

# $S_p$ -上阶梯知识挖掘及状态识别算法

刘若慧<sup>1</sup>, 刘保仓<sup>2</sup>

LIU Ruo-hui<sup>1</sup>, LIU Bao-cang<sup>2</sup>

1. 黄淮学院 计算机科学系, 河南 驻马店 463000

2. 黄淮学院 数学科学系, 河南 驻马店 463000

1. Department of Computer Sciences, Huanghuai University, Zhumadian, Henan 463000, China

2. Department of Mathematics Sciences, Huanghuai University, Zhumadian, Henan 463000, China

E-mail: lruohui@126.com

LIU Ruo-hui, LIU Bao-cang. Mining of  $S_p$ -upper ladder knowledge and algorithm of state recognition. Computer Engineering and Applications, 2010, 46(1): 45-47.

**Abstract:** In one direction  $S_p$ -rough sets, the knowledge  $[x](R\text{-element equivalence class } [x])$  with the attribute set  $\alpha$  has such characteristic: If there are new attributes supplemented to  $\alpha$ , then the elements in  $[x]$  will decrease; By using these characteristics and considering the reliability characteristics of the element transference random, the concepts of  $S_p$ -upper ladder knowledge, the generation of depending on reliability of  $S_p$ -upper ladder knowledge, and attribute dependence of  $S_p$ -upper ladder knowledge are presented. Attribute dependence mining theorem of  $S_p$ -upper ladder knowledge, and attribute dependence algorithm are also given.

**Key words:**  $S_p$ -upper ladder knowledge; attribute dependence; generation of depending on reliability; dependence mining; unit circle theorem; state recognition; mining algorithm

**摘 要:** 单向  $S_p$ -粗集中, 具有属性集  $\alpha$  的知识  $[x](R\text{-元素等价类 } [x])$  具有这样的特征: 若  $\alpha$  内被补充属性, 则  $[x]$  内的元素个数被减少。利用这一特征, 考虑属性补充的随机性, 给出  $S_p$ -上阶梯知识,  $S_p$ -上阶梯知识的依信度生成,  $S_p$ -上阶梯知识属性依赖的原理, 给出  $S_p$ -上阶梯知识的属性依赖挖掘定理,  $S_p$ -上阶梯知识的状态识别算法。

**关键词:**  $S_p$ -上阶梯知识; 属性依赖; 依信度生成; 依赖挖掘; 单位圆定理; 状态识别; 挖掘算法

DOI: 10.3778/j.issn.1002-8331.2010.01.015 文章编号: 1002-8331(2010)01-0045-03 文献标识码: A 中图分类号: O159

文献[1-2]提出了  $S$ -粗集。在静态、动态意义下,  $S$ -粗集是 Z.Pawlak 粗集<sup>[3]</sup>的推广。  $S$ -粗集是用具有动态特征的  $R$ -元素等价类  $[x]$  (知识  $[x]$ ) 定义的。文献[4-9]提出  $S_p$ -粗集和变异  $S_p$ -粗集并对其进行了诸多讨论, 使得具有动态特征的  $R$ -元素等价类  $[x]$  (知识  $[x]$ ) 及其相对应的属性集  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$  不仅具有动态性而且具有随机性。  $S_p$ -粗集、变异  $S_p$ -粗集均具有两种形式: 单向  $S_p$ -粗集、双向  $S_p$ -粗集, 单向变异  $S_p$ -粗集、双向变异  $S_p$ -粗集。  $S_p$ -粗集和变异  $S_p$ -粗集具有依赖关系。结果是利用单向  $S_p$ -粗集和单向变异  $S_p$ -粗集所具有的依赖关系得到的。在单向  $S_p$ -粗集中, 知识  $[x]$  的变化, 依赖于单向变异  $S_p$ -粗集中的  $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$  的变化, 知识  $[x]$  具有上阶梯特征。如何利用这一特征, 挖掘人们还不知道的知识是主题。

约定: 在讨论中,  $R$ -元素等价类  $[x]$ , 知识  $[x]$  两个概念不加区别, 直接使用,  $f \in F$  是属性迁移<sup>[6-7]</sup>,  $V$  是有限属性论域,  $P_f(f)$  为属性迁移的信度函数,  $S_p$ -粗集和变异  $S_p$ -粗集所涉及的概念和符号在文献[4-8]中均能找到。

## 1 $S_p$ -上阶梯知识与它的依信度生成

**定义 1** 称  $[x]_p^f$  是  $[x]$  的依信度  $p$  生成的  $S_p$ -上阶梯知识, 如

果  $[x]_p^f$  的属性集  $\alpha_p^f$  与  $[x]$  的属性集  $\alpha$  满足

$$\alpha \subseteq \alpha_p^f \quad (1)$$

称  $\eta_p^f$  是  $[x]_p^f$  关于  $[x]$  的  $S_p$ -上阶梯度, 简称  $[x]_p^f$  的  $S_p$ -阶梯度, 如果

$$\eta_p^f = \frac{\text{card}(\alpha_p^f)}{\text{card}(\alpha)} \quad (2)$$

这里:  $\alpha_p^f = \alpha \cup \{f(\beta_i) = \alpha_i', i=1, 2, \dots, m, P_f(f) \geq p\}$ ,  $\beta_i \in V, \beta_i \in \alpha, f(\beta_i) = \alpha_i' \in \alpha, \alpha = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ ,  $\text{card}(\alpha)$  是  $\alpha$  的基数。

**定义 2** 设  $[x]_p^f$  是  $[x]$  的依信度  $p$  生成的  $S_p$ -上阶梯知识, 称  $[x]$  是  $[x]_p^f$  的依信度  $p$  生成的  $S_p$ -上阶梯属性还原知识, 简称  $S_p$ -上阶梯还原知识, 如果存在属性差集  $\nabla \alpha_p^f$  满足

$$\nabla \alpha_p^f = \alpha_p^f - \alpha \quad (3)$$

称  $\varphi_p^f$  是  $[x]$  关于  $[x]_p^f$  的  $S_p$ -上阶梯还原系数, 简称  $\varphi_p^f$  是  $[x]$  的  $S_p$ -上阶梯还原系数, 如果

$$\varphi_p^f = \frac{\text{card}([x]_p^f)}{\text{card}([x])} \quad (4)$$

基金项目: 河南省基础与前沿技术研究重点资助项目 (No.082300410040); 河南省教育厅自然科学基金 (No.2008C120002)。

作者简介: 刘若慧 (1965-), 女, 副教授, 研究方向: 软件工程, 数据挖掘; 刘保仓 (1965-), 男, 教授, 主要研究方向: 模糊系统、粗系统理论与应用研究。

收稿日期: 2008-07-22 修回日期: 2008-10-17

这里:  $\alpha_p^f, \alpha$  分别是  $[x]_p^f, [x]$  的属性集,  $\nabla \alpha \neq \phi, \text{card}([x])$  是  $[x]$  的基数。

由定义 1, 定义 2, 直接得到

**命题 1**  $[x]_p^f$  的属性集  $\alpha_p^f$  与  $[x]$  的属性集  $\alpha$  满足  $\alpha_p^f - \alpha \neq \phi$ ,  $S_{p-}$  上阶梯知识  $[x]_p^f$  存在。

**命题 2** 若  $\forall f \in F, P_r(f)=0$ , 则  $[x]_p^f$  的属性集  $\alpha_p^f$  与  $[x]$  的属性集  $\alpha$  满足  $\nabla \alpha_p = \phi$ ,  $S_{p-}$  上阶梯知识  $[x]_p^f$  不存在。

**命题 3** 若  $\forall f \in F, P_r(f)=1$ , 则  $[x]_p^f$  的属性集  $\alpha_p^f$  与  $[x]$  的属性集  $\alpha$  满足  $\nabla \alpha_p \neq \phi$ , 则  $S_{p-}$  上阶梯知识  $[x]_p^f$  存在。

**定理 1** ( $S_{p-}$  上阶梯度单位圆定理) 若  $[x]_p^f = \{[x]_{p_i}^f | i=1, 2, \dots, n-1, n; p \leq p_1 \leq p_2 \leq \dots \leq p_{n-1} \leq p_n\}$  是  $S_{p-}$  上阶梯知识  $[x]_{p_i}^f$  构成的上阶梯知识集,  $I(\eta_p^f)$  是上阶梯度  $\eta_p^f=1$  为半径构成的单位圆, 则  $\forall [x]_{p_i}^f \in [x]_p^f$  的上阶梯度  $\eta_{p_i}^f$  落在  $I(\eta_p^f)$  之外,  $k=1, 2, \dots, n$ ; 或者

$$\eta_{p_i}^f \bar{\in} I(\eta_p^f) \quad (5)$$

这里: 是一个记号  $\eta_{p_i}^f \bar{\in} I(\eta_p^f)$ , 它表示  $\forall k, \eta_{p_i}^f \geq \eta_p^f, k=1, 2, \dots, n$ 。

**证明**  $[x]_p^f = \{[x]_{p_i}^f | k=1, 2, \dots, n\}$  是上阶梯知识集,  $\forall [x]_{p_i}^f \in [x]_p^f, k \in (1, 2, \dots, n)$ , 或者  $[x]_{p_i}^f \subseteq [x]_{p_{i-1}}^f \subseteq \dots \subseteq [x]_{p_{i-1}}^f \subseteq [x]_{p_i}^f \subseteq [x]$  由定义 1 得到  $\alpha \subseteq \alpha_{p_i}^f \subseteq \alpha_{p_{i-1}}^f \subseteq \dots \subseteq \alpha_{p_1}^f \subseteq \alpha_p^f, 1=\eta_p^f \leq \eta_{p_i}^f \leq \eta_{p_{i-1}}^f \leq \dots \leq \eta_{p_1}^f \leq \eta_p^f$ ; 显然,  $\forall \eta_{p_i}^f$  都在以  $\eta_p^f=1$  为半径的圆  $I(\eta_p^f)$  之外。这里:  $1=\eta_p^f = \text{card}(\alpha) / \text{card}(\alpha)$  是  $[x]$  关于  $[x]$  的上阶梯度。

**推论 1** 若  $[x]_p^f = \{[x]_{p_i}^f | k=1, 2, \dots, n\}$  是  $S_{p-}$  上阶梯知识  $[x]_{p_i}^f$  构成的  $S_{p-}$  上阶梯知识集,  $I(\varphi_p^f)$  是  $S_{p-}$  上阶梯还原系数  $\varphi_p^f=1$  为半径构成的单位圆, 则  $\forall [x]_{p_i}^f \in [x]_p^f$  的  $S_{p-}$  上阶梯还原系数  $\varphi_{p_i}^f$  落于  $I(\varphi_p^f)$  之内,  $\lambda=1, 2, \dots, n$ ; 或者

$$\varphi_{p_i}^f \in I(\varphi_p^f) \quad (6)$$

## 2 属性依赖 $S_{p-}$ 上阶梯知识挖掘定理

**定义 3** 设  $\alpha, \alpha_p^f$  分别是  $[x], [x]_p^f$  的属性集, 称  $[x]_p^f$   $f$ -属性依信度  $p$  单依赖于  $[x]$ , 记作  $[x]_p^f \Rightarrow [x]$ , 如果

$$\alpha \Rightarrow \alpha_p^f \quad (7)$$

**定义 4** 设  $\alpha, \alpha_p^f$  分别是  $[x], [x]_p^f$  的属性集, 称  $[x]_p^f$   $f$ -属性依信度  $P$  双依赖于  $[x]$ , 记作  $[x]_p^f \Leftrightarrow [x]$ , 如果

$$\alpha \Leftrightarrow \alpha_p^f \quad (8)$$

由定义 3~4, 得到

**定理 2** (属性依信度单依赖  $S_{p-}$  上阶梯度序定理) 若  $[x]_p^f, [x]_{p_i}^f, [x]_{p_j}^f \in [x]_p^f$ , 它们的属性集  $\alpha_p^f, \alpha_{p_i}^f, \alpha_{p_j}^f$  满足

$$\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f \Rightarrow \alpha_p^f \quad (9)$$

则

$$\eta_{p_i}^f \leq \eta_{p_j}^f \leq \eta_p^f \quad (10)$$

这里:  $\eta_{p_i}^f, \eta_{p_j}^f, \eta_p^f$  分别是  $S_{p-}$  上阶梯知识  $[x]_{p_i}^f, [x]_{p_j}^f, [x]_p^f$  的上阶梯度。

事实上,  $\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f \Rightarrow \alpha_p^f$  则有  $\alpha_{p_i}^f \subseteq \alpha_{p_j}^f \subseteq \alpha_p^f$ , 或  $\text{card}(\alpha_{p_i}^f) \leq \text{card}(\alpha_{p_j}^f) \leq \text{card}(\alpha_p^f)$ ; 由式(2)得到式(10), 证明略。

**定理 3** (属性依信度单依赖  $S_{p-}$  上阶梯还原系数序定理)

若  $[x]_{p_i}^f, [x]_{p_j}^f, [x]_{p_k}^f \in [x]_p^f$ , 它们的属性集  $\alpha_{p_i}^f, \alpha_{p_j}^f, \alpha_{p_k}^f$  满足

$$\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f \Rightarrow \alpha_{p_k}^f \quad (11)$$

则  $\varphi_{p_i}^f \leq \varphi_{p_j}^f \leq \varphi_{p_k}^f$  (12)

这里:  $\varphi_{p_i}^f, \varphi_{p_j}^f, \varphi_{p_k}^f$  分别是  $[x]$  关于的  $[x]_{p_i}^f, [x]_{p_j}^f, [x]_{p_k}^f$  的  $S_{p-}$  上阶梯还原系数。

**证明** 因为  $\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f \Rightarrow \alpha_{p_k}^f$ , 或者  $\alpha_{p_i}^f \subseteq \alpha_{p_j}^f \subseteq \alpha_{p_k}^f$ , 则有  $[x]_{p_i}^f \subseteq [x]_{p_j}^f \subseteq [x]_{p_k}^f$  或者  $\text{card}([x]_{p_i}^f) \leq \text{card}([x]_{p_j}^f) \leq \text{card}([x]_{p_k}^f), \text{card}([x]_{p_i}^f) / \text{card}[x] \leq \text{card}([x]_{p_j}^f) / \text{card}[x] \leq \text{card}([x]_{p_k}^f) / \text{card}[x]$ ; 由式(6)得式(21)。

**定理 5** (属性依信度双依赖  $S_{p-}$  上阶梯知识不可分辨定理)

若  $[x]_{p_i}^f, [x]_{p_j}^f \in [x]_p^f$ , 它们的属性集  $\alpha_{p_i}^f, \alpha_{p_j}^f$ , 满足

$$\alpha_{p_i}^f \Leftrightarrow \alpha_{p_j}^f \quad (13)$$

则

$$\text{IND}([x]_{p_i}^f, [x]_{p_j}^f) \quad (14)$$

这里:  $\text{IND} = \text{indiscernibility}$

事实上, 若  $\alpha_{p_i}^f \Leftrightarrow \alpha_{p_j}^f$ , 则有  $\alpha_{p_i}^f = \alpha_{p_j}^f$ , 显然有  $\text{IND}([x]_{p_i}^f, [x]_{p_j}^f)$ , 证明略。

利用定理 2~4, 定义 1~4, 得到

**定理 5** (属性依信度单依赖  $S_{p-}$  上阶梯知识内挖掘定理)

若  $[x]_{p_i}^f$  的属性集  $\alpha_{p_i}^f$  与  $[x]_{p_j}^f$  的属性集  $\alpha_{p_j}^f$  满足

$$\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f \quad (15)$$

则  $[x]_{p_i}^f$  在  $[x]_{p_j}^f$  之内被挖掘, 而且

$$[x]_{p_i}^f \Rightarrow [x]_{p_j}^f \quad (16)$$

**证明** 设  $[x]_{p_i}^f, [x]_{p_j}^f$  是  $[x]$  上的  $S_{p-}$  上阶梯知识;  $\alpha_{p_i}^f, \alpha_{p_j}^f$  分别是  $[x]_{p_i}^f, [x]_{p_j}^f$  的属性集, 而且  $\alpha_{p_i}^f \Rightarrow \alpha_{p_j}^f$ , 或者  $\alpha_{p_i}^f \subseteq \alpha_{p_j}^f$ , 则有  $\text{card}([x]_{p_i}^f) \leq \text{card}([x]_{p_j}^f), [x]_{p_i}^f$  的粒度  $\text{GRD}([x]_{p_i}^f)$  与  $[x]_{p_j}^f$  的粒度  $\text{GRD}([x]_{p_j}^f)$  满足  $\text{GRD}([x]_{p_i}^f) \leq \text{GRD}([x]_{p_j}^f)$ , 或者  $[x]_{p_i}^f \subseteq [x]_{p_j}^f, [x]_{p_i}^f$  在  $[x]_{p_j}^f$  内被挖掘。

**推论 2** 若  $[x]_{p_i}^f$  是在  $[x]_{p_j}^f$  之内被挖掘的  $S_{p-}$  上阶梯知识, 则  $[x]$

关于  $[x]_{p_i}^f, [x]_{p_j}^f$  的上阶梯还原系数  $\varphi_{p_i}^f, \varphi_{p_j}^f$  满足

$$\varphi_{p_i}^f - \varphi_{p_j}^f \leq 0 \quad (17)$$

由定理 5, 推论 2 直接得到。

## 3 基于属性依赖的 $S_{p-}$ 阶梯知识挖掘及识别算法

**定义 5** 称  $g(x)$  是知识  $[x]$  的状态函数, 简称  $g(x)$  是  $[x]$  的状态, 而且

$$g(x) = \sum_{j=1}^n y_j \prod_{\substack{i=1 \\ i \neq j}}^n \frac{x-x_i}{x_j-x_i} = a_n x^{n-1} + a_{n-1} x^{n-2} + \dots + a_1 x + a_0 \quad (18)$$

这里:  $g(x)$  是过数据点  $(x_1, y_1), (x_2, y_2), \dots, (x_{n+1}, y_{n+1})$  的插值多项

式规<sup>[9]</sup> $\{y_1, y_2, \dots, y_{n+1}\}$ 是知识 $[x]=\{x_1, x_2, \dots, x_{n+1}\}$ 的特征数据集,  
 $\forall y_k \in R$  是  $x_k \in [x]$  的特征值,  $k=1, 2, \dots, n+1, R$  是实数集。

**定义 6** 称  $\rho_{p_i, p_i}^f$  是  $S_{F-}$  上阶梯知识  $[x]_{p_i}^f$  关于  $[x]_{p_i}^f$  的识别模,  
 而且

$$\rho_{p_i, p_i}^f = \frac{\|y\|_{p_i}^f}{\|y\|_{p_i}^f} \quad (19)$$

这里:  $\|y\|_{p_i}^f$  是  $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f = b_{m-1}x^{m-1} + b_{m-2}x^{m-2} + \dots + b_1 + b_0$   
 的系数构成的向量  $(b_{m-1}, b_{m-2}, \dots, b_1, b_0)^T$  2-范数,  $\|y\|_{p_i}^f =$   
 $(b_{m-1}^2 + b_{m-2}^2 + \dots + b_1^2 + b_0^2)^{1/2}$ ,  $\|y\|_{p_i}^f$  是  $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f = a_{n-1}x^{n-1} +$   
 $b_{n-2}x^{n-2} + \dots + a_1 + a_0$  的系数构成的向量  $(a_{n-1}, a_{n-2}, \dots, a_1, a_0)^T$  2-范  
 数,  $\|y\|_{p_i}^f = (a_{n-1}^2 + a_{n-2}^2 + \dots + a_1^2 + a_0^2)^{1/2}$ 。

**定理 6**(属性依信度单依赖  $S_{F-}$  上阶梯知识状态识别定理)  
 $S_{F-}$  上阶梯知识  $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f$  与  $S_{F-}$  上阶梯知识  $[x]_{p_i}^f$  的状态  
 $g(x)_{p_i}^f$  满足

$$DIS(g(x)_{p_i}^f, g(x)_{p_i}^f) \quad (20)$$

的充分必要条件是

$$[x]_{p_i}^f \Rightarrow [x]_{p_i}^f \quad (21)$$

这里:  $DIS=discernibility$

**证明**(1) 若  $[x]_{p_i}^f \Rightarrow [x]_{p_i}^f$  或者  $[x]_{p_i}^f = \{x_1, x_2, \dots, x_m\} \subseteq \{x_1, x_2, \dots, x_n\} =$   
 $[x]_{p_i}^f$ 。令  $y_{p_i}^f = \{y_1, y_2, \dots, y_m\}$  是  $[x]_{p_i}^f$  的特征值集合,  $y_j \in R$  是  $x_j \in [x]_{p_i}^f$   
 的特征值,  $R$  是实数集, 令  $y_{p_i}^f = \{y_1, y_2, \dots, y_n\}$  是  $[x]_{p_i}^f$  的特征值集  
 合,  $y_i \in R$  是  $x_i \in [x]_{p_i}^f$  的特征值;  $g(x)_{p_i}^f, g(x)_{p_i}^f$  分别是  $[x]_{p_i}^f, [x]_{p_i}^f$  的  
 状态, 而且  $g(x)_{p_i}^f = a_{m-1}x^{m-1} + a_{m-2}x^{m-2} + \dots + a_1x + a_0, g(x)_{p_i}^f = b_{n-1}x^{n-1} +$   
 $b_{n-2}x^{n-2} + \dots + b_1x + b_0$ , 因为  $m < n$ , 则有  $DIS(g(x)_{p_i}^f, g(x)_{p_i}^f)$ , 或者  
 $[x]_{p_i}^f$  关于  $[x]_{p_i}^f$  可分辨, 而且  $\alpha_{p_i}^f \Rightarrow \alpha_{p_i}^f$ 。

(2) 若  $g(x)_{p_i}^f, g(x)_{p_i}^f$  分别是  $[x]_{p_i}^f, [x]_{p_i}^f$  的状态, 而且  $DIS(g(x)_{p_i}^f,$   
 $g(x)_{p_i}^f)$ , 则  $[x]_{p_i}^f$  在  $[x]_{p_i}^f$  之内, 而且  $[x]_{p_i}^f \Rightarrow [x]_{p_i}^f; [x]_{p_i}^f \neq [x]_{p_i}^f$ 。

**推论 3**  $S_{F-}$  上阶梯知识  $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f$  关于上阶梯知识  
 $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f$  是可识别的, 必有

$$\rho_{p_i, p_i}^f \leq 1 \quad (22)$$

由 1~3 章的讨论, 得到  $S_{F-}$  上阶梯知识属性依赖顺序挖掘  
 及状态识别准则:

给定具有属性集  $\alpha$  的知识  $[x]$ , 若存在  $\nabla \alpha_i \neq \phi, i=1, 2, \dots, n,$   
 而且

$$card(\nabla \alpha_{p_i}) \leq card(\nabla \alpha_{p_{i+1}}) \leq \dots \leq card(\nabla \alpha_{p_1}) \quad (23)$$

则  $S_{F-}$  上阶梯知识  $[x]_{p_i}^f$ , 依  $p_1 \leq p_2 \leq \dots \leq p_{n-1} \leq p_n$  的顺序依次在  
 $[x]$  内被挖掘,  $[x]_{p_i}^f$  的状态  $g(x)_{p_i}^f$  关于  $[x]$  的状态  $g(x)$ , 依  $\rho_{p_i}^f - 1 \leq 0$   
 被识别。

$S_{F-}$  上阶梯知识挖掘及状态识别算法:

**输入** 信息系统  $[x]$  的数据分布与具有特征(属性)入侵的信  
 息系统  $[x]_{p_i}^f (k=1, 2, \dots, n)$  的数据分布

**输出** 被挖掘的阶梯知识  $[x]_{p_i}^f$

**步骤 1** 初始化  $k=1, find=true;$

**步骤 2** 输入信息系统  $[x]$  的数据分布;

**步骤 3** 计算  $[x]$  的状态函数  $g(x);$

**步骤 4** 具有特征(属性)入侵的信息系统  $[x]_{p_k}^f (k=1, 2, \dots,$   
 $n)$  的数据分布;

**步骤 5** 计算阶梯知识  $[x]_{p_k}^f$  的状态函数  $g(x)_{p_k}^f;$

**步骤 6** 计算上阶梯知识  $[x]_{p_i}^f$  关于  $[x]_{p_{i-1}}^f$  的识别模,  $\rho_{p_{i-1}, p_i}^f =$   
 $\|y\|_{p_i}^f / \|y\|_{p_{i-1}}^f$ , 其中  $\|y\|_{p_i}^f$  是  $[x]_{p_i}^f$  的系数构成的向量的 2-  
 范数;

**步骤 7** if  $\rho_{p_{i-1}, p_i}^f < 1$  then 输出  $[x]_{p_i}^f$  在  $[x]$  内被挖掘,  $[x]_{p_i}^f$  的状  
 态  $g(x)_{p_i}^f$  被识别

else  $find=false;$

**步骤 8**  $k=k+1;$

**步骤 9** 重复步骤 4~步骤 9, 直到  $k=n$  或者  $find=false$ 。

## 4 结束语

$S_{F-}$  粗集是改进了  $S-$  粗集被提出的,  $S_{F-}$  粗集中没有改变  
 $S-$  粗集中的等价关系,  $S_{F-}$  粗集与  $S-$  粗集相比, 不仅具有动态  
 性而且具有个随机特征, 正是这些特征为研究动态信息系统提  
 供了诸多方便, 因为一般的信息系统都具有动态性又具有随机  
 性, 实际的信息系统具有动态特性和随机性, 构成信息系统的  
 知识  $[x]$  ( $R-$  元素等价类  $[x]$ ) 也具有动态特性及随机性。利用单  
 向  $S_{F-}$  粗集, 变异单向  $S_{F-}$  粗集的关系, 给出  $S_{F-}$  上阶梯知识的  
 依信度生成, 给出知识状态识别及其算法的讨论, 给出一些基  
 本结果, 这些结果对于挖掘-识别动态信息系统中的知识(知识  
 状态)研究, 有较好的应用前景。

## 参考文献:

- [1] Shi Kai-quan. S-rough sets and its applications in diagnosis-recog-  
 nition for disease[C]//IEEE Proceedings of the 1st International  
 Conference on Machine Learning and Cybernetics, 2002: 50-54.
- [2] 史开泉, 崔玉泉. S-粗集和它的两类基本形式[J]. 计算机科学, 2004  
 (10A): 24-27.
- [3] Pawlak Z. Rough sets[J]. International Journal of Computer and In-  
 formation Sciences, 1982, 11: 341-356.
- [4] 刘保仓, 史开泉. S-粗集的信度特征[J]. 山东大学学报: 理学版, 2006,  
 41(5): 26-32.
- [5] 刘保仓, 卢昌荆. 副集的随机特征[J]. 山东大学学报: 理学版, 2007,  
 42(2): 77-82.
- [6] 刘保仓, 张环理. 单向变异 S-粗集的信度特征[J]. 南阳师范学院学  
 报, 2006, 5(9): 11-13.
- [7] 刘保仓, 刘若慧. 双向变异 S-粗集的信度特征[J]. 海南师范学院学  
 报: 自科版, 2006, 19(2): 116-119.
- [8] 刘若慧, 刘保仓. S-粗集的  $F_p-$  分解与  $\bar{F}_p-$  还原[J]. 天中学刊, 2007,  
 22(2): 17-19.
- [9] 刘保仓, 刘若慧. 函数 S-粗集的随机刻画[J]. 信阳师范学院学报: 自  
 科版, 2007, 20(2): 143-146.
- [10] 史开泉, 刘保相. S-粗集与动态信息处理[M]. 北京: 冶金工业出版  
 社, 2005.