知识的粗识别及其评判*)

管延勇[1,2] 王洪凯[1] 史开泉[1]
(山东大学数学与系统科学院 济南250100)[1] (济南大学理学院 济南250022)[2]

摘 要 讨论了在近似空间中知识的粗识别,对不可分辨关系 R 对范畴 X 的粗识别能力进行了量化,给出了相应的度量评判标准,提出了知识的多重粗识别及知识再挖掘的概念,对多个不可分辨关系对论域的共同作用进行了讨论。

关键词 粗糙集,知识颗粒,正域,负域,粗识别

Rough Recognition of Knowledge and its Assessment

GUAN Yan-Yong[1,2] WANG Hong-Kai[1] SHI Kai-Quan[1]
(School of Mathematics and System Science, Shandong University, Jinan 250100)[1] (School of Science, Jinan University, Jinan 250022)[2]

Abstract Discuss the rough recognition of knowledge in approximation space (U,R). Introduce the recognition of indiscernible relation R to the objects of U by using the knowledge granule and positive region and negative region of the category X in order to discover which objects belong to X completely and which objects do not belong to X completely, propose the concept of multiple recognition for knowledge and the concept of knowledge re-mining to discuss the recognition of more indiscernible relations on U, and quantify the capability of all kinds of recognitions of indiscernible relations.

Keywords Rough sets, Knowledge granule, Positive region, Negative region, Rough recognition

1 引言

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,它在知识发现、数据挖掘及模式识别等方面取得了广泛的应用。

粗糙集理论处理问题的基本出发点是对研究对象进行分类,它把对研究对象的分类视为对研究对象的一种认识,即知识。

设有一个有限论域 U,利用有关属性,可将 U 中元素进行分类,得到 U 的一种划分。根据集合论的思想,集合 U 的一种划分与 U 上的一个等价关系是相互唯一确定的,故 U 上元素的属性,U 的划分及 U 上的等价关系是相互对应的。在下面的行文中,为了方便,我们交叉使用这三个概念而不加区分。

设 R 是有限论域 U 上的一个等价关系,R 将 U 中的元素分成有限个等价类。对某个等价类 $[x]_R$ 中的元素来说,它们之间是 R 不可分辨的。例如,设 U 是一些积木的集合,R 对应的属性是"颜色",$[x]_R$ 表示"红色积木"形成的等价类,则 $[x]_R$ 中的元素关于"颜色"这个属性是无法加以分辨的。由此可见,R 对 U 的分类体现出一种"知识"—"颜色"。

定义1[1] 称序对 (U,R) 是一个知识库,其中 U 是一有限论域,R 是 U 上的一个等价关系。任一等价类 $[x]_R$ 称为一个知识颗粒,也称为原子知识或初等知识。几个等价类的并集称为基本知识;U 的任一子集 X 称为 U 的一个范畴。有时为了方便,也称 R 是 U 上的知识。

设 $U=\{x_1,x_2,\cdots,x_{10}\}$ 是10个积木的集合,按属性"颜色"分类为:

$[x_1]_R=\{x_1,x_3\}$ 红色积木

$[x_2]_R=\{x_2,x_4,x_7\}$ 黄色积木

$[x_5]_R=\{x_5,x_{10}\}$ 黑色积木

$[x_6]_R=\{x_6,x_8,x_9\}$ 蓝色积木

这四个等价类体现了 R 对 U 的认识,是 R 认识 U 的基本单位。

对 U 中的某些范畴,R 对其认识是明确的,如 $[x_1]_R=\{x_1,x_3\}\subseteq U$,是初等知识"红色积木";$X_1=\{x_1,x_3,x_5,x_{10}\}=[x_1]_R\cup[x_5]_R\subseteq U$ 是基本知识"红色或黑色积木";而对范畴 $X_2=\{x_1,x_2,x_3,x_5,x_6,x_{10}\}$,$R$ 如何认识它呢?显然 X_2 不是知识颗粒的并集,X_2 含着知识颗粒 $[X_1]_R$ 与 $[x_5]_R$,但仅含着知识颗粒 $[x_2]_R$ 与 $[x_6]_R$ 的一部分。由于知识颗粒是 R 认识 U 的最小单位,故 R 对 X_2 的认识是不清楚的,是模糊的。Pawlak 粗集理论对像 X_2 这样的范畴的认识,是利用 X_2 的所谓上近似和下近似来进行的。

定义2[2,3] 设 (U,R) 是一个知识库,$X\subseteq U$,令

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

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

分别称 $\underline{R}(X)$、$\overline{R}(X)$ 为 X 的下近似和上近似。亦称 (U,R) 为近似空间,记为 $apr(U,R)$。

显然 $\underline{R}(X)$ 与 $\overline{R}(X)$ 是 (U,R) 中的基本知识且 $\underline{R}(X)\subseteq X\subseteq\overline{R}(X)$。若 $\underline{R}(X)=\overline{R}(X)(=X)$,则 X 是基本知识,也称 X 为可定义集;若 $\underline{R}(X)\neq\overline{R}(X)$,则称 X 是不可定义集,也称 X 为粗糙集。不可定义集可由其上下近似 $\overline{R}(X)$ 与 $\underline{R}(X)$ 来界定。

定义3[2,4] 近似空间 $apr(U,R)$ 中,对 $X\subseteq U$,$\underline{R}(X)$ 是据知识 R 判断必在 X 中的知识颗粒的并集,称为 X 的 R 正域,亦记为 X^{+R};$U-\overline{R}(X)$ 是据知识 R 判断必不在 X 中的知识

*)国家自然科学基金资助项目(70271048)。**管延勇** 副教授,博士生,主要研究领域为粗糙集理论及其应用、模糊集理论及其应用。**王洪凯** 博士生,主要研究领域为粗糙集理论及其应用、模糊集理论及其应用。**史开泉** 教授,博士生导师,主要研究领域为粗糙集理论及其应用、模糊集理论及其应用。

颗粒的并集,称为 X 的 R 负域,亦记为 X^{-R};$\overline{R}(X)-\underline{R}(X)$ 是据 R 判断不完全在 X 中也不完全在 X 中的知识颗粒的并集,称为 X 的边界,记为 $Bn_R(X)$。

定义4[5] $\underline{\Delta}_R(X)=X-\underline{R}X$ 称为 X 的下边界;$\overline{\Delta}_R(X)=\overline{R}X-X$ 称为 X 的上边界。

显然有,$\underline{\Delta}_R(X)=B_{n_R}(X)\cap X$,$\overline{\Delta}_R(X)=B_{n_R}(X)\cap(U-X)$,$\underline{\Delta}_R(X)\cup\overline{\Delta}_R(X)=B_{n_R}(X)$;$\underline{\Delta}_R(X)=\phi\Leftrightarrow\overline{\Delta}_R(X)=\phi\Leftrightarrow\overline{R}(X)=\underline{R}(X)$。

2 正域、负域的性质

令 $Def(U,R)$ 表示近似空间 $apr(U,R)$ 中的所有可定义集(基本知识)的集合。

(1)$\forall\ X\subseteq U$,X^{+R},X^{-R} 由 X 唯一确定;反之不真,即有:

(2)$\forall\ A,B\in Def(U,R)$,$A\cap B=\phi$,存在 $X\subseteq U$ 使 $X^{+R}=A$ 且 $X^{-R}=B$,但这样的 X 不唯一。事实上,令

$$Def(U,R)\otimes Def(U,R)=\{(A,B)\mid A,B\in Def(U,R),A\cap B=\phi\}$$

作映射

$$f:\mathscr{P}(U)\to Def(U,R)\otimes Def(U,R)$$
$$X\to(X^{+R},X^{-R})$$

则 f 是满射但不是单射,且易证:

$$\forall\ (A,B)\in Def(U,R)\otimes Def(U,R)$$
$$f^{-1}(A,B)=\{X\mid A\subseteq X\subseteq U\text{ 且}\forall\ x\in U-A-B,\text{有 }\phi\neq[x]_R\cap X\neq[x]_R\}$$

即 U 中有多个范畴 X 满足 $(X^{+R},X^{-R})=(A,B)$。

(3)若 $X\subseteq Y$,则 $X^{+R}\subseteq Y^{+R}$,$X^{-R}\supseteq Y^{-R}$(见文[1]);反之不真,即有:

(4)对 $X,Y\subseteq U$,即使有

$X^{+R}\subseteq Y^{+R}$,$X^{-R}\subseteq Y^{-R}$,也不一定有 $X\subseteq Y$。

(5)对 $X,Y\subseteq U$,$(X\cap Y)^{+R}=X^{+R}\cap Y^{+R}$,$(X\cap Y)^{-R}=X^{-R}\cup Y^{-R}$。

由文[1]中结论及德摩根律易证之。

(6)对 $X,Y\subseteq U$,$(X\cup Y)^{+R}\supseteq X^{+R}\cup Y^{+R}$,$(X\cup Y)^{-R}=X^{-R}\cap Y^{-R}$,见文[1]。

(7)$(\sim X)^{+R}=X^{-R}$,$(\sim X)^{-R}=X^{+R}$,见文[1]。

3 知识的粗识别

定义5 设 (U,R) 为一近似空间,$X\subseteq Y$。称序对 (X^{+R},X^{-R}) 为知识 R 关于范畴 X 对论域 U 的 R-粗识别。

R-粗识别以知识颗粒为认识的最小单位,明确表示出 U 中元素哪些必在 X 中,哪些必不在 X 中。显然,若 $X^{+R}\cup X^{-R}=U$,则 X 是可定义集,U 中所有元素与 X 的关系可 R 由完全识别;若 $X^{+R}\cup X^{-R}\neq U$,则 $B_{n_R}(X)\neq\phi$,说明 U 中有元素不能被知识 R 依知识颗粒识别是否在 X 中。

在近似空间 $apr(U,R)$ 中,知识 R 对论域 U 的认识程度取决于它对 U 的分类能力。等价关系 R(或理解为属性集)将 U 划分得越细,则每一等价类中的元素越少,其所具有的属性越多,从而说明 R 对 U 的认识越深刻。下面给出刻画知识 R 对 U 的识别能力的几个概念。

定义6 设 (U,R) 为一近似空间,$X\subseteq U$。称 $Drec_R^+(X)=\dfrac{|X^{+R}|}{|X|}$ 为知识 R 对 X 的正识别度;称 $Drec_R^-(X)=\dfrac{|X^{-R}|}{|U-X|}$ 为知识 R 对 X 的负识别度;称 $Drec_R(X)=\dfrac{|X^{+R}\cup X^{-R}|}{|U|}$ 为知识 R 关于范畴 X 对 U 的 R-识别度。

定义7 若近似空间 (U,R) 中的范畴 X^* 满足:$\phi\neq X^*\neq U$,且 $\forall\ x\in U$,$[x]_R\cap X^*\neq\phi$,且 $[x]_R\not\subset X^*$,则称 X^* 为 R-泛范畴。

例如,设 R 是 U 上的任一个等价关系,其任意等价类的基数大于1。我们从其每一等价类中仅取一部分而不是全部元素,构成一个集合 X,则 X 即是一个泛范畴。显然,若近似空间 (U,R) 中存在泛范畴,则它不存在单元素知识颗粒。

定理1 设 (U,R) 为一近似空间,$X\subseteq U$,则有:

(1)$0\leqslant Drec_R^+(X)\leqslant 1$,$0\leqslant Drec_R^-(X)\leqslant 1$,$0\leqslant Drec_R(X)\leqslant 1$。

(2)$Drec_R^+(x)=1\Leftrightarrow Drec_R^-(X)=1\Leftrightarrow Drec_R(X)=1\Leftrightarrow X\in Def(U,R)$。

(3)$Drec_R^+(X)=0\Leftrightarrow\forall\ x\in U,[x]_R\not\subset X$;$Drec_R^-(X)=0\Leftrightarrow\forall\ x\in U,[x]_R\cap X\neq\phi$;$Drec_R(X)=0\Leftrightarrow Drec_R^+(X)=Drec_R^-(X)=0\Leftrightarrow X$ 是泛范畴。

证:(1)由 $X^{+R}\subseteq X$ 及 $X^{-R}\subseteq U-X$ 易证。

(2)$Drec_R^+(X)=1\Leftrightarrow X^{+R}=X\Leftrightarrow X\in Def(U,R)\Leftrightarrow X^{+R}\cup X^{-R}=U\Leftrightarrow Drec_R^-(X)=1\Leftrightarrow Drec_R(X)=1$

(3)①$Drec_R^+(X)=0\Leftrightarrow X^{+R}=\phi\Leftrightarrow\forall\ x\in U,[x]_R\not\subset X$

②$Drec_R^-(X)=0\Leftrightarrow X^{-R}=\phi\Leftrightarrow\forall\ x\in U,[x]_R\cap X\neq\phi$

③$Drec_R(X)=0\Leftrightarrow X^{+R}\cup X^{-R}=\phi\Leftrightarrow X^{+R}=\phi$ 且 $X^{-R}=\phi\Leftrightarrow Drec_R^+(X)=0$ 且 $Drec_R^-(X)=0\Leftrightarrow X$ 是泛范畴(由①②)。

$Drec_R^+(X)$ 表示 X 中那些由 R 依知识颗粒识别也必在 X 中的元素所占的比例;$Drec_R^-(X)$ 表示 U 中由 R 依知识颗粒识别必不在 X 中的元素在 $U-X$ 中所占的比例;$Drec_R(X)$ 表示 U 中由 R 依知识颗粒识别能够精确识别是否在 X 中的元素在 U 中所占的比例。一般地,这三种识别度既与 R 有关又与 X 有关。下面考虑几种特殊情况,识别度仅与一个因素有关。

(1)对 U 上恒等关系 I_U,$\forall\ X\subseteq U$,$Drec_{I_U}^+(X)=Drec_{I_U}^-(X)=Drec_{I_U}(X)=1$。事实上,对 I_U,U 上每一元素即为一知识颗粒,I_U 对 U 的认识最清楚,从而它对任意范畴的识别度为1。

(2)对 U 上论域关系 $R=U\times U$,$\forall\ X\neq U$,有 $Drec_{U\times U}^+(X)=Drec_{U\times U}^-=Drec_{U\times U}(X)=0$。事实上,$U$ 上只有一个关于知识 $U\times U$ 的知识颗粒,U 上元素在 $U\times U$ 之下不可分辨,从而得不到 U 的任何知识,故 $U\times U$ 关于任意范畴 $X(X\neq U)$ 识别度为0。

定义8 设 U 为一论域,U 上元素的两个属性集 P_1、P_2 诱导的不可分辨关系分别为 R_1、R_2,即 $R_1=IndP_1$,$R_2=IndP_2$。若 $R_2\subseteq R_1$,则称知识 R_2 是知识 R_1 的加深。

若 R_2 是 R_1 的加深,则 R_2 关于 U 的划分 U/R_2 是 R_1 关于 U 的划分 U/R_1 的加细,即 U 关于 R_1 的等价类是 U 关于 R_2 的等价类的并;换言之,U 关于 R_2 的等价类必包含在 U 关于 R_1 的等价类中,从而 R_2 的知识颗粒比 R_1 的知识颗粒小,因此具有较多的属性,可以认为 R_2 对 U 的认识比 R_1 深刻。由上面分析,知识 I_U 是任何一个知识 R 的加深,任何一个知识 R 是 $U\times U$ 的加深。

定理2 设 U 为有限论域,R_1 与 R_2 是 U 上的两个知识,且 R_2 是 R_1 的加深,则 $\forall\ X\subseteq U$,有

(1)$X^{+R_1}\subseteq X^{+R_2}$,$X^{-R_1}\subseteq X^{-R_2}$;

(2)$Drec_{R_1}^+(X)\leqslant Drec_{R_2}^+(X)$,$Drec_{R_1}^-(X)\leqslant Drec_{R_2}^-(X)$,$Drec_{R_1}(X)\leqslant Drec_{R_2}(X)$。

证:(1)因 R_2 是 R_1 的加深,故 $R_2\subseteq R_1$。对任意 $[x]_{R_2}$,$\forall\ y\in[x]_{R_2}$,则 yR_2x,$\langle y,x\rangle\in R_2\subseteq R_1$,所以 $y\in[x]_{R_1}$。从而证得 $\forall\ x\in U$,$[x]_{R_2}\subseteq[x]_{R_1}$。

$\forall z \in X^{+R_1}=U\{[u]_{R_1} \mid [u]_{R_1} \subseteq X\}$，$\exists u_0$ 使 $[u_0]_{R_1} \subseteq X$ 且 $z \in [u_0]_{R_1}$，从而 $[z]_{R_1}=[u_0]_{R_1} \subseteq X$，$[z]_{R_2} \subseteq [z]_{R_1} \subseteq X$，$z \in X^{+R_2}$。由 $z \in X^{+R_1}$ 的任意性，证得 $X^{+R_1} \subseteq X^{+R_2}$。

$\forall \omega \in X^{-R_1}=\bigcup\{[v]_{R_1} \mid [v]_{R_1} \cap X=\phi\}$，$\exists v_0$ 使 $\omega \in [v_0]_{R_1}$ 且 $[v_0]_{R_1} \cap X=\phi$。又 $[\omega]_{R_1}=[v_0]_{R_1}$，故 $[\omega]_{R_1} \cap X=\phi$，$[\omega]_{R_2} \cap X \subseteq [\omega]_{R_1} \cap X=\phi$，$[\omega]_{R_2} \cap X=\phi$，即 $\omega \in X^{-R_2}$。由 $\omega \in X^{-R_1}$ 的任意性知 $X^{-R_1} \subseteq X^{-R_2}$。

(2)由(1)及定义6易得。

4 知识的多重粗识别

实际生活中，不同的人对同一事物往往有不同的认识，不管是根据相同的属性还是不同的属性来判断。设有 n 个不同的专家，他们对论域 U 的知识分别为 $R_1,R_2,\cdots,R_n$。现要请这 n 位专家对 U 上的范畴 X 进行粗识别，有以下两个方法：

(1)n 个专家首先经过磋商，后达成共识，得到知识 $\bigcap_{i=1}^{n} R_i$，利用知识 $\bigcap_{i=1}^{n} R_i$ 对 X 进行识别。这个方法类似于医疗中的专家会诊。

(2)n 个专家先各自对 X 进行粗识别，分别得到 X 的粗识别 (X^{+R_1},X^{-R_1})，(X^{+R_2},X^{-R_2})，$\cdots$，$(X^{+R_n},X^{-R_n}$，然后再取共识，得到 $(\bigcap_{i=1}^{n} X^{+R_i},\bigcap_{i=1}^{n} X^{-R_i})$ 作为他们对 X 的粗识别。这个方法类似于体操比赛中的打分法。

这两个方法的效果显然不同。下面给出相应的概念及每个专家在团体中的作用之评判。

定义9 设 U 为有限论域，$R_1,R_2,\cdots,R_n$ 为 U 上的 n 个不同的等价关系(知识)。

称 $(\bigcap_{i=1}^{n} X^{+R_i},\bigcap_{i=1}^{n} X^{-R_i})$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的 n 重共同识别；称 $(X^{+\bigcap_{i=1}^{n} R_i},X^{-\bigcap_{i=1}^{n} R_i})$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的 n 重共商识别。

称 $Dcr^{+}_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|\bigcap_{i=1}^{n} X^{+R_i}|}{|X|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共同正识别度；称 $Dcr^{-}_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|\bigcap_{i=1}^{n} X^{-R_i}|}{|U-X|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共同负识别度；称 $Dcr_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|(\bigcap_{i=1}^{n} X^{+R_i}) \cup (\bigcap_{i=1}^{n} X^{-R_i})|}{|U|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共同识别度。

称 $Dnr^{+}_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|X^{+\bigcap_{i=1}^{n} R_i}|}{|X|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共商正识别度；称 $Dnr^{-}_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|X^{-\bigcap_{i=1}^{n} R_i}|}{|X|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共商负识别度；称 $Dnr_{(R_1,R_2,\cdots,R_n)}(X)=\dfrac{|X^{+\bigcap_{i=1}^{n} R_i} \cup X^{-\bigcap_{i=1}^{n} R_i}|}{|U|}$ 为 $R_1,R_2,\cdots,R_n$ 关于 X 的共商识别度。

定理3 设 U 为有限论域，R_1、R_2 为 U 上的等价关系。$\forall X \subseteq U$，有

(1)$X^{+R_1} \cup X^{+R_2} \subseteq X^{+(R_1 \cap R_2)}$

(2)$X^{-R_1} \cup X^{-R_2} \subseteq X^{-(R_1 \cap R_2)}$

(3)$Drec^{+}_{R_1}(X)+Drec^{+}_{R_2}(X) \leqslant Drec^{+}_{R_1 \cap R_2}(X)$

$Drec^{-}_{R_1}(X)+Drec^{-}_{R_2}(X) \leqslant Drec^{-}_{R_1 \cap R_2}(X)$

$Drec_{R_1}(X)+Drec_{R_2}(X) \leqslant Drec_{R_1 \cap R_2}(X)$

证：由定理2的结论易证。

可见，$R_1 \cap R_2$ 的识别度大于 R_1 与 R_2 的识别度之和。由定理3结论之(1)(2)易证下面定理

定理4 设 $R_1,R_2,\cdots,R_n$ 为 U 上的 n 个知识，则有

(1)$Dcr^{+}_{(R_1,R_2,\cdots,R_n)}(X) \leqslant Dnr^{+}_{(R_1,R_2,\cdots,R_n)}(X)$

(2)$Dcr^{-}_{(R_1,R_2,\cdots,R_n)}(X) \leqslant Dnr^{-}_{(R_1,R_2,\cdots,R_n)}(X)$

(3)$Dcr_{(R_1,R_2,\cdots,R_n)}(X) \leqslant Dnr_{(R_1,R_2,\cdots,R_n)}(X)$

在 n 重识别中，每个专家的识别相对于团体的共同识别与共商识别是有偏差的，为了评判个体对团体的可信程度，引入以下概念。

定义10 称 $Dcrel^{+}_{R_i}(X)=\dfrac{|\bigcap_{i=1}^{n} X^{+R_i}|}{|X^{+R_i}|}$，$Dcrel^{-}_{R_i}(X)=\dfrac{|\bigcap_{i=1}^{n} X^{-R_i}|}{|X^{-R_i}|}$ 分别为知识 R_i 相对于 $(R_1,R_2,\cdots,R_n)$ 关于 X 的共同识别正可信度、共同识别负可信度。

称 $Dnrel^{+}_{R_i}(X)=\dfrac{|X^{+R_i}|}{|X^{+\bigcap_{i=1}^{n} R_i}|}$，为知识 R_i 相对于 $(R_1,R_2,\cdots,R_n)$ 关于 X 的共商识别正可信度；

称 $Dnrel^{-}_{R_i}(X)=\dfrac{|X^{-R_i}|}{|X^{-\bigcap_{i=1}^{n} R_i}|}$ 为知识 R_i 相对于 $(R_1,R_2,\cdots,R_n)$ 关于 X 的共商识别负可信度。

(1)对共商识别，若 $R_{i_0}=\bigcap_{i=1}^{i=n} R_i$，则

$\forall X \subseteq U$，$Dnrel^{+}_{R_{i_0}}(X)=Dnrel^{-}_{R_{i_0}}(X)=1$；

反之不真，见下例2④。

(2)在共同识别中，有以下两个极端情况：

①n 个专家的知识虽略有不同，但没有重大差别。此时利用他们的共同识别来识别范畴 X，得到的结果较为可信，每个 R_i 的可信度较高；

②n 个专家的知识差别较大，此时共同识别度较小，每个专家的可信度亦小。

例：设 $U=\{x_1,x_2,\cdots,x_{10}\}$，$X=\{x_1,x_2,x_3,x_4,x_6,x_7,x_9\}$，专家1、2、3对 U 的分类分别为：

$U \mid R_1=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5\},\{x_6\},\{x_7,x_9\},\{x_8,x_{10}\}\}$

$U \mid R_2=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5,x_6\},\{x_7,x_9\},\{x_8\},\{x_{10}\}\}$

$U \mid R_3=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5\},\{x_6\},\{x_7,x_{10}\},\{x_8\},\{x_9\}\}$

从而

①$X^{+R_1}=\{x_1,x_2,x_3,x_4,x_6,x_7,x_9\}$，$X^{-R_1}=\{x_5,x_8,x_{10}\}$

$Drec^{+}_{R_1}(X)=Drec^{-}_{R_1}(X)=Drec_{R_1}(X)=1$。

②$X^{+R_2}=\{x_1,x_2,x_3,x_4,x_7,x_9\}$，$X^{-R_2}=\{x_8,x_{10}\}$，$Drec^{+}_{R_2}(X)=\dfrac{6}{7}$，$Drec^{-}_{R_2}(X)=\dfrac{2}{3}$，$Drec_{R_2}(X)=\dfrac{8}{10}=\dfrac{4}{5}$。

③$X^{+R_3}=\{x_1,x_2,x_3,x_4,x_6,x_9\}$，$X^{-R_3}=\{x_5,x_{10}\}$，$Drec^{+}_{R_3}(X)=\dfrac{6}{7}$，$Drec^{-}_{R_3}(X)=\dfrac{2}{3}$，$Drec_{R_3}(X)=\dfrac{8}{10}=\dfrac{4}{5}$。

可见，三者中 R_1 对 X 的识别度最高，R_2、R_3 对 X 的识别度相等。

④$U \mid \bigcap_{i=1}^{3} R_i=\{\{x_1,x_3\},\{x_2,x_4\},\{x_5\},\{x_6\},\{x_7\},\{x_8\},\{x_9\},\{x_{10}\}\}$，$X^{-\bigcap_{i=1}^{3} R_i}=\{x_5,x_8,x_{10}\}$，$X^{+\bigcap_{i=1}^{3} R_i}=\{x_1,x_2,x_3,x_4,x_6,x_7,x_9\}$，$Dnr^{+}_{(R_1R_2R_3)}(X)=\dfrac{7}{7}=1$，$Dnr^{-}_{(R_1R_2R_3)}(X)\dfrac{3}{3}=1$，$Dnr_{(R_1R_2R_3)}(X)=\dfrac{10}{10}=1$，$Dnrel^{+}_{R_1}(X)=\dfrac{7}{7}=1$，$Dnrel^{-}_{R_1}(X)=$

$\frac{3}{3}=1$，$Dnrel^+_{R_2}(X)=\frac{6}{7}$，$Dnrel^-_{R_1}(X)=\frac{2}{3}$，$Dnrel^+_{R_3}(X)=\frac{6}{7}$，$Dnrel^-_{R_3}(X)=\frac{2}{3}$。

可见，相对于 X，R_1的共商可信度最高，R_2、R_3的共商可信度相等。

⑤ $\bigcap_{i=1}^{3}X^{+R_i}=\{x_1, x_2, x_3, x_4, x_9\}$，$\bigcap_{i=1}^{3}X^{-R_i}=\{x_{10}\}$，$Dcr^+_{(R_1R_2R_3)}(X)=\frac{5}{7}$，$Dcr^-_{(R_1R_2R_3)}(X)=\frac{1}{3}$，$Dcr_{(R_1R_2R_3)}(X)=\frac{6}{10}=\frac{3}{5}$，$Dcrel^+_{R_1}(X)=\frac{5}{7}$，$Dcrel^-_{R_1}(X)=\frac{1}{3}$，$Dcrel^+_{R_2}(X)=\frac{5}{6}$，$Dcrel^-_{R_2}(X)=\frac{1}{2}$，$Dcrel^+_{R_3}(X)=\frac{5}{6}$，$Dcrel^-_{R_3}(X)=\frac{1}{2}$。

可见，R_1的共同识别正可信度较高而负可信度较低，R_2、R_3关于 X 的共同识别可信度相同。

另外，由④⑤可见共商识别度比共同识别度高。

5 知识的再挖掘

设 R_1、R_2是论域 U 上的两个等价关系。由定理3，$\forall X\subseteq U$，有 $X^{+R_1\cap R_2}\supseteq X^{+R_1}$，$X^{-R_1\cap R_2}\supseteq X^{-R_1}$。

定义11 令 $X^{+(R_2\to R_1)}=X^{+R_1\cap R_2}-X^{+R_1}$，称之为知识 R_2 相对于知识 R_1对 X 的正再挖掘；称 $X^{-(R_2\to R_1)}=X^{-R_1\cap R_2}-X^{-R_1}$为知识 R_2相对于知识 R_1对 X 的负再挖掘。

正再挖掘事实上是在 X 的 R_1下边界 $\underline{\Delta}_{R_1}(X)$中利用 $R_1\cap R_2$的知识颗粒再来识别属于 X 的元素，负再挖掘事实上是在 X 的 R_1上边界 $\overline{\Delta}_{R_1}(X)$中利用 $R_1\cap R_2$的知识颗粒再来识别不属于 X 的元素。

定理5 $\forall X\in Def(U,R_1)$，$X^{+(R_2\to R_1)}=X^{-(R_2\to R_1)}=0$

证：因 $X\in Def(U,R_1)$，故 $X^{+R_1}=X$。又 $X^{+R_1}\subseteq X^{+(R_1\cap R_2)}\subseteq X$，故 $X^{+R_1\cap R_2}=X$，$\overline{R_1\cap R_2}(X)=X$。从而 $X^{+(R_2\to R_1)}=0$，$X^{-(R_2\to R_1)}=X^{-R_1\cap R_2}-X^{-R_1}=(U-\overline{R_1\cap R_2}(X))-(U-\overline{R_1}(X))=0$。

由此可见，讨论知识 R_2对 R_1的再挖掘，仅限于对 R_1-粗集。

定义12 设 R_1、R_2为论域 U 上的知识，若 $R_1\cap R_2\cap X\times X=I_X$，则称知识 R_1与 R_2关于 X 互补；若 $R_1\cap R_2=I_U$，则称知识 R_1与 R_2互补。

定理6 知识 R_2与 R_1互补$\Leftrightarrow\forall X\subseteq X$，$X^{+(R_2\to R_1)}$达到最大值 $X-X^{+R_1}\Leftrightarrow\forall X\subseteq U$，$X^{-(R_2\to R_1)}$达到最大值 $U-X-X^{-R_1}$。

证：①先证知识 R_2与 R_1互补$\Leftrightarrow\forall X\subseteq U$，$X^{+(R_2\to R_1)}$达到最大值 $X-X^{+R_1}$。

必要性：若 R_1与 R_2互补，则 $R_1\cap R_2=I_U$，此时 U 的任意子集 $R_1\cap R_2$是可定义集，从而 $X^{+(R_1\cap R_2)}=X$，$X^{+(R_2\to R_1)}$达到最大值 $X-X^{+R_1}$。

充分性：若$\forall X\subseteq U$，$X^{+(R_2\to R_1)}$达到最大值 $X-X^{+R_1}$，则 $X^{+(R_1\cap R_2)}=X$。可证 $U/R_1\cap R_2$的每个等价类基数全为1。(否则，存在基数大于1的等价类$[x]_{R_1\cap R_2}$。从每个等价类中仅取一个元素组成 U 的真子集 Y。对于$[x]_{R_1\cap R_2}$，由 Y 的构造显然$[x]_{R_1\cap R_2}\cap Y\neq\phi$，且$[x]_{R_1\cap R_2}\not\subset Y$，即$[x]_{R_1\cap R_2}\subseteq Bn_{R_1\cap R_2}(Y)=\phi$，从而 $Y^{+R_1\cap R_2}\neq Y$，矛盾。)

$\forall\ (x,y)\in R_1\cap R_2$，则 $y\in[x]_{R_1\cap R_2}$。由于$|[x]_{R_1\cap R_2}|=1$，故必有 $x=y$，由此证得 $R_1\cap R_2=I_U$。

②$X^{+(R_2\to R_1)}$达到最大值 $X-X^{+R_1}\Leftrightarrow X^{+R_1\cap R_2}=X\Leftrightarrow X\in Def(U,R_1\cap R_2)\Leftrightarrow X^{-R_1\cap R_2}=U-X\Leftrightarrow X^{-(R_2\to R_1)}$达到最大值 $U-X-X^{-R_1}$。

由①②定理得证。

定义13 若 $X^{+(R_2\to R_1)}=\phi$，则称 R_2为相对于 R_1对 X 的平凡正挖掘；若 $X^{-(R_2\to R_1)}=\phi$，则称 R_2为相对于 R_1对 X 的平凡负挖掘。

若$\forall X\subseteq U$，R_2相对于 R_1对 X 是平凡挖掘，则称 R_2是相对于 R_1的平凡挖掘。

定义14 设 R_1、R_2为 U 上的两个知识，$X\subseteq U$，称 $DM^+_{R_2\to R_1}(X)=\frac{|X^{+R_1\cap R_2}-X^{+R_1}|}{|\underline{\Delta}_{R_1}(X)|}$为 R_2相对于 R_1对 X 的正挖掘度；称 $DM^-_{R_2\to R_1}(X)=\frac{|X^{-R_1\cap R_2}-X^{-R_1}|}{|\overline{\Delta}_{R_1}(X)|}$为 R_2相对于 R_1对 X 的负挖掘度。

定理7 设 R_1、R_2为 U 上的两个知识，则有：

(1)$\forall X\in\mathscr{P}(U)-Def(U,R_1)$，$0\leqslant DM^+_{R_2\to R_1}(X)\leqslant 1$，$0\leqslant DM^-_{R_2\to R_1}(X)\leqslant 1$，

(2)$DM^+_{R_2\to R_1}(X)=1\Leftrightarrow DM^-_{R_2\to R_1}(X)=1\Leftrightarrow R_1\cap R_2\cap(X\times(U-X))=\phi$

证：(1)因 $X^{+R_1}\subseteq X^{+R_1\cap R_2}\subseteq X$，$X^{-R_1}\subseteq X^{-R_1\cap R_2}\subseteq U-X$，$U-X-X^{-R_1}=\overline{\Delta}_{R_1}(X)$，故不等式成立。

(2)$DM^+_{R_2\to R_1}(X)=1\Leftrightarrow X^{+R_1\cap R_2}=X\Leftrightarrow X\in Def(U,R_1\cap R_2)\Leftrightarrow X^{-(R_1\cap R_2)}=U-X\Leftrightarrow DM^-_{R_2\to R_1}(X)=1$

下面仅需证：

$DM^+_{R_2\to R_1}(X)=1\Leftrightarrow R_1\cap R_2\cap(X\times(U-X))=\phi$

必要性：若 $DM^+_{R_2\to R_1}(X)=1$，则 $X^{+R_1\cap R_2}=X$，$X\in Def(U,R_1\cap R_2)$。

假设 $R_1\cap R_2\cap(X\times(U-X))\neq\phi$，则存在$\langle x,y\rangle\in R_1\cap R_2\cap(X\times(U-X))$，从而 $xR_1\cap R_2y$，$x\in X$ 且 $y\notin X$，故$[x]_{R_1\cap R_2}\not\subset X$，$[x]_{R_1\cap R_2}\not\subset X^{+R\cap R_2}$，但$[x]_{R_1\cap R_2}$有 X 中的元素 x，故$[x]_{R_1\cap R_2}\subseteq B_{nR_1\cap R_2}(X)\neq\phi$，与 $X\in Def(U,R_1\cap R_2)$矛盾。

充分性：假设 $DM^+_{R_2\to R_1}(X)\neq 1$，则 $X^{+R_1\cap R_2}\neq X$，$X\notin Def(U,R_1\cap R_2)$，从而 $B_{nR_1\cap R_2}(X)\neq\phi$，即存在$[x]_{R_1\cap R_2}$，$[x]_{R_1\cap R_2}\subseteq B_{nR_1\cap R_2}(X)$。故存在 $y,z\in[x]_{R_1\cap R_2}$但 $y\in X$，$z\notin X$，可见：

$\langle y,z\rangle\in R_1\cap R_2\cap(X\times(U-X))$，与条件 $R_1\cap R_2\cap(X\times(U-X))=\phi$ 矛盾。由此证得 $DM^+_{R_2\to R_1}(X)=1$。

结束语 本文利用知识颗粒及正域、负域的概念，在 Pawlak 近似空间(U,R)中对论域 U 中的对象与 U 中范畴 X 的关系进行粗识别，对不可分辨关系 R 对范畴 X 的粗识别能力进行了量化，给出了相应的度量评判标准，并进一步提出了知识的多重粗识别及知识再挖掘的概念，对多个不可分辨关系对论域的共同作用进行了讨论。这些工作将有助于论域上知识的发现与挖掘。

参考文献

1 张文修，吴伟志，梁吉业，等. 粗糙集理论与方法. 北京：科学出版社，2001. 1～12

2 Pawlak Z. Rough Sets. Internation Journal of Information and computer Sciences, 1982, 11(15):341～356

3 Pawlak Z. Rough classification. International Journal of Human-computer studies, 1995, 51: 369～383

4 Mousavi A, Jabedar-Maralani P. Double-faced rough sets and rough communication. Information Sciences, 2002, 148:41～53

5 Chan C-C. A rough set approach to attribute generation in data miming. Journal of Information Sciences, 1998,107:169～176