining three-way decision regions with gini coefficients//Proceedings of the 9th International Conference on Rough Sets and Current Trends in Computing, Granada and Madrid. Spain, 2014: 160-171

[66] Azam N, Yao J T. Analyzing uncertainties of probabilistic rough set regions with game-theoretic rough sets. International Journal of Approximate Reasoning, 2014, 55(1): 142-155

[67] Jia Xiu-Yi, Li Wei-Wei, Shang Lin, Chen Jia-Jun. An adaptive learning parameters algorithm in three-way decision-theoretic rough set model. Acta Electronica Sinica, 2011, 39(11): 2520-2525(in Chinese)
(贾修一，李伟清，商琳，陈家骏. 一种自适应求三枝决策中决策阈值的算法. 电子学报，2011, 39(11): 2520-2525)

[68] Jia X Y, Tang Z M, Liao W H, Shang L. On an optimization representation of decision-theoretic rough set model. International Journal of Approximate Reasoning, 2014, 55(1): 156-166

[69] Deng X F, Yao Y Y. A multifaceted analysis of probabilistic three-way decisions. Fundamenta Informaticae, 2014, 132(3): 291-313

[70] Liu D, Li T R, Liang D C. Incorporating logistic regression to decision-theoretic rough sets for classifications. International Journal of Approximate Reasoning, 2014, 55(1): 197-210

[71] Liang D C, Liu D, Pedrycz W, Hu P. Triangular fuzzy decision- theoretic rough sets. Journal of Approximate Reasoning, 2013, 54(8): 1087-1106

[72] Li T J, Yang X P. An axiomatic characterization of probabilistic rough sets. International Journal of Approximate Reasoning, 2014, 55(1): 130-141

[73] Qian Y H, Zhang H, Sang Y L, Liang J Y. Multigranulation decision-theoretic rough sets. International Journal of Approximate Reasoning, 2014, 55(1): 225-237

[74] Yao J T, Herbert J P. Web-based support systems with rough set analysis//Proceedings of the International Conference on Rough Sets and Intelligent Systems Paradigms. Warsaw, Poland, 2007: 360-370

[75] Zhao W Q, Zhu Y L, Gao W. Information filtering model based on decision-theoretic rough set theory. Computer Engineering and Applications, 2007, 43(7): 185-187

[76] Zhou B, Yao Y Y, Luo J G. Cost-sensitive three-way email spam filtering. Journal of Intelligent Information Systems, 2014, 42(1): 19-45

[77] Jia X Y, Zheng K, Li W W, et al. Three-way decisions solution to filter spam email: An empirical study//Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing. Chengdu, China, 2012: 287-296

[78] Li W, Miao D Q, Wang W, Zhang N. Hierarchical rough decision theoretic framework for text classification//Proceedings of the 9th IEEE International Conference on Cognitive Informatics. Shanghai, China, 2010: 484-489

[79] Liu D, Li T R, Liang D C. Three-way government decision analysis with decision-theoretic rough sets. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2012, 20(1): 119-132

[80] Ayad R A, Liu J. Supporting E-learning system with modified Bayesian rough set model//Yu W, He H B, Zhang N eds. Advances in Neural Networks-ISNN 2009. Berlin: Springer, 2009: 192-200

YU Hong, born in 1972, Ph. D., professor. Her research interests include three-way decisions, three-way clustering, rough sets, interval sets, intelligence information processing, Web intelligence, data mining etc.

WANG Guo-Yin, born in 1970, Ph. D., professor. His research interests include rough sets, granular computing, machine learning, data mining, knowledge technology, cognitive computing etc.

YAO Yi-Yu, born in 1962, Ph. D., professor. His research interests include three-way decisions, rough sets, interval sets, granular computing, information retrieval, Web intelligence, data mining etc.

Background

The advances of rough set theory significantly enhance the capabilities for data analysis, and the theory has been applied successfully in many fields such as decision analysis, machine learning, data mining, knowledge discovery, pattern recognition, etc. Because the lower and upper approximations are defined by a pair of definable sets with respect to Pawlak's rough set theory, the rules generated don't allow any tolerance of errors. For this purpose, probabilistic rough set models are generalized to overcome the weakness. Yao and colleagues proposed the Decision-Theoretic Rough Sets (DTRS) model in 1990s' by combing the Bayesian decision theory. In recent years, the model received much attentions and produced applications in uncertain information processing.

In this paper, we explain motivations of the decision-theoretic rough sets model and discuss the features and challenges of the model. We provide a review of the main theoretical developments and applications of the DTRS theory. Finally, we point out the existing challenge problems, and explore the research directions in future.

This work is partially supported by the National Natural Science Foundation of China (Nos. 61379114, 61272060) and the Natural Science Foundation of Chongqing of China (No. cstc2013jjB40003).