粗交流的一种新定义形式*)

王洪凯[1,2]　赵树理[3]

(济南大学理学院　济南 250022)[1]　(山东大学数学与系统科学学院　济南 250100)[2]

(商丘师范学院数学系　商丘 476000)[3]

摘　要　利用上近似、下近似等概念给出粗交流的一种新定义形式，得出该定义形式下粗交流与原定义形式下粗交流之间的关系定理。最后通过一个具体实例给出直观分析。

关键词　粗集，粗交流，关系定理，实例，直观分析

A New Definition Form of Rough Communication

WANG Hong-Kai[1,2]　ZHAO Shu-Li[3]

(School of Science, Jinan University, Jinan 250022)[1]　(School of Mathematics and System Sciences, Shandong University, Ji'nan 250100)[2]

(Department of Mathematics, Shangqiu Normal College, Shangqiu 476000)[3]

Abstract　A new definition form of rough communication is proposed by means of the concepts of lower approximation and upper approximation. The relation theorem between old-rough communication and new-rough communication is obtained. Using a specific example, visual analysis of the new definition form is given.

Keywords　Rough sets, Rough communication, Relation theorem, Example, Visual analysis

1　引言

1982年波兰数学家 Z. Pawlak 教授提出粗集理论[1]，作为一种处理含糊和不精确性问题的新型数学工具，粗集在机器学习，知识发现，数据挖掘，模式识别等诸多领域得到了广泛的应用[2~4]。

作为粗集理论应用的一个新发展，A. Mousavi 等在文[5]中提出粗交流(roughcommunication)的概念，即在多个 agents 间进行信息交流时，由于各个 agent 拥有不同的知识(即等价划分)，使得相互之间不能提供准确的信息而产生粗糙性，并将这一信息交流称为粗交流。粗交流过程中，一个概念在多个 agents 间传递时总会丢失部分信息，从而得到精度更差或更加粗糙的概念传递结果。我们注意到：文[5]中定义的粗交流是借助于正域、负域的概念而给出的，由此我们自然会想到：是否可以直接由粗集的其它基本概念，例如上近似、下近似等给出粗交流的定义？如果可以，那么新定义形式下粗交流与原定义形式下粗交流之间有什么联系？

基于上面的问题，本文给出粗交流的一种新定义形式(为叙述方便，不妨称之为新-粗交流)，并对新-粗交流进行简单分析，得出新-粗交流与原定义形式下粗交流(为叙述方便，不妨称之为原-粗交流)的关系定理。最后通过一个具体实例给出直观分析。

2　粗交流

2.1　粗集[1]

设 U 为论域，R 为 U 上的等价关系，集合 $X\subseteq U$ 的下近似，上近似分别为：

$$\underline{R}X=\{x\mid[x]_R\subseteq X\} \tag{2.1}$$

$$\overline{R}X=\{x\mid[x]_R\cap X\neq\phi\} \tag{2.2}$$

其中：$[x]_R$ 表示包含 x 的等价类。称 $(\underline{R}X,\overline{R}X)$ 为 $X\neq U$ 的粗集。

记 $POS_RX=\underline{R}X$，表示 X 的 R-正域；

$NEG_RX=U-\overline{R}X$，表示 X 的 R-负域；

$BND_RX=\overline{R}X-\underline{R}X$，表示 X 的 R-边界域。

若 $BND_R(X)=\phi$，则称 X 在近似空间 (U,R) 中是可定义的。

2.2　粗交流[2]

设 U 为论域，$A\subseteq U$，R 为 U 上的等价关系；$Def(apr)$ 表示近似空间 (U,R) 中所有可定义集的全体，$A^+,A^-\in Def(apr)$，$A^+\cap A^-=\phi$，称 (A^+,A^-) 为近似空间 (U,R) 中的粗集，而且

$$A^+=\underline{R}A \tag{2.3}$$

$$A^-=U-\overline{R}A \tag{2.4}$$

设 n 个近似空间 $(U,R_1),(U,R_2),\cdots,(U,R_n)$ 分别表示 n 个 agents 各自的知识，某确定概念 A 按 agent1，agent2，agent n 的有向顺序传递(图1)。集合对 (A_{i-1}^+,A_{i-1}^-) 在近似空间 (U,R_i) 中的粗集表示为：

$$(A_i^+,A_i^-)=(\underline{R_i}A_{i-1}^+,\underline{R_i}A_{i-1}^-)\quad 2\leqslant i\leqslant n \tag{2.5}$$

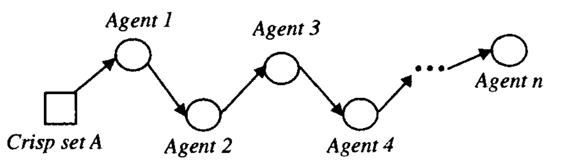

图1　概念 A 的传递顺序

*)国家自然科学基金资助项目(70271048)，山东省自然科学基金资助项目(Y2004A04)。**王洪凯**　博士生，主要研究领域为粗集、模糊集理论及其应用。**赵树理**　讲师，主要研究领域为粗集理论及其应用。

对同一概念 A,得到粗交流传递序列 SEQ_A:

$$SEQ_A=\{(A_1^+,A_1^-),(A_2^+,A_2^-),\cdots,(A_n^+,A_n^-)\},(A_k^+,A_k^-)\in \mathrm{Rough}(apr_k),k\in\{1,2,\cdots,n\} \quad (2.6)$$

其中:$\mathrm{Rough}(apr_k)$表示近似空间(U,R_k)中所有粗集的全体。

由于,$\underline{R_{i+1}}A_i^+\subseteq A_i^+$,$\underline{R_{i+1}}A_i^-\subseteq A_i^-$,因此,$(A_{i+1}^+,A_{i+1}^-)\subseteq_k(A_i^+,A_i^-)$,其中,“$\subseteq_k$”是知识序,表示$(A_{i+1}^+,A_{i+1}^-)$所代表的信息或概念准确度低于$(A_i^+,A_i^-)$。知识序“$\subseteq_k$”的定义:$(A_i^+,A_i^-)\subseteq_k(B_i^+,B_i^-)\Leftrightarrow A_i^+\subseteq B_i^+,A_i^-\subseteq B_i^-$,其中,$(A_i^+,A_i^-)$,$(B_i^+,B_i^-)$分别表示在不同近似空间中,对某一固定概念 A 的粗集表示。

显然,概念 A 在传递过程中,丢失了部分信息从而使概念的准确度下降。

3 新-粗交流

设U为论域,$A\subseteq U$,R 为 U 上的等价关系;设 n 个近似空间(U,R_1),(U,R_2),$\cdots$,(U,R_n)分别表示 n 个 agents 各自的知识,某确定概念 A 按 agent1, agent2,$\cdots$, agent n 的有向顺序传递(图 1)。对同一概念 A,得到新-粗交流传递序列 $NSEQ_A$:

$$NSEQ_A=\{(NA_1^+,NA_1^-),(NA_2^+,NA_2^-),\cdots,(NA_n^+,NA_n^-)\} \quad (3.1)$$

其中:

$$\begin{cases}(NA_1^+,NA_1^-)=(\underline{R_1}A,\overline{R_1}A)\\(NA_i^+,NA_i^-)=(\underline{R_i}NA_{i-1}^+,\overline{R_i}NA_{i-1}^-)\end{cases}\quad 2\leqslant i\leqslant n \quad (3.2)$$

(NA_i^+,NA_i^-) $1\leqslant i\leqslant n$ 是在近似空间(U,R_i)中对某一固定概念 A 的粗集形式的表示,显然有:$NA_i^+\subseteq A\subseteq NA_i^-$。

一般情况下(参与粗交流的 agents 拥有的知识不满足一定条件时),随着 agents 人数的增加,原-粗交流中(A_i^+,A_i^-)趋向于(ϕ,ϕ),新-粗交流中(NA_i^+,NA_i^-)趋向于(ϕ,U),它们的最终结果都是趋向于对概念 A 的一无所知。

下面给出原-粗交流和新-粗交流的关系定理:

引理[6] 设 U 为论域,$X\subseteq U$,则:$\underline{R}(U-X)=U-\overline{R}X$。

证明(略)

定理 1 设 n 个近似空间(U,R_1),(U,R_2),$\cdots$,(U,R_n)分别表示 n 个 agents 各自的知识,原-粗交流传递序列 $SEQ_A=\{(A_1^+,A_1^-),(A_2^+,A_2^-),\cdots,(A_n^+,A_n^-)\}$,新-粗交流传递序列 $NSEQ_A=\{(NA_1^+,NA_1^-),(NA_2^+,NA_2^-),\cdots,(NA_n^+,NA_n^-)\}$,则:

$$\begin{cases}A_i^+=NA_i^+\\A_i^-=U-NA_i^-\end{cases}\quad 1\leqslant i\leqslant n \quad (3.3)$$

证明:显然 $A_i^+=NA_i^+$ $1\leqslant i\leqslant n$。

数学归纳法证明:$A_i^-=U-NA_i^-$ $1\leqslant i\leqslant n$。

$i=1$ 时,$A_1^-=U-\overline{R_1}A=U-NA_1^-$;

假设 $i=k$ 时成立,$A_k^-=U-NA_k^-$,即

$\underline{R_k}(\underline{R_{k-1}}(\cdots\underline{R_2}(U-\overline{R_1}A)\cdots))=U-\overline{R_k}(\overline{R_{k-1}}(\cdots\overline{R_2}(\overline{R_1}A)\cdots))$;

则 $i=k+1$ 时,

$$\begin{aligned}A_{k+1}^-&=\underline{R_{k+1}}(R_k(\underline{R_{k-1}}(\cdots\underline{R_2}(U-\overline{R_1}A)\cdots)))\\&=\underline{R_{k+1}}(U-\overline{R_k}(\overline{R_{k-1}}(\cdots,\overline{R_2}(\overline{R_1}A)\cdots)))。\end{aligned}$$

由引理,则 $A_{k+1}^-=U-\overline{R_{k+1}}(\overline{R_k}(\overline{R_{k-1}}(\cdots\overline{R_2}(\overline{R_1}A)\cdots)))=U-NA_{k+1}^-$。

结论得证。

4 实例

设有 5 个 agents 参与某玩具集合 A 的一个粗交流,论域 U 由 16 个玩具构成,即 $U=\{t_1,t_2,\cdots,t_{16}\}$。agent1, agent2, agent3, agent4, agent5 分别根据颜色、形状、大小、价格、产地等对玩具进行划分:

agent1:$\{t_1,t_3\},\{t_2,t_4,t_8\},\{t_5,t_6\},\{t_7,t_9\},\{t_{10},t_{11},t_{12},t_{13},t_{14}\},\{t_{15},t_{16}\}$;

agent2:$\{t_1,t_3\},\{t_2,t_4\},\{t_5,t_6\},\{t_7,t_9\},\{t_8,t_{10},t_{12}\},\{t_{11},t_{13}\},\{t_{14},t_{15},t_{16}\}$;

agent3:$\{t_1,t_3\},\{t_2,t_4\},\{t_5\},\{t_6,t_7\},\{t_8,t_9,t_{10}\},\{t_{11},t_{12},t_{13},t_{14}\},\{t_{15}\},\{t_{16}\}$;

agent4:$\{t_1,t_2,t_3\},\{t_4\},\{t_5,t_6,t_7\},\{t_8\},\{t_9\},\{t_{10},t_{11},t_{12}\},\{t_{13},t_{14}\},\{t_{15},t_{16}\}$;

agent5:$\{t_1,t_2,t_4\},\{t_5,t_6\},\{t_3,t_7,t_9\},\{t_8\},\{t_{10},t_{11},t_{12}\},\{t_{13},t_{14}\},\{t_{15},t_{16}\}$。

取 $A=\{t_1,t_3,t_5,t_6,t_7,t_9,t_{10},t_{12},t_{15},t_{16}\}$,则得新-粗交流传递序列 $NSEQ_A=\{(NA_1^+,NA_1^-),(NA_2^+,NA_2^-),\cdots,(NA_5^+,NA_5^-)\}$,其中:

$(NA_1^+,NA_1^-)=(\{t_1,t_3,t_5,t_6,t_7,t_9,t_{15},t_{16}\},\{t_1,t_3,t_5,t_6,t_7,t_9,t_{10},t_{11},t_{12},t_{13},t_{14},t_{15},t_{16}\})$;

$(NA_2^+,NA_2^-)=(\{t_1,t_3,t_5,t_6,t_7,t_9\},\{t_1,t_3,t_5,t_6,t_7,t_9,t_{10},t_{11},t_{12},t_{13},t_{14},t_{15},t_{16}\})$;

$(NA_3^+,NA_3^-)=(\{t_1,t_3,t_5,t_6,t_7\},\{t_1,t_3,t_5,t_6,t_7,t_8,t_9,t_{10},t_{11},t_{12},t_{13},t_{14},t_{15},t_{16}\})$;

$(NA_4^+,NA_4^-)=(\{t_5,t_6,t_7\},\{t_1,t_2,t_3,t_5,t_6,t_7,t_8,t_9,t_{10},t_{11},t_{12},t_{13},t_{14},t_{15},t_{16}\})$;

$(NA_5^+,NA_5^-)=(\{t_5,t_6\},\{t_1,t_2,t_3,t_4,t_5,t_6,t_7,t_8,t_9,t_{10},t_{11},t_{12},t_{13},t_{14},t_{15},t_{16}\})$。

由式(2.3)~(2.5)可得原-粗交流传递序列 $SEQ_A=\{(A_1^+,A_1^-),(A_2^+,A_2^-),\cdots,(A_5^+,A_5^-)\}$,在此不再赘述。不难验证,$NSEQ_A$ 与 SEQ_A 的计算结果与上节定理结论是一致的。

结束语 文[5]借助于粗集理论中正域、负域的概念给出粗交流的定义,本文则直接由上、下近似的概念给出粗交流的一种新定义形式,简单直观。在对原-粗交流,新-粗交流进行分析后,给出了两种定义形式下粗交流的关系定理。最后通过一个具体实例给出直观分析。

参考文献

1 Pawlak Z. Rough sets. International Journal of Computer and Information Sciences, 1982, 11(15):341~356

2 Pawlak Z. Rough sets, decision algorithm and Bayes' theorem. European Journal of Operational Research, 2002, 136(1):181~189

3 Dimitras A I, Slowinski R, Susmaga R, Zopounidis C. Business failure prediction using rough sets. European Journal of Operational Research, 1999, 114(2): 263~280

4 Hashemi R R, Blanc L A L, Rucks C T, Rajaratnam A. A hybrid intelligent system for predicting bank holding structures. European Journal of Operational Research, 1998, 109(2):390~402

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

6 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