

基于 Rough Set 理论的“数据浓缩”

王 珏 王 任 苗夺谦 郭 萌 阮永韶 袁小红 赵 凯

(中国科学院自动化研究所 北京 100080)

摘 要 本文讨论了基于 Rough Set (RS) 理论数据浓缩的几个问题. 首先, 介绍了一个基于差别矩阵的属性约简策略, 并给出了数据浓缩的测量; 然后分析了对 UCI 机器学习数据库 40 余个例子的数据浓缩的结果; 最后, 我们强调了在数据浓缩中例外的重要性, 并讨论了不一致数据浓缩.

关键词 Rough Set 理论, KDD, 数据浓缩, 数据挖掘.

分类号: TP18

DATA ENRICHING BASED ON ROUGH SET THEORY

WANG Jue WANG Ren MIAO Duo-Qian GUO Meng

RUAN Yong-Shao YUAN Xiao-Hong ZHAO Kai

(Institute of Automation, Chinese Academy of Sciences, Beijing 100080)

Abstract Data enriching is the attempt to transfer a incomprehensible database, which is made up of a large amount of data, into a concise and natural representation which is more comprehensible for users. Several problems related to data enriching based on Rough Set (RS) are discussed in this paper. A reduction strategy based on discernibility matrix, which defines the attribute significance as the number of the attribute appeared in discernibility matrix, is employed to analyze more than 40 databases from UCI repository. In accordance with the principle of RS, the enriched database is equivalent to the one given under the consistent decision constraint. Such an enriching process is called consistent enriching. But there are some databases, if some instances are deleted from the database, higher evaporation rate of attribute and instance can be obtained. These instances are the so-called exceptions and the enriching process of considering exception is called inconsistent enriching. This paper presents a method, called modified discernibility matrix, to make inconsistent enriching, it is applied to enrich databases with continuous attributes, the results demonstrate that it is very efficient.

Keywords Rough Set theory, KDD, data enriching, data mining.

本文 1997-05-28 收到, 修改文 1998-01-14 收到. 本课题得到国家 863 高科技基金和国家自然科学基金资助. 王 珏, 男, 1948 年生, 研究员, 博士生导师, 目前主要从事人工智能等方面的研究, 包括 ANN、GA、多 Agent 系统、机器学习、知识表示与 DataMining 等. 王 任, 女, 1971 年生, 硕士研究生, 主要从事 DataMining 的研究. 苗夺谦, 男, 1964 年生, 获博士学位, 目前在山西大学数学系任教, 研究兴趣为 Rough Set 理论与 DataMining 理论. 郭 萌, 男, 1973 年生, 硕士研究生, 主要从事 DataMining 图文互补解释的研究. 阮永韶, 男, 1971 年生, 获硕士学位. 研究兴趣为多 Agent 系统理论与 DataMining 理论. 袁小红, 女, 1971 年生, 获博士学位, 研究兴趣为图文理解、Case-based 推理及 DataMining 等. 赵 凯, 男, 1970 年生, 博士研究生, 主要从事遗传程序设计、人工生命及 DataMining 的研究.

1 引言

所谓“数据浓缩”就是将用户难以理解的、由大量数据所构成的数据库,在一定条件下,蒸发那些无意义的数,使其变换为相对自然简洁的表示,以便使用户(或计算机)对这个数据库内容的理解更容易.本文使用“数据浓缩”而未使用“数据挖掘(Data mining)”,一方面是因为本文将讨论限制在基于 RS 约简的方法上,因此可能未能概括数据挖掘这个术语的全部含义;另一方面这种方法似乎使用“浓缩”比“挖掘”可以更形象地说明其本质.

对任务而言,“数据浓缩”与机器学习的区别仅仅在于前者强调整用户(或计算机)对给定数据库理解方式的改变,而不强调它的外延性质,即将外延性质交给人或人-机系统来判断,这个考虑更适合经济、军事等这类不能完全依赖机器决策的领域的需求.

Pawlak 提出的 Rough Set 的约简方法^[4],在机器学习与机器发现中的应用已有报道^[7,9],应用方面更为详细的介绍,见文献[11].本文则是将这个方应用于“数据浓缩”.为此,首先描述了作者提出的一个基于差别矩阵^[5]的有效约简策略,然后使用这个策略对来自 UCI 机器学习数据库的 24 个离散数据库进行数据浓缩.

为了使 RS 理论中的约简方法可以用于浓缩带有连续属性的数据库,我们给出了两个方案:其一,设计了一个将连续属性离散化的方法,本文示出了使用这个方法对 UCI 机器学习数据库中的 17 个带有连续属性数据库的数据浓缩结果;其二,受启于 Ziarko 提出的变精确性约简方法^[6],设计了一个修改差别矩阵的约简方法,并使用这个方法对数据库进行不一致数据浓缩,其本质是根据领域知识优化出合理的离散化方案,这个方法被用于数据库某些连续属性的离散化上.根据对本文定义的数据浓缩测量指标的分析,说明 RS 的约简方法是“数据浓缩”的一个有效工具.

2 数据浓缩的测量

为了说明在领域无关的条件下,基于 RS 约简的数据浓缩的有效性,定义数据浓缩效果的测量是必要的.本文提出了以数据库中所包含有用信息为基础的三个测量:(1)属性蒸发率;(2)实例蒸发率;(3)数据蒸发率.

令原数据库条件属性个数为 N_a , 浓缩后的属性个数为 N_a' ,

$$\text{属性蒸发率 } E_a = \left(1 - \frac{N_a'}{N_a}\right) \times 100\%.$$

令原数据库实例的个数为 N_r , 浓缩后产生的规则数为 N_r' ,

$$\text{实例蒸发率 } E_r = \left(1 - \frac{N_r'}{N_r}\right) \times 100\%.$$

令原数据库数据量为 N , 浓缩后产生的数据量为 M ,

$$\text{数据蒸发率 } E_w = \left(1 - \frac{M}{N}\right) \times 100\%.$$

属性蒸发率表示数据浓缩之后问题所涉及的因素的减少程度,实例蒸发率则说明了给定数据库所需规则的减少程度,而数据蒸发率则描述了这个数据库所需信息的减少程度,由于 $N = N_a \times N_r$, 如果 $M < N_a \times N_r'$, 则间接地说明了规则的简洁程度.对“数据浓缩”而言,我们的经验是,如果 E_a 大于 30% 已可以令人满意,而用户只需要考虑一半的属性那将是相当好了.对

E , 则一般需要大于 60% 才可以使用户满意. 而 E_w 则一般需要大于 85—90%, 甚至更高. 由于数据浓缩效果与数据库的领域或构造过程有十分密切的关系, 因此上述数据只是经验而已.

3 基于差别矩阵的属性约简策略

由于属性蒸发率表示了数据浓缩之后用户 (或计算机) 需要考虑因素的减少程度, 因此它对数据浓缩的效果影响极大, 这样, 获得最小属性集合 (称为最小约简) 将是十分重要的. 目前, 已存在一些多项式时间求属性约简的策略^[1,2], 不幸的是, 我们已在理论上证明它们对最小约简是不完备的^[6]. 尽管本文以下所提出的基于差别矩阵原理的策略也是如此, 但这个策略从理论上所构造的反例均十分特殊, 因此我们猜测, 以下所描述的属性约简策略, 对本文所列出的绝大多数数据库所获得的属性约简集合均是最小的.

基于差别矩阵的属性约简策略可以描述如下 (有关 RS 的基本概念与差别矩阵原理, 分别请参见文献[4,5]):

令 M 是信息表 S (或决策表 T) 的差别矩阵, M 的元素表示为 A_{ij} , 称为项, 它是 S (或 T) 中第 i 个实例与第 j 个实例有差别的所有属性的集合. $A = \{a_1, a_2, \dots, a_n\}$ 是 S (如果是 T , 则只包括条件属性) 所有属性的集合, $A_{ij} \subseteq A, a_k \in A_{ij}$.

令 $p(a_k)$ 是在 M 中属性 a_k 的属性频率函数, 它定义为属性 a_k 在 M 中出现的次数^①, 则 $p(a_k)$ 将定义为属性 a_k 的重要性. 根据这个定义, 属性约简策略可以设计如下:

C_0 是 M 的核集合, 并且 $R = C_0$,

(1) $Q = \{A_{ij}; A_{ij} \cap R \neq \emptyset, i \neq j, i, j = 1, 2, \dots, n\}, M = M - Q, B = A - R$;

(2) 对所有 $a_k \in B$, 计算在 M 中的 $p(a_k)$ 并且令 $p(a_q) = \max_k \{p(a_k)\}$;

(3) $R \leftarrow R \cup \{a_q\}$;

(4) 重复上述过程, 直到 $M = \emptyset$.

集合 R 是信息表 S (或决策表 T) 的一个属性约简.

4 对 UCI 离散数据库的数据浓缩

我们从 UCI 机器学习数据库中挑选了 24 个属性离散数据库, 使用上述方法对它们作数据浓缩, 其结果如表 1.

表 1

数据库名称	原属性数	浓缩后属性数	原实例数	规则数	属性蒸发率	实例蒸发率	数据蒸发率
Balance Scale Weight & Distance(1)	4	4	625	303	0.00	51.52	57.28
Balance Scale Weight & Distance(2)	4	3	20	16	25.00	20.00	36.00
Balloon Databases(1)	4	2	20	3	50.00	85.00	93.00
Balloon Databases(2)	4	4	16	6	0.00	62.50	77.50
Balloon Databases(3)	4	3	20	16	25.00	20.00	36.00
BUPA Liver Disorders	6	3	345	290	50.00	15.94	64.55

① 如果是决策表, 则 M 是删除了关于等价类的项之后所构成的差别矩阵.

续 表

数据库名称	原属性数	浓缩后属性数	原实例数	规则数	属性蒸发率	实例蒸发率	数据蒸发率
Challenger Space Shuttle O-Ring Data	6	3	15	9	50.00	40.00	78.09
Chess End-Game	36	27	3196	100	25.00	96.87	99.47
Fitting Contact Lenses	4	4	24	9	0.00	62.5	71.67
German Credit Data	24	5	999	776	79.17	23.32	88.32
Kinship Domain(Relational)	2	2	112	112	0.00	0.00	0.00
Lung Cancer Data	56	4	32	24	92.86	25.00	95.34
The Monk's Problems(1)	6	3	432	22	50.00	94.91	97.26
The Monk's Problems(2)	6	6	432	254	0.00	41.20	46.23
The Monk's Problems(3)	6	3	432	12	50.00	97.22	98.84
Moral Reasoner	23	1	202	4	95.65	98.02	99.83
Mushroom Database Domain	22	4	8124	23	81.82	99.72	99.96
Postoperative Patient Data	8	8	90	64	0.00	28.89	63.46
Solar Flare Database	10	9	1066	282	10.00	73.55	89.62
Soybean	35	9	307	143	74.29	53.42	93.53
Standardized Audiology Database	38	7	798	69	81.58	91.35	99.18
Space Shuttle Autoland Domain	6	3	15	9	50.00	40.00	78.09
Tic-Tac-Toe Endgame Database	9	8	958	344	11.11	64.09	80.39
Voting database	16	9	435	44	43.75	89.89	97.63

上表中的三个蒸发率表明,大多数数据库均有不同程度的数据蒸发效果.根据 RS 理论可知,给定数据库 W 与蒸发后的数据库 W' 对决策等价.这样,如果蒸发率不为 0,用户(或计算机)就可以从更为简洁的 W' ,而不是从复杂的 W 中获得信息.但是,在上表中恰恰有一个数据库 Kinship,它的三个蒸发率都为 0.在这个数据库中,记录了各种人与人之间的亲戚与朋友关系,这种关系是没有冗余信息的,因此,这个数据库的蒸发率为 0 就是正确的结果了.这个例子说明,数据浓缩的效果与数据库所表述的问题或构造过程(数据收集与整理)有关,如果给定数据库所表述的问题(像 Kindship)没有冗余,数据蒸发率为 0 就是自然的结果;另外,如果数据库是经过仔细整理或刻意构造,这时由于在数据库中信息冗余少,其蒸发率不高也是可能的.总之,一致性数据浓缩只能蒸发在决策一致条件下的那些冗余信息,而不能在此意义之外提高数据蒸发率,那需要不一致的数据浓缩的方法.

5 对 UCI 连续数据库的数据浓缩

对带有连续属性的数据库的浓缩将更为复杂,根据 RS 中等价类的定义可知,在属性值域中的每一个连续量值与一个等价类一一对应,如果将这些等价类使用不同整数标记,这些连续量值将与一个整数集合中的整数一一对应,这意味着对这类数据库存在一个同构映射,使得映射两边的数据库具有相同的约简.这样对带有连续属性数据库的浓缩问题,就可以转换为对整数数据库的浓缩问题.

根据 RS 原理可知,由于这些整数可以理解为等价类的标记,因此这个整数集合越大,等价类就越多,这样对这个属性的划分将越细,其蒸发率越低.为了提高数据的蒸发率,就需要在决策一致的前提下,将这个数据库的连续属性值域映射到一个较小的合理整数集合,这有时也称为离散化.

对连续属性的离散化,本文采用了两种方法:(1)在保证决策一致的前提下,将连续属性

离散化,进而获得它的一致数据浓缩,对这个考虑使用对数据库预处理的方法是自然的,换句话说,根据事先给定的原则将原始数据库变换为一个新的离散表示的数据库;(2) 假设属性值域被离散的区间根据先验知识事先确定,则可以使用本文提出的修改差别矩阵方法,由于这种方法不能保证对决策的一致性,本文称为不一致数据浓缩. 本节将讨论前个问题,而后一个问题将留在下一节讨论.

对连续属性的一致性数据浓缩,我们设计了一个经验公式,它涉及了三个与领域无关的因素:(1) 实例数量 S , (2) 属性个数 A , (3) 分类的类别数量 C (即,决策属性的值域集合的元素个数),如果连续属性映射到的整数集合的元素个数为 I ,则:

$$I = aC + b \log(S)/A + D$$

$$D = 0.382[E - (aC + b \log(S)/A)]$$

其中, a 与 b 是两个可以调整的系数,而第二项考虑的是,如果属性个数越多,产生矛盾的可能性越小,而实例个数越多产生矛盾的可能性越大. 第三项考虑的是,不同属性的数据分布可能不同,令属性值之间距离平均值作为初始阈值, E 则定义为中间距离分类法所得到的属性等价类的个数,而 0.382 是黄金分割原理所决定的系数,即 $(1 - 0.618)$. 这个经验公式可以较好地完成对连续属性的离散化. 对 UCI 机器学习数据库中 17 个连续属性数据库一致数据浓缩的结果如表 2 所示.

表 2

数据库名称	原属性数	浓缩后属性数	原实例数	规则数	属性蒸发率	实例蒸发率	数据蒸发率
Annealing Data	38(6)	8	798	42	78.38	94.74	99.56
Blocks Classification	11(4)	4	5473	668	60.00	84.01	93.55
Credit Approval	16(6)	13	690	263	13.33	61.88	89.31
Cylinder Bands	40(17)	7	512	292	82.05	42.97	93.67
Echocardiogram Data	13(7)	7	132	80	41.67	39.39	82.40
Glass Identification Database	11(9)	7	214	118	30.00	44.86	78.44
Heart Disease Database	14(5)	8	270	108	38.46	60.60	87.30
Hepatitis Domain	20(2)	3	155	96	84.21	38.06	92.00
Johns Hopkins Ionosphere	35(32)	5	351	109	85.29	68.95	96.91
Horse Colic Database	28(7)	3	300	42	88.89	86.00	98.88
Image Segmentation	20(19)	5	210	124	73.68	40.95	89.50
Iris Plants Database	5(4)	3	150	56	25.00	62.67	80.00
Pima Indians Diabetes Database	9(7)	7	767	346	12.50	54.89	74.00
thyroid Disease Database	30(5)	12	2796	459	58.62	83.61	97.85
Wine Recognition Database	14(13)	4	178	78	69.23	56.18	90.09
Wiscons Diagnostic Breast Cancer	32(30)	7	569	92	77.42	83.83	98.01
Wiscons Prognostic Breast Cancer	35(32)	5	198	110	85.29	44.44	94.28

对 UCI 中带有连续属性数据库的数据浓缩效果要比离散属性数据库的效果稍好.

6 修改的差别矩阵方法与不一致数据浓缩

提高蒸发率的有效方法是减少等价类的数量,在上节中对连续属性所采用的方法其本质就是基于这种考虑. 如果希望获得更高的蒸发率就需要进一步减少等价类,但是,其风险是使新产生的数据库存在矛盾. 事实上,在数据库中产生矛盾的实例可以理解为例外. 如果根据先

验知识可以事先推测连续属性的离散区间,那么是否可以发展一个方法,通过有限次的测试,以达到例外数量最小而蒸发率最大的目标呢?这类似一种优化.

受启于 Ziarko 提出的变精确性约简的概念,本文描述了一个修改的差别矩阵方法,其原理如下:

Skowron 提出的差别矩阵,定义为^[5]:

令 $S = (U, A)$ 是一个信息系统(数据库), $U = \{x_1, x_2, \dots, x_n\}$, A 是属性集合, $a(x)$ 是 x 在属性 a 上的值,差别矩阵可定义为

$$(c_{ij}) = \{a \in A: a(x_i) \neq a(x_j)\}, \quad i, j = 1, 2, \dots, n$$

其中 c_{ij} 是 x_i 与 x_j 存在差别的所有属性所构成的集合.

令 E_j 为区间 $[e_{1j}, e_{2j}]$, $j = 1, 2, \dots, k$, e_{1j} 与 e_{2j} 为实数. 修改的差别矩阵定义为

$$(c_{ij}) = \{a \in A: e_{1j} \leq a(x_i) \leq e_{2j} \text{ and } e_{1j} \leq a(x_j) \leq e_{2j}, \text{ 不同时满足}\}, \quad i, j = 1, 2, \dots, n$$

即:如果在数据库中两个实例的属性 a 的值不同时落在同一区间 E_j ,则这两个实例对属性 a 存在差别,否则就是相同的.

给定数据库 $S = (U, A)$ 是一致的,根据领域知识,将它的任一属性 $a \in A$ 的值域 V_a ,根据领域(或经验)知识划分为不同的区间 E_1, E_2, \dots, E_n ,则基于 RS 的不一致数据浓缩如下:

(1) 根据上述对属性值域的划分,检查 S 的一致性,如果存在矛盾,则选出导致矛盾的实例,构成例外集合 Q ,剩余实例构成一致的实例集合 S_q ;

(2) 令 $Card(Q)$ 为计算集合 Q 元素个数的函数, p 为一个领域知识决定的门限值,如果 $Card(Q) < p$,则利用修改的差别矩阵方法,求 S_q 差别矩阵 D ,停机;否则根据领域知识修改区间划分,并重复这个过程.

这个过程可以理解为在领域知识指导下的知识发现过程. 以下是一个选自 UCI 机器学习数据库的例子:

Thyroid 数据库的领域是软骨病,它有 30 个属性,2796 个病例,其中四个连续属性,首先使用在上一节预处理的方法对除了年龄之外的三个属性离散化,可以获得年龄段为 1 的一致性约简,然后,根据对年龄段划分的经验,对年龄属性使用年龄段分别为 10, 15, 20, 25, 30, 分别求其约简,其试验结果及各项测量见表 3.

表 3

年龄段	例外个数	约简后的属性个数	规则个数	数据蒸发率(%)
1	0	12	459	97.85
10	13	15	245	98.63
15	15	16	212	98.76
20	6	16	199	98.85
25	17	15	185	98.94
30	18	15	182	98.91

因为除了年龄之外,在获得上表时未使用任何领域知识,因此蒸发率与例外数量随着年龄段的增大而增加是可以预想到的. 但是,有趣的是,在上表中,当年龄段取 20 时,例外数量骤然减少,而规则数量继续减少,并且蒸发率保持增加,这似乎说明对诊断软骨病(Thyroid)以 20 作为年龄段更为合理. 这个例子说明,不一致的数据浓缩可以获得一致数据浓缩无法获得的机器发现的结果. 这种方法也被应用于 UCI 的其它数据库,并获得了一些有趣的结果. 另外,如果使用“规则使用频率”方法^[10]与上述方法整合,其数据浓缩效果将更为明显.

7 讨论与问题

在 UCI 机器学习数据库中,大约有 100 余个数据库,我们的本意是试图使用 RS 方法对它们全部作一次数据浓缩,从而为进一步的研究作准备.但是,当我们打开这些数据库时,我们发现,在它们当中,一类只有说明没有数据;一类的数据格式非标准或过于凌乱,而其所包含的内容与我们已浓缩过的数据库大同小异;还有一类则是决策属性是连续的,使得不具备专门领域知识将无法处理这类数据库,这三类只能被排除.在 UCI 中有 4 个数据库的实例个数超过 10^4 ,尽管目前我们设计的系统已可以浓缩它们,但是将需要较大的计算开销,而使用并行处理则易如反掌,由于我们这项研究是以实例个数在 10^6 以上并且属性个数在 10^3 以上的数据浓缩为目标,因此,将这类数据库留在对数据浓缩的并行算法的研究中也许更为合理.就我们目前已完成的研究来看,数据库所包含的实例个数在 10^4 并属性个数在 10^2 以内,使用 PC586,一般可在大约 3—5 小时之内完成,至于列在两个表中的关于 UCI 的 41 个数据库中的大多数数据库,几乎均可以在瞬间完成数据浓缩任务.因此,在技术上,这个数量级上领域无关的数据浓缩可以说基本已无问题.

一般地说,本文所讨论的基于 RS 的数据浓缩方法是与领域无关的,尽管目前这种方法对 UCI 机器学习数据库中的大多数例子已获得了很好的蒸发率,但是领域知识还是十分重要的.在数据浓缩中,领域知识的作用大致可以分为两个方面:其一是利用领域知识选择保留特定属性的数据浓缩,或规定领域中属性重要性条件下的数据浓缩,这类问题不难解决;其二是对不一致数据浓缩中例外的选择,例如,在第 6 节中的例子对年龄段划分的选择,就需要领域(或经验)知识,这是我们目前重点研究的问题.

我们认为例外在数据浓缩中起着重要的作用,这个考虑与心理学家对在概念学习时的观察是一致的^[3,10].心理学认为例外是影响对复杂问题理解的重要障碍之一,其表现为需要记忆更复杂且更多的规则,而对数据浓缩任务的解释就是较低的数据蒸发率与更大的计算开销.更深入地研究例外对数据浓缩的影响及对例外的解释,即不一致数据浓缩方法,对建立实用的数据浓缩系统具有重要的意义.

另外,由于本文将讨论聚集在基于 RS 约简的数据浓缩上,因此对建立数据浓缩系统所需考虑的各种问题未作进一步的讨论.事实上,对例外的分析、产生自然语言说明的报告以及与其它方法的整合等均是设计数据浓缩系统所必须考虑的问题.

致 谢 涂序彦教授曾建议作者将本文所描述的方法称为数据浓缩,以与数据压缩相区别.本文使用了涂序彦教授的这个建议,为此我们对涂序彦教授表示衷心的感谢.

参 考 文 献

- 1 Hu X H, Cercone N. Learning in relational databases: a Rough Set approach. *Computational Intelligence*, 1995, 11(2):323—338
- 2 Jelonek J *et al.* Rough Set reduction of attributes and their domains for neural networks. *Computational Intelligence*, 1995, 11(2):339—347
- 3 Nosofsky M, Palmeri J, McKinley C. Rule-plus-exception model of classification learning. *Psychological Review*,

1994, 101(1):53-79

- 4 Pawlak Z. Rough Set—Theoretical Aspects of Reasoning about Data. Dordrecht, Boston, London; Kluwer Academic Publishers, 1991
- 5 Skowron A, Rauszer C. The discernibility matrices and functions in information systems. In: Slowinski R ed. Intelligent Decision Support—Handbook of Applications and Advances of the Rough Sets Theory, Kluwer Academic Publishers, 1992. 331-362
- 6 Wang J, Miao D. Analysis on attribute reduction strategies of rough set. *J Computer Science and Technology*, 1998, 13(2):189-192
- 7 Ziarko W. The discovery, analysis, and representation of data dependencies in databases. In: Piatetsky-Shapiro G, Frawley W J eds. Knowledge Discovery in Databases, AAAI Press/MIT Press, 1991. 177-195
- 8 Ziarko W. Variable precision rough set model. *J Computer and System Sciences*, 1993, 46(1):39-59
- 9 Ziarko W ed. Rough Set, Fuzzy Sets and knowledge discovery. In: Proc International Workshop on Rough Set and Knowledge Discovery, Canada, 1993
- 10 周育健. “规则+例外”的学习与机器学习[硕士学位论文]. 中国科学院自动化研究所, 北京, 1996
- 11 王 珏, 苗夺谦, 周育健. 关于 Rough Set 理论与应用的综述. *模式识别与人工智能*, 1996, 9(4):337-344

全国第十届计算机辅助设计与图形学学术会议征文通知

第十届全国计算机辅助设计与图形学学术会议将于 1998 年 10 月中旬在广西桂林举行. 这次会议由中国计算机学会 CAD 及 CG 专业委员会主办, 中国航天工业总公司二院二十三所等单位承办. 欢迎大家积极投稿并参加会议.

本次论文集中的优秀论文将刊载于《软件学报》增刊、《计算机辅助设计与图形学学报》和《系统工程与电子技术》上.

征文内容如下:

- 数字系统设计自动化
- 集成电路设计自动化
- 计算机图形学; 科学计算可视化
- 系统集成
- 工程数据库
- CSCW
- 机械 CAD
- CIMS 研究、开发与应用
- 并行工程理论、方法、技术和应用
- 智能 CAD
- 虚拟现实
- 多媒体技术及应用
- 拟实制造方法、技术与工具
- INTERNET/INTRANET 在 CAD/CAM 中的应用

重要日期:

截稿日期:1998 年 6 月 30 日
 录用通知:1998 年 7 月 31 日
 正式文稿:1998 年 8 月 31 日

征文要求:

每篇论文附一中文首页, 首页内容包括: 题目、摘要、通讯地址、邮编、作者简介、工作单位、电话/传真/E-mail 地址.

应征论文概不退稿, 敬请谅解.

投寄地址:

通讯地址:北京 142 信箱 203 分箱 32 号, 邮政编码:100854
 电 话:010-68385647
 传 真:010-66819538
 Email: clquan@public.intercom.co.cn
 联系人:刘 嘉, 肖 玲