

文章编号 1000-5013(2003)01-0033-06

粗糙集理论在模糊隶属函数 辨识中的应用

方 千 山

(华侨大学信息科学与工程学院, 福建 泉州 362011)

摘要 提出应用粗糙集理论辨识模糊隶属函数的可行性. 它对机械化立窑水泥煅烧过程操作工的经验知识和专家知识进行综合考虑, 归纳为知识系统决策表. 根据粗糙集理论, 求出决策表的最小约简算法. 依此属性, 重新划分模糊空间, 并定义边界以确定隶属函数. 经模型有效性确认, 表明该方法是有用的.

关键词 粗糙集理论, 隶属函数, 辨识

中图分类号 TB 112 TB 114.2 TQ 172.6

文献标识码 A

粗糙集(RS, Rough Sets)理论^[1]是一种刻画不完整性和不确定性的数学工具. 它能有效地分析和处理不精确、不一致和不完整等各种不完备信息, 并从中发现隐含的知识, 揭示潜在的规律, 因而具有较强的实用性^[2]. 从其诞生到现在虽然只有十几年的时间, 但已经在模式识别、决策分析、专家系统及控制等领域取得显著成果. 目前, 国际上对于 RS 理论及其应用的研究已成为智能领域中一个新的学术热点, 但在国内有关应用成果的报导不多. 本文尝试应用粗糙集理论, 通过对现场相关数据和专家经验知识的采集, 有效地分析和处理这些不精确、不完备的信息. 从而, 实现隶属函数的优化, 并在实验中检验其有效性, 仿真结果说明该方法是切实可行的.

1 粗糙集理论基本思想

1.1 知识与分类

在 RS 理论中, 知识被认为是一种对对象进行分类的能力, 并使用属性(Attribute)及其值(Value)来描述论域中的对象. 一个知识表达系统定义为

$$S = (U, C, D, V, f). \quad (1)$$

在式(1)中, U 是对象的集合; $C \cup D = A$ 是属性集合, 子集 C 和 D 分别称为条件属性和结果属性; $V = \bigcup_{a \in A} V_a$ 是属性值的集合, V_a 表示属性 $a \in A$ 的范围; $f: U \times A \rightarrow V$ 是一个信息函数, 它指定 U 中每一个对象的属性值. 设一个对象的论域 U , 对任何子集 $X \subseteq U$ 称为 U 中的一个概念, U 中的任何概念称为关于 U 的对象知识. 这样, 知识库表达了一个或一组智能机构的各种

基本分类方式, 它构成所属的定义与环境或其本身的关系的基础构件. 当用 R 表示论域 U 上的等价关系, 并记为 U/R , 则 U/R 为 R 的所有等价类族. 若 $P \subset R$, 且 $P \sim Q$, 则 P (P 中全部等价关系的交集) 也是一等价关系. 它称为 P 上的不可分辨关系 (Indiscernibility Relation), 且记为 $\text{ind}(P)$, 有

$$X_{\text{ind}(P)} = X_R, \quad R \subset P. \quad (2)$$

不可分辨关系概念是 RS 理论的基础, 它揭示出论域知识的颗粒状结构.

1.2 边界与粗糙度

1.2.1 边界 RS 理论中的不确定性和模糊性是一种基于边界的概念, 即一个模糊的概念具有模糊的边界. 每一个不确定概念由一对称为上近似、下近似的精确概念来表示. 设给定知识库 $P = (U, R)$. 对于每个子集 $X \subset U$ 和一个等价关系 $R \in \text{ind}(P)$, 可以根据 R 的基础集合描述来划分集合 X , 有

$$\left. \begin{aligned} R_-(X) &= \{Y \in U/R \mid Y \subseteq X\}, \\ R^-(X) &= \{Y \in U/R \mid Y \cap X \neq \emptyset\}, \\ R_{NR}(X) &= R^-(X) - R_-(X). \end{aligned} \right\} \quad (3)$$

在式(3)中, $R^-(X)$ 和 $R_-(X)$ 分别称为 X 的 R 上近似 (Upper Approximation) 和 R 的下近似 (Lower Approximation), $R_{NR}(X)$ 称为 X 的边界 (Boundary Region). 图 1 描述了粗糙集的基本概念. 可见, $R_-(X)$ 包含给定集合中的所有基本的集合, $R^-(X)$ 包含给定集合元素中所有基本集的最小集合. 显然, 当 $R_{NR}(X) = \emptyset$ 时, X 是一个确定的概念. X 对象所组成的最大的集合, 称为 X 的正区 (Positive Region), 记为 $\text{POS}(X)$; 不属于对象所组成的集合, 称为 X 的负区 (Negative Region) 记 $\text{NEG}(X)$.

1.2.2 粗糙度 集合的不确定性是由于边界的存在而引起的. 集合的边界域越大, 其精确性越低. 为了更准确地表达这一点, 引入粗糙度概念. 令 card 为一个求集合成员个数的函数, 称为集合的基数或势 (Cardinality). $X \subset U$ 则 U 中的一个子集合 X 与被关系 R 划分后的 U 的相对程度可定义为粗糙度 (Roughness), 即

$$P_R(X) = 1 - \frac{\text{card}(R_-(X))}{\text{card}(R^-(X))}. \quad (4)$$

显然, $P_R(X) = 0$ 时, 称集合 X 相对 R 是清晰的; $0 < P_R(X) < 1$, 称集合 X 相对 R 是粗糙的. 由此可见, RS 与概率论和模糊集合论不同. 不精确性的数值不是事先假定的, 而是通过表达

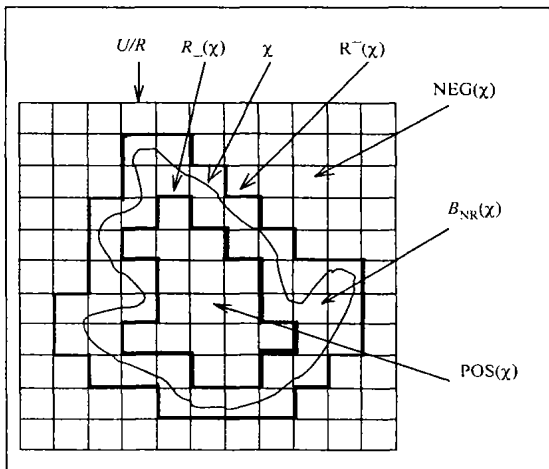


图 1 粗糙集的基本概念

知识不精确性的概念, 近似计算得到这样不精确性的数值表示, 它是有限知识的结果^[6]. 这里, 我们不需要用一个机构来指定精确的数值去表达不精确的知识, 而且采用最优概念 (分类) 来处理不精确性的数值特性来表示概念的精确度.

1.2.3 知识约简 RS 理论在知识表达系统的基础上定义了约简与核的概念, 它是 RS 理论

的精华. 约简(Reduct) 定义为不含多余属性, 并保证分类正确的最小条件属性集. 约简的不可省略关系的集合, 称为核(Core) . 它一般是通过对决策表(Decision Table) 中的属性值的处理来加以实现的.

综上所述, RS 理论的基本框架^[6]经简要归纳, 认为它以不可分辨关系划分所研究论域的知识, 形成知识的表达系统. 然后, 利用上、下近似描述对象, 通过知识约简, 从而获得最简知识的表达.

2 隶属函数辨识

在模糊控制系统中, 隶属函数用来表示确定量的特征函数. 它将过程操作的模糊语言值, 转化成控制的精确值的中间环节. 由于每个人在专家知识、实践经验、判断能力等各方面的差异, 即使对同一模糊概念的认识和理解也有差异, 因此隶属函数的确定又带有一定的主观性. 正因为观念上的模糊性, 对同一个模糊概念, 不同的人会使用不同的确定隶属函数的方法, 建立不完全相同的隶属函数. 从统计学的角度分析, 正态分布的隶属函数更能体现人的直觉推理方式. 因为人们在实践中对客观事物的认识和判断往往用正态分布的思维特点, 故目前在工程上模糊控制设计大都采用正态分布的隶属函数. 常见的基于正态分布的隶属函数, 有三角形隶属函数、梯形隶属函数、高斯形隶属函数和钟形隶属函数.

2.1 立窑水泥煅烧工艺及操作^[6]

水泥立窑是一个连续进出料的高温热工设备, 其中理化反应复杂, 热工参数离散, 是一个非线性、多变量、大滞后、强干扰的复杂系统. 它难以获得精确的数学模型, 长期以来都依赖于人工操作. 立窑水泥生产, 从生料成球喂入窑内至熟料由窑底卸出, 经历了 8.5 ~ 12.0 m 的高度. 按其设计产量计算, 一般下沉速度为 $1.2 \sim 1.3 \text{ m} \cdot \text{h}^{-1}$. 生料球一般是全黑料或半黑料, 入窑后自上而下运动, 而窑底鼓入的高压空气则由下而上, 二者相对运动. 随着料球向下运动, 煤在料球中由表及里燃烧, 其温度渐渐升高. 根据沿窑高度分布的温度和进行的理化反应, 按窑高可分为预热带、高温带、冷却带. 预热带和高温带, 一般在 1.5 ~ 2.0 m 左右. 熟料在这一高度的时间内要求基本形成, 其中最关键的是高温带反应, 其温度为

$$1\ 300 \quad 1\ 500 \quad 1\ 300 \quad .$$

此时, 部分物料出现熔融, 在上部物料的挤压下, 粘结成厚实的烧结层形成所谓的“底火”. 当料球进入“底火”, 其表面温度达 1 300 左右, 开始并逐步完成了 CaS 吸收 CaO 而形成 CaS 的反应. 立窑要稳质高产, 必须保持窑高温带的厚度, 确保窑“底火”以群体形式均匀地向下沉落. 可见, 立窑内它具有形状一定、但边界不明、温度分布不等的传热过程. 由于底风的作用高温料层向上发展, 又随卸料过程而向下运动, 因此难以用一精确的数学方程来描述. 但作为窑面操作工, 在长期的实践中积累大量的经验知识, 摸索并总结出一套切实可行的操作方法. 一般地, 在模糊控制器的设计中, 模糊规则是在总结人工操作经验的基础上经优化处理而形成的, 而隶属函数则取决于设计者的经验. 这样设计出的模糊控制器与该领域专家的经验有密切的关系. 工业过程控制工况复杂多变, 如何从复杂的实际问题中发现蕴含的模糊知识来修正专家的经验不足, 是人们一直为之探讨的问题.

2.2 知识系统的表达与分析

我们将水泥立窑操作工的的经验归纳成为一个知识表达系统. 在系统知识中, 论域 U 有

104 个对象、条件属性 $C= \{a, b, c, d\}$ 、决策属性 $D= \{e, f\}$. 条件属性与属性值的对应表, 如表 1 所示.

表 1 属性与属性值对应表

属 性	属 性 值				
	1	2	3	4	5
a -燃烧区的温度/()	1 300 ~ 1 350	1 350 ~ 1 400	1 400 ~ 1 450	1 450 ~ 1 500	-
b -燃烧区火焰	深红色	鲜红色	深桃红色	浅桃红色	淡红色
c -物料烧结情况	细粒+ 粉末	细颗粒	粗颗粒	粗颗粒+ 结块	
d -物料颜色	第 1 班经验	第 2 班经验	第 3 班经验		
e -经验知识类型	操作工	专家			
f -隶属函数类型	三角形	梯形	高斯形	钟形	

这样, 我们可以建立一个有关确立隶属函数的知识表达系统的决策表, 如表 2 所示. 那么, 应用 RS 理论的来对该决策表进行简化, 具体

表 2 确立隶属函数知识
表达系统决策表

有如下 5 个步骤. (1) 进行条件属性的简化, 将决策表中信息相同的对象及其信息删除, 只留下其中一个. (2) 删除多余的属性, 即对每一决策规则进行冗余属性值的简化. (3) 求出最小约简. (4) 求出最小解的约简算法. (5) 据此算法的属性值重新划分模糊空间, 并依边界定义确定隶属函数. 表 3 即为决策表的最小约简, 其最小解的约简算法为

U	a	b	c	d	e	f
1	3	2	2	2	2	4
28	3	2	2	3	2	3
46	4	3	3	3	2	2
68	4	3	3	2	2	1
104	3	3	2	2	2	4

$a_3d_2 \Rightarrow e_2f_4,$ $a_3d_1 \Rightarrow e_2f_4,$
 $a_2 \Rightarrow e_1f_4,$ $c_3d_3 \Rightarrow e_2f_3,$
 $c_3 \Rightarrow e_2f_2.$

由此可见, 在隶属函数的知识表达系统决策中, 主要采用高斯形隶属函数(F). 据此, 重新划分模糊空间, 以粗糙度来辩识模糊隶属函数. 其辩识后的隶属函数, 如图 2 所示.

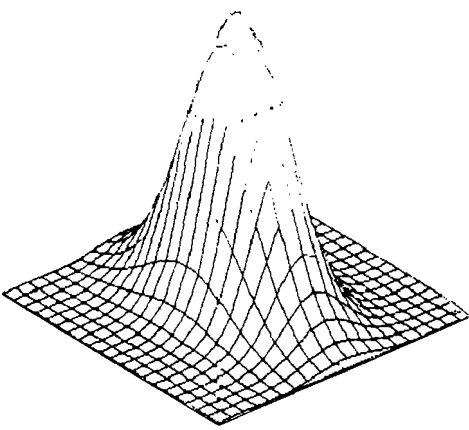
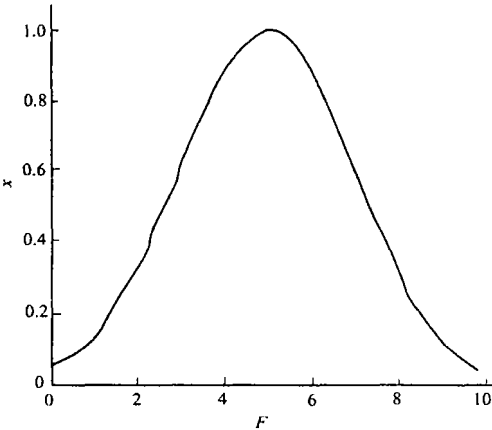


图 2 辩识后的隶属函数

表 3 隶属函数知识表达系统最小约简决策表

U	1	2	3	4	5	5	6	6	7	8	8	8
a	3	3	2	2	χ	3	4	χ	χ	4	χ	4
c	χ	χ	χ	χ	2	χ	2	2	3	3	3	χ
d	2	1	χ	χ	3	3	χ	3	χ	χ	2	2
e	2	2	1	1	2	2	2	2	2	2	2	2
f	4	4	4	4	3	3	3	3	2	2	2	2

3 实验与仿真

为了检验文中提出方法的有效性,我们以二阶加纯滞后的模型来近似模拟窑炉的特性⁶⁾. 下面介绍模糊控制器的结构和参数.

两输入一输出的二维模糊控制器,取 E, EC, U 的论域为 $\{-6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6\}$. 其模糊子集为 $\{NB, NM, NS, ZO, PS, PM, PB\}$. 49 条模糊控制规则,模糊决策采用 M amdani's (min-max), 而解模糊则采用 Centroid. 仿真系统框图,如图 3 所示.

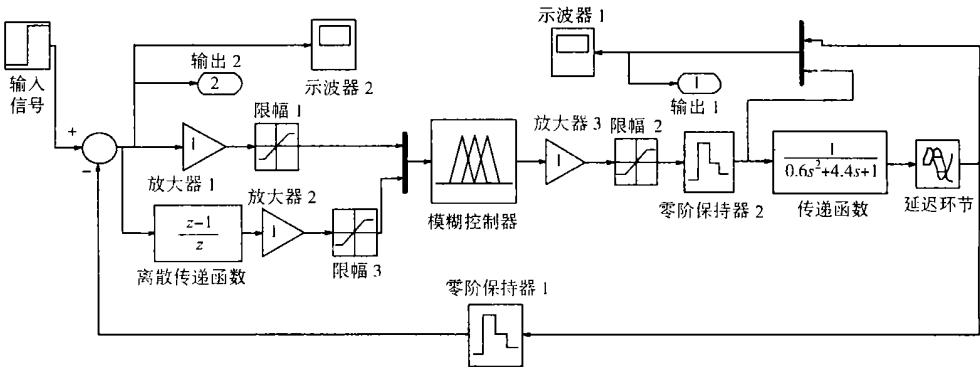


图 3 SIMULINK 环境下的系统结构

为了便于比较,还采用了三角形(trimf)、梯形(trapmf)和钟形隶属函数(gbell). 其系统响应(Y),如图 4 所示. 仿真结果表明,采用高斯隶属函数比三角形、梯形和钟形隶属函数,更符合

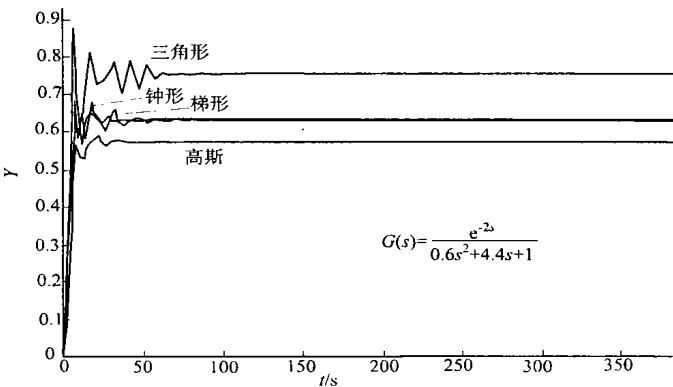


图 4 仿真系统响应

合立窑操作的知识表达,并能取得良好的效果.

4 结束语

模糊控制系统设计中,对于模糊变量隶属函数的确定和规则的选取,一直是设计中的瓶颈.迄今为止,还没有发现一种完全可靠的、系统的解决此问题的方法.近年来,人们已经引入了神经网络、进化算法等试图解决这个难题.但是,这些方法都存在一定的局限和不足.例如,训练时间长、易出现早熟和局部最优而不能保证算法的收敛等,使得理论研究和实际应用价值之间存在较大的差距.本文应用粗糙集理论,通过对机械化立窑水泥煅烧过程中操作工的经验知识和专家知识的综合,以知识系统决策表的形式,求出最小约简算法.然后依此属性值来重新划分模糊空间,并定义边界来确定隶属函数,从而实现隶属函数的优化.该方法简单易行,避免了专家知识的主观性和机器学习的盲目性,做到具体问题具体分析.仿真结果表明,应用本文提出的方法来辨识模糊隶属函数是可行的.

参 考 文 献

- 1 Pawlak Z. Rough sets[J]. International Journal of Information and Computer Science, 1982, (11): 341 ~ 356
- 2 韩帧祥, 张琦, 文福拴. 粗糙集理论及其应用[J]. 信息与控制, 1998, 27(1): 37 ~ 45
- 3 曾黄麟. 粗集理论及其应用——关于数据推理的新方法[M]. 重庆: 重庆大学出版社, 1996. 1 ~ 10
- 4 谢克明, 杨静. 粗糙集理论及其在智能控制领域的应用前景[J]. 太原理工大学学报, 1999, 30(4): 338 ~ 342
- 5 方千山, 王永初. 立窑水泥煅烧过程的控制策略[J]. 华侨大学学报(自然科学版), 2000, 21(4): 357 ~ 360
- 6 王永初, 任秀珍. 工业过程控制系统设计范例[M]. 北京: 科学出版社, 1986. 56 ~ 78

Application of Rough Set Theory to the Identification of Fuzzy Membership Function

Fang Qianshan

(College of Info. Sci. & Eng., Huaqiao Univ., 362011, Quanzhou, China)

Abstract The feasibility of applying theory of rough sets to the identification of fuzzy membership function is confirmed in this paper. A decision table of knowledge system is concluded by comprehensively considering knowledge, from experience and experts, on the operation of cement calcination in a mechanized shaft kiln. Based on the theory of rough sets, the minimal reduced algorithm of the decision table is solved. The fuzzy space is then redvided by this attribute value; and the membership function is determined by defining the boundary. This is an effective method, as confirmed by the effectiveness of the model.

Keywords theory of fuzzy sets, membership function, identification