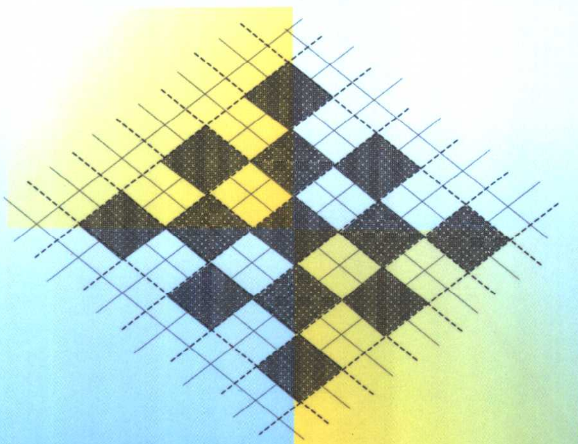
S-粗集与动态信息处理

S-CUJI YU DONGTAI XINXI CHULI

史开泉 刘保相 著



冶金工业出版社

内 容 简 介

本书共分8章，主要内容包括：Z. Pawlak粗集与应用、S-粗集、S-粗集与它的遗传、S-粗集与它的记忆、函数S-粗集、模糊粗糙集与模糊决策、双枝模糊决策集对分析、不完备信息系统集对S-粗集模型和参考文献。本书适合于工程系统中的科技人员参考，也可供正在攻读工程各个领域的硕士学位、博士学位的研究生阅读使用。

图书在版编目（CIP）数据

S-粗集与动态信息处理/史开泉，刘保相著.
—北京：冶金工业出版社，2005.2
ISBN 7-5024-3757-6

Ⅰ.S…　Ⅱ.①史…　②刘…　Ⅲ.数值计算
Ⅳ.0241

中国版本图书馆CIP数据核字（2005）第048710号

出版人　曹胜利（北京沙滩嵩祝院北巷39号，邮编100009）
责任编辑　杨盈园　美术编辑　李　心
责任校对　王永欣　李文彦　责任印制　牛晓波
北京市铁成印刷厂印刷；冶金工业出版社发行；各地新华书店经销
2005年2月第1版，2005年2月第1次印刷
850mm×1168mm　1/32；8.75印张；233千字；267页
26.00元

冶金工业出版社发行部　电话：(010)64044283　传真：(010)64027893
冶金书店　地址：北京东四西大街46号(100711)　电话：(010)65289081

引　言

在经济、金融、工程的众多领域中，人们经常接收到这样的两类信息，这些信息不能用精确的信息集合表示。从它们的特征上看，一类信息是粗糙的，另一类信息是模糊的。因为不能用有效的数学方法表示这两类信息，使得人们对它们的认识能力与使用能力降低，人们为此而丢失了重要的信息资源。在计算机与计算技术广泛使用的今天，因为没有合适的数学方法，利用计算机去识别这些信息遇到了困难，一些重要的应用研究因此而搁浅。L. A. Zadeh 教授 1965 年提出模糊集（Fuzzy Sets），给出模糊集的一般性研究；L. A. Zadeh 教授的工作给人们研究模糊信息和它的特性提供了理论支持；由此，对模糊信息的理论研究与应用研究得到广泛的开展并取得了很多优秀成果。Z. Pawlak 教授 1982 年提出粗集（Rough Sets），给出粗集的一般性研究；Z. Pawlak 教授的工作给人们研究粗信息和它的特性提供了理论支持；由此，人们对粗信息的理论研究与应用研究产生了浓厚的兴趣，研究成果引人注目。模糊集已成为模糊系统研究的理论基础，粗集已成为粗系统研究的理论基础。粗信息、模糊信息同属于不确定信息领域中的两类不确定信息，这两类不确定信息在系统管理、系统识别、系统稳定、系统决策及金融系统、投资系统中表现得非常活跃。如何有效地利用这两类信息资源研究我们面前的问题是本书的主题。本书涉猎的研究思想和讨论的主题是：

利用 Z. Pawlak 教授的工作，2002 年史开泉教授提出 S-粗集（Singular Rough Sets）。S-粗集具有两类形式：单向 S-粗集

(One Direction Singular Rough Sets) 和双向 S-粗集 (Two Direction Singular Rough Sets)，给出 S-粗集的一般性讨论。因为 S-粗集具有动态特性，对于动态数据挖掘，动态知识发现和系统动态粗特性的研究，S-粗集提供了理论支持。利用 Z. Pawlak 粗集，S-粗集对系统中的规律挖掘及规律发现的研究遇到了困难，这是因为 Z. Pawlak 粗集，S-粗集都是以 R-元素等价类 $[x]_R$ 定义的。什么是规律？函数就是规律。为了能够把潜藏在系统中的某个规律挖掘出来，能够发现隐藏在系统中的某个规律，必须把 S-粗集给出进一步讨论，2005 年史开泉教授提出函数 S-粗集 (Function Singular Rough Sets)。函数 S-粗集具有两类形式：函数单向 S-粗集 (Function One Direction Singular Rough Sets)，函数双向 S-粗集 (Function Two Direction Singular Rough Sets)。对函数 S-粗集给出一般性讨论，函数 S-粗集对系统中的规律挖掘、规律发现、规律生成、系统的规律特征的研究提供了理论支持。在系统中对一个数据的挖掘，一个知识的发现固然有重要的理论与应用价值，但我们应该看到，在一个系统中去寻找（挖掘、发现）不被我们事先知道的规律更重要，特别是对于变化莫测的金融系统、风险投资系统、系统故障预警分析系统更是如此。本书 1 ~ 5 章，对 S-粗集，函数 S-粗集的结构、特性、应用进行了讨论，给出了 S-粗集与 Z. Pawlak 粗集、函数 S-粗集与 Z. Pawlak 粗集之间关系。这些讨论开启了读者思考问题的思路，启迪了读者的创新能力，拓宽了读者对粗集认识的视角。从 S-粗集、函数 S-粗集的结构不难看到：Z. Pawlak 粗集分别是 S-粗集、函数 S-粗集的特例。

应该给读者回答一个问题：在 Z. Pawlak 粗集 $(R_-(X), R^-(X))$ 中，集合 $X \subset U$ 是静态的。在工程、经济、金融诸多系统中，人们遇到的集合 $X \subset U$（信息集合）大多是动态的。为了找回集合的动态特性并把动态特性返还给静态集合

$X \subset U$，这里用“Singular”一词表示这种“返还”。因为这个原因，才有本书的S-粗集、函数S-粗集这两个概念。

在L. A. Zadeh教授的工作基础上，史开泉教授分别于1998年提出双枝模糊集、2003年提出双枝模糊决策-识别。这些研究的直接背景是：一个成熟的商人，为了换取投资的回报，他要考虑哪些因素（构成因素集X^{+}）支持投资回报，哪些因素（构成因素集X^{-}）抑制投资回报，经过盘算，决定是投资还是不投资。显然，这是一个正、反权衡考虑$X^{+} \cup X^{-}$上的双枝模糊决策。利用L. A. Zadeh模糊集研究这类问题遇到了困难，这是因为在L. A. Zadeh模糊集中，隶属函数$A(x)$定义在［0，1］上。容易看到：双枝模糊集比L. A. Zadeh模糊集具有更宽广的应用空间，特别是在经济、金融、管理、系统识别等多个应用系统中。

我们面前的实际系统，往往粗信息、模糊信息相互交织，互相掺杂。在本书的6~8章中，把粗集与模糊集、S-粗集与模糊集进行交叉、渗透、互补共享，给出了一些既有理论性又有应用性的重要讨论。这些讨论无疑给应用工作者带来一些新的思考问题的方法和求解问题的思路，提高读者综合运用多个边缘学科解决问题的能力。

本书由山东大学数学与系统科学学院史开泉教授与河北理工大学理学院刘保相教授合作撰著。史开泉教授撰著1~5章，刘保相教授撰著6~8章。全书的取材是两位作者近几年已发表和一些待发表的学术论文。本书仅是对S-粗集、函数S-粗集研究的开始，书中的内容具有良好的后继理论研究和应用研究的前景。或许利用本书中的讨论，读者将获得一些新的、富有创意的结果。我们祈盼着这些新结果的问世。书中涉猎的研究领域和研究思想，对于工程系统中诸多领域的应用工作者具有重要的参考价值，对于那些正在攻读工程各个领域的硕士学位、博士学位的年轻一代，为完成他们的学

业具有启迪性。

本书在撰著过程中，得到了我的学生：胡海清博士、王洪凯博士、李健博士、何童博士、张萍硕士、崔明辉硕士及河北理工大学理学院的张春英副教授的热情帮助，是他们不辞劳苦地完成本书的组稿和校对，向他们表示感谢！本书的出版，山东大学数学与系统科学学院、河北理工大学的有关领导给予了热情的支持，这里向他们表示深深的感谢！冶金工业出版社为本书的问世，提供了多方面的支持与帮助，这里向他们致谢！

因为时间仓促，加上两位作者的学识浅陋，书中的不足与错误，敬请同行给予批评指正。

史开泉　刘保相

2004 年 12 月

目　录

第 1 章　Z. Pawlak 粗集与应用

我们先看一个通俗的例子.

图 1.1 是一块长方形的牛皮，它们由许多块小长方形组成.

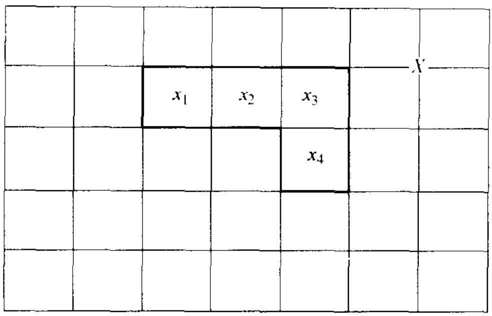


图 1.1　集合 X 的边界规则

如果图 1.1 中的每一块长方形的牛皮可做 1 双皮鞋，图 1.1 中粗线所包围的牛皮可做 4 双皮鞋，它们分别用 x_1,x_2,x_3,x_4 表示,如图 1.1 所示. 所做成的皮鞋用集合 X 表示,则有 $X=\{x_1,x_2,x_3,x_4\}$. 这里的集合 X 是我们在数学分析，高等代数，高等数学等课程中经常遇到的，它是一个精确集. 换句话说，粗线所包围的牛皮能做而且只能做 4 双皮鞋，所用的皮革不多不少. 在生活与实际中，我们见到的牛皮是否都是方方正正的? 回答是否定的. 我们给出图 1.2.

图 1.2 是一块自然形状的牛皮.

如果制作皮鞋的用皮尺寸不变，图 1.2 中曲线围成的牛皮（白色方块）只能做 4 双皮鞋，它们用 x_1,x_2,x_3,x_4 表示（长方形），如图 1.2 所示；图中带有阴影的部分是做 4 双皮鞋剩下来

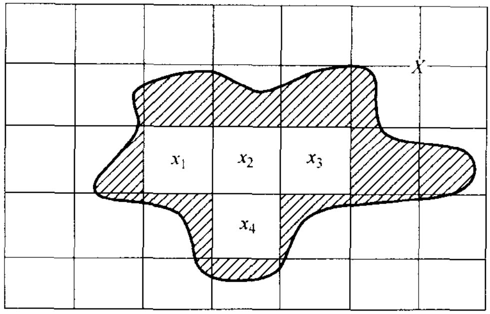


图1.2 集合 X 的边界非规则

的牛皮（俗称下脚料）．人们自然想到，阴影部分可以通过拼接的方式，使它成为长方形的牛皮，它们也可以做皮鞋使用．例如，阴影部分通过拼接，大约可做2.53双皮鞋．因此，图1.2中的曲线边界围成的牛皮大约可做6.53双皮鞋．

比较图1.1和图1.2，容易得到这样的事实：图1.1能用精确的整数表示皮鞋的数目（4双）；图1.2不能用精确的整数表示皮鞋的数目（6.53双）．图1.1能用普通集（精确集）表示，或者 $X = \{x_1, x_2, x_3, x_4\}$；图1.2不能用普通集（精确集）表示．这个通俗的例子，一般人都能接受．

如果我们把图1.1或图1.2中的每一个长方形看成是一个由有限点（元素）构成的 R- 元素等价类 $[x]$，则图1.1可以用 R- 元素等价类 $[x]$ 精确表示，图1.2不可以用 R- 元素等价类 $[x]$ 精确表示．等价类是数学中的一个普通概念．

对于图1.2，如何用集合的概念来描述，或者，用集合的概念如何去表达图1.2？波兰数学家 Z. Pawlak 教授对此给出讨论，于1982年提出粗集（Rough Sets）[34]的概念；由此，数学中又增加了一个新的理论与应用分支——粗集理论及其应用．在最近几年中，粗集成为国际国内研究的热点，有很多重要的理论与应用研究问世．

1.1 Z. Pawlak 粗集与它的结构

设 U 是一个有限论域,X 是 U 上的集合,$X \subset U$,R 是 U 上的等价关系,$[x]$ 是 R- 元素等价类，见图 1.3.

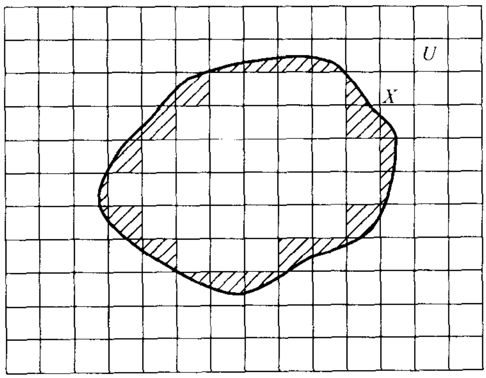


图 1.3 图中的每一个小方块是 R- 元素等价类$[x]$，
阴影中的白色方块是 $X \subset U$ 的下近似 $R_-(X)$

定义 1.1.1 称 $R_-(X)$ 是集合 $X \subset U$ 的下近似，而且

$$R_-(X) = \cup [x]_R = \{x \mid x \in U, [x]_R \subseteq X\} \quad (1.1.1)$$

定义 1.1.2 称 $R^-(X)$ 是集合 $X \subset U$ 的上近似，而且

$$R^-(X) = \cup [x]_R$$

$$= \{x \mid x \in U, [x]_R \cap X \neq \phi\} \quad (1.1.2)$$

定义 1.1.3 由 $R_-(X)$,$R^-(X)$ 构成的集合对,称作 $X \subset U$ 的 R- 粗集,简称 $X \subset U$ 的粗集，而且

$$(R_-(X), R^-(X)) \quad (1.1.3)$$

定义 1.1.4 称 $B_{nR}(X)$ 是 $X \subset U$ 的 R- 边界 ，而且

$$B_{nR}(X) = R^{-}(X) - R_{-}(X) \tag{1.1.4}$$

例 1.1　设论域 $U = \{x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$，$U$ 上的 R- 元素等价类 $[x]_1, [x]_2, [x]_3, [x]_4, [x]_5$，而且

$$\begin{aligned}
[x]_1 &= \{x_0, x_1\} \\
[x]_2 &= \{x_2, x_6, x_9\} \\
[x]_3 &= \{x_3, x_5\} \\
[x]_4 &= \{x_4, x_8\} \\
[x]_5 &= \{x_7, x_{10}\}
\end{aligned} \tag{1.1.5}$$

取 U 上的子集 $X = \{x_0, x_3, x_4, x_5, x_8, x_{10}\} \subset U$，则有 $X \subset U$ 的下近似 $R_{-}(X)$ 和上近似 $R^{-}(X)$，而且

$$\begin{aligned}
R_{-}(X) &= \cup [x]_R = [x]_3 \cup [x]_4 \\
&= \{x_3, x_4, x_5, x_8\}
\end{aligned} \tag{1.1.6}$$

$$\begin{aligned}
R^{-}(X) &= \cup [x]_R = [x]_1 \cup [x]_3 \cup [x]_4 \cup [x]_5 \\
&= \{x_0, x_1, x_3, x_4, x_5, x_7, x_8, x_{10}\}
\end{aligned} \tag{1.1.7}$$

$X \subset U$ 的粗集 $(R_{-}(X), R^{-}(X))$ 是

$$\begin{aligned}
&(R_{-}(X), R^{-}(X)) \\
&= \{\{x_3, x_4, x_5, x_8\}, \{x_0, x_1, x_3, x_4, x_5, x_7, x_8, x_{10}\}\}
\end{aligned} \tag{1.1.8}$$

1.2　集合 X 的下近似与上近似关系[36]

给定集合 $X, Y \subset U$，$R_{-}(X), R^{-}(X)$ 分别是 $X \subset U$ 的下近似和上近似；$R_{-}(Y), R^{-}(Y)$ 分别是 $Y \subset U$ 的下近似，上近似；则有

(1) $R_{-}(X) \subseteq X \subseteq R^{-}(X)$ 　　(1.2.1)

(2) $R_{-}(\phi) = R^{-}(\phi) = \phi$ (1.2.2)

$R_{-}(U) = R^{-}(U) = U$

(3) $R^{-}(X \cup Y) = R^{-}(X) \cup R^{-}(Y)$ (1.2.3)

(4) $R_{-}(X \cap Y) = R_{-}(X) \cap R_{-}(Y)$ (1.2.4)

(5) $X \subseteq Y \Rightarrow R_{-}(X) \subseteq R_{-}(Y)$ (1.2.5)

(6) $X \subseteq Y \Rightarrow R^{-}(X) \subseteq R^{-}(Y)$ (1.2.6)

(7) $R_{-}(X \cup Y) \supseteq R_{-}(X) \cup R_{-}(Y)$ (1.2.7)

(8) $R^{-}(X \cap Y) \subseteq R^{-}(X) \cap R^{-}(Y)$ (1.2.8)

(9) $R_{-}(\sim X) = \sim R^{-}(X)$ (1.2.9)

(10) $R^{-}(\sim X) = \sim R_{-}(X)$ (1.2.10)

(11) $R_{-}(R_{-}(X)) = R^{-}(R_{-}(X)) = R_{-}(X)$ (1.2.11)

(12) $R^{-}(R^{-}(X)) = R_{-}(R^{-}(X)) = R^{-}(X)$ (1.2.12)

证明:

(1) 若 $x \in R_{-}(X)$,则$[x] \subseteq X$;而 $x \in [x]$, 因此

$$x \in X \text{ 而且 } R_{-}(X) \subseteq X$$

若 $x \in X$,因为 $x \in [x] \cap X$,则$[x] \cap X \neq \phi$, 因此

$$x \in R_{-}(X) \text{ 而且 } X \subseteq R_{-}(X)$$

(2) 因为空集包含在每一个集合中, $R_{-}(\phi) \subseteq \phi$ 而且 $\phi \subseteq R_{-}(\phi)$, 所以

$$R_{-}(\phi) = \phi$$

假设 $R^{-}(\phi) \neq \phi$. 则存在 x,而且 $x \in R^{-}(\phi)$, 则有

$$[x] \cap \phi \neq \phi$$

因为 $[x] \cap \phi = \phi$, 与假设矛盾, 所以

$$R^{-}(\phi)=\phi$$

由 (1), $R_{-}(U)\subseteq U$,若使 $U\subseteq R_{-}(U)$,对于 $x\in U$,因$[x]\subseteq U$,则 $x\in R_{-}(U)$, 所以

$$R_{-}(U)=U$$

由 (1), $R^{-}(U)\supset U$,但 $R^{-}(U)\subseteq U$ 所以

$$R^{-}(U)=U$$

(3) 若 $x\in R^{-}(X\cup Y)$, 当且仅当$[x]\cap(X\cup Y)\neq\phi\Rightarrow([x]\cap X)\cup([x]\cap Y)\neq\phi\Rightarrow[x]\cap X\neq\phi\vee[x]\cap Y\neq\phi\Rightarrow x\in R^{-}(X)\vee x\in R^{-}(Y)\Rightarrow x\in R^{-}(X)\cup R^{-}(Y)\Rightarrow R^{-}(X\cup Y)=R^{-}(X)\cup R^{-}(Y)$

(4) 若 $x\in R_{-}(X\cap Y)$,当且仅当$[x]\subseteq X\cap Y\Rightarrow[x]\subseteq X\wedge[x]\subseteq Y\Rightarrow x\in R_{-}(X)\cap R_{-}(Y)\Rightarrow R_{-}(X\cap Y)=R_{-}(X)\cap R_{-}(Y)$

(5) 当且仅当 $X\cap Y=X$,由(4)得到 $R_{-}(X\cap Y)=R_{-}(X)$. 因为 $R_{-}(X)\cap R_{-}(Y)=R_{-}(X)\Rightarrow R_{-}(X)\subseteq R_{-}(Y)$

(6) 当且仅当 $X\cup Y=Y\Rightarrow X\subseteq Y\Rightarrow R^{-}(X\cup Y)=R^{-}(Y)$. 由(3) 知 $R^{-}(X)\cup R^{-}(Y)=R^{-}(Y)\Rightarrow R^{-}(X)\subseteq R^{-}(Y)$

(7) 因为 $X=X\cup Y$ 而且 $Y\subseteq X\cup Y\Rightarrow R_{-}(X)\subseteq R_{-}(X\cup Y)$,又因为 $R_{-}(Y)\subseteq R_{-}(X\cup Y)\Rightarrow R_{-}(X)\cup R_{-}(Y)\subseteq R_{-}(X\cup Y)$

(8) 因为 $X\cap Y=X$ 而且 $X\cap Y=Y\Rightarrow R^{-}(X\cap Y)\subseteq R^{-}(X)$,又因为 $R^{-}(X\cap Y)\subseteq R^{-}(Y)\Rightarrow R^{-}(X\cap Y)\subseteq R^{-}(X)\cap R^{-}(Y)$

(9) 当且仅当 $[x]\subseteq X,[x]\cap(\sim X)=\phi\Rightarrow x\,\bar{\in}\,R^{-}(\sim X)$ 而且 $x\in R_{-}(X)$, $\Rightarrow x\in\sim R^{-}(\sim X)\Rightarrow R_{-}(X)=\sim R^{-}(\sim X)$

(10) 在(9)中用 $\sim X$ 代替 X,则有 $R^{-}(X)=\sim R_{-}(\sim X)$

(11) 由 (1) 得到: $R_-(R_-(X)) \subseteq R_-(X), X \in R_-(X) \Rightarrow [x] \subseteq X \Rightarrow R_-([x]) \subseteq R_-(X)$, 因为 $R_-([x]) = [x] \Rightarrow [x] \subseteq R_-(X)$ 而且 $X \in R_-(R_-(X)) \Rightarrow R_-(X) \subseteq R_-(R_-(X))$. 由(1) 得到 $R_-(X) \subseteq R^-(R_-(X))$, 当 $X \in R^-(R_-(X)) \Rightarrow [x] \in R_-(X) \neq \phi$,存在 $Y \in [x], Y \in R_-(X) \Rightarrow [y] \subseteq X$,因此$[x] = [y] \Rightarrow [x] \subseteq X$而且$x \in R_-(X) \Rightarrow R_-(X) \subseteq R^-(R_-(X))$

(12) 若$x \in R^-(R^-(X)), [x] \cap R^-(X) \neq \phi$,对于某些$y \in [x], y \in R^-(X) \Rightarrow [y] \cap X \neq \phi$,而且$[y] = [x] \Rightarrow [x] \cap X \neq \phi \Rightarrow x \in R^-(X) \Rightarrow R^-(X) \supset R^-(R^-(X)) \Rightarrow R_-(R_-(X)) = R^-(R_-(X)) = R_-(X)$

由 (1) 得到$R_-(R^-(X)) \subseteq R^-(X)$,若$R_-(R^-(X)) \supset R^-(X)$ 成立,当$x \in R^-(X), [x] \cap X \neq \phi \Rightarrow [x] \subseteq R^-(X)$. 因为,若$y \in [x]$,则$[y] \cap X = [x] \cap X \neq \phi$,即$y \in R^-(X) \Rightarrow x \in R_-(R^-(X))$,因此$R_-(R^-(X)) \supset R^-(X)$.

1.3 知识与依赖属性的知识发现

Z. Pawlak 粗集是以 R- 元素等价类$[x]$ 定义的,R 是属性集. 例如,属性 α_1 = 红色,α_2 = 甜味,属性集 $R = \{\alpha_1, \alpha_2\}$;具有属性 α_1, α_2 的苹果 x_1, x_2, x_3, x_4 构成关于属性 α_1, α_2 的 R- 元素等价类 $[x]_{(\alpha_1,\alpha_2)} = \{x_1, x_2, x_3, x_4\}$;元素 x_1, x_2, x_3, x_4 关于属性 α_1, α_2 不可分辨, 记作

$$\underset{\alpha_1,\alpha_2}{\mathrm{IND}}(\{x_1, x_2, x_3, x_4\}) \tag{1.3.1}$$

因此, 等价关系 R 称作不可分辨关系.

如果在属性集 R 中再增加一个属性 α_3 = 产地山东,则有属性集,$R = \{\alpha_1, \alpha_2, \alpha_3\}$,$R$- 元素等价类$[x]_{(\alpha_1,\alpha_2,\alpha_3)} = \{x_2, x_3\}$, 或者

$$\underset{\alpha_1,\alpha_2,\alpha_3}{\mathrm{IND}}(\{x_2, x_3\}) \tag{1.3.2}$$

如果属性集 $R = \{\alpha_1, \alpha_2, \alpha_3\}$ 中删除属性 α_2, α_3,则有属性集