基于形式概念分析的不完备电子病历系统粗糙挖掘研究

丁卫平[1] 顾春华[2] 石振国[1,3] 陈建平[1] 管致锦[1,4]

(南通大学计算机科学与技术学院 南通 226019)[1] (南通市中医院 南通 226006)[2]

(上海大学计算机工程与科学学院 上海 200072)[3]

(南京航空航天大学信息科学与计算机学院 南京 210016)[4]

摘 要 形式概念分析与粗糙集理论是近年来获得飞速发展的两种数据挖掘工具。充分利用概念格在形式概念表示和粗糙集在知识约简等方面的独特优势,提出了基于形式概念分析的不完备电子病历系统粗糙挖掘算法(FCRM)。该算法利用决策规则格进行不完备知识的形式概念表示和粗糙正域近似约简,并能较好地提取相应一致的决策规则。最后构建不完备中医电子病历方剂挖掘专家系统,实验结果表明该算法在不完备电子病历系统约简和挖掘方面均具有较好性能。

关键词 不完备知识,形式概念分析,粗糙近似约简,电子病历挖掘,决策规则格

中图法分类号 TP301.6 **文献标识码** A

Research of Formal Concepts Rough Mining under Incomplete Electronic Patient Record Knowledge System

DING Wei-ping[1] GU Chun-hua[2] SHI Zhen-guo[1,3] CHEN Jian-ping[1] GUAN Zhi-jing[1,4]

(School of Computer Science and Technology, Nantong University, Nantong 226019, China)[1]

(Nantong Chinese Medicine Hospitial, Nantong 226006, China)[2]

(School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China)[3]

(College of Information Science and Technology, Nanhang University, Nanjing 210016, China)[4]

Abstract Formal concepts analysison and rough sets theory are two different fast developing tools for data mining. The advantages of both the concept lattice in formal concepts representation and rough sets in knowledge reduction were enough taken, and the algorithm of formal concepts rough mining(FCRM) under the incomplete electronic patient record knowledge system was put forward. The algorithm can carry on incomplete knowledge formal concept representation and rough positive approximate reduction with the decision rule lattice, and the corresponding consistent decision rule was extracted. Finally the expert system of the traditional chinese medicine(TCM) patient record was designed. The experimental results show that algorithm of FCRM is better on the knowledge reduction and mining capability.

Keywords Incomplete knowledg, Formal concepts analysison, Rough positive approximate reduction, Electronic patient record mining, Decision rule lattice

1 引言

电子病历挖掘(Electronic Patient Record Mining,简称EPRM)是近年来随着医学信息和人工智能技术发展而产生的一个新的研究方向,它是指从大量的、不完全的、有噪声的、模糊的医学数据库中提取隐含在其中的医学诊断规则和模式,用来在电子病历数据库记录中识别相关医学知识联系,形成相应预报和分类决策支持模型,从而为疾病的诊断和治疗提供辅助决策等[1]。但由于电子病历数据库系统是一个特殊的不完备知识系统,其存储的医学数据有其特有的属性,这给电子病历挖掘带来了较大困难。目前国内外对不完备电子病历系统的知识发现有很多方法,其中粗糙集理论和形式概念分析是两种较为有效的方法。

粗糙集理论是由波兰学者 Z. Pawlak 于 1982 年提出的一个用于分析数据的数学理论,它能够分析处理不精确、不一致和不完备信息,因此作为一种具有极大潜力和有效的知识获取工具而受到人工智能工作者的广泛关注[2]。由于在不完备信息系统进行知识发现时删除和补齐空值处理方法的局限

到稿日期:2008-12-16 返修日期:2009-03-23 本文受国家自然科学基金-微软亚洲研究院联合资助(60873069),江苏省高校自然科学研究项目(09KJD520008),南通市应用研究计划项目(K2008031),南通大学自然科学基金项目(05Z061),南通大学通信与信息系统学科科技创新基金资助。

丁卫平(1979—),男,硕士,讲师,主要研究方向为粗糙集、概念格和电子病历挖掘等,E-mail:dwp9988@163.com;顾春华(1976—),男,工程师,主要研究方向为中医电子病历;石振国(1964—),男,博士,副教授,硕士生导师,主要研究方向为网格计算、分布式智能信息处理等;陈建平(1960—),男,教授,硕士生导师,主要研究方向为数字信号处理、快速算法等;管致锦(1962—),男,博士,教授,硕士生导师,主要研究方向为人工智能和量子可逆计算等。

性、不确定性值的出现使得不能够在对象集上找到符合实际需要的等价关系，因此对不完备信息系统进行知识发现可以从对象之间粗糙相似性来考虑对象之间的复杂关系，从而为完成不完备信息系统粗糙分析等方面提出相关方法。目前国内外已有部分专家开展了这方面的研究，如 M. Kryszkiewicz 提出基于相容关系的不完备信息系统的知识约简研究方法[3]；J. Stefanowski 等人提出利用粗糙分类的方法进行不完备知识系统约简而获取最简规则[4]；周献中等人将最大分布、分配约简引入到不完备决策系统中提出分配序约简方法等[9]。但上述这些研究工作都是假设信息系统中的未知属性值仅仅是被遗漏的，但又是确实存在的，也就是说所使用的粗糙集模型是建立在容差关系（自反性对称性）基础上的，然而不完备电子病历系统中存在大量丢失型未知属性值的数据和信息。如何对电子病历系统中这些信息进行知识约简，是目前粗糙集理论研究面临的挑战性课题。

形式概念分析，也称概念格（Galois 格），是由德国教授 R. Wille 于 1982 年首先提出的，主要用于概念的发现、排序和显示，它也是一种支持数据分析的有效工具[5]。利用形式概念分析对粗糙集进行近似，可较好地解决上述不完备信息系统利用粗糙集知识约简存在的问题。目前国内外已有专家开展这方面的研究，如 J. Saquer 等人给出了一个不完备知识的属性确定算法，它比较确切地对不完备知识系统进行形式概念分析研究，并能提取未知背景中有效属性隐含知识的最大信息量[6]；J. W. Grzymala-Busse 等人提出类似于完备信息系统的等价不可分辨关系描述方法，较好地刻画了正对象集（下近似）与不确定对象集（上近似），有效解决了知识中的不确定值和缺失值问题[7]。但是上述方法在利用形式概念粗糙约简和规则挖掘时，往往不能处理大量非线性的、不精确的、模糊的医学数据，得到的规则数目较大。且目前这方面主要是理论研究，实际应用系统较少。

粗糙集理论与形式概念分析各自在不完备信息知识挖掘方面都有一定的优势，可以开展形式概念分析与粗糙集相融合的研究与应用，结合不完备电子病历系统中数据特有属性，提出基于形式概念分析的不完备电子病历系统粗糙挖掘算法（FCRM），并构建相应智能挖掘专家系统，提取隐含的医学诊断规则和模式，为疾病的诊断和治疗提供辅助决策等。

2 不完备电子病历系统特征分析

电子病历系统作为医院信息系统的一个重要组成部分，它是将传统的纸质病历电子化，并超越纸质病历的管理模式，提供查询、统计、分析、信息交换等功能。电子病历挖掘的主要对象是临床医疗中隐含着大量有用且可提取的诊断规则相关病历信息，但是电子病历系统是一个特殊的不完备知识系统，这使得电子病历挖掘与普通数据挖掘存在较大差异，其数据主要特点如下[1]：

（1）电子病历数据的多样性。不完备电子病历数据库中医疗数据对象异常丰富，可能含有各种病理参数、化验与测量结果、诊断记录、相关的参数数据（如年龄、性别、病史、入出院时间）以及相关医学声音、图像等，这是它区别于其它不完备系统数据的最显著特征，这种数据多样性加大了电子病历挖掘的难度。

（2）电子病历数据的不完整性和不确切性。病例和病案的有限性使电子病历数据库不可能对任何一种疾病信息都能作出全面的反映，疾病信息所体现出的客观不完整和描述疾病的主观不确切，都是形成电子病历存储数据不完备性的主要因素。

（3）电子病历数据的动态变化性。电子病历中一些医学检测的波形、图像都是时间的函数，它们是不断更新和变化的，具有较强的时效性，日常病历中记录的只是一小部分动态数据。

（4）电子病历挖掘前复杂的预处理过程。电子病历系统的不完备性，其存储数据中含有大量模糊的、冗余的和有噪音的数据，在进行数据挖掘前必须首先对这些不完备病历数据进行复杂的预处理：过滤、清洗、转换等，使数据统一化和规范化。

以上这些医学数据特有属性，给不完备电子病历系统挖掘带来了较大困难。要保证电子病历挖掘的效率和质量，必须寻找一种较好的规则挖掘算法。而传统的关联规则挖掘算法效率不高，先进的粗糙集理论不能较好处理系统中已经丢失型未知属性值信息，形式概念分析又不能处理大量非线性的、不精确的、模糊的医学数据，且得到的规则数目较大等。大量实验证明，不完备电子病历系统中以往直接删除空值对象和填充值方法均不同程度地改变了原电子病历系统的信息成分，因此需要寻求一些其他方法来解决不完备电子病历信息系统的相关不完备病历信息的处理问题。

3 基本定义与定理

定义 1（粗糙集[2]） 任意给定一个集合 $X\subseteq U$，如果使用 R 等价类无法精确描述 X，则 X 就是 R 的粗糙集；反之 X 是 R 的精确集。包含在 X 中的最大可定义集称为 X 的 R 下近似（Lower Approximation）

$$\underline{R}(X)=\{x\in U|[x]_R\subseteq X\}$$

包含 X 的最小可定义集称为 X 的 R 上近似（Upper Approximation）

$$\overline{R}(X)=\{x\in U|[x]_R\cap X\neq\emptyset\}$$

定义 2（概念格[5]） 给定形式背景 K 上的一个序偶 (X,S) $X\subseteq U,S\subseteq D$，如果满足 $f(X)=S$ 并且 $X=g(S)$，则称 (X,S) 为一个概念。X 称为概念 (X,S) 的外延，S 称为概念 (X,S) 的内涵。形式背景 K 中的所有概念及概念之间的偏序关系构成的结构称为概念格，记作 $L(U,D,R)$，其中每个概念被看作概念格中的一个节点。

定义 3（全局决策表） 全局决策表 DT 是一个四元组 $\langle U,C\cup D,V,f\rangle$，其中 U 是一组对象的非空有限集合，称为论域；设有 n 个对象，则 U 可表示为 $U=\{x_1,x_2,\cdots,x_n\}$；C 为条件属性集，D 为决策属性集，$V=\bigcup_{a\in(C\cup D)}V_a$，$V_a$ 为属性 a 的值域集；f 是 $U\times(C\cup D)\to V$ 的映射。

定义 4（P-正域） 在近似空间 $K=(U,R)$ 中，$P,Q\subseteq R$，若概念 (A,B) 的外延 A 是 $IND(P)$ 的等价类，并且满足 $A\subseteq POS_P(Q)$，则称概念 (A,B) 为 $IND(P)$ 的 P-正域概念，将 $IND(Q)$ 的所有 P-正域概念的集合记作 $POSC_P(Q)$。

定理 1[10] 在近似空间 $K=(U,R)$ 中，$P\subseteq R$ 且 $P\neq\emptyset$，L 为相应的概念格，(A,B) 是 L 中的概念，则存在

$$IND(P)=\bigcup\sup\{A|B_i\subseteq B,attrib(B_i)=P,(A,B)\in L\}$$

其中，$attrib(B_i)$ 表示内涵 B_i 所对应的属性集。

定义 5(约简与核) 对于形式背景(U,A,I),如果存在属性集$D\subseteq A$,使得$L(U,D,I_D)\cong L(U,D,I)$,则称D是(U,A,I)的协调集。

若进一步$\forall d\in D, L(U,D-\{d\},I_{D-\{d\}})\neq L(U,A,I)$,则称$D$是$(U,A,I)$的约简。所有$(U,A,I)$的交集称为$(U,A,I)$的核心。

定理 2 在决策表$DT=\langle U,C\cup D,V,f\rangle$对应的概念格中,对于概念$(A,B)$,$attrib(B)\subset C$且$|C-attrib(B)|=1$,当且仅当$C-attrib(B)$是条件属性的核元素。

证明:充分性。若$attrib(B)\subset C$且$|C-attrib(B)|=1$,则$C-attrib(B)$是当A对象的决策属性不同时唯一取值不同的条件属性,即唯一能分辨决策属性的条件属性,不可缺少。所以$C-attrib(B)$是决策表的核;

必要性。当$C-attrib(B)$是核元素时,显然$|C-attrib(B)|=1$,且$C-attrib(B)$是D-不可缺的,即$POS_C(D)\supseteq POS_{attrib(B)}(D)$,因此存在$attrib(B)$的等价类对应的概念$(A,B)$,$A\subset POS_C(D)$且$A$中的对象属于不同的决策表,即概念$(A,B)$满足$D\not\subset attrib(B)$,即$attrib(B)\subset C$。

定理 3 在一个给定的背景(U,D,R)上,概念格的每一个结点是粗糙集在此背景属性和对象的分类并集。即对背景上的概念C上的概念(X,Y),有

$$X=\bigcup_{x\in X}[x]R_M, Y=\bigcup_{y\in Y}[y]R_G$$

证明:显然,$A,\varnothing$是在分割簇中。对$x\in X$,因为R_M是自反的,有$X\in[x]R_M$即$\supseteq\bigcup_{x\in X}[x]R_M$。反之,因为$X$是概念的对象集,那么

$$X=(\bigcup_{x\in X})^{**}=(\bigcap_{x\in X}x^*)^*=(\bigcap_{x\in X}[x]R_M{}^*)^*$$
$$=(\bigcup_{x\in X}[x]R_M)^{**}\supseteq\bigcup_{x\in X}[x]R_M$$

故$X=\bigcup_{x\in X}[x]R_M$,同理可证$Y=\bigcup_{y\in Y}[y]R_G$

定义 6[11,12](决策规则) 设$S=(U,A)$是一决策表,$A=C\cup D, C\cap D=\varnothing$,其中$C$为条件属性集,$D$为决策属性集。令$X_i$和$Y_j$分别表示$U/IDN(C)$和$U/IDN(D)$中的各个等价类。$Des_C(X_i)$表示对等价类$X_i$的描述,$Des_C(Y_i)$表示对等价类$Y_j$的描述。

$$r_{ij}: Des_C(X_i)\rightarrow Des_C(Y_i), Y_j\cap X_i\neq\varnothing$$

如果$Y_j\cap X_i=X_i$,则决策规则r_{ij}是确定的,即$r_{ij}: Des_C(X_i)\rightarrow Des_C(Y_i)\Leftrightarrow Y_j\subseteq X_i$;否则,$r_{ij}$是不确定的,即

$$r_{ij}: Des_C(X_i)\xrightarrow{conf} Des_C(Y_i)\Leftrightarrow Y_j\cap X_i\neq\varnothing$$

其中,$conf$为规则r_{ij}的信度。

定义 7(决策规则格) 给定一个概念格结点(O_1,A_1),把O_1看成是事务集,A_1看成属性集,一个概念格的节点就表示具有属性集A_1的最大事务集,在决策规则挖掘中,为了提高效率,在格中使用O_1的基数$||O_1||$代替O_1。

决策规则格节点用$(||O_1||, C_1, sup_dcs(O_1), conf(sup_dcs(O_1)))$表示,其中$C_1$是条件属性的集合,称为$O_1$的内涵,即$O_1$对各条件属性的特定取值;$||O_1||$是$O_1$的基数,

定理 4 对于任一概念$H=(O_1,C_1)$和一个特征子集$C_2\subseteq C_1$,$g(C_2)=g(C_1)=O_1$,当且仅当对于每个父概念$H_p=(O_p,C_p)$有$C_2\cap(C_1-C_p)\neq\varnothing$。所以

$$H_1=(O_1,C_1,sup\_des(O_1)), conf(sup\_des(O_1)))$$

H的父节点为

$$H_p=(O_p,C_p,sup\_des(O_p)), conf(sup\_des(O_p)))$$

定理 5[10] 若$\varphi\rightarrow(d=d_i)$是一个最简决策规则,则在$B(U,M,I)$中存在概念$(\alpha,\beta)$,使得

(1) $\bar{\alpha}\subseteq\beta$; (2) $d_i\in\beta$; (3) $\bar{\varphi}\in\langle\alpha,\beta\rangle^{**}$。

4 基于形式概念分析的粗糙挖掘算法(FCRM)

4.1 FCRM 算法思想

在FCRM算法中,综合运用了形式概念分析、粗糙集和不完备知识系统有关思想,对不完备知识系统的全局决策表进行P-正域近似集合约简,并利用决策规则格模型进行形式概念的不完备知识表示,从而得到不完备决策规则格的一致医学决策规则等。设计的FCRM算法由形式概念格的构造和决策规则格粗糙的挖掘两部分组成。相关实验结果表明,此算法能够提取隐含在不完备电子病历系统中的医学一致决策规则,在运行速度和挖掘性能上都是高效的。

4.2 FCRM 算法核心步骤

将不完备知识系统初始化为最顶端节点$(||O||,\varnothing,default,0)$的格$L$,并通过$P$-正域近似集合进行部分约简。算法核心步骤描述如下:

Step1: $L=(||O||,\varnothing,default,0)$
$queue=(||O||,\varnothing,default,0)$
$Ag=$attribute$\in\{$ number$>\varepsilon\}$

Step2: $Candset\leftarrow FindSubNodes(||O_k||, C_k, sup\_des(O_k), conf(sup\_des(O_k)), A\varepsilon)$

Step3: for each Node
$(||O_{ki}||, C_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))\in Candset$

Step4: if $(||O_{ki}||, C_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))\notin L$

Step5: if $(||O_{ki}||>\varepsilon)$
$L=L\cup(||O_{ki}||, C_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))$
$(||O_k||, C_k, sup\_des(O_k), conf(sup\_des(O_k)))\rightarrow(||O_{ki}||, C_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))$

Step6: if $(conf(sup\_des(O_{ki}))<1$
$(||O_{ki}||, C_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))$

Step7: else
$(||O_k||, D_k, sup\_des(O_k), conf(sup\_des(O_k)))$ to $(||O_{ki}||, D_{ki}, sup\_des(O_{ki}), conf(sup\_des(O_{ki})))$

Step8: $queue=L$ first Node
$result\leftarrow\varnothing$

Step9: while $queue\neq\varnothing$
$H=$ $queue$ first Node

Step10: $ruleforwardset=\varnothing$
for each H_p
$H_3=H-H_p$
$ruleforwardset=ruleforwardset$
$\cup combine\ of\ H_3$ and H_p

Step11: while $ruleforwardset=\varnothing$
for $C_2\in ruleforwardset$
if $g(C_2)\subseteq IND(sup\_des(O_1))$
$ruleset=ruleset+\{C_2\xrightarrow{conf(sup\_dcs(O_1))}sup\_dcs(O_1)\}$
$ruleforwardset=ruleforwardset-C_2$

Step12: output consistent decision rule L

4.3 FCRM 算法实验分析

实验数据选取作者已经建立的4个分布式存储在SQL Server2000中部分典型疾病电子病历数据库，实验条件是在Intel 1.5GHz CPU的PC机上运行，其内存为1024M，运行环境为Window XP。

实验一　利用文献[9]算法和本文提出的FCRM算法对不完备电子病历系统随条件属性增加算法约简的运行时间进行比较实验，结果如图1所示。

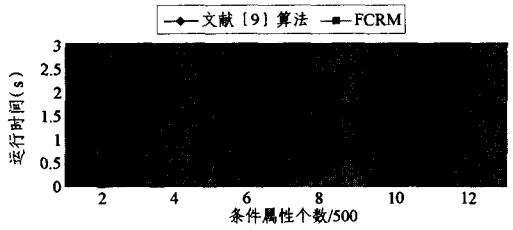

图1　随条件属性增加算法运行时间测试比较

实验二　选取3种不完备率(10%，20%，30%)的电子病历数据库系统对FCRM算法和文献[13]GENRED_GROWTH算法随不完备率变化规则提取的准确率进行比较实验，结果如图2所示。

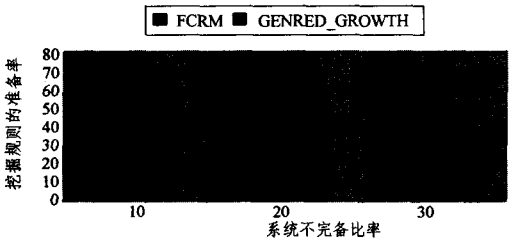

图2　在不同不完备率下挖掘准确率测试比较

从图1和图2可以看出，随着电子病历疾病数据库中条件属性递增，FCRM算法运行时间明显少于经典算法。且随着不完备率的增加，该算法规则提取准确率明显较高，具有较好的挖掘性能。

5　FCRM算法在中医电子病历方剂挖掘系统中应用

不完备电子病历系统中，尤其是中医病历方剂含有大量模糊的、非量化数据，其中隐藏着很多中医方剂组用药的配伍规律，这对于进行证型和症状间错综交叉、辨证认治困难的中医领域开展多角度、多层次和量化研究以及形成中医病历方剂有关技术的规则和处理程序分析具有重要意义，但目前一般的数据挖掘算法很难挖掘出其中特殊的配伍规律和模式[13,14]。

在FCRM算法研究基础上，结合医院现有HIS特点，设计了中医病历方剂配伍规则挖掘专家系统，将FCRM算法用JAVA语言编程后通过API函数的方法嵌套在HIS系统，以实现挖掘功能。中医病历方剂数据存储在SQL Server2000数据库中。选取中医病历脾胃方剂中症状频数大于80的方剂，挖掘脾胃方中医方剂之间的"方、药、症、因"之间的关联关系。实验结果如表1所列。实验数据来源于南通中医院2008年6月1日至2008年10月30日中医病历方剂部分数据库。

通过FCRM算法在中医病历方剂挖掘实验，可以更深刻地理解脾胃方中医方剂之间"方、药、症、因"之间的关联关系和中医病历方剂之间存在的配伍规律，这对辅助医生临床诊断和有关疾病治疗，为中医病历的现代化研究开发提供了重要的辅助决策功能。

表1　中医方剂"方、药、症、因"关系挖掘部分实验数据

代表方剂	药物	主要症状	基本病因
异功散	半夏、干姜	心下痞满	脾胃不和
四君子汤	白术、干姜	泄泻	脾气虚
半夏干姜散	白蔻、橘皮	呕吐	脾胃虚寒
茯苓汤	橘皮、半夏、茯苓	恶心、胸闷	胃气上逆
……			

结束语　不完备电子病历系统挖掘是一个新兴的、有着美好应用前景但又充满挑战的研究方向，目前国内外只有少数人涉足该领域。形式概念分析与粗糙集理论是近年来获得飞速发展的数据挖掘工具，其各自在数据挖掘方面都有一定的优势。本文充分利用粗糙集在知识约简和概念格的形式概念表示等方面的独特优势，提出基于形式概念分析的不完备电子病历系统粗糙挖掘算法(FCRM)，并构建不完备中医电子病历方剂挖掘专家系统，能较好地提取出电子病历中的关联关系和中医病历方剂的配伍规律。

本文的研究工作为不完备电子病历系统诊断治疗的知识发现提供了辅助决策，为实现医学信息多层面的综合智能分析提供了有效途径。

参考文献

[1] 丁卫平，管致锦，施佺，等. 电子病历挖掘：概念、技术及应用[J]. 计算机工程与设计，2008，29(2)：405-407

[2] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Sciences，1982，11：341-356

[3] Kryszkiewicz M. Rough set approach to incomplete information-systems [J]. Information Sciences，1998，112：39-49

[4] Grzymala-Busse J W. Data with missing attribute values：generalization of indiscernibility relation and rule induction[J]. Transactions on Rough Sets I，2004：78-95

[5] Wille R. Restructuring lattice theory：an approach based on hierarchies of concepts[M]// Rival I，eds. Ordered Sets Reidel. Dordrecht，1982：445-470

[6] Saquer J，Deogun J. Concept approximations based on rough sets and similarity measures Int [J]. Appl Math and Comp Sci，2001，11(3)：655-674

[7] Stefanowski J，Tsoukias A. Incomplete information tables and rough classification [J]. Computational Intelligence，2001，17：545-566

[8] Kent R E. Rough concept analysis. Knowledge Discovery (RS-KD'93)[C]// Ziarko W P. Rough Sets，and Fuzzy Sets. London：Springer-Verlag，1994：248-255

[9] 周献中，黄兵. 基于粗集的不完备知识系统属性约简[J]. 南京理工大学学报，2003，27(5)：630-635

[10] 张文修，魏玲，祁建军. 概念格的属性约简理论与方法[J]. 中国科学E辑：信息科学，2005，35(6)：628-639

[11] 王国胤. Rough集理论在不完备知识系统中的扩充[J]. 计算机研究与发展，2002，39(10)：1238-1243

[12] 刘宗田，强宇，周文，等. 一种模糊概念格模型及其渐进式构造算法[J]. 计算机学报，2007，30(2)：184-188

[13] 丁卫平，管致锦，顾春华. 基于Rough Sets的中医指症挖掘研究与应用[J]. 计算机工程与应用，2008，44(7)：234-237

[14] 曾令明，唐常杰，阴小雄. 基于位图矩阵和双支持度的中药配伍挖掘技术[J]. 四川大学学报：自然科学版，2005，42(1)：57-62