

高炉专家系统的知识求精^{*}

刘金琨 王树青

(浙江大学工业控制技术国家重点实验室, 浙江 杭州 310027)

摘要 针对高炉专家系统中知识表达的特点, 探讨了高炉异常炉况判定中最常出现的规则出错类型, 给出了具体的知识求精策略并实现了相应的算法。实际应用结果表明, 文中所采用的方法具有较好的效果。

关键词 高炉 专家系统 知识库 求精

The Knowledge Base Refinement of Blast Furnace Expert System

Liu Jinkun Wang Shuqing

¹(Zhejiang University, Hangzhou 310027)

²(Northeastern University, Shenyang 110006)

Abstract This paper discusses the mistake types of the rule often happened in judging blast furnace abnormal conditions, proposes the concrete refinement tactics and realizes the relevant algorithms according to the characteristics of blast furnace conditions. The practical application indicates that the method used by this paper has good result.

Keywords blast furnace; expert system; knowledge base; refinement

1 前言

利用专家系统技术实现高炉的计算机控制, 是当今高炉计算机控制的发展方向^[1]。高炉冶炼工艺复杂, 生产过程中影响炉况的因素很多, 要实现高炉的动态控制, 高炉专家系统的预报起着重要的作用。高炉操作是一个十分复杂的过程, 铁矿石和焦炭从炉顶加入, 鼓风机由底部吹风, 这是一个非常重要的生产环节, 生铁冶炼的质量将直接影响炼钢过程与钢的质量。为了监视炉况, 高炉上设有几百个乃至上千个检测点, 从采集的数据中观察高炉内的状况^[2]。传感器采集的数据由 PLC 全部送到上位机进行处理。数据采集系统将量测参数值自动采入, 经预处理后传递给高炉专家系统用于推理使用。高炉炉况判断专家系统可分为两个主要部分: 一个是炉身料柱非正常下降——崩料(或悬料)和煤气流“管道”预报专家系统, 一个是以出铁口铁水温度为代表的热状态预测用的炉热控制专家系统。这两个子系统要求预先在几分钟到几十分钟前觉察出信息来^[3], 以便及时进行处理。因此, 高炉专家系统是基于传感器型的在线实时专家系统, 高炉专家系统是由数据采集与处理系统及专家系统相结合的产物。

高炉专家系统是具有多目标、多影响因素的系统, 高炉炉况知识具有复杂性、动态性和不确定性, 知识库质量是高炉专家系统性能是否优越的关键因素。由机器学习得到的初始知识常有些问题, 如知识不完全、知识之间不一致、有的知识不正确等^[4]。因此要进行调试、修改和补充, 即对知识库进行求精, 其目的在于能消除知识库的弱点和失误, 提高知识库的性能。实践证明^[5], 初始知识库求精后可显著地提高专家系

^{*} 本文于 1997 年 5 月 19 日收到
国家自然科学基金资助项目(批准号: 69874037)
工业控制技术国家重点实验室开放课题基金(编号: K97M 05)
清华大学智能技术与系统国家重点实验室开放课题基金资助

统的运行性能。目前,专家系统在高炉中已经得到了广泛的应用,但对高炉专家系统的知识求精问题尚未进行研究。本文针对高炉专家系统知识库的特点,具体探讨了高炉专家系统的知识求精策略及其方法。

2 高炉专家系统知识求精的理论基础

高炉经验知识多属因果性判断,这些知识用产生式规则表达比较合适。这种表达方式具有模块性、自然性、结构化和通用性的特点。在高炉专家系统中,可利用产生式规则来描述高炉专家的判断性知识和过程性知识,产生式规则的一般形式为:

IF 前提 THEN 动作 (或 结论)

在高炉专家系统中,存在下述三种类型的产生式规则:

- 1) IF A THEN B ;
- 2) IF A_1 AND A_2 AND ... A_n THEN B ;
- 3) IF A_1 OR A_2 OR ... A_n THEN B

其中 A_1, A_2, \dots, A_n 为前提, B 为结论。

在高炉专家系统中,有数百条规则,这些规则构成单个规则或规则关系树,存在于知识库中。例如,有下述规则:

IF 透气性差 AND 料速呆滞 THEN 悬料;

IF 料速增大 OR 负荷增大 THEN 向凉

因此,高炉专家系统是基于产生式规则的系统,应针对产生式规则来进行求精。

专家系统的知识求精问题,可采用如下方法^[6]:

已知: (1) 专家系统的初始知识库; (2) 专家系统的运行实例;

求解: 用已知实例检测专家系统初始知识库,并对它进行修改、删除和补充,使求精后的知识库达到预期的运行效果。

专家系统的知识求精过程主要由下列三个阶段组成:

第一阶段: 调试初始知识库,获取求精信息,即找出知识库毛病,明确知识库求精方向。

第二阶段: 对初始知识库求精。

第三阶段: 测试求精后的知识库。如已合乎要求,停止求精,否则重复做第二、三阶段,直到知识库合乎要求为止。

在高炉炉况判定中,经常出现高炉专家系统判定的结果与高炉实际出现的结论不同,说明高炉专家系统知识库中的规则有错误,这时就需要对规则进行求精。在高炉异常炉况判定中常出现的规则出错类型是漏判和误判,二者的定义如下:

1) 漏判: 应该是结论为 c 的实例,在给定前提条件下推不出 c 来。

2) 误判: 不该是结论为 c 的实例,在给定前提条件下却推出了 c 。

作者根据高炉专家系统规则的出错类型,将错误规则定义为以下几种类型:

定义 1 规则的使用次数 N ,漏判次数 M ,当 $M/N > k_1$ 时,则该规则定义为漏判规则。

定义 2 规则的使用次数 N ,误判次数 M ,当 $M/N > k_2$ 时,则该规则定义为误判规则。

其中 k_1, k_2 为常数,具体数值由专家根据经验确定。

定义 3 规则与实例的前提相同,由于规则结论可信度过大,使规则出现误判情况,则定义该规则为误判规则(强规则)。

定义 4 规则与实例的前提相同,由于规则结论可信度过小,使规则出现漏判情况,则定义该规则为漏判规则(弱规则)。

例如,现有完整的对异常炉况“悬料”的判断规则如下:

IF 料速下降 AND 风量升高 AND 炉顶煤气平均温度上升 THEN 悬料(0.80)

其中规则的前提(工艺参数)通过传感器采集来获得,工艺参数由传感器采集后,存在着不正确、不完全、不

一致等问题,需经过过程计算机进行数据预处理(包括特征参数的计算及参数的模糊处理)^[7],才能用于高炉专家系统的推理。

在高炉专家系统中,该条规则的前提是结论成立的充要条件,如果将此规则的前提增加或减少,以及对结论的可信度增加或降低,都会造成此规则的漏判或误判,以如下几条条件规则为例:

1) IF 料速下降 AND 风量升高 AND 炉顶煤气平均温度上升 AND 风温 THEN 悬料(0.80)

由于该条件规则增加了额外条件,当规则前几个条件得到满足而第四个条件“风温”得不到满足时,使本应得出“悬料”结论而未得出,造成规则漏判。避免这种情况发生的方法是去掉额外条件“风温”。

2) IF 料速下降 AND 风量升高 THEN 悬料

假如该条规则缺少了悬料判断的一个必要条件,易造成悬料判断的征兆没有得到充分满足而错误地得出“悬料”的结论,造成规则误判。避免这种情况发生的方法是加上条件“炉顶煤气平均温度”。

3) IF 料速下降 AND 风量升高 AND 炉顶煤气平均温度上升 THEN 悬料(0.60)

假如,该条规则结论的可信度过低,在同样条件下,本来发生炉况异常,应采取措施而未采取,造成漏判。同理,如果规则的可信度过高,则易造成误判。

通过以上分析,可得到以下几点结论:

造成规则漏判的原因有如下两种:1) 规则的前提过多;2) 规则结论的可信度过大。以上两种情况都会使规则的结论难于满足,从而出现规则的漏判情况。

造成规则误判的原因有如下两种:1) 规则的前提过少;2) 规则结论的可信度过小。以上两种情况都会使规则的结论易于满足,从而出现规则的误判情况。

3 高炉专家系统的知识求精方法

3.1 建立实例库

在高炉运行过程中,由于高炉参数及炉况状态的动态变化会产生大量的实例,通过实例与知识库中的规则相比较,可判断出规则的出错类型,从而对知识库进行求精。用于求精的实例应该具有准确性和典型性,错误的或不典型的实例会使知识求精发生错误^[8]。由于高炉炉况处在相对稳定的动态变化过程中,因此应对高炉最近一段时间发生的典型实例做为实例库,来对知识库进行求精。

实例库中的实例应能针对性地探测到知识库在某个特定方面是否有缺陷。本研究选择测试实例的方法是:针对当前求精目标(如“悬料”),选择若干相关的实例来对当前规则进行求精。所选择的实例应该是典型的和正确的。采用“爬山法”来处理错误的实例,即当通过知识求精得到一些规则后,用更多的相关实例来对这些规则进行检测,以淘汰由于局部实例错误而得到错误的规则。

3.2 查找有毛病的规则

实例数是正判数和漏判数之和,针对每种判断异常炉况的规则,可算出下列三个比率:

- 1) 误判率 = $\frac{\text{误判数}}{\text{实例数}}$
- 2) 漏判率 = $\frac{\text{漏判数}}{\text{实例数}}$
- 3) 规则 R 的可信度 = $\frac{\text{满足 } R \text{ 的前提且和 } R \text{ 的结论一致的实例数}}{\text{满足 } R \text{ 的前提的实例数}}$

通过上述三个比率,可判断出规则是否出错,以及规则的出错类型。

在实际高炉应用中,高炉专家系统初始知识库中含有大量规则,推理中往往运用了几条、几十条规则,当系统在实际应用中出现了错误的结论,如何对错误规则进行快速搜索,并判定出规则的出错类型、规则的使用次数和使用失败次数很重要。为此,采用了如下两种方法:

1) 要很快地找出有漏判、误判的规则,必须开发一种机制,它能在知识库中找到一些子集,只需把这些子集做为求精对象,其它的规则不必考虑,这样就大大提高了知识求精的速度和效率。为此,采用了“分而治之”的策略,即针对高炉各种异常炉况类型分别进行处理。例如,在某一时刻要对判断为“悬料”的规则进行求精,则只需考虑规则右端为“悬料”的规则,以及规则右端结论(中间结论)出现在结论为“悬料”的规则的左端者,或者以链接方式与以“悬料”为结论的规则间接有关的规则。找到相应规则后,将规则与实例库

中典型的实例相比较,从而确定规则是否有毛病。

2) 由于在高炉推理过程中规则的使用情况记录在黑板中,实例的使用情况记录在数据库中,可采用下述的跟踪比较法来判定规则出错的类型,并可判定出规则的使用次数及使用失败的次数。具体地说,就是对推理过程中所使用的规则进行分析,并在实例库中搜索导致本次结论相对应的实例,进行逐步比较,从而确定规则是漏判还是误判。

3.3 对错误规则的处理方法

设高炉异常炉况类型为 A ,根据高炉专家系统“宁肯误判也不漏判”的原则^[9],当以 A 为结论的规则漏判率等于误判率时,按漏判进行处理,则对高炉专家系统错误规则的处理原则如下:

- 1) 若 A 的漏判率大于误判率,且漏判率大于20%,则对以 A 为结论的规则作泛化处理;
- 2) 若 A 的误判率大于漏判率,且误判率大于20%,则对以 A 为结论的规则作特化处理;
- 3) 若 A 的误判率等于漏判率,且都大于20%,则对以 A 为结论的规则作泛化处理。

针对造成漏判和误判的原因,采用如下方法进行规则的泛化处理和漏判处理:

对泛化处理,减弱规则作判断时所用的条件,包括两种情况:(1)减少规则前提条件的数目;(2)增加规则的可信度。

对特化处理,增强规则作判断时所用的条件,包括两种情况:(1)增加规则前提中条件的数目;(2)降低规则的可信度。

下面分别对规则的泛化处理和特化处理进行探讨。

1) 规则的泛化处理

统计造成规则漏判的实例数,统计有多少实例因不满足多少条件而被漏判,选择漏判率 M/N 最大的规则,将漏判规则实例与该规则进行比较,分析造成漏判的主要因素,找出造成规则漏判实例最多的前提,然后以该条件作为泛化处理突破口,具体进行如下处理:

若规则 R 的漏判实例数最多,则选择 R 作为泛化处理对象。

在规则 R 的各漏判因素中,选择造成漏判实例数最多的因素做为泛化处理突破口,处理的办法是去掉规则 R 中造成漏判实例数最多的规则的前提。

重复和,直至规则 R 的前提满足充要条件。

如果规则实例的条件与规则 R 条件相同,说明不存在漏判因素,该规则为弱规则,需要增加规则的可信度。

2) 规则的特化处理

具体作特化处理时,不直接分析那些被误判的实例是由于什么原因造成的,而采用间接的方法,即分析那些被正确判断的实例所满足的条件是否被规则前提完全使用,然后对误判规则进行特化处理,具体的方法如下:

若规则 R 的误判率最大,则选择 R 作为特化处理对象。

选择与误判规则 R 具有相同条件的被正确判断的实例,并将实例的条件与规则 R 的前提进行比较,如果前者的条件数过多,具体分析造成误判实例的各因素,然后从中选择最多的因素,对误判规则 R 进行特化处理,即在规则 R 中增加该前提。

重复和,直至规则 R 的前提满足充要条件。

如果规则实例的条件与规则 R 条件相同,说明该规则为强规则,需要降低误判规则 R 的可信度。

3.4 检测求精后的知识库

利用统计学方法对求精后的知识库进行检测,针对高炉专家系统知识库存在多目标、实例量大、实例随高炉运行动态变化的特点,对高炉专家系统知识库进行检测的具体方法是:在实例库中随机抽出与当前求精目标(如“悬料”)相关的两个子实例集 N_1 和 N_2 ,用 N_1 做训练集对知识库求精,把求精后的知识库用 N_2 做测试集进行检测,将最大失误差保存起来(即最大误判率和最大漏判率两者之较大者)。再以 N_2 为训练集,对知识库进行求精,用 N_1 做测试集对求精后的知识库检测,并将其最大失误差与第一个最大失误差

求平均值。用此方法选取第二组 N_1 和 N_2 ,又可得到一个最大失误率。如此进行多次,最终可求得一个总的最大失误率,用该失误率就可以评价求精后的知识库水平。本研究针对高炉实际,选取阈值0.20,如果求精后的知识库最大失误率大于该阈值,则需要继续对知识库进行求精,否则知识库求精结束。

3.5 建立知识库求精模型

通过上述分析后,可建立高炉专家系统的知识库求精模型,如图1所示。该模型的基本思想是:通过适当而准确的设计并选取典型实例,发现知识库中的错误,对错误规则进行求精,从而提高知识库的性能。

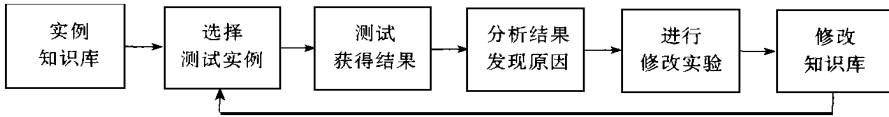


图1 高炉专家系统知识库求精模型

4 结论

本文针对高炉专家系统知识库的特点,探讨了它的知识求精策略及其规则的泛化处理和特化处理方法,并给出了查找有毛病规则及检测求精效果的方法,从而建立起知识库求精模型。将本文所述方法应用于首钢炼铁厂实际高炉专家系统中,取得了良好的效果。

高炉炉况知识具有动态性和不确定性,因此,高炉专家系统的知识库必须不断改进,才能对操作环境的变化作出迅速而准确的反映,提高系统的命中率和实时响应能力。为此,高炉专家系统的知识求精应随着高炉专家系统的运行而不断进行。

参考文献

- 1 Amano S, Takarabe T, Nakamori T, et al Expert system for blast furnace operation at Kimitsu works ISIJ International, 1990, 30(2): 105~ 110
- 2 Hirata T, Yamamura K, Morimoto S, et al Blast furnace operation system using neural networks and knowledge-base Proceedings of the sixth International Iron and Steel Congress, Nagoya, ISIJ, 1990: 23~ 27
- 3 Matsuda K, Tamura N, Konishi M, et al Application of artificial intelligence to operation control of Kobe No. 3 Blast furnace Proceedings of the Sixth International Iron and Steel Congress, Nagoya, SIIJ 1990: 1~ 7
- 4 Ginsberg A, Weiss SM, Politakis P. Automatic knowledge base refinement for classification system, AI 1988, 35(2): 197~ 226
- 5 Musa J, Bahill A T. Interactive verification of knowledge based systems IEEE Expert, 1993, 8(1): 25~ 32
- 6 Mengshoel O J, Delab S Knowledge validation: principles and practice IEEE Expert, 1993, 8(3): 62 ~ 68
- 7 刘金琨, 邓守强, 苏士权等. 高炉专家系统的数据采集与处理. 东北大学学报, 1996, 17(5): 485~ 489
- 8 Li M F. Building an expert system that can learn Proceedings of IEEE Conference on AI Applications, 1986, 120~ 135
- 9 刘金琨, 徐心和, 苏士权等. 高炉操作指导专家系统的设计与实现. 中国控制会议, 青岛, 1996: 300~ 304