

基于数据库的产生式知识表达及其在发酵过程中的应用

安 莉 王建林*

(北京化工大学 信息科学与技术学院, 北京 100029)

摘 要: 结合数据库技术的优势,提出了一种基于数据库的产生式知识表达方法,并说明了产生式规则的树型结构及其数据库表达。结合发酵过程生物量软测量专家系统的应用,给出了发酵过程生物量软测量建模知识的产生式规则定义及其知识库设计。实验表明,与一般顺序组织结构的产生式知识表达方法相比,基于数据库的产生式知识表达方法有利于提高专家系统的推理效率。

关键词: 知识表达; 产生式规则; 数据库

中图分类号: TP182

引 言

产生式知识表达方法是目前应用较多的一种知识表达方法,它用接近于人类思维特点的形式获取和表示知识,直观且便于推理。但是,产生式规则之间相互作用的限制可能导致降低推理效率。薛冬娟等^[1]提出了一种称为分类产生式规则的知识表达方法,有效地解决了知识组合爆炸和搜索匹配代价高的问题,但不具备自学习功能,不能添加新知识。陈亚兵等^[2]提出了一种基于知识库的专家咨询系统,实现了知识的提取、知识的存储与传递、知识的检索及知识的应用,然而在知识组织和推理方面还不太完善,有待于进一步研究。知识表达方法已成为近年来众多国内外学者研究的热点问题^[3-7]。

数据库技术具有数据结构化、低冗余度、高独立性、易扩充性的特点。基于数据库技术的知识表达的优势在于:强大、广泛的知识表示能力;易于管理;强大灵活的可扩展能力;与高效的推理方法的配合^[8]。

本文结合数据库技术的优势,提出了一种基于数据库的产生式知识表达方法,并将其应用于发酵过程生物量软测量专家系统中,有利于提高专家系

统的推理效率。

1 基于数据库的产生式知识表达方法

1.1 产生式规则的树型结构

产生式规则以条件和结论的方式表达知识,并且在推理过程中,上一条规则的结论又作为下一条规则的条件,形成规则链,将其表示为树型逻辑结构如图1所示。

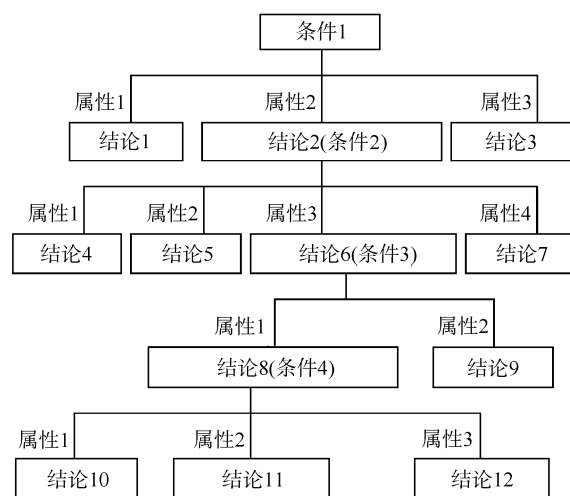


图1 产生式规则的树型结构

Fig. 1 Tree-structure of production rules

在图1的树型结构中,每一组条件节点到结论节点对应一条产生式规则,连接条件节点与结论节点的树枝上的属性值,对应该条规则的约束,从根节点到叶节点经由属性值的约束的全部路径即为一条产生式规则链。

收稿日期: 2008-09-16

基金项目: 国家自然科学基金(20476007/20676013)

第一作者: 女,1981年生,博士生

* 通讯联系人

E-mail: wangjl@mail.buct.edu.cn

1.2 产生式规则的数据库表达

知识是以数据库表的形式存储在知识库中的,产生式规则的数据库表的定义如下:

rules (rule-id, rule-cond-id, rule-conclu-id, description) (1)

rule-cond (rule-cond-id, condition) (2)

rule-conclu (rule-conclu-id, conclusion) (3)

表达式(1)为规则表的定义, rule-id 为规则的编号, rule-cond-id 为规则条件的编号, rule-conclu-id 为规则结论的编号, description 为规则的约束;表达式(2)为规则条件表的定义, rule-cond-id 为规则条件的编号, condition 为规则条件的描述;表达式(3)为规则结论表的定义, rule-conclu-id 为规则结论的编号, conclusion 为规则结论的描述。

2 基于数据库的产生式知识表达的应用

发酵过程的控制,通常需要状态变量,如菌体、基质和产物的浓度,由于生物传感器的缺乏,这些变量很难在线测量^[9],将专家系统应用于发酵过程生物量软测量建模,将已有的发酵过程生物量软测量建模方面的知识总结到专家系统的知识库中,依靠专家系统有效的推理,可以完成从已知的参数到未知的待测生物量之间的估计。

2.1 产生式规则定义

在发酵过程生物量软测量专家系统中,定义“目标规则”为“符号”、“变量”、“参数”和“方案”构成的序列:

$$TR_i(c, x, \theta, f_i), i = 1, \dots, p.$$

其中, $c = [c_1, \dots, c_k]^T \in \Omega^k$, Ω 为可列有限符号集合,是发酵对象的文字和特征性说明的符号集合; $x = [x_1, \dots, x_n]^T \in X^n$, X 为所有发酵过程机理模型中的变量集合; $\theta = [\theta_1, \dots, \theta_m]^T \in \theta^m$, θ 为所有发酵过程机理模型中的参数集合; $f_i \in F^l$, F 为软测量模型方案集合。

定义“元规则”为“方案”和“操作”构成的序列: $MR_j(f, A_j), j = 1, \dots, q$ 。其中, $f = [f_1, \dots, f_l]^T \in F^l$; A_j 表示专家系统对“事实”的操作。

针对发酵过程的软测量建模,对规则中存在的变量退化为参数,参数进化为变量的情况,通过加入调整因子 α 扩展目标规则为: $TR_i(c, \alpha_x \cdot x, \alpha_\theta \cdot \theta, f_i), i = 1, \dots, p$ 。其中, $\alpha = [\alpha_x, \alpha_\theta]^T = [\alpha_1, \dots, \alpha_{n+m}]^T, \alpha_i \in \{0, 1\}, i = 1, \dots, n + m$ 。

相应的元规则扩展为: $MR_j(f, \alpha, A_j), j = 1, \dots, q$ 。

扩展后的规则能够很好地处理参数进化和变量退化的情形。

将专家系统用于建模不同于常规的专家系统,为了得到无法事先确定的建模方案,还需要加入置信因子 β 。扩展后的产生式规则为:

$$TR_i(c, \alpha_x \cdot x, \alpha_\theta \cdot \theta, f_i, \beta_i), i = 1, \dots, p,$$

$$MR_j(f, \alpha, \beta, A_j), j = 1, \dots, q.$$

2.2 知识库设计

采用基于数据库的产生式知识表达的发酵过程生物量软测量专家系统的知识库表的定义如下:

rules (rule-id, rule-cond-id, rule-conclu-id, description) (4)

rule-cond (rule-cond-id, S_v, S_p, S_c, α) (5)

rule-conclu (rule-conclu-id, F, β) (6)

表达式(4)为规则表的定义, rule-id 为规则的编号, rule-cond-id 为规则条件的编号, rule-conclu-id 为规则结论的编号, description 为规则的约束/描述;表达式(5)为规则条件表的定义, rule-cond-id 为规则条件的编号, S_v 为变量, S_p 为参数, S_c 为符号, α 为调整因子;表达式(6)为规则结论表的定义, rule-conclu-id 为规则结论的编号, F 为模型方案, β 为置信因子。

发酵过程生物量软测量专家系统中的知识库设计如图2所示。

分层结构的特点:每一层的结构简单,可以根据需要继续细化,便于知识库的扩充;这种分层结构实际上已经对已知事实和机理进行了模块化处理,在查找时定位准确,能够缩短搜索目标的时间,提高搜索效率;便于知识库更新,在向知识库添加或删除规则时,只需对某一具体层次的内容进行操作,易于移植。

2.3 实验分析

推理任务描述如下:已知可测变量为温度、压力、pH值、溶解氧浓度;发酵方式为分批发酵;限制性基质数量为单一限制性基质;菌体浓度增加对自身生长有抑制作用;高浓度底物对菌体的生长也有抑制作用;待求生物量为菌体浓度。

推理过程采用基于匹配度的广度优先推理方法进行推理,并采用单步调试方式以统计推理步长,推理过程如图3~5所示。

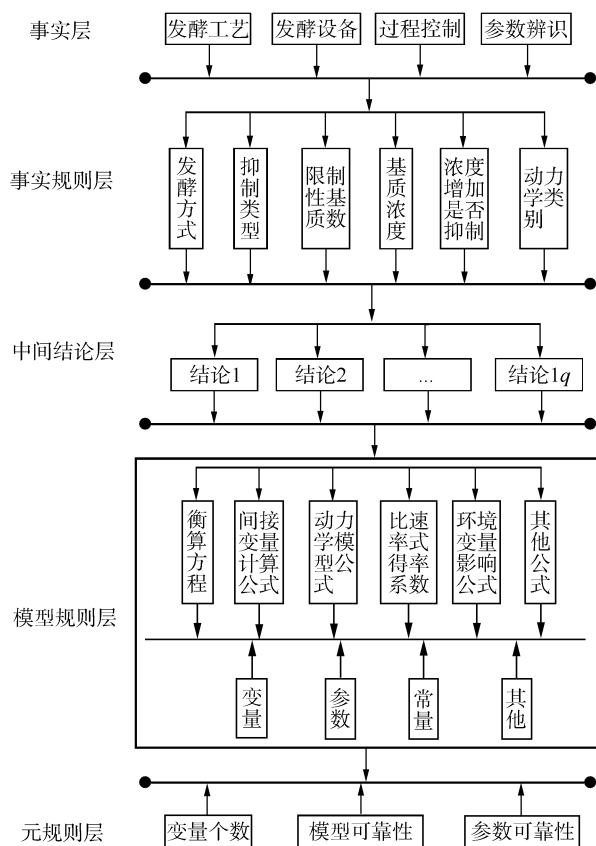


图 2 知识库的分层结构

Fig. 2 Layered structure of knowledge-base

通过实验可知,采用一般顺序组织结构的产生式知识表达方法时的推理步长为 51,采用基于数据库的产生式知识表达方法时的推理步长为 48。基于数据库的产生式知识表达方法与一般顺序组织结构的产生式知识表达方法相比,推理步长减少了 3,推理效率提高了 5.88%。



图 3 初始条件的输入

Fig. 3 Input of conditions



图 4 约束属性的输入

Fig. 4 Input of descriptions

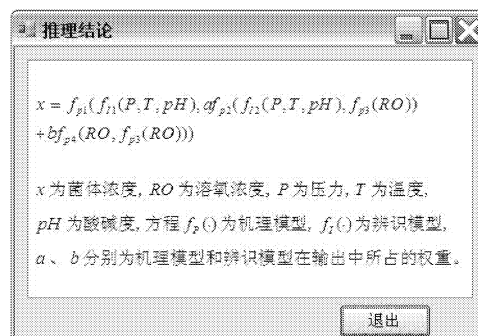


图 5 推理结果

Fig. 5 Reasoning results

3 结论

专家系统的行为效果在很大程度上取决于知识的表达方法。本文提出了一种基于数据库的产生式规则知识表达方法,该方法与一般顺序组织结构的产生式知识表达方法相比,使发酵过程生物量软测量专家系统的推理效率提高了 5.88%。

参考文献:

- [1] 薛冬娟, 张冬冬, 张彦峰, 等. 农业专家系统中分类产生式规则的知识表示方法[J]. 河北农业大学学报, 2004, 27(3): 104-108.
- Xue D J, Zhang D D, Zhang Y F, et al. Knowledge expression to classifying the producing type rule in the agricultural expert system[J]. Journal of Agricultural Uni-

- versity of Hebei, 2004, 27(3): 104 – 108. (in Chinese)
- [2] 陈亚兵, 孙济庆. 基于知识库的专家咨询系统设计与实现[J]. 计算机工程, 2007, 33(16): 196 – 198.
Chen Y B, Sun J Q. Design and implementation of expert referring system based on knowledge base[J]. Computer Engineering, 2007, 33(16): 196 – 198. (in Chinese)
- [3] 潘玲琳. 基于产生式规则的专家系统的研究实现[J]. 计算机技术与发展, 2007, 17(5): 66 – 68.
Pan L L. Application of expert system based on generative formula computer technology and development[J]. 2007, 17(5): 66 – 68. (in Chinese)
- [4] 陈璐璐, 刘成瑞, 张庆振, 等. 基于数据库技术的故障诊断知识库管理系统设计[J]. 航天控制, 2007, 25(4): 82 – 86.
Chen L L, Liu C R, Zhang Q Z, et al. Knowledge base management system designed for fault diagnosis expert system based on database technology[J]. Aerospace Control, 2007, 25(4): 82 – 86. (in Chinese)
- [5] 高毅如. 基于知识库及专家系统的智能信息交换系统[J]. 计算机工程, 2007, 33(21): 166 – 168.
Gao Y R. Intelligent information exchange system based on knowledge base and expert system[J]. Computer Engineering, 2007, 33(21): 166 – 168. (in Chinese)
- [6] 吴晓知, 李兴明. 网络故障管理专家系统中知识库的构造[J]. 微计算机信息, 2008, 24(2/3): 116 – 118.
Wu X Z, Li X M. The construction of object-oriented knowledge base in network fault management expert system[J]. Microcomputer Information, 2008, 24(2/3): 116 – 118. (in Chinese)
- [7] Laskey K B. MEBN: A language for first-order bayesian knowledge bases[J]. Artificial Intelligence, 2008, 172(2/3): 140 – 178.
- [8] 王小军, 周昌盛, 吕汉兴. 一种基于关系数据库的知识表示和推理方法[J]. 控制工程, 2005, 12(1): 40 – 43.
Wang X J, Zhou C S, Lv H X. Knowledge representation and reasoning method based on relation database[J]. Control Engineering of China, 2005, 12(1): 40 – 43. (in Chinese)
- [9] 隋青美, 王正欧. 发酵过程混合神经网络模型及其仿真[J]. 系统仿真学报, 2002, 14(4): 415 – 417.
Sui Q M, Wang Z O. A hybrid neural network model and simulation for fermentation processes[J]. Journal of System Simulation, 2002, 14(4): 415 – 417. (in Chinese)

A production knowledge representation based on a database and its application in a fermentation process

AN Li WANG JianLin

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: A production rules knowledge representation based on a database, which introduces the advantages offered by databases, is presented. The tree-structure of the production rules and their expression in the database are described. When combined with the application of a biomass soft-sensor expert system in a fermentation process, definitions of production rules concerning biomass soft-sensor modeling knowledge and design of the knowledge base can be obtained. Tests show that, compared with an ordinal production knowledge representation, the production knowledge representation based on the database leads to an enhanced reasoning efficiency of the expert system.

Key words: knowledge representation; production rules; database