在不完备信息系统中上、下近似的模糊化

杨霁琳[1] 秦克云[1] 裴 峥[2]
(西南交通大学数学学院 成都 610031)[1] (西华大学数学与计算机学院 成都 610039)[2]

摘 要 在不完备信息系统中,对象存在不确定性。基于对象的不确定性,讨论一种不完备信息系统中上、下近似的模糊化,即,上、下近似是论域上的一对模糊集合。一方面,在不完备信息系统中,集合的上、下近似由相似关系确定;另一方面,由于"*"的存在,使得对象具有不确定性。因此,首先定义了对象的不确定性程度和对象之间的相似度;然后利用对象之间的相似度,通过逻辑关系和相应的函数运算,分别给出了对象隶属于上、下近似集合的隶属度,形式地,该过程可理解为上、下近似集合的模糊化。实例分析说明在不完备信息系统中,考察对象对于上、下近似的隶属度能更直观、合理地反应对象隶属于某一集合的上、下近似的情况。

关键词 不完备信息系统,粗糙集,模糊集,相似关系,相似度,隶属度
中图法分类号 TP18 **文献标识码** A

Fuzzed of Lower and Upper Approximation in Incomplete Information System

YANG Ji-lin[1] QIN Ke-yun[1] PEI Zheng[2]
(School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China)[1]
(School of Machematics & Computer Science, Xihua University, Chengdu 610039, China)[2]

Abstract The uncertainty of objects is considered in the incomplete information system. Based on the uncertainty of objects, the method in which lower and upper approximation are fuzzed, i. e., the lower and upper approximation is a pair of fuzzy sets, was discussed in the incomplete information system. One hand, lower and upper approximations are decided by the similarity relation; on the other hand, the objects with uncertainty due to "*" is existent. Therefore, the uncertainty degree of an object and similarity degree among the objects were defined. Then the membership degree of the object belonging to lower and upper approximation was given by similarity degree. Formally, the process is considered as fuzzed of the lower and upper approximation sets. Finally, an example shows that membership degrees more visualy and reasonably describe that objects belong to the lower and upper approximation of the set.

Keywords Incomplete information system, Rough set, Fuzz set, Similarity relations, Similar degree, Membership degree

1 引言

粗糙集理论[1,2]由Pawlak于1982年提出,是经典集合论的一种推广。随着粗糙集理论研究的深化,粗糙性——信息系统中知识和概念的不确定性,逐渐被人们认识和接受[7-10]。粗糙集在理论和应用上都得到了很大的发展[11-14]。

Pawlak粗糙集模型的主要思想是以等价关系(自反性、对称性、传递性)为基础,通过两个可定义的精确集合,即上、下近似集合来描述知识的不确定性。这种模型被成功地运用在完备信息系统中[7]。在现实生活中,不完备信息系统广泛存在。所谓不完备信息系统是指信息系统中某些对象的若干属性值无法确定,借鉴经典粗糙集理论的学术思想,Kryszkiewicz[3]提出了相容关系,Stafanowski[4]提出了相似关系。目前大多数研究者主要是基于相容关系对不完备信息系统进行研究和讨论[3-5,7]。

在不完备信息系统中,某些对象存在一定程度的描述不完全,即对象自身存在不确定性,且每个对象的不确定性程度各不相同。因此,基于对象的不确定性,本文在相似关系下讨论了一种不完备信息系统中的上、下近似的模糊化,上、下近似是论域上的一对模糊集合。本文主要从两方面考虑:一方面,在不完备信息系统中,集合的上、下近似由相似关系确定,即由对象的相似类组成;另一方面,由于"*"的存在,使得对象具有不确定性。因此,本文首先定义了对象的不确定性程度和对象之间的相似度,然后利用对象之间的相似度,通过逻辑关系和相应的函数运算,分别给出了对象隶属于上、下近似集合的隶属度,由该隶属度,可确定论域中任意集合的上、下近似这一对模糊集合。形式地,整个过程可理解为上、下近似集合的模糊化。最后实例分析说明,在不完备信息系统下的粗糙集模型中,考察对象隶属于上、下近似集合的隶属度,能更直观地刻画对象隶属于某一集合的上、下近似的情况,且能更合理地描述一个粗糙集合。

2 对象的不确定性程度和对象之间的相似度

定义1 一个完备信息系统是一个四元组$\Omega=(U,A,V,$

到稿日期:2009-05-05 返修日期:2009-07-16 本文受四川省杰出青年基金(No. 06ZQ026-037),国家自然科学基金(No. 60875034)资助。
杨霁琳(1981—),女,博士生,主要研究方向为智能信息处理,E-mail:yjl524@163.com;**秦克云**(1962—),男,教授,主要研究方向为多值逻辑、粗糙集理论与方法;**裴 峥**(1968—),男,教授,主要研究方向为语言值处理、智能信息处理。

f)，其中 U 是表示对象的非空有限集合，称为论域；A 是属性的非空有限集合，即 $\forall a\in A: U\to V_A, V_A=\bigcup_{a\in A}V_a$，$V_a$ 是属性 a 的值域，f 表示 $U\times A\to V_A$ 的一个映射，它为每个对象在每个属性上所赋予的一个信息值 $f(x,a)$，即 x 在属性 a 上的取值。若存在 $x\in U, a\in A, f(x,a)$ 是未知的，记为"$*$"，则把这样的信息系统叫做不完备的信息系统，也可简写成二元组 (U,A)。假设 x 在属性集 A 中，至少有一个存在，即：$B\subseteq A$，$\exists a\in B, f(x,a)$ 是已知的，那么该信息系统的值域就是 $V=V_A\cup *$。

定义 2[4] 设 (U,A) 是一个不完备信息系统，其中 U 是对象集，A 是属性集。相似关系 S 定义为：$\forall x,y\in U, S(x,y)$ 是指 $\forall a\in A, f(x,a)\neq *, f(x,a)=f(y,a)$。

定义 2 中的"$*$"值被理解为当前不存在或者无法确定，因此，"$*$"值不参与关系的确定[5]。在基于相似关系 S 的不完备信息系统中，$f(x,a)=*$ 表示取值范围具有不确定性，即 $f(x,a)$ 是否属于 V_a 无法确定，但 $\forall x\in U$，至少 $\exists a\in A$，使得 $f(x,a)$ 已知。在相似关系 S 下，属性值域 $V=V_A\cup *$ 具有不确定性，因此，本文从属性集 A 出发，首先考察每个对象在属性集 A 中被刻画的不确定性程度；然后进一步得到对象的确定性程度。利用对象的确定性程度，定义对象之间的相似度。

2.1 对象的不确定性

在一个不完备信息系统 (U,A) 中，对象对应的属性取值中的"$*$"值越多，它被描述得越不精确，其不确定性程度越高。相反，对象所对应的属性取值中的"$*$"值越少，它被描述得越是精确，其不确定性程度越低。因此，一个对象在属性集 A 中的"$*$"值所对应的属性个数的多少可用来度量该对象的不确定性程度。形式地，定义如下：

定义 3 设 (U,A) 是一个不完备信息系统，U 是对象集，A 是属性集。$\forall x\in U$，对象 x 在属性集 A 下的不确定性程度为：

$$l_A(x)=|N_A(x)|/|A| \tag{1}$$

其中，$N_A(x)=\{a\in A; f(x,a)=*\}$，表示属性集合 A 中对象 x 属性值未知的属性集合；$|N_A(x)|$ 为集合的势；$l_A(x)\in[0,1)$，$l_A(x)$ 是对象 x 为"$*$"值的属性个数与属性集总个数的比值。

$l_A(x)$ 的值越接近 1，表示在不完备信息系统中描述对象 x 的信息越不完全，它的不确定性程度越高；反之，$l_A(x)$ 的值越接近 0，表示对象 x 被描述得越完全，其不确定性程度越低。当 $l_A(x)=0$ 时，即，$|N_A(x)|=0$，对象 x 的信息是完整的，因此对象 x 在属性集 A 下的不确定性程度为 0。直观地，根据定义 3，对象 x 在属性集 A 下的确定性程度为：

$$h_A(x)=1-l_A(x)=\frac{|A|-|N_A(x)|}{|A|}=\frac{|M_A(x)|}{|A|} \tag{2}$$

其中，$M_A(x)=\{a\in A; f(x,a)\neq *\}$，即是对象 x 属性值已知的属性集合；$|M_A(x)|$ 为集合的势。显然，$|M_A(x)|+|N_A(x)|=|A|$，$h_A(x)\in(0,1]$。$h_A(x)$ 的值越接近 0，表示对象 x 的确定性程度越低；反之，$h_A(x)$ 的值越接近 1，表示对象 x 的确定性程度越高。当 $h_A(x)=1$ 时，对象 x 的信息是完整的，即对象 x 在属性集 A 下是完全确定的。

性质 1 设 (U,A) 是一个不完备信息系统，$\forall x,y\in U$，若有 $S(x,y)$，则 $h_A(x)\leqslant h_A(y)$。

证明：$S(x,y)$ 成立，则 $M_A(x)\subseteq M_A(y)$，$|M_A(x)|\leqslant|M_A(y)|$，根据定义 4 可知，$h_A(x)\leqslant h_A(y)$。

性质 2 设 (U,A) 是一个不完备信息系统，$\forall x,y,z\in U$，若有 $S(x,y)$ 和 $S(y,z)$，则 $h_A(x)\leqslant h_A(y)\leqslant h_A(z)$。

证明：$S(x,y)$ 和 $S(y,z)$ 成立，根据相似关系 S 的传递性，有 $M_A(x)\subseteq M_A(y)\subseteq M_A(z)$，则 $|M_A(x)|\leqslant|M_A(y)|\leqslant|M_A(z)|$，所以 $h_A(x)\leqslant h_A(y)\leqslant h_A(z)$。

例 1 在不完备信息系统 (U,A) 中，3 个对象分别是：$x_1(a,b,c,d,e)$，$x_2(a,*,*,*,*)$，$x_3(a,b,c,*,e)$，且属性集 $|A|=5$。

对象 x_1,x_2,x_3 在属性集合 A 下的不确定性程度分别为：$l_A(x_1)=0$，$l_A(x_2)=4/5$，$l_A(x_3)=1/5$；则 3 个对象的确定性程度分别为：$h_A(x_1)=1$，$h_A(x_2)=1/5$，$h_A(x_3)=4/5$。根据定义 2，显然有 $S(x_2,x_3)$ 和 $S(x_3,x_1)$，因此 $h_A(x_2)\leqslant h_A(x_3)\leqslant h_A(x_3)$。

2.2 对象之间的相似度

由定义 2，显然相似关系 S 满足自反性和传递性，但不满足对称性，因此也叫做不对称的相似关系。在不完备信息系统 (U,A) 中，根据 U 上的一个相似关系 S，对任意对象 $x\in U$，可得如下两个集合[4]：

$$I(x)=\{y\in U\mid S(y,x)\} \tag{3}$$

$$I^{-1}(x)=\{y\in U\mid S(x,y)\} \tag{4}$$

式(3)表示相似于 x 的对象构成的集合，式(4)表示与 x 相似的对象构成的集合。显然 $I(x)$ 和 $I^{-1}(x)$ 是两个不同的集合。

对于任意 $B\subseteq A$，显然有[4,6]：

$$I_A(x)\subseteq I_B(x) \tag{5}$$

$$y\in I_B^{-1}(x)\Leftrightarrow I_B^{-1}(y)\subseteq I_B^{-1}(x) \tag{6}$$

$$I_B^{-1}(x)=\bigcup\{I_A^{-1}(y): y\in I_B^{-1}(x)\} \tag{7}$$

例 2 继续例 1，基于相似关系 S，根据式(2)和式(3)可得：$I(x_1)=\{x_1,x_2,x_3\}$，$I^{-1}(x_1)=\{x_1\}$；$I(x_2)=\{x_2\}$，$I^{-1}(x_2)=\{x_1,x_2,x_3\}$；$I(x_3)=\{x_2,x_3\}$，$I^{-1}(x_3)=\{x_1,x_3\}$。

显然 x_2,x_3 都和 x_1 满足相似关系，但它们之间相似的程度不一样。直观地，x_3 比 x_2 更相似于 x_1，因为 $x_3(a,b,c,*,e)$ 有 4 个属性已知属性值，且都与 $x_1(a,b,c,d,e)$ 相等；而 $x_2(a,*,*,*,*)$ 仅有 1 个属性已知属性值，且与 x_1 相等。因此，对象相似类中各元素与该对象之间的相似程度是完全有可能不一样的。同时注意到，相似程度的大小与对象的确定性程度是密切相关的。因此，基于相似关系的对象之间的相似度定义如下：

定义 4 设 (U,A) 是一个不完备信息系统，U 是对象集，A 是属性集，$\forall x,y\in U$，基于相似关系 $S(x,y)$，其相似度定义为：

$$g(x,y)=\begin{cases}h_A(x)/h_A(y), & S(x,y)\text{成立}\\ 0, & S(x,y)\text{不成立}\end{cases} \tag{8}$$

其中，$h_A(x)$ 是对象 x 的确定性程度，$g(x,y)\in[0,1]$，即是相似关系下对象之间的相似度。

$g(x,y)$ 越接近 1 时，表示两个对象在相似关系下的相似程度越大，当 $g(x,y)=1$ 时，表示两个对象已知属性值的属性完全相等；反之 $g(x,y)$ 越接近 0，则表示两个对象在相似关系下的相似程度越低，当 $g(x,y)=0$ 时，说明两个对象不存在相似关系。

性质 3 设(U,A)是一个不完备信息系统，$\forall x\in U$，$g(x,y)=1$。

证明：根据相似关系 S 的自反性，必有 $S(x,x)$成立，即证 $g(x,x)=1$。

性质 4 设(U,A)是一个不完备信息系统，$\forall x,y\in U$，若 $g(x,y)\neq 0$，则 $h_A(x)\leqslant h_A(y)$。

证明：$g(x,y)\neq 0$，则有 $S(x,y)$，根据性质 1，显然 $h_A(x)\leqslant h_A(y)$。

性质 5 设(U,A)是一个不完备信息系统，$\forall x,y\in U$，若 $g(x,y)=g(y,x)$，则$\forall a\in A$，有 $f(x,a)=f(y,a)$。

证明：$g(x,y)=g(y,x)$，则 $S(x,y)$和 $S(y,x)$成立，根据定义 2，有 $M_A(x)=M_A(y)$，因此，$\forall a\in A$，$f(x,a)=f(y,a)$。

性质 5 说明，若两个对象之间的相似度相等，则两个对象完全相同。

性质 6 设(U,A)是一个不完备信息系统，$\forall x,y,z\in U$，若 $g(x,y)\neq 0$，$g(y,z)\neq 0$，则 $g(x,z)\neq 0$，且 $g(x,y)\geqslant g(x,z)$。

证明：根据相似关系 S 的传递性，易证 $g(x,z)\neq 0$，由性质 2 知 $h(x)\leqslant h(y)\leqslant h(z)$，且属于$(0,1]$，因此 $h(x)/h(y)\geqslant h(x)/h(z)$，即 $g(x,y)\geqslant g(x,z)$。

例 3 继续例 1 和例 2，$I^{-1}(x_1)=\{x_1\}$，有 $g(x_1,x_1)=1$，$g(x_1,x_2)=g(x_1,x_3)=0$；$I^{-1}(x_2)=\{x_1,x_2,x_3\}$，$g(x_2,x_1)=1/5$，$g(x_2,x_2)=1$，$g(x_2,x_3)=1/4$；$I^{-1}(x_3)=\{x_1,x_3\}$，$g(x_3,x_1)=4/5$，$g(x_3,x_3)=1$。

针对对象 x_2，显然 $g(x_2,x_1)\neq g(x_2,x_2)\neq g(x_2,x_3)$。在不完备信息系统中，由于对象存在不确定性，因此在同一对象的相似类中，各元素与该对象的相似程度也各不相同。自然地，在相似关系下确定的上、下近似，即由相似类组成的上、下近似集合中，各对象对于上、下近似的贡献也各不相同。

3 粗糙集上、下近似的模糊化

在 Pawlak 粗糙集模型中，不精确集由两个精确集近似地定义，即由粗糙集的上近似和下近似来描述。在不完备信息系统下的粗糙集模型中，对象存在不确定性。在相似关系下，对象的相似度也各不相同。因此，各对象对于上、下近似的贡献也各不相同，直观地，上、下近似可以看成是一对模糊集合。当一个不精确集通过上、下近似集合描述时，就存在上、下近似集合中的元素是否更逼近该不精确集合且是否能被描述得更准确、合理的问题。

针对以上问题，本节利用对象之间的相似度，通过逻辑关系和相应的函数运算，分别给出了对象隶属于上、下近似集合的隶属度，形式地，该过程可理解为上、下近似集合的模糊化。最后通过实例分析，说明在不完备信息系统下的粗糙集模型中，考察对象对于上、下近似集合的隶属度，能更直观地刻画对象隶属于某一集合的上、下近似的情况，且能更合理地描述一个粗糙集合。

3.1 对象对于上、下近似集合的隶属度

定义 5 设(U,A)是一个不完备信息系统，U 是对象集，A 是属性集，基于相似关系 S，任意对象集 $Z\subseteq U$ 的上、下近似分别定义为[4,6]：

$$\underline{R}(Z)=\{x\in U;I^{-1}(x)\subseteq Z\} \tag{9}$$

$$\overline{R}(Z)=\{x\in U;I^{-1}(x)\cap Z\neq\phi\} \tag{10}$$

用逻辑上的析取、合取、蕴涵和非等关系以及选取恰当的函数运算把上、下近似翻译成另一种表达式。

(1) “非”：函数 N，通常的运算是 $N(x)=1-x$；

(2) 合取：函数 T-norm，通常的运算是 $T(x,y)=xy$；

(3) 析取：函数 T-conorm，通常的运算是 $S(x,y)=x+y-xy$；

(4) 蕴涵：函数 I，通常的运算是 $I(x,y)=1-x+xy$。

若有对象集合 $Z,W\subseteq U$，则根据上面的逻辑关系和相关的函数[5]有：

$$\forall xZ(x)\overset{\text{def}}{=}T_xZ(x);\ \exists xZ(x)\overset{\text{def}}{=}S_xZ(x);$$

$$Z\subseteq W\overset{\text{def}}{=}T_x(I(\mu_Z(x),\mu_W(x)));$$

$$Z\cap W\neq\phi=\exists xZ(x)\wedge W(x)\overset{\text{def}}{=}S_x(T(\mu_Z(x),\mu_W(x)))$$

带入式(9)和式(10)可得：

$$\mu_{\underline{R}(Z)}(x)=T_{y\in S^{-1}(x)}(I(g(x,y),\hat{y})) \tag{11}$$

$$\mu_{\overline{R}(Z)}(x)=S_{y\in S^{-1}(x)}(T(g(x,y),\hat{y})) \tag{12}$$

其中，$\mu_{\underline{R}(Z)}(x)$表示对象 x 隶属于任意对象子集 Z 的下近似集合的程度；$\mu_{\overline{R}(Z)}(x)$表示对象 x 隶属于任意对象子集 Z 的上近似集合的程度；$I^{-1}(x)$是对象 x 的相似类；T,S,I 是相对应的逻辑函数；$g(x,y)$是在相似关系下对象 x 与 y 的相似度；$\hat{y}$是对象 y 在集合 Z 上的特征函数值，$\hat{y}\in\{0,1\}$，$\hat{y}=0$ 表示 y 不属于集合 Z，$\hat{y}=1$ 表示 y 属于集合 Z。把 T,S,I 对应的函数运算带入到式(11)和式(12)，可定义任意对象 x 对于上、下近似集合的隶属度。

定义 6 设(U,A)是一个不完备信息系统，U 是对象集，A 是属性集，$\forall x\in U$，任意对象集 $Z\subseteq U$，则 x 隶属于 Z 上、下近似集合的隶属度为：

$$\mu_{\underline{R}(Z)}(x)=\prod_{y\in S^{-1}(x)}(1-g(x,y)+g(x,y)\hat{y}) \tag{13}$$

$$\mu_{\overline{R}(Z)}(x)=1-\prod_{y\in S^{-1}(x)}(1-g(x,y)\hat{y}) \tag{14}$$

其中，$g(x,y)$是对象 x 与 y 在相似关系下的相似度。$\mu_{\underline{R}(Z)}(x)$，$\mu_{\overline{R}(Z)}(x)\in[0,1]$。$\mu_{\underline{R}(Z)}(x)$越接近 1，表示 x 隶属于 Z 的下近似集合的程度越大；$\mu_{\underline{R}(Z)}(x)$越接近 0，表示 x 隶属于 Z 的下近似集合的程度越低。$\mu_{\overline{R}(Z)}(x)$越接近 1，表示 x 隶属于 Z 的上近似集合的程度越大；$\mu_{\overline{R}(Z)}(x)$越接近 0，表示 x 隶属于 Z 的上近似集合的程度越低。

在一个不完备信息系统(U,A)中，对象集 $U=\{x_1,x_2,\cdots,x_n\}$，利用定义 6 的隶属度就可确定论域 U 上“Z 的下近似”和“Z 的上近似”这一对模糊集合$\underline{R}(Z)$和 $\overline{R}(Z)$，利用 Zadeh 表示法，可分别表示为：

$$\underline{R}(Z)=\mu_{\underline{R}(z)}(x_1)/x_1+\mu_{\underline{R}(z)}(x_2)/x_2+\cdots+\mu_{\underline{R}(z)}(x_n)/x_n \tag{15}$$

$$\overline{R}(Z)=\mu_{\overline{R}(z)}(x_1)/x_1+\mu_{\overline{R}(z)}(x_2)/x_2+\cdots+\mu_{\overline{R}(z)}(x_n)/x_n \tag{16}$$

若$\underline{R}(Z)=\overline{R}(Z)$，则称 Z 是可定义的；若$\underline{R}(Z)\neq\overline{R}(Z)$，则称 Z 是不可定义的。

3.2 实例分析

下面用一个具体的例子，分析说明对象隶属于上、下近似的隶属度能更直观、准确地刻画粗糙集合。

例 4 设(U,A)是一个不完备信息系统，如表 1 所列对象集 $U=\{x_1,\cdots,x_{10}\}$，属性集 $A=\{a_1,\cdots,a_4\}$。

表 1　不完备的信息系统(U,A)

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
a_1	3	3	2	2	*	*	3	3	2	2
a_2	2	*	3	3	0	2	2	2	3	*
a_3	1	*	0	2	0	*	1	1	*	*
a_4	0	3	1	1	*	*	*	3	1	1

在属性集 A 下，由式(4)可得各对象的相似类：$I^{-1}(x_1)=\{x_1\}$；$I^{-1}(x_2)=\{x_2,x_8\}$；$I^{-1}(x_3)=\{x_3\}$；$I^{-1}(x_4)=\{x_4\}$；$I^{-1}(x_5)=\{x_5\}$；$I^{-1}(x_6)=\{x_1,x_6,x_7,x_8\}$；$I^{-1}(x_7)=\{x_1,x_7,x_8\}$；$I^{-1}(x_8)=\{x_8\}$；$I^{-1}(x_9)=\{x_3,x_4,x_9\}$；$I^{-1}(x_{10})=\{x_3,x_4,x_9,x_{10}\}$。

对象子集 $Z=\{x_1,x_2,x_4,x_6,x_9\}$，$W=\{x_1,x_2,x_4,x_7,x_{10}\}$，显然两个集合不相同。由式(9)和式(10)定义的上、下近似来描述这两个不精确集合，可知：$\underline{R}(Z)=\underline{R}(W)=\{x_1,x_4\}$，$\bar{R}(Z)=\bar{R}(W)=\{x_1,x_2,x_4,x_6,x_7,x_9,x_{10}\}$。用式(9)和式(10)刻画一个不精确集合，显然在很大程度上并不能准确、合理地进行描述。因此，将上、下近似模糊化，用定义 6 中对象对于上、下近似的隶属度来刻画对象子集 Z 和 W。

首先，根据定义 4 得到对象之间的相似度，如表 2 所列。

表 2　对象之间的相似度 $g(x,y)$

x \ y	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	1	0	0	0	0	0	0	0	0	0
x_2	0	1	0	0	0	0	0	1/2	0	0
x_3	0	0	1	0	0	0	0	0	0	0
x_4	0	0	0	1	0	0	0	0	0	0
x_5	0	0	0	0	1	0	0	0	0	0
x_6	1/4	0	0	0	0	1	1/3	1/4	0	0
x_7	3/4	0	0	0	0	0	1	3/4	0	0
x_8	0	0	0	0	0	0	0	1	0	0
x_9	0	0	3/4	3/4	0	0	0	0	1	0
x_{10}	0	0	1/2	1/2	0	0	0	0	2/3	1

然后，根据式(13)和式(14)，$\forall x\in U$，分别对于 Z,W 的下近似和上近似的隶属度，如表 3 所列。其中，各对象在集合 Z,W 下的特征值分别是：$\hat{y}_Z=(1,1,0,1,0,1,0,0,1,0)$，$\hat{y}_W=(1,1,0,1,0,0,1,0,0,1)$。

表 3　对象对于两个不精确集合的下近似和上近似的隶属度

U	$\mu_{\underline{R}(Z)}(x)$	$\mu_{\underline{R}(W)}(x)$	$\mu_{\bar{R}(Z)}(x)$	$\mu_{\bar{R}(W)}(x)$
x_1	1	1	1	1
x_2	0.5	0.5	1	1
x_3	0	0	0	0
x_4	1	1	1	1
x_5	0	0	0	0
x_6	0.5	0	1	0.5
x_7	0	0.25	0.75	1
x_8	0	0	0	0
x_9	0.25	0	1	0.75
x_{10}	0	0.17	0.83	1

根据式(15)和式(16)，Z 的下近似集合和上近似集合可分别表示为：

$$\underline{R}(Z)=1/x_1+0.5/x_2+0/x_3+1/x_4+0/x_5+0.5/x_6+0/x_7+0/x_8+0.25/x_9+0/x_{10}$$

$$\bar{R}(Z)=1/x_1+1/x_2+0/x_3+1/x_4+0/x_5+1/x_6+0.75/x_7+0/x_8+1/x_9+0.83/x_{10}$$

$\underline{R}(Z)\neq\bar{R}(Z)$，显然 Z 是不可定义的。类似地，W 的下近似和上近似也可根据式(15)和式(16)表示，且 W 也是不可定义的。

通过以上分析，说明了在不完备信息系统中，将上、下近似模糊化，通过对象对于某一集合上、下近似的隶属度，能更直观、合理地反应每个对象隶属于该集合上、下近似的具体情况，同时也能更准确地描述这一集合。

结束语　针对基于相似关系的不完备信息系统，考虑到，一方面，在不完备信息系统中，集合的上、下近似由相似关系确定；另一方面，由于“*”的存在使对象具有不确定性，因此，本文基于对象的不确定性，讨论了一种不完备信息系统中上、下近似的模糊化，即上、下近似是论域上的一对模糊集合。定义了对象的不确定性程度和对象之间的相似度，然后利用对象之间的相似度，分别给出了对象隶属于上、下近似集合的隶属度，由该隶属度可确定论域中任意集合的上、下近似这一对模糊集合。形式地，该过程可理解为上、下近似集合的模糊化。实例分析说明在不完备信息系统中，考察对象对于上、下近似的隶属度能更直观、合理地反应对象隶属于某一集合的上、下近似的情况。

参 考 文 献

[1] Pawlak Z. Rough sets and intelligent data analysis [J]. Information Sciences, 2002, 147: 1-12

[2] Pawlak Z, Skowron A. Rough sets: some extensions [J]. Information Sciences, 2007, 177: 41-73

[3] Kryszkiewicz M. Rough set approach to incomplete information systems [J]. Information Sciences, 1998, 112: 39-49

[4] Stefanowski J, Tsoukias A. On the extension of rough sets under incomplete information[C]// Zhong N, Ohsuga S, eds. New Directions in Rough sets, Data Mining and Granular-Soft Computing, Lecture Notes in Artificial Intelligence. vol. 1711, Berlin: Springer-Verlag, 1999: 73-82

[5] Stefanowski J, Tsoukias A. Incomplete information tables and rough classification[J]. Computational Intelligence, 2001, 17: 545-566

[6] Slowinski R, Vanderpooten D. A generalized definition of rough approximations based on similarity[J]. IEEE Transactions on knowledge and data engineering, 2000, 12(2): 1041-4347

[7] 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

[8] 裴峥，秦克云. 粗糙集的直觉模糊特殊集表示及其在属性约简中的应用[J]. 模式识别与人工智能，2004，17(3)：262-266

[9] Qin Keyun, Pei Zheng. On the topological properties of fuzzy rough sets[J]. Fuzzy sets and systems, 2005, 114(3): 601-613

[10] Qin Keyun, Yang Jilin, Pei Zheng. Generalized rough sets based on reflexive and transitive relations[J]. Information Sciences, 2008, 178(21): 4138-4141

[11] 米据生，吴伟志，张文修. 粗糙集的构造与公理化方法[J]. 模式识别与人工智能，2002，15(3)：280-284

[12] 吴伟志，张文修，徐宗本. 粗糙模糊集的构造与公理化方法[J]. 计算机学报，2004，27(2)：197-203

[13] 程昳，莫智文. 模糊粗糙集及粗糙模糊集的模糊度[J]. 模糊系统与数学，2001，15(3)：15-18

[14] 唐丽，王燕，莫智文. 广义模糊粗糙集的广义不确定性[J]. 四川师范大学学报：自然科学版，2007，30(3)：288-290