

以层次结构形式实现可拓分类的研究

陈智斌^{1,2}, 余永权¹

(1. 广东工业大学 计算机学院 智能工程研究所, 广东 广州 510090;

2. 广东技术师范学院 计算机科学系, 广东 广州 510665)

摘要: 在讨论可拓分类特点的基础上, 给出了关于可拓集合的层次结构形式. 提出以层次结构形式实现可拓分类, 并研究了以层次结构形式实现可拓分类的方法, 最后给出了具体的分类实例.

关键词: 可拓分类; 可拓集合; 层次结构

中图分类号: N94; TP18

文献标识码: A

文章编号: 1007-7162(2005)04-0124-05

分类是人们认识事物的一个基本手段, 是进行识别、检索、策划和控制的基础. 在客观世界中, 事物总是处于不停的运动和变化中, 可拓学从事物随变换而不断开拓的角度出发, 提出了对事物进行动态分类的思想, 即可拓分类思想. 本文从研究可拓分类的特点出发, 提出以层次结构作为实现可拓分类的结构形式, 并对此作出了相关的研究.

1 可拓分类的特点

分类是指对研究对象集进行的子类划分^[1]. 可拓分类对事物进行分类, 其实现的理论工具是可拓集合. 根据可拓集合的定义^[2], 设有论域 U , k 是 U 到实域 $(-\infty, +\infty)$ 的一个映射, $T = (T_U, T_k, T_u)$ 为给定的变换, 则称

$$A(T) = \{(u, y, y') \mid u \in T_U U, y = k(u) \in (-\infty, +\infty), y' = T_k k(T_u u) \in (-\infty, +\infty)\}$$

为论域 U 上关于变换 T 的一个可拓集合, $y = k(u)$ 为 $A(T)$ 的关联函数. 由此可知, 可拓集合的划分是以变换为依据的. 在本文中假定分类所处的论域是不变的, 即设 $T_U = e$, 并且只讨论对元素的变换.

可拓分类能对事物进行一种动态的分类. 例如对于同一个人, 他的角色会因为各种变换, 包括了时间和空间的变换而改变. 如一个人白天作为职员在单位上班而晚上作为学生在学校上课这是一个普遍现象. 传统分类方法无法对这些不断运动变化的事物和现象进行动态分类. 然而由于可拓集合是根据引起事物发生变化的各种变换来对论域作出划分的, 显然这种划分具有动态性, 因此可拓分类本身能够实现动态分类.

2 可拓集合的层次结构形式

当变换 $T = e$ 时, 可拓集合 A 根据关联函数值划分为正域 A_+ 、负域 A_- 及零界 J_0 3 部分, 而当 $T \neq e$ 时, 根据元素变换后所对应的关联函数值的改变情况, 把 A 划分 5 部分, 把

$$A_+(T) = \{(u, y, y') \mid u \in U, y = k(u) \leq 0, y' = k(T_u u) \geq 0\}$$

称为 $A(T)$ 的正可拓域;

$$A_-(T) = \{(u, y, y') \mid u \in U, y = k(u) \geq 0, y' = k(Tu) \leq 0\}$$

称为 $A(T)$ 的负可拓域;

$$A_+(T) = \{(u, y, y') \mid u \in U, y = k(u) \geq 0, y' = k(Tu) \geq 0\}$$

称为 $A(T)$ 的正稳定域;

$$A_- = \{(u, y, y') \mid u \in U, y = k(u) \leq 0, y' = k(Tu) \leq 0\}$$

称为 $A(T)$ 的负稳定域;

$$J_0(T) = \{(u, y, y') \mid u \in U, y' = k(Tu) = 0\}$$

称为 $A(T)$ 的拓界。

本文以层次结构形式对上述可拓集合划分的各个部分进行组织. 如图 1 所示, 当 $T = e$ 时, 图中层次结构从根结点集合 A 中生成 A_+ 、 A_- 及 J_0 三结点集合. 这时可拓集合为静态可拓集合, 而所作的分类为静态分类. 当实际环境或条件发生变化而引起变换 T_0 后, 则可从 A_+ 、 A_- 、 J_0 中分别对应地生成结点集合 $A_+(T_0)$ 、 $A_-(T_0)$ 、 $J_0(T_0)$ 及 $A_+(T_0)$ 、 $A_-(T_0)$ 及 $J_0(T_0)$. 此时有 $T_0 u = u_0$. 若对某一结点集合中的元素再实施变换 T_1 , 即使 u_0 在变换 T_1 下得到改变, 有 $T_1 u_0 = u_1$. 这时 A 可作出进一步的划分, 其层次结构则在对应的结点下生长出新的结点集合.

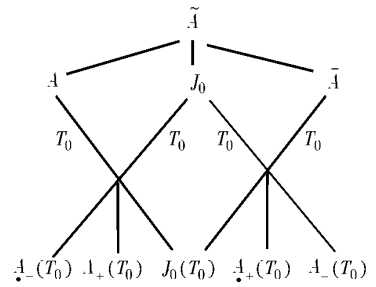


图 1 可拓集合的层次结构形式

由此可见, 可拓集合是动态变化的, 体现在不同的变换下对可拓集合的不同划分. 可拓集合的层次结构形式也是动态变化的, 体现在根据这些不同的变换对自身进行调整. 因此, 可拓分类可借助可拓集合的层次结构形式的动态变化特性以实现对事物进行动态分类.

3 以层次结构形式实现可拓分类

3.1 关联度的计算方法

有界区间 $X = \langle a, b \rangle$ 的模定义为 $|X| = |b - a|$, 则某点 x 到区间 $X = \langle a, b \rangle$ 的距^[3]为:

$$\rho(x, X) = |x - \frac{a+b}{2}| - \frac{1}{2}(b - a).$$

设特征 $c_i (i = 1, 2, \dots, m)$ 为描述某类事物 $P_j (j = 1, 2, \dots, N)$ 的其中一个特征, 有待判别对象 p , 其 c_i 的量值为 v_i . 计算关联函数值^[4]

$$k_j(v_i) = \begin{cases} -\frac{\rho(v_i, X_{0ji})}{|X_{0ji}|}, & v_i \in X_{0i}, \\ \frac{\rho(v_i, X_{0ji})}{\rho(v_i, X_{pi}) - \rho(v_i, X_{0ji})}, & v_i \notin X_{0i}, \end{cases}$$

则 $k_j(v_i)$ 从特征 c_i 的角度反映了待判别对象 p 隶属于 P_j 的程度. 其中 X_{pi} 为特征 c_i 的节域, X_{0ji} 为特征 c_i 对于类事物 P_j 的经典域. 在某些情况下, X_{0ji} 会表现为多个有界区间之和, 如 $\langle 2, 0, 3, 0 \rangle \cup \langle 4, 0, 5, 0 \rangle$, 则应分别计算并取其中最大值作为 $k_j(v_i)$. 计算 p 关于 P_j 的关联

度有: $K_j(p) = \sum_{i=1}^m \lambda_i k_j(v_i)$, 其中 λ_i 为特征 c_i 的权系数, $\sum_{i=1}^m \lambda_i = 1$.

3.2 类别的判别方法

对论域中元素的类别判别是在具体的可拓集合的层次结构形式中进行的. 层次结构中每一结点分别代表了某一类事物, 其中叶结点代表类别判别过程中最终所确定的类别. 类别判别的过程, 就是从层次结构的根结点出发, 选择分支路径, 直至某一叶结点的过程.

对于待判别对象 p , 在层次结构中某一层次选择分支路径时, 需分别计算 p 对于待选择结点的关联度. 设有待选择类事物 $P_j (j= 1, 2, \dots, N)$, 则 p 对于 P_i 的关联度为 $K_j(p)$, 若 $K_{j_0}(p) = \max K_j(p) (j_0 \in \{1, 2, \dots, N\})$, 则判 p 属于类 P_{j_0} . 若对所有 j , 有 $K_j(p) \leq 0$, 则 p 不在已知类型范围中. 若对于某变换 T , 使得有 $K_j(Tp) \geq 0$, 则认为 p 的所属类型从不在已知类型转变为属于已知类型. 此时根据变换 T 在可拓集合的层次结构中生长出对应的类事物结点, 并将 p 归于该类.

3.3 分类实例

在实际生产过程中, 常由于机器和人为原因而造成加工工件具有一定的尺寸误差, 如某工件的尺寸规格可表示为 $\Phi 50^{+0.1}_{-0.1}$. 在加工条件受到限制的情况下, 只能对工件作合格品与不合格品的分类. 设有关于论域为工件的可拓集合 A , 此时 A 的层次结构形式如图 2 所示, 其中结点以物元形式表示某类事物特征的经典域.

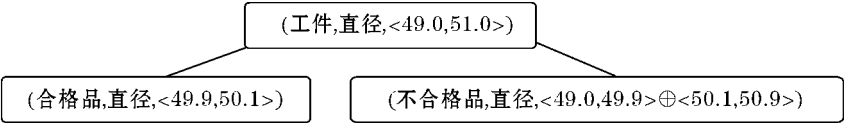


图 2 加工条件限制下的工件集合层次结构

建立关于合格品的关联函数 $k(u)$ 并且以 $k(u)$ 为准则对 A 作出划分, 显然有正域 A 代表合格品而负域 A 代表不合格品, 且零界 J_0 代表直径为 49.9 或 50.1 的工件. 为简化问题, 在本文中零界工件视为合格品工件. 设现有 5 个不同尺寸的待分类工件, 可计算出每个工件对合格品类型和不合格品类型的隶属程度, 从而对它们作出分类. 5 个工件的关联度计算结果和类型判别结果如表 1 所示. 当加工条件提高后, 可将部分的不合格品重新返工, 设有变换 T_0 , 使得 $T_0 u = u_0$, 相应地在 A 的层次结构形式中由不合格品结点生成可返工品结点及废品结点, 如图 3 所示. 可知

表 1 加工条件限制下的工件分类

	50.0	49.8	50.3	50.8	49.4
合格品	0.50	- 0.11	- 0.22	- 0.67	- 0.56
不合格品	- 0.09	0.11	0.25	0.125	0.44
所属类型	合格品	不合格品	不合格品	不合格品	不合格品

由不合格品结点生成可返工品结点及废品结点, 如图 3 所示. 可知

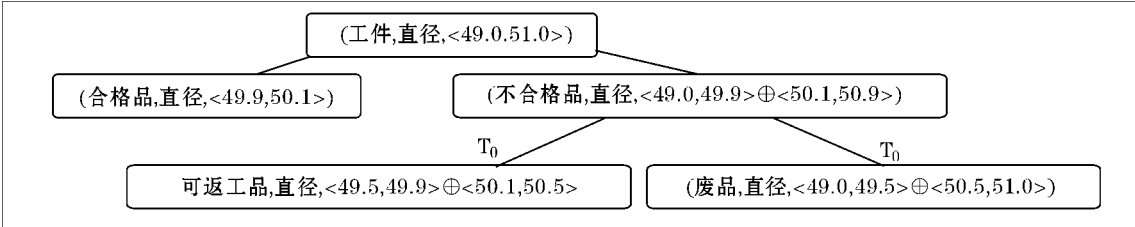


图 3 基于再加工变换 T_0 的工件集合层次结构

该两结点分别代表正可拓域 $A_+(T_0)$ 和负稳定域 $A_-(T_0)$. 对不合格品重新分类的关联度计算结果和类型判别结果如表 2 所示.

可返工品与废品是本质不同的两种不合格品, 后者在当前加工环境下并不具备条件再加工成为合格品, 而可返工品却可借助一定加工条件而重新加工成为合格品. 设现在加工环境中已具备电镀及车工条件, 对应于变换 T_1 及 T_2 , 使得 $T_1 u_0 = u_1, T_2 u_0 = u_2$, 相应地在 A 的树结构形式中由可返工品结点生成电镀工件结点及车工工件结点, 如图 4 所示. 电镀工件结点和车工工件结点代表了 $A_+(T)$ 中不同的两个部分. 对可返工品重新分类的关联度计算结果和类型判别结果如表 3 所示.

表 2 不合格品的再分类

	49. 8	50. 3	50. 8	49. 4
可返工品	0. 25	0. 50	- 0. 60	- 0. 20
废品	- 0. 27	- 0. 22	0. 40	0. 20
所属类型	可返工品	可返工品	废品	废品

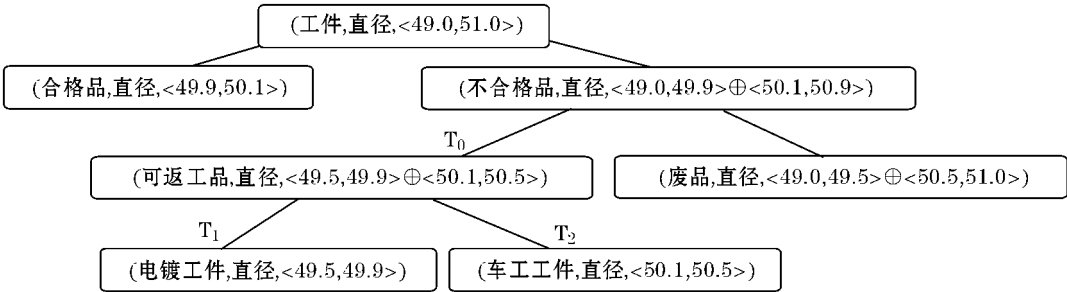


图 4 加工条件提高后的工件集合层次结构

同理, 当生产条件发生改变, 如电镀的工艺条件不允许时, 变换不成立, 则将电镀工件结点从层次结构中删去. 由此可见, 工件分类过程是一个按照实际生产条件的变化而不断调整的分类过程, 也即为一个动态分类过程. 这种动态分类的实现, 是依靠可拓集合层次结构形式中的结点生长和修剪这种动态变化来实现的.

表 3 可返工品的再分类

	49. 8	50. 3
电镀工件	0. 25	- 0. 36
车工工件	- 0. 27	0. 50
所属类型	电镀工件	车工工件

4 结束语

每一种分类思想和方法的应用是通过设计相应的分类器来加以实现的. 分类器是指这样的一种机器, 它根据外界从待分类对象中提取的与分类任务相关的特征量, 给待分类对象赋予一个类别标记^[5]. 以可拓分类思想指导设计的分类器本文称为可拓分类器. 如何有效地设计合适的可拓分类器的结构形式是实现可拓分类应用的重要环节, 有待进一步讨论和研究.

参考文献:

[1] 何斌, 张应利. 可拓集合的分类性质和信息开发的可拓方法[J]. 系统工程理论与实践, 1999, (7): 63-68.
[2] 蔡文, 杨春燕, 何斌. 可拓逻辑初步[M]. 北京: 科学出版社, 2003. 69-75.
[3] 蔡文. 物元模型及其应用[M]. 北京: 科学技术文献出版社, 1992. 22-23.
[4] 蔡文, 杨春燕, 林伟初. 可拓工程方法[M]. 北京: 科学出版社, 1997. 203-206.
[5] 吴逸飞. 模式识别——原理、方法及应用[M]. 北京: 清华大学出版社, 2002. 5.

Study on the Implementation of Extension Classification by Hierarchy Structure

CHEN Zhi-bin^{1,2}, YU Yong-quan¹

(1. Institute of Intelligent Engineering, Faculty of Computer, Guangdong University of Technology, Guangzhou 510090, China;
2. Department of Computer Science, Guangdong Polytechnic Normal University, Guangzhou 510665, China)

Abstract: On the basis of the discussion about the extension classification, the hierarchy structure of the extension set is given. In this paper, the realization of extension classification by the hierarchy structure is presented, and the method of extension classification is studied. The instance of extension classification is given.

Key words: extension classification; extension set; hierarchy structure

(上接第 123 页)

Extension and Rough Set

LU Dan, WU Meng-da, ZHANG Xue-zhi

(Dept. of Mathematics and System Science, National University of Defense Technology, Changsha 410073, China)

Abstract: In this paper, the concepts of rough set and extension set are compared. Any rough set can be expressed by means of a dependent function $U^{\rightarrow} \{-1, 0, 1\}$. It is shown that these two concepts are different, and this representation can be extended to union and intersection. At last, the concept of extension field is used to solve this problem effectively.

Key words: rough set; extension set; rough dependent function