

获取知识的一种新方法——粗糙集(Rough Set)

董彩凤, 王天宇

(哈尔滨工业大学 航天学院, 黑龙江 哈尔滨 150001)

摘要: 旋转机械故障诊断的一个困难问题是诊断规则的获取。提出获取知识的一种方法——粗糙集(RS), RS能自动地从旋转机械的大量信息中有效地获取诊断知识,并能减少误诊与漏诊现象。文中介绍了RS的原理与方法并给出应用实例。

关键词: 故障诊断; 旋转机械; 粗糙集(RS)

中图分类号: O231 文献标识码: B

1 引言

在旋转机械故障诊断中,影响诊断结果的关键技术之一是诊断规则。无论从理论上还是从实际中提炼出来的诊断规则,在实际应用中都有一定的局限性。由于环境的不断变化,设备的诊断规则不会去自适应。研究一种使得设备在不同的应用条件下能灵活的进行诊断的方法,一直是人们追求的目标。在故障诊断中应用智能技术是提高诊断精度的有效措施。专家系统是目前广泛应用的智能诊断技术之一,但是,专家系统的“瓶颈”是“知识”,自动获取知识是提高诊断精度的重要途径。粗略集(RS)能自动地从大量数据中提炼与获取知识,并有自学习的功能,因而能适应环境的变化,提高诊断结果的精确性,还扩大了专家系统的应用范围。RS在故障诊断中的应用,将推进诊断技术的研究与发展。

2 RS理论概述

RS是1982年由波兰科学家Z. Poulund提出来的。直到1992年,Z. Poulund出版了RS的应用专集才引起世界的广泛注意。RS在处理模糊性、不确定性知识,分析不精确性、不一致性和冗余数据方面都

有其独到之处。

2.1 粗糙集的概念

粗糙集是刻划不确定性数据的有力工具。在RS理论中,“知识”被认为是一种根据特征属性对现实中抽象的对象进行分类的能力。RS理论把用于分类的知识嵌入集合体本身,是经典集合论的延拓。集合的划分依赖于所掌握关于论域(研究对象的全体“ U ”)的知识,是相对的。论域 U 被一族等价关系 R 分割为互不相交的子集 $E_i, i = 1, 2 \dots n$,称 E_i 为基本等价族。 E_i 中每个对象之间为不可分辨关系,不可分辨的关系就是等价的关系。如果再进一步分类就要加入一定的条件。RS将集合概念用不可分辨的对象组成的集合来表示,这种对象是组成论域知识的颗粒。RS的这一概念深刻揭示出知识的颗粒状结构,表示成为:

定义:在信息系统 S 中,对于一个属性子集 E ,定义二元关系 $\text{ind}(E)$ 为不可分辨关系或等价关系(如果 u, v 属于 E ,只考虑 E 时, u, v 无法区别)。

2.2 集合的下逼近、上逼近、边界区与粗糙集函数

设 $X, Y \subseteq U$ 的集合, R 是定义在 U 上的等价关系,集合 X 关于 R 的下逼近为

$$R_*(X) = U\{Y \in U/R; Y \subseteq X\} \quad (1)$$

$R_*(X)$ 是根据现有知识判断肯定属于 X 的对象组成的最大的集合,称为正区,记为 $\text{POS}(X)$ 。而根据现有知识判断肯定不属于 X 的对象组成的集合称为负区,记为 $\text{NEG}(X)$ 。

集合上逼近为:

$$R^*(X) = U\{Y \in U/R; Y \cap R \neq \Phi\} \quad (2)$$

其中: Φ 表示为空集。 $R^*(X)$ 是由所有集合 X 相交非空的等效类的并集,是那些可能属于 X 的对象组

成的最小集合。集合边界区定义为:

$$\text{Bnd}(X) = R^*(X) - R_*(X) \quad (3)$$

如果 $\text{Bnd}(X)$ 是空集, 则称 X 关于 R 是清晰的; 反之, 如果 $\text{Bnd}(X)$ 不是空集, 则称 X 为关于 R 的粗糙集(RS), 见图 1。 $\text{Bnd}(X)$ 是逼近最后结果的集合。

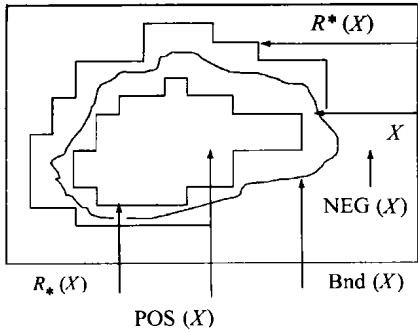


图 1 粗略集示意图

上下逼近及边界区等概念刻划了一个边界含糊集合的逼近特性。粗糙度定义为:

$$\alpha_R(X) = \frac{\text{Card}[R_*(X)]}{\text{Card}[R^*(X)]}$$

Card 为求集

合成员个数的函数。

显然 $0 \leq \alpha_R(X) \leq 1$, 如果 $\alpha_R(X) = 1$, 则称集合 X 相对于 R 是清晰的; $\alpha_R(X) \leq 1$, 则称集合 X 是粗糙的。 $\alpha_R(X)$ 可认为是在等效关系 R 下逼近集合 X 的精度。

2.3 粗糙集约简与核的概念

2.3.1 粗糙集的约简

人类要认识自然界, 分析自然界, 就必须对自然界的现象进行分类, 于是产生了多种分类形式。由于各种分类的概念不同, 分类的程度也不同。在描述同一现象的各种信息中, 有些是相近的, 有些是重复的, 这样就存在着模糊性。如: 在旋转机械故障诊断中, 描述故障信息表与决策表中, 存在着很多近似与冗余。在这些信息中可以进行约简, 从中找出描述现象的简单而骨干的数据, 使之尽量真实的揭示现象的实质。

如果抽象地看, 把各种描述看成是一种记录关系。记录关系用真假来表示, 用一个二维信息表, 即可明确的表达现象之间的关系。

假设下述各表为二维表格, 列为对象的种类, 行为条件属性, 最后一列为决策属性。

表 1 条件属性与决策属性

	a_1	a_2	a_3	a_4	D
1	0	1	0	0	1
2	0	0	1	1	1
3	0	0	0	0	0
4	1	0	0	1	0
5	0	1	0	1	1

表 2 一次约简

	a_2	a_3	a_4	D
1	1	0	0	1
2	0	1	1	1
3	0	0	0	0
4	0	0	1	0
5	1	0	1	1

表 3 二次约简

	a_2	a_3	D
1	1	0	1
2	0	1	1
3	0	0	0
4	0	0	0
5	1	0	1

表 4 三次约简

	a_2	a_3	D
1	1	0	1
2	0	1	1
3	0	0	0

从表 1 看, a_1 与 a_4 列的第 4 行对结论为 0 无影响。而对 1、2、5 行的结论为 1 来说, a_1 列也是无影响的。因此, a_1 列的条件属性是多余的, 可以删去。约简为表 2。对于表 2 来说, a_4 列第 1 行为 0, 结果为 1; 第 5 行为 1, 结果也为 1, 自相矛盾, 故 a_4 也可以删去。约简为表 3。在表 3 上, 3 与 4 行是相同的, 1 与 5 行也是相同的, 都取其中一行。因此, 最后只剩下表 4。从而得出的结论: a_2 为 1 与 a_3 为 0 结果为 1; 或 a_3 为 1 与 a_2 为 0 结果也为 1。从一个复杂的信息表中可以得出简单的结论, 这就是粗集优越之处。

2.3.2 粗集的核(Core)

核属性是描述对象的条件属性不可缺少的属性。在条件属性中核以外的属性可以约简, 核属性应取为无法约简的属性, 核是所有属性的公共部分。在表 1 中的核即是 a_2, a_3 。

核的属性在集合中的定义:

设 S 为一信息系统, U 为论域, $S = \{U, D, B, C\}$

定义: (1) 若 B 中每个属性关于 D 是不可缺少的, 则 B 关于 D 是正交的; (2) 若 $B \subset C, B$ 与 D 正交, 且 $\text{POS}_C(D) = \text{POS}_B(D)$, 则 B 为 S 的约简; (3) 所有属性集属于 C , 而 C 是关于 D 全部约简的正交, 则称 C 为核, $\text{CORE}(C, D)$ 。

定义清楚地说明了核是正交的公共部分, 核可以是空集, 核在约简中起到了提高精度与加快速度的作用。

3 知识的获取规则

在旋转机械故障诊断中自动获取诊断规则是非常困难的工作。通常是通过实验或者从已经发生的事实中积累经验总结规则。应用粗集技术可以在大量的数据中自动进行约简处理,进而得到简练的规则。这种方法弥补并解决了专家系统知识的“瓶颈”问题。

根据粗集的约简原理,可以进行属性约简与值约简,约简的过程就是自动获取规则的过程。故障诊断的约简过程主要是一个值约简过程,就是除掉冗余,保留条件独立的属性的过程。

4 旋转机械故障诊断自动获取知识的例子

以某机组的工况为例,研究规则获取的步骤(转子碰磨故障示例)。

表5 记录的信息表

	a_1	a_2	a_3	a_4	D
1	0	1	0	0	0
2	1	0	1	0	0
3	0	1	1	1	1
4	1	1	1	1	1
5	0	1	1	1	1
6	1	1	0	0	0
7	1	0	0	1	0
8	1	1	0	1	1
9	1	1	0	1	1
10	1	0	1	1	1

序号列表示各类对象, $a_1 \sim a_4$ 列表示条件属性值, D 列表示决策属性值。由条件属性与决策属性,表示某一对象的特征。某一对象由条件属性进行分类。表内的数值 1 表示事件发生,数值 0 表示事件不发生。表 5 内 1~10 行表示测量的各次记录, $a_1 \sim a_4$ 列表示属性: a_1 表示振动一倍频幅值, a_2 表示振动高频幅值, a_3 表示振动的相位变化, a_4 表示轴心轨迹的变化。

从表 5 可见, 3、4、5、8、9、10 行的结论 D 为 1(发生)。1、2、6、7 行的结论 D 为 0(不发生)。第 1 行的 a_1 列为 0, 结论为 0, 第 3 行的 a_1 列为 0, 而结论为 1, 是矛盾的。第 5 行与第 6 行也存在同样问题, 故 a_1 列是矛盾的, 可以删除。

a_2 、 a_3 、 a_4 列的数据中 3 与 4 行以及 8 与 9 行是重复的, 因此, 可以去掉两行 4 与 9 行, 见表 6。

表 6 中 3 与 5 行以及 1 与 6 行也是重复的, 3 行

的结论足以代表 10 行的结论, 因此可以去掉 5、6、10 行。第 1 行与第 2 行的 a_2 、 a_3 为 0 或 1, 对结果无影响故保留一行。最后的结果记录在表 7 中。

表6 记录的信息表

	a_2	a_3	a_4	D
1	1	0	0	0
2	0	1	0	0
3	1	1	1	1
5	1	1	1	1
6	1	0	0	0
7	0	0	1	0
8	1	0	1	1
10	0	1	1	1

表7 记录的信息表

	a_2	a_3	a_4	D
1	*	0	0	0
3	1	*	1	1
7	0	0	*	0
8	1	0	1	1

从表 6 中可以看到, 第 3 与第 8 行 a_3 值为 0 或 1 对结论无影响故计为 *, 第 1 行与第 7 行的 a_2 、 a_4 值, 对结果也无影响计为 *, 见表 7。最后只剩下三条规则:

- (1) 若 a_2 、 a_4 同时发生, 则碰磨发生(从第 3 行和第 8 行看);
- (2) 若 a_4 不发生, 则碰磨不发生(从第 1 行);
- (3) 若 a_2 、 a_3 不发生, 则碰磨不发生(从第 7 行)。

* 表示发生与不发生对结果无影响。

5 结束语

粗集是处理不确定性和模糊性数据的有力工具。

- (1) 它能有效地提取最少、最佳分类特征, 运算量小精度高。
- (2) 消除数据量的冗余, 浓缩有用信息含量。
- (3) 对数据可做进一步挖掘, 提炼诊断规则, 解决专家系统的“瓶颈”问题。

参考文献:

- [1] KOMOROWSKI J, OHRN A. Modelling prognostic power of cardiac using rough sets[J]. *Artificial Intelligence in Medicine*, 1999, 15(2): 167-191.
- [2] SHUSAKU TSUMOTO. Automated extraction of medical expert system rules from clinical databases based on rough set theory[J]. *Information Sciences* 1998 112(4): 67-84.
- [3] 张国斌. 旋转机械故障诊断知识获取的理论与方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2000.
- [4] 侯祥林, 李永强, 虞和济. 神经网络快速算法及其在故障诊断中的应用[J]. *振动工程学报*, 2000, 13(增刊).

(辉 编辑)

speed-measuring pump with a floating type seal construction and reliable conclusions obtained. All the above has provided a theoretical and experimental basis for the improvement of an overspeed protection device for a main steam turbine. **Key words:** main steam turbine over-speed protection, speed-measuring pump, seal, test

寒区太阳能—土壤源热泵系统太阳能保证率的确定 = **Determination of the Solar Energy Assurance Factor of a Solar Energy - Ground Soil-source Heat Pump System in Frigid Regions** [刊, 汉] / YU Yan-shun, LIAN Le-ming (Department of Architectural Thermal Energy Engineering, Harbin Institute of Technology, Harbin, China, Post Code: 150090) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(4): 393 ~ 395

In frigid regions the use of solar energy and ground soil-source heat as a low-level heat source of heat pumps is of great significance from the perspective of environmental protection and the development and utilization of new and renewable energy sources. The authors have for the first time put forward the concept of ground-soil temperature restoration rate of the ground soil-source heat pump system under different ratios of operation/shutdown. With the above-cited restoration rate serving as an index the restoration degree of ground soil temperature field is evaluated of the ground soil-source heat pump under various operation/shutdown ratios. On this basis, the optimum operation/shutdown ratio of the said heat pump and solar energy assurance factor were identified, and thereby the capacity of solar energy heat-collection device was determined. **Key words:** frigid region, solar energy - ground soil-source heat pump, solar energy assurance factor, ground soil temperature restoration rate

大型火电站高压加热器可靠性增长模型验证及预测 = **Verification of a Reliability Enhancement Model and its Prediction for High-pressure Heaters at a Large-sized Thermal Power Plant** [刊, 汉] / CAO Xian-chang, ZHONG Zhi-qiang, JIANG An-zhong (Department of Energy Sources, Shanghai Jiaotong University, Shanghai, China, Post Code: 200240), SHI Jin-yuan (Shanghai Power Equipment Design and Research Institute, Shanghai, China, Post Code: 200240) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(4). // Journal of Engineering for Thermal Energy & Power. — 2002, 17(4): 396 ~ 398

A reliability enhancement model is proposed for high-pressure heaters at a large-sized thermal power plant along with a detailed discussion of the model parameter estimation and fitting inspection. According to the statistics data of Electric Power Reliability Management Center the model fully conforms with the reliability-enhancement change tendency of high-pressure heaters under statistical evaluation. The results of analysis also indicate that the reliability of the heaters has been gradually enhanced through the introduction of such measures as eradication of defects in product design and fabrication, enhancement of personnel quality, improvement of operational level, production facilities and operating conditions. **Key words:** high-pressure heaters, reliability, growth model, prediction

移动—流化型组合阀工作特性的研究 = **A Study Concerning the Test of the Operating Characteristics of a Shifting-fluidized type of Combination Valve and Its Model** [刊, 汉] / ZHANG Jing-yuan (Power Engineering Department, Taiyuan Electric Power High Technical School, Taiyuan, China, Post Code: 030013), YE Gui-zhen (Guangdong Provincial Electric Power Test Research Institute, Guangzhou, China, Post Code: 510006) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(4): 399 ~ 401, 413

A shifting-fluidized type of combination valve is a kind of novel non-mechanical valve for transporting solid-particle raw materials. The key technique affecting its operating characteristics involves the provision of a sort of slackening air at the bottom of the shifting bed. The present paper mainly focuses on an experimental investigation regarding the provision of the slackening air device and a mathematical model for calculating raw material flow rate. **Key words:** valve, slackening air, non-Newton fluid

获取知识的一种新方法——粗糙集(Rough Set) = **“Rough Set” as a New Method of Machine Self-learning** [刊, 汉] / DONG Cai-feng, WANG Tian-yu (Aerospace Institute under the Harbin Institute of Technology, Harbin, China, Post Code: 150001) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(4): 402 ~ 404

A difficult issue encountered during the diagnosis of rotating machine failures consists in the acquisition of diagnostic

rules. The author has come up with a method called "rough set" (or RS for short) to serve as a new method of machine self-learning. The RS can automatically acquire diagnosis knowledge from a huge quantity of rotating machine data, decrease erroneous diagnoses and phenomena of missing diagnosis. Also described are RS working principles, its operating method and practical usage examples. **Key words:** fault diagnosis, rotating machine, machine self-learning, rough set

电子束湿法烟气脱硫工艺 = **Electronic Beam and Wet Method-based Flue Gas Desulfurization Technology** [刊, 汉] / CHOU Qiao-li (Dalian Bingshan Group Co. Design and Research Institute, Dalian, China, Post Code: 116033) // Journal of Engineering for Thermal Energy & Power. —2002, 17(4). 405~407

A concept involving an electronic beam and wet method-based technology is proposed for flue gas desulfurization, which comprises three vertical, wet-type equipment items. Among them, a pressure-type fog-spray drying device is used to dry ammonium salt solution, rendering it into grain particles. It also enables flue gases containing SO₂ and NO_x to be cooled and humidified. Then, the flue gases after being cooled, humidified and electron beam radiated by a vertical cooling and radiation tower are made into aerosol micro-particles of ammonium salt to be separated by a sieve-plate type adsorbing tower and washed clean to form an ammonium salt solution. The results of a feasibility study and preliminary design of the process indicate that the latter can attain a desulfurization rate of 98%, NO_x removal rate of 70% and a dust collection rate of 99%. Moreover, the electron beam radiation dosage, circulating solution flow-rate and flue gas flow resistance are all reduced simultaneously. The technology under discussion also features a simple flow process, ease of control, a low requirement for equipment, power consumption and space, and a large particle size of by-products. **Key words:** electronic beam, wet method, flue gas desulfurization

电站锅炉掺烧煤泥可行性研究 = **Feasibility Study of Burning Peat as a Subsidiary Fuel in Utility Boilers** [刊, 汉] / DUAN Shi-fang (Xuzhou Ducheng Electric Power Co. Ltd., Xuzhou, China, Post Code: 221142) // Journal of Engineering for Thermal Energy & Power. —2002, 17(4): 408~409

A feasibility study of burning peat as a subsidiary fuel in utility boilers indicates such an approach is realistic. A normal operation of boilers with rated parameters can be achieved without any modification of related equipment and systems. This will result in sizable economic and social benefits as well as good environmental conditions. The burning of peat as an added fuel conforms to State policy regarding the comprehensive utilization of energy resources and has wide prospects of development in addition to playing an exemplary role in energy use. The author has also pointed out some issues worthy of due attention during the design and operation of boilers firing peat as a subsidiary fuel. **Key words:** boiler, peat, mixed burning, drying system

轻油预蒸发燃烧中的稳焰技术 = **Flame Stabilization Techniques during the Pre-evaporative Combustion of Light Oil** [刊, 汉] / HE Hong-zhou (Mechanical Engineering Institute under the Jimei University, Xiamen, China, Post Code: 361021) // Journal of Engineering for Thermal Energy & Power. —2002, 17(4): 410~413

The technical principles of pre-evaporative combustion of liquid fuels are described and the problem of flame stabilization during the pre-evaporative combustion of light oil is analyzed. Also highlighted are several measures currently adopted for flame stabilization during the pre-evaporative burning of light oil. **Key words:** light oil, pre-evaporation, combustion, flame stabilization

万丰热电厂2号燃油锅炉改烧水煤浆工程实例 = **The Retrofitting of No. 2 Oil-fired Boiler at Wangfeng Power Plant for Burning Water-coal Slurry** [刊, 汉] / YOU Xiao-bo QIU Zhou-wei, LIU Xue-gui, et al (Wangfeng Thermal Power Plant, Shantou, China, Post Code: 515000) // Journal of Engineering for Thermal Energy & Power. —2002, 17(4): 414~417

The specific features of and difference between oil-fired boilers and coal-fired ones are briefly described and analyzed. Meanwhile, a detailed account is given of the retrofitting of a No. 2 oil-fired boiler at Wangfeng Thermal Power Plant for burning water-coal slurry along with a description of its main operating parameters. **Key words:** oil-fired boiler, water-coal slurry, design study