

## 基于粗糙集理论的故障诊断系统知识库设计

刘 军, 王 萍

(浙江海洋学院 信息学院, 浙江 舟山 316004)

**摘 要:**故障诊断是与有效决策密切相关的复杂而困难的问题。粗糙集理论可以有效地分析、处理不完备信息。知识库是整个故障诊断系统的核心,利用基于粗糙集的知识约简和决策规则提取算法,将柴油机故障信息值进行约简,求出其决策规则。知识库由事实库和规则库组成。在知识库中采用链表数据结构,以数据文件形式存储,完成知识库设计的程序。采用粗糙集方法进行故障条件属性约简十分有效,得到简化的决策规则,使得知识库的设计更加方便快捷。

**关键词:**粗糙集;知识约简;知识库

**中图分类号:**TP302

**文献标识码:**A

**文章编号:**1673-629X(2008)09-0095-03

Design of Knowledge Base for Fault Diagnosis  
System Based on Rough Set Theory

LIU Jun, WANG Ping

(Information College, Zhejiang Ocean University, Zhoushan 316004, China)

**Abstract:** Fault diagnosis is a complex and difficult problem that concerns effective decision-making. Incomplete information can be effectively analyzed and processed by rough set theory. Knowledge base is the core of fault diagnosis system. A decision-rule can be deduced based on arithmetic of rough set theory, in which the fault information of diesel engine was extracted and reduced. A program was designed for knowledge base composed of fact-base and rule-base, in which the data was arranged in linking table and saved in the form of document. It is very efficiency to reduce the attributes of fault condition based on the method of rough set. The simplified decision-making rules were carried out and the design for knowledge base became more convenient.

**Key words:** rough set; knowledge reduction; knowledge base

## 0 引 言

船舶动力设备故障诊断中,故障征兆之间的关系不确定。在故障诊断系统中知识获取是一个难题,粗糙集理论能有效地分析和处理不精确、不完整、不一致数据,可以发现数据间隐藏的关系,提取有用信息,简化信息处理。粗糙集理论主要思想是在保持分类能力不变的前提下,通过对知识的约简导出概念的分类规则,其中知识被认为是一种对抽象或现实的现象进行分类的能力<sup>[1]</sup>。粗糙集理论已被应用在许多方面,如知识发现、概念分类、机器学习和数据挖掘<sup>[2-4]</sup>。

## 1 粗糙集描述和全描述及决策规则形式化

在信息系统  $S = (U, AT)$  中,任意  $a \in A, A \subseteq$

$AT$ , 任意  $v \in V_a$ , 属性-值  $(a, v)$  称为  $A$  的基本元。所有的  $A$  的基本元或它们的“ $\wedge$ ”连接称为  $A$  描述<sup>[5]</sup>。

在  $S = (U, AT)$  中,任意  $A \subseteq AT$ , 设  $DES(A) = \{t \mid t \text{ 是 } A \text{ 的描述, 且 } \|t\| \neq \emptyset\}$ 。对任意的  $t \in DES(A)$ , 如果  $A(t) = A$ , 则  $t$  称为  $A$  的全描述。记  $FDES(A) = \{t \in DES(A) \mid t \text{ 是 } A \text{ 的全描述}\}$ 。在不完备信息系统  $S = (U, AT)$  中,  $A \subseteq AT, t \in FDES(A)$ , 存在  $t' \in DES(AT)$ , 满足  $t' \geq t, \|t'\| = \|t\|$ , 而且  $\forall t'' > t', \|t''\| \neq \|t\|$ , 则  $t'$  称描述  $t$  的约简<sup>[5]</sup>。任何决策表可以看作为如下形式的(广义)决策规则集:

$\Lambda(c, v) \rightarrow \vee(d, w)$ , 其中  $c \in AT, v \in V_c, w \in (V_d \wedge (c, v))$  称为规则的条件部分,  $\vee(d, w)$  称为规则的决策部分。

在不完备决策表  $T = (U, AT \cup \{d\})$  中, 决策规则  $t \rightarrow s$ , 其中  $t = \Lambda(a, v), s = \vee(d, w)$ , 如果  $\|t\| \neq \emptyset$ , 且  $\|t\| \subseteq \|s\|$ , 则称决策规则  $t \rightarrow s$  是真的, 记作  $t \Rightarrow s$ 。

收稿日期:2007-12-30

基金项目:国家自然科学基金项目(60373078);浙江省教育厅资助项目(20050113, 20050125)

作者简介:刘 军(1968-),女,黑龙江密山人,副教授,研究方向为智能系统。

在不完备决策表  $T = (U, AT \cup \{d\})$  中,  $t = \wedge(a, v), s = (d, w)$ 。

(1) 如果决策规则  $t \rightarrow s$  是真的, 则称决策规则  $t \Rightarrow s$  是确定性规则;

(2) 如果决策规则  $t$  是  $T$  中的一个广义约简描述, 则称决策规则  $t \rightarrow s$  是广义决策规则;

(3) 如果决策规则  $t$  是  $T$  中的一个  $u$  约简描述, 则称决策规则  $t \rightarrow s$  是  $u$  决策规则。

## 2 船舶动力设备故障诊断系统知识库设计

系统知识库由事实库和规则库两部分组成。事实库为从船舶柴油机采集的数据及其特征值的集合; 规则库为经过粗糙集推导后的决策规则的集合。面向对象的方法根据同一类对象的特点, 提炼成类(class), 即对实体对象进行抽象。一个抽象类能反映同一类型的所有实体对象的信息结构, 以系统中柴油燃油系故障为例, 对产生式规则的结构进行描述:

```
class rule Object { 属性: 规则条件: 规则结论: 操作: 规则推理: }
```

根据这个描述结构, 柴油机燃油系故障现象的规则:

```
rule Object rule GuanDao { 属性:
  规则条件:  $P_0 \wedge V$  低于正常值,  $I_1$  等于正常值;  $P_m \wedge V$  低于正常值,  $I_1$  等于正常值;
  规则结论: 燃油系故障现象(1, 发生);
  操作: 规则推理; }
```

本系统中知识获取采用基于粗糙集描述的理念实现, 系统中集合以链表结构存储, 以样本集为例, 其结点结构为:

```
struct set
{ char sample[20];
  struct set * next; };
```

根据粗糙集基于描述的支持集、全描述集及约简的定义, 已求得描述  $t$  的支持集  $T$ , 对  $t$  约简, 约简的程序设计流程如图 1 所示。

根据基于描述的决策规则形式表示的定义, 求出确定性决策规则, 广义决策规则及真规则, 规则集的结点结构为:

```
struct cer
{ char describe[500];
  char descision[100];
  struct cer * next; };
```

由于船舶动力设备各部件故障属性不同, 其参数多样, 为便于实现存储及推理, 在本系统中采用数据文件的形式。根据船舶柴油机故障的主要分布, 故障类别分为燃油喷射部件、气缸部件、增压器及空冷器、“三

漏”故障。故障状态信息和基本事实的数据结构保持一致, 故障状态信息是用户要进行故障诊断的部件状态信息, 基本事实中存储了上述部件正常运转时的参数值。故障属性根据专家经验获得, 气缸部件由测得的波形经离散处理后得主要属性频域波形合成、中心频率、时域波形合成、非周期组合、时间系列变化及时间系列峰值。增压器及空冷器主要属性为机油压力、冷却水压力、柴油压力、扫气压力、气缸冷却水进出口温度和活塞冷却水进出口温度。“三漏”故障属性值为油、气、水的泄量, 机械零件温度、振动。每一部件数据结构中第一个属性值为其部件编号, 获取故障属性值后存入数据文件时, 以“#”为一个部件参数值结束标识符。诊断故障信息与基本事实文件结构相同, 信息获取后存储于数据文件, 此文件为一临时文件, 当一次故障诊断结束时将删除。

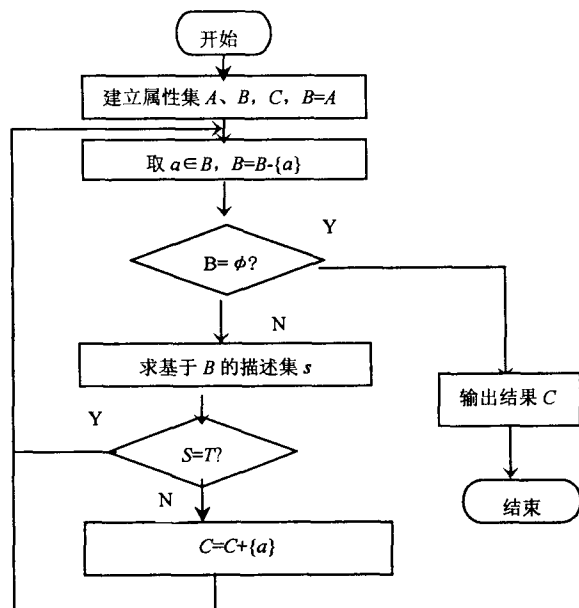


图 1 基于描述的知识约简

将描述及其对应的决策值采用字符数组方式存储, 整个确定性决策规则集为链表数据结构, 在求解过程中决策值存放在字符型数组中。根据故障求得全部确定性规则集后, 将其保存在数据文件中, 在每条规则后使用“#”作为标识符。规则库采用规则链形式, 每个规则条件往往是由几个事实组成, 在每个规则中可使用条件支链。在知识库中, 可以使用类对象的形式存储所有记录, 生成面向对象数据库。规则文件结构为规则号、事实、结论, 每条规则的结束仍以“#”为标识符。应用基于粗糙集描述理论的支持集、全描述集、属性约简及确定性决策规则集、广义决策规则集、真规则集提取算法<sup>[6]</sup>进行相关知识处理, 基本思想与图 1 相似, 流程图在此略。知识推理方法是知识处理中的

重要部分,在本系统的研究中,知识推理的算法为:

- (1) 根据部件编号打开相应的规则文件。
- (2) 从第一条规则起将规则中的前提事实与故障事实比较,记录匹配的个数。
- (3) 匹配个数与规则前提事实数比较,相等则输出结论转第五步,判断结束;否则记下匹配比率,转下条规则。
- (4) 判断规则文件是否结束,是求匹配比率最大值,得出可能性判断结论,如无匹配输出无判断结论;否转第三步。
- (5) 诊断推理结束。

以船舶燃油系统故障为例,应用粗糙集理论求出其决策规则。

船舶燃油系统故障决策表如表 1 所示,应用粗集描述理论约简后的决策表如表 2 所示。表中 E 代表等于正常值,H 代表高于正常值,L 代表低于正常值。

表 1 燃油系统故障决策表

U	a	b	c	d	e	f
1	E	H	H	H	H	0
2	E	H	H	H	H	0
3	E	H	H	H	H	0
4	E	H	H	H	H	0
5	L	L	E	H	H	1
6	L	L	L	H	H	1
7	L	L	E	*	H	1
8	L	L	E	H	H	1
9	L	L	L	H	L	2
10	L	L	L	E	L	2
11	L	L	E	E	L	2
12	L	L	L	E	L	2

表 2 燃油系统故障约简决策表

U	A(b)	e	f
1-4	E(H)	H	0
5-8	L	H	1
9-12	L	L	2

进行决策规则提取,结果为:

确定性决策规则:

$r_1:(a,E)\rightarrow(f,0)$

$r_2:(b,H)\rightarrow(f,0)$

$r_4:(e,L)\rightarrow(f,2)$

U 决策规则为:

$r_5:(a,L)\wedge(e,L)\rightarrow(f,2)$

$r_6:(b,L)\wedge(e,L)\rightarrow(f,2)$

$r_7:(b,L)\wedge(e,H)\rightarrow(f,1)$

$r_8:(a,L)\wedge(e,L)\rightarrow(f,1)$

由上述 4 条 U 决策规则可得两条真规则:

$r_9:(a,L)\wedge(e,L)\vee(b,L)\wedge(e,L)\rightarrow(f,2)$

$r_{10}:(a,L)\wedge(e,H)\vee(b,L)\wedge(e,H)\rightarrow(f,$

1)

由此得决策规则集为:

$(a,L)\wedge(e,H)\rightarrow(f,1)$

$(e,L)\rightarrow(f,2)$

或者

$(b,H)\rightarrow(f,0)$

$(b,L)\wedge(e,H)\rightarrow(f,1)$

$(e,L)\rightarrow(f,2)$

3 结束语

应用粗糙集描述的理论进行知识约简并导出决策规则,建立故障诊断系统的知识库。粗糙集理论在故障诊断知识的表示、属性约简及决策规则提取中非常有效,特别是当所面对的系统特征参数值不能完全、确定取得时,即不完备信息系统,粗糙集理论对不完备信息系统约简更简单、易行,使得故障诊断系统中知识获取更有效。采用链表数据结构实现知识的约简及知识集合的建立,便于系统知识库的建立及维护。

参考文献:

[1] 李干目,戚 湧,张 宏,等.基于粗糙集神经网络的网络故障诊断新方法[J]. 计算机研究与发展,2004,41(10): 1696-1702.

[2] Khoo L P, Tor S B, Zhai L Y. A rough-set based approach for classification and rule induction[J]. International Journal of Advanced Manufacturing Technology,1999,15:438-444.

[3] Tsumoto S. Automated Induction of Medical Expert System Rules from Clinical Databases based on Rough Set Theory[J]. Journal of Information Sciences,1998,112: 67-84.

[4] 王伟辉,蔡台明. 船舶泵系统信号处理与诊断模拟[J]. 海运研究学刊,2003,15:95-122.

[5] Wu Wei-Zhi, Mi Ju-Sheng, Zhang Wen-Xiu. A new rough set approach to knowledge discovery in incomplete information systems[C]// IEEE Proceedings of the Second International Conference on Machine Learning and Cybernetics. Xi'an:[s. n. ],2003:1713-1718.

[6] 顾沈明,吴伟志,高 济. 不完备信息系统中知识获取算法[J]. 计算机科学,2005,32(9):149-152.