

粗集和多关系学习综述^{*})

孙成敏 刘大有 孙舒杨

(吉林大学计算机科学与技术学院符号计算与知识工程教育部重点实验室 长春 130012)

摘 要 文中给出了多关系学习的产生、实质以及任务,指出多关系学习具有狭义和广义两个层面。对粗集和多关系学习给出了简单综述,表明粗集理论在多关系学习中占有重要地位。分析研究了粗集和归纳逻辑程序设计方法用于多关系学习的几种结合途径,尤其重点介绍了 RSILP 系列模型,并在文中给出其一般模型。文中对用于多关系学习中其它粗集方法也作了简单介绍。RSILP 模型的完善扩展以及粗集方法在多关系学习中的进一步应用是今后的工作方向。

关键词 多关系学习,粗集,归纳逻辑程序设计,RSILP 模型

The Overview on Rough Set and Multi-Relational Learning

SUN Cheng-Min LIU Da-You SUN Shu-Yang

(College of Computer Science and Technology, Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012)

Abstract Give the generation, essence and task of multi-relational learning, and bring forward the viewpoint that multi-relational learning shows narrow and broad sense. The overview on rough set and multi-relational learning is given, which indicates importance of rough set in multi-relational learning. Several approaches combining rough set and inductive logic programming used for multi-relational learning are studied, especially RSILP models are introduced and general model is given. Other rough set methods are briefly presented. The improvement of RSILP models and further application of rough set in multi-relational learning are future focus.

Keywords Multi-relational learning, Rough set, Inductive logic programming, RSILP model

1 引言

数据挖掘方法是在数据中寻找模式。传统的机器学习和数据挖掘方法都是命题化的,集中于扁(flat)数据,在单一的表中寻找模式。多关系学习(MRDM, MRL)方法,也称为关系学习(RDM, RL),是从关系数据库中寻找涉及多表(多关系)的模式^[1]。

从文[1~7]中我们可以看到,多关系学习具有两个层面的含义。狭义地讲,多关系学习指学习的对象是多关系的数据;广义地讲,多关系学习就是统计关系学习^[2]。

狭义的多关系学习系指挖掘的数据对象是关系数据,即是在多表组成的关系数据库中,处理数据挖掘的多学科领域^[7]。多关系学习的数据对象一般由数据库中的几张表组成,通常并非一张表。多关系学习能直接从多关系数据库中分析数据,不需要先将其转化为单表。关系或者外延的表示为关系数据库,或者内含的表示为逻辑。

统计关系学习,又称概率逻辑学习。所谓“统计(或概率)”是指应用统计学习和推理技术及基于概率论的概率表示和推理机制,如贝叶斯网、(隐)马尔科夫模型、随机文法、马尔科夫网等,这些表示方法已经成功应用于很多领域,并由此得到了许多不确定性推理模型;所谓“关系(或逻辑)”系指关系和一阶逻辑表示,使用这些方法的好处是能够很好地表示包

括多个对象及对象间关系在内的复杂情况;所谓“学习”,在这里与数据挖掘等同,系指在数据基础上得到概率关系模型^[2]。

事实上,这两个层面的多关系学习存在着密切的关系。前者侧重于数据对象的描述,后者侧重于所使用的理论技术的刻画。统计关系学习方法所研究的数据一般都是多关系的,而狭义的多关系学习又经常使用统计关系学习方法中的理论技术。在本文中,正如开始所描述的,我们更多的是研究狭义的多关系学习,但也会涉及到统计关系学习。

粗糙集(粗集)理论,形成于 20 世纪 80 年代早期^[8,9],是处理模糊和不确定性的数学工具。粗集理论定义了不可识别关系,用以将论域中的对象划分入初等集合。当至少存在一个初等集,其中既包含正实例,又包含负实例时,我们说该概念是粗的。

经过多年的研究,应用于单表进行数据挖掘的粗集理论已经相当成熟和有效^[6,9~15]。多关系学习框架下,在对数据进行预处理和挖掘中,粗集理论也发挥着重要的作用。本文主要给出了粗糙集理论在多关系学习中的方法和应用。一方面,粗集可以作为预处理阶段来帮助或者提高学习算法,另一方面也可以将粗集融入入学习算法中,以有效地处理不确定性^[7]。

正像许多数据挖掘算法来源于机器学习领域一样,许多 RDM 算法来自于归纳逻辑程序设计领域^[1]。ILP 方法在多

^{*} 基金项目:国家自然科学基金重大项目(60496321);非规范知识处理的基本理论和核心技术;国家自然科学基金项目(60373098,60173006);国家 863 高技术研究发展计划项目(2003AA118020);吉林省科技发展计划重大项目(20020303);吉林省科技发展计划项目(20030523)。

孙成敏 博士研究生,讲师,从事机器学习、数据挖掘、粗集的研究。

关系学习中占有重要的地位,但不可否认,ILP方法只能进行确定推理。对于出现不良数据的情形,单独使用ILP有时显得无能为力。将粗糙集和ILP结合,利用ILP的强大表示功能和推理机制,以及粗集处理不确定性的能力,对于多关系数据进行挖掘是一个比较有效的途径。本文第3节将是对该方向工作的综合阐述^[16~22]。当然,对于关系数据,许多学者也尝试使用单独的粗集方法进行学习。第4节即是对在在多关系学习中,其它粗集方法^[5,6,23]的介绍。第2节是关于多关系学习、ILP、RS背景知识的简介;最后是对本文工作的总结。

2 相关背景

2.1 多关系学习 MRDM

典型的关系数据由几个表(关系)组成,可以内涵表示(使用逻辑规则)或者外延表示(使用表形式)。使用传统的数据挖掘方法,需要对数据进行预处理。预处理一般是通过联结或者聚集操作将多表中的数据整合为一张表,这要求艰巨的工作,且可能导致信息丢失。多关系数据挖掘要求能够直接分析来自于多表中的数据,不需要将其转换为单表。

由于多关系学习涉及多表,因此所挖掘的关系模式需要具有更强表达力的语言。单表挖掘的模式一般使用命题逻辑即可表达,但多关系学习却需要建立在一阶逻辑(或称谓词逻辑、关系逻辑)框架下,通常是一阶逻辑的子集。MRDM方法考虑所有的主要数据挖掘任务,包括关联分析、分类、聚类、学习概率模型和回归等。

由于能够处理结构数据和领域知识,MRDM的应用领域得到丰富,这是单表方法很难解决的。MRDM已被用于不同的领域,从商务数据分析、环境和交通工程,到Web挖掘,尤其是在生物信息学上得到了成功的应用^[1]。

2.2 粗糙集 RS 的知识背景

20世纪80年代早期,Pawlak提出粗糙集理论^[8]。粗集理论有着较严格的数学体系,是处理模糊和不确定性的有效方法。经过众多学者的研究、完善和扩展,该理论已经日趋成熟,并被广泛应用于相关领域。此处给出粗集理论的一般问题背景。

设 $U \neq \emptyset$ 是我们感兴趣的对象组成的有限集合,称为论域。任何子集 $X \subseteq U$ 称为 U 中的一个概念。设 R 是 U 上的一个等价关系, U/R 表示 R 的所有等价类构成的集合, $[x]_R$ 表示包含元素 $x \in U$ 的 R 等价类。一个知识库就是一个关系系统 $K=(U, R)$,其中 U 为论域,是非空有限集合, R 是 U 上的一个等价关系族。

不可识别关系是一个等价关系,若 $\emptyset \neq P \subseteq R$,则 $\bigcap (P$ 中所有等价关系的交集)称为 P 上的不可识别(indiscernibility)关系,记为 $IND(P)$,且有

$$[x]_{IND(P)} = \bigcap_{R \in P} [x]_R$$

这样, $U/IND(P)$ 表示与等价关系族 P 相关的知识,称为 K 中关于 U 的 P 基本知识, $IND(P)$ 的等价类称为知识 P 的基本概念。

给定知识库 $K=(U, R)$,对于每个子集 $X \subseteq U$ 和一个等价关系 $R \in IND(k)$,存在两个子集:

$$\underline{RX} = \bigcup \{Y \in U/R \mid Y \subseteq X\}$$

$$\overline{RX} = \bigcup \{Y \in U/R \mid Y \cap X \neq \emptyset\}$$

分别称它们为 X 的 R 下近似和 R 上近似。集合 $BN_R(X) = \overline{RX} - \underline{RX}$ 称为 X 的 R 边界域; $POS_R(X) = \underline{RX}$ 称为 X

的 R 正域; $NEG_R(X) = U - \overline{RX}$ 称为 X 的 R 负域。显然有

$$\overline{RX} = POS_R(X) \cup BN_R(X)$$

$\underline{RX}$ 或 $POS_R(X)$ 是由那些根据知识 R 判断肯定属于 X 的 U 中元素组成的集合; $\overline{RX}$ 是那些根据知识 R 判断可能属于 X 的 U 中元素组成的集合; $BN_R(X)$ 是那些根据知识 R 既不能判断肯定属于 X 又不能判断肯定属于 $\neg X$ 的 U 中元素组成的集合; $NEG_R(X)$ 是那些根据知识 R 判断肯定不属于 X 的 U 中元素组成的集合。

粗集理论一个重要的应用是用于决策。含有决策属性的信息表称为决策表,一般记作 $T=(U, CUD)$,其中 C 为条件属性集合, D 为决策属性集合。在决策表 T 中,根据决策属性集 D 得到知识库 $K_D=(U, D)$,根据条件属性集获得知识库 $K_C=(U, C)$ 。两个知识库中的等价类分别形成被近似的知识粒子和用于近似的知识粒子。决策过程,一般就是使用条件属性所获得的知识粒子,来近似决策属性得到的知识粒子。如果某个来自于条件的知识粒子蕴含了某个来自于决策的知识粒子,则形成决策规则。

2.3 ILP 一般问题背景

归纳逻辑程序设计(ILP)是统计关系学习中最有潜力的一种方法。ILP形成于逻辑编程和机器学习的交叉处。它使用背景知识、正负实例集来推导描述实例的逻辑程序。推导的逻辑程序由原始的背景知识和归纳的假说组成。本文对于ILP基本知识的简介基于文^[24,25],限于篇幅,我们只给出ILP最一般的问题背景。

子句是构成背景知识的主要形式,一般形式的子句称为程序子句,程序子句的集合称为常规程序;恰包含一个正定文字的子句称为正定子句,其集合称为正定程序;一个谓词定义是逻辑子句的集合,这些逻辑子句的首部具有相同的谓词符号(以及维数)。

ILP一般的问题背景描述如下:

给定正例集 E^+ 和反例集 E^- ,以及背景知识 B ,ILP致力于找到一个理论假说 H ,使得正例集 E^+ 和反例集 E^- 都正确,即覆盖所有正例,排斥所有反例。

在常规语法中,背景知识、证据和假设可以是任何形式的逻辑公式,需要ILP系统满足的条件是,

先验满足性: $B \wedge E^- \models \square$;

后验满足性: $B \wedge H \wedge E^- \wedge \models \square$;

先验必要性: $B \not\models E^+$;

后验充分性: $B \wedge H \not\models E^+$ 。

然而,作为常规语义的一类特殊情形考虑,正定语法限制背景知识和假说是正定子句。ILP问题一般也简化为单谓词的学习。根据学习方式,ILP分为经验式ILP和交互式ILP;根据假说空间的搜索方向,ILP分为自顶向下ILP和自底向上ILP;根据学习任务,ILP分为预测型ILP和描述型ILP。

3 多关系学习中 ILP 和 RS 结合的方法

从知识发现的角度讲,ILP作为多关系学习强有力的工具,在多关系学习技术和方法中占有重要地位。相比于传统的面向属性值表示的挖掘方法,ILP方法具有更强的表达能力和更有效的处理背景知识的能力。然而,ILP一个较大的弱点是不擅长处理存在于实例、背景中的不良数据。基于此,ILP中引入RS方法,用于处理模糊性和不确定性。对于不完备的背景知识(缺失必要的谓词、子句)以及不可识别数据(某些实例既属于正例集,又属于反例集),提出粗配置概念。粗配置放松了常规ILP解决问题的严格要求,从而能从不良数

据中推导出粗的但是有用的假说^[7]。因此许多研究者围绕着 RS、ILP 进行了一定的研究尝试,并提出了若干关系学习模型。

Chunnian Liu 和 Ning Zhong 对于应用 ILP 粗配置处理不良数据进行了较早的研究。1999 年在文[16]中,作者将不良数据分为五种,针对缺失谓词(或子句)的情形,建议了几种 ILP 粗问题配置,但没有提供切实可行的解决方案。2001 年在文[21]中,作者对两种不良数据(不完备背景知识和缺失分类)进行研究。文中较早提出了混合的 RS-ILP 方法,利用扩大训练集的方法推导规则以处理现实应用中的不良数据。即推理时,将类别不明的实例既放入正例集,又放入反例集,期望获得更加泛化的推理结果。但文中设计的工作仍在初级阶段,提出的思想在将来的研究中可能获得较具体的结果。

另一种处理 ILP 中不完美数据的粗集方法见文[18]。由于语言的通用性,该方法基于近邻系统。文中引进了一阶决策系统,给出了寻找规则(子句)的贪心算法。文中进一步描述了使用该方法处理的两类问题。

基于以往的工作,1999 年 T. Y. Lin 和 Qing Liu 对不具有完备性的一阶粗逻辑进行了扩充,提出了一致且完备的一阶粗逻辑。在文[26]中,将等价关系的上、下近似转换为逻辑术语,构建了一阶粗逻辑。虽然该逻辑模型是一致的,但并不完备。在文[27]中,作者修正了一阶粗逻辑,使之是一致且完备的。

对于粗集和 ILP 结合方法,更加形式化的定义以及扩展工作见文[7,17,19,20]。在研究过程中,R. S. Milton, V. Uma Maheswari 和 Arul Siromoney 提出了系列 RS-ILP 模型。在通用模型 gRS-ILP^[20]中,给出了该模型获得两个子句集合(一个子句集合用于说明正实例,另一个集合用于说明反实例)的形式化方法。集合由一个一般 ILP 系统获得。该方法基于 ILP,但并不局限于 ILP。如果某命题问题仅有两个决策类,该方法也可以很容易地应用于命题问题。可变精度模型 VPRSILP^[17],继承了通用模型的所有基本特性,且利用 VPRS 模型特点,允许可控程度的错误分类。cVPRSILP 模型^[7]来源于 VPRSILP 模型,利用属性定义等价关系,这些属性基于用户感兴趣的有限数目子句。

为便于理解,下面我们给出 RSILP 的一般模型:

实例配置的 ILP 系统:一个实例配置的 ILP 系统是一个元组 $S_e = (E_e, B)$, 此处

(1) $E_e = E_e^+ \cup E_e^-$ 是论域, E_e^+ 是正实例(真的基事实)集合, E_e^- 是反实例(假的基事实)集合。

(2) B 是给定的关于正定子句背景知识,满足 (i) 对所有的 $e^- \in E_e^-$, $B \not\models e^-$, (ii) 对某些 $e^+ \in E_e^+$, $B \not\models e^+$ 。

设 $S_e = (E_e, B)$ 是一个实例配置的 ILP 系统, $H(S_e)$ 记为能从 E_e 和 B 推得的所有可能的正定子句空间,称为从 S_e 推得的假说空间。进一步,设 $\Pi(S_e)$ (也记为 $\Pi(E_e, B) = \{P = B \wedge H \mid H \in H(E_e, B)\}$) 表示从 E_e 和 B 推得的所有程序集合,被称为从 S_e (或 E_e 和 B) 推得的程序空间。

我们的目标是寻找程序 $P \in \Pi(E_e, B)$, 满足如下两个条件成立: (iii) 对所有的 $e^- \in E_e^-$, $\Pi \not\models e^-$, (iv) 对所有的 $e^+ \in E_e^+$, $B \models e^+$ 。

实例配置的 RSILP 系统:一个实例配置的 RSILP 系统(简记为 RSILP-E 系统)是一个实例配置的 ILP 系统, $S_e = (E_e, B)$ 中不存在程序 $P \in \Pi(E_e, B)$ 满足条件 (iii) 和条件 (iv)。同时称这样的 ILP 系统是粗配置的。

显然,只有 RSILP 系统,即粗配置的 ILP 系统,才需要借助不确定性处理方法。

称实例配置的 RSILP 系统为 RSILP-E, 单谓词文法的 RSILP-E 系统为 RSILP-ES, 具有声明偏向的 RSILP-ES 系统为 RSILP-ESD。

不可识别关系 $R(S)$: 设 $S = (E, B, L)$ 是一个 RSILP-ESD 系统, S 的不可识别关系,记为 $R(S)$, 是定义如下的 E 上的一个关系: $\forall x, y \in E, (x, y) \in R(S)$ 当且仅当对于任何 $P \in \Pi(S)$, $(P \vdash x \Leftrightarrow P \vdash y)$ (即当且仅当 x 和 y 是关于 $P(S)$ 中的任何逻辑程序 P 固有不可识别的)。

结论 1 对任何 RSILP-ESD 系统 S , $R(S)$ 是等价关系。

结论 2 RSILP-E 系统中,一定存在 $e^+ \in E^+$, $e^- \in E^-$, 满足 $((e^+, e^-) \in R(S))$ 。

VPRSILP 系统:一个可变精度 RSILP-ESD 系统(简记为 VPRSILP-ESD 系统),是一个元组 $S = (S', \beta)$, 其中

(1) $S' = (E, B, L)$ 是一个 RSILP-ESD 系统;

(2) β 是介于区间 $(0.5, 1]$ 间的实数。

S 的 β 正区域: S 的 β 正区域 $Pos_\beta(S)$ 定义为 S 的组合集,该组合集的任一个初等类 $[x]_{R(S)}$ 的条件概率 $P(E^+ | [x]_{R(S)})$ 大于等于 β

$$Pos_\beta(S) = \bigcup_{P(E^+ | [x]_{R(S)}) \geq \beta, \text{ for all } [x]_{R(S)} \in R(S)} \{[x]_{R(S)}\}.$$

S 的 β 负区域: S 的 β 负区域 $Neg_\beta(S)$ 定义为 S 的组合集,该组合集的任一个初等类 $[x]_{R(S)}$ 的条件概率 $P(E^- | [x]_{R(S)})$ 大于等于 β

$$Neg_\beta(S) = \bigcup_{P(E^- | [x]_{R(S)}) \geq \beta, \text{ for all } [x]_{R(S)} \in R(S)} \{[x]_{R(S)}\}.$$

S 的 β 边界区域: S 的 β 边界区域 $Bndr(S)$ 定义为 S 中既不属于 $Pos_\beta(S)$, 又不属于 $Neg_\beta(S)$ 的剩余对象集合

$$Bndr(S) = S - Pos_\beta(S) - Neg_\beta(S)$$

4 多关系学习中的其他粗集方法

不与 ILP 相结合,多关系学习中还存在其它的粗集方法。这些方法主要是模拟单表中的粗集学习机制,对多表情形进行模式挖掘。下面给出一些相关工作:

在文[6]中,研究了基于多决策表的可变精度粗集模型。该模型在每个表中,都能控制可允许的分类错误,即定义了支持决策表与所有决策表和反对决策表与所有决策表两个比率。提出的可变精度粗集模型在多决策表分析中有一定的作用。该文章使用模式刻画对象集合,并利用双重可变精度。本文虽是对多决策表的研究,但该多决策表指的是结构一致,模式未必相同的多表。类似于多专家系统,所获得的关于多表 T 的上下近似是这种情形,说模式 w 属于上下近似,当且仅当模式 w 存在于相当比例表的上下近似中。

文[5]研究多表信息系统(MTIS),在 MTIS 中使用粗集理论和可变精度粗集的方法。在 MTIS 中近似区域的概念根据单表中的概念而定义。基于该实例所属单表中的初等集合,可用于分类 MTIS 中的实例。

该文的研究并不是很深入。第一部分命题情形研究,所涉及到的多表实际上是原来单表的物理分割,即按照列分割原来的表。这样所有表涉及到的对象是完全相同的,只不过是每张表描述这批对象的不同方面。最后的分析实质上或者是将多表合并成一个单表,在这个单表上进行;或者是在每个单表上进行分析,对获得的结果分析并综合。

文[23]研究了 VPRS 模型中,由下近似、上近似和边界区域定义的几种简式,并讨论了它们之间的关系。文中研究了

以保存下近似、上近似和边界区域等为特色的逻辑函数。模式可以作为这些逻辑函数的主要蕴含而精确的或者近似的获得。

相对而言,单纯基于粗集理论的多关系学习方法,研究的广度和深度进展不是很大,因此,理论价值和应用价值不如 RSILP 系列模型,这方面需要进一步的工作。

结束语 本文的工作围绕粗集和多关系学习展开。文章简单给出了多关系学习的产生、实质以及任务,并指出多关系学习具有狭义和广义两种层面。本文最主要的工作是对于粗集和多关系学习进行了简单综述,表明粗集方法在多关系学习中非常重要。对于多关系学习中粗集方法的应用,介绍了其与 ILP 结合的几种途径,重点研究了系列 RSILP 模型。这些模型在多关系学习中具有一定的理论研究价值,并且能够很好地处理不良数据。另外,在多关系学习中存在着其它的粗集方法,但相对而言这些方法限制较严,通用性不高,可利用价值不大。

多关系学习是近年来人工智能领域一个新的研究热点,大量的统计关系学习方法已被提出,新方法也正在不断涌现。粗集方法在多关系学习中,尤其在 ILP 理论结合方向,已经取得一定的进展。但不容否认,RSILP 模型还存在许多有待完善之处,并且将 RS 与其它统计(概率)模型或者关系(逻辑)模型结合,用以处理多关系学习中的问题和任务,也是今后的工作方向。

参 考 文 献

- 1 Dzeroski S. Multi-relational data mining: an introduction. ACM SIGKDD Explorations Newsletter, 2003, 5(1): 1~16
- 2 Raedt L D, Kersting K. Probabilistic logic learning. ACM-SIGKDD Explorations: Special issue on Multi-Relational Data Mining, 2003, 5(1): 31~48
- 3 Domingos P. Prospects and challenges for multi-relational data mining. SIGKDD Explorations, 2003, 5(1): 80~83
- 4 Blockeel H, Sebag M. Scalability and efficiency in multi-relational data mining. SIGKDD Explorations, 2003, 5: 17~30
- 5 Milton R S, et al. Studies on Rough Sets in Multiple Tables. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. In: 10th International Conference, RSFDGrC'2005, Regina, Canada, 2005
- 6 Inuiguchi M, Miyajima T. Variable Precision Rough Set Approach to Multiple Decision Tables. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. In: 10th International Conference, RSFDGrC'2005, Regina, Canada, 2005
- 7 Milton R S, et al. Rough Sets and Relational Learning. Transactions on Rough Sets, 2004, 1: 321~337
- 8 Pawlak Z. Rough Sets. Journal of Computer Science, 1982, 11(5): 341~356
- 9 Pawlak Z. Rough sets: theoretical aspects of reasoning about data. Norwell, MA, USA: Kluwer Academic Publishers, 1992. 1~248
- 10 Liu W. Reasoning about Knowledge Using Rough Sets. Symbolic and Quantitative Approaches to Reasoning with Uncertainty. In: 6th European Conference, ECSQARU 2001, Toulouse, France, 2001
- 11 Pawlak Z. Rough Sets and Decision Algorithms. Rough Sets and Current Trends in Computing. In: Second International Conference, RSCTC 2000, Banff, Canada, 2000
- 12 Greco S, et al. Rough sets theory for multicriteria decision analysis. European Journal of Operational Research, 2001, 129(1): 1~47
- 13 Sakai H, Okuma A. Algorithms and Tools for Rough Non-deterministic Information Analysis. Transactions on Rough Sets I, 2004, 3500: 209
- 14 Siromoney A, Inoue K. Consistency and Completeness in Rough Sets. Journal of Intelligent Information Systems, 2000, 15: 207~220
- 15 Ziarko W. Probabilistic Rough Sets. Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing. In: 10th International Conference, RSFDGrC'2005, Regina, Canada, 2005
- 16 Liu C, Zhong N. Rough Problem Settings for Inductive Logic Programming. New Directions in Rough Sets, Data Mining, and Granular-Soft Computing. In: 7th International Workshop, RSFDGrC'99, Canada, 2005
- 17 Maheswari V U, et al. The variable Precision Rough Set Inductive Logic Programming Model and Web Usage Graphs. New Frontiers in Artificial Intelligence. In: Joint JSAI 2001 Workshop Post-Proceedings, Matsue, Shimane, Japan, 2001
- 18 Midelfart H, Komorowski J. A Rough Set Approach to Inductive Logic Programming. In Rough Sets and Current Trends in Computing. In: Second International Conference, RSCTC 2000, Banff, Canada, 2000
- 19 Milton R S, et al. The Variable Precision Rough Set Inductive Logic Programming Model - a Statistical Relational Learning Perspective. In: Workshop on Learning Statistical Models from Relational Data(SRL2003), IJCAI-2003, Acapulco, Mexico, 2003
- 20 Siromoney A, Inoue K. The Generic Rough Set Inductive Logic Programming Model and Motifs in Strings. RSFDGrC'1999, Yamaguchi, Japan, 1999
- 21 Liu C, Zhong N. Dealing with Imperfect Data by RS-ILP. New Frontiers in Artificial Intelligence. In: Joint JSAI 2001 Workshop Post-Proceedings, London, UK, 2001
- 22 Tamaddoni-Nezhad A, Muggleton S. A Genetic Algorithms Approach to ILP. Inductive Logic Programming. In: 12th International Conference, ILP 2002, Sydney, Australia, 2002
- 23 Inuiguchi M. Several Approaches to Attribute Reduction in Variable Precision Rough Set Model. Modeling Decisions for Artificial Intelligence. In: Second International Conference, MDAI 2005, Tsukuba, Japan, 2005
- 24 Muggleton S, Raedt L D. Inductive Logic Programming: The ory and Methods. Journal of Logic Programming, 1994, (19-20): 681~714
- 25 Lavrac N, Dzeroski S. Inductive Logic Programming: Techniques and Applications. New York: Ellis Horwood, 1994
- 26 Lin T Y, Liu Q. First Order Rough Logic I: Approximate Reasoning via Rough Sets. Fundamenta Informaticae, 1996, 27: 137~153
- 27 Lin T Y, Liu Q. First Order Rough Logic-Revisited. New Directions in Rough Sets, Data Mining, and Granular-Soft Computing. In: 7th International Workshop, RSFDGrC'99, Yamaguchi, Japan, 1999