

基于知识网约束的合理性检验方法

刘楠楠,严洪森,季海

(东南大学复杂工程系统测量与控制教育部重点实验室,江苏南京 210096;
东南大学自动化学院,江苏南京 210096)

摘要:文中首先引入知识网相关的概念和定义,然后通过分析知识网的完整性特性,引入合理性检验理论,根据知识网完备性理论,提出了基于知识网约束的合理性检验方法。然后针对知识网中存在的父子功能约束、功能信息流约束特点,采用自动选择知识点的方法,实现添加知识点到已选知识点中,构成完整的知识网。最后,再通过实例,显示基于知识网约束的知识点选择的合理性检验过程,并根据所选知识点,生成动态知识网模型,表明基于知识网约束的合理性检验方法是可行的。

关键词:知识网;合理性检验;父子功能约束;功能信息流约束

中图分类号:TP391;TH166

文献标识码:A

文章编号:1673-629X(2012)10-0043-04

Method of Rationality Checking Based on Restrictions of Knowledge Mesh

LIU Nan-nan, YAN Hong-sen, JI Hai

(Ministry of Education Key Laboratory of Measurement and Control of CSE, Southeast University,
Nanjing 210096, China;
School of Automation, Southeast University, Nanjing 210096, China)

Abstract:Introduce the related conception and definition about knowledge mesh firstly, and then through the characteristic analysis of the integrity of knowledge mesh, a rationality checking theory is introduced, and according to the theory of the integrity of knowledge mesh, a method of rationality checking based on the restrictions of knowledge mesh is proposed. Considering the characteristics of the restriction between father and son functions, and the restriction between function and information flow, the method of selecting knowledge points automatically is used in this paper, realizing that knowledge points are added to the selected knowledge points to construct an entire knowledge mesh. Finally, the example shows the process of rationality checking of knowledge point selection based on the restrictions of knowledge mesh, and generates a model of dynamic knowledge mesh using the selected knowledge points. And it shows that the method of rationality checking based on the restrictions of knowledge mesh is feasible.

Key words:knowledge mesh; rationality checking; the restriction between father and son functions; the restriction between function and information flow

0 引言

知识网理论是东南大学严洪森教授在2000年提出的理论,该理论是从现行制造系统模式出发,通过考虑其存在的不足问题^[1],另辟蹊径,提出了知识网理论,并在制造行业已经取得一定的成绩。知识网理论已趋近成熟,如知识网自重构、多重集表达式优化、知识网数据库^[2-9]等理论已获得实际应用。除了将知识

网理论应用在制造领域,它还应用到了其他领域,如基于知识网最简约生成的大气环境分析系统中^[10],文中根据大气环境分析系统用户的目标和需求,设计了系统,并将大气环境分析系统表示成知识网,研究并实现了知识网的最简约生成理论与算法,将其改进应用到大气环境分析系统中。基于知识网的大气环境特征分析系统的实现^[11]一文中重点论述了基于知识网的大气环境特征分析系统的软件系统自动生成和界面重组的实现,并在反射技术的基础上论述了程序集的动态调用。

知识网最简约生成算法是采取分层实现的策略,从根知识点出发,一直到底层知识点(不包括底层知识点),通过联系和约束等信息重构出上层知识点,然

收稿日期:2012-02-20;修回日期:2012-05-22

基金项目:国家自然科学基金资助项目(60934008)

作者简介:刘楠楠(1988-),男,安徽淮南人,硕士研究生,研究方向为知识化制造;严洪森,博士,教授,博士生导师,研究方向为知识化制造、生产计划与调度、预测等。

后按照用户的功能需求,从知识网数据库中找到最能匹配用户功能需求的知识点,最后根据知识网模型,重构出整个知识网。在实际软件实现时需要手动选择底层知识点,手动选择可能会造成所选知识点不能满足生成知识网的条件,造成知识网生成失败,从而不能成功自动生成软件系统。

合理性检验是根据某一正确理论对事物进行的符合这一理论规则的检验和检查,动态测量数据的合理性检验方法研究^[12]一文中以异常点剔除的自适应算法作为检验的标准,找到异常点并判断异常点位置,实现对采样数据的合理性检验,从而鉴别采样数据中可能含有的异常值。区域物流系统动力学模型的建立与合理性检验^[13]一文中根据系统动态学与区域物流系统的特点,建立区域物流系统的动态学模型,并通过模型行为试验和模型参数的灵敏度试验来检验整个模型的合理性。合理性检验的方法研究^[14]一文中对距离测量数据的合理性检验是用多项式拟合预报方法来进行。这些方法都取得了很好的效果。

然而,以上合理性检验方法都不兼具知识网的特殊性,不适合知识点选择的合理性检验。为此,文中提出一种基于知识网约束的合理性检验方法来解决最简约生成时可能会出现知识网生成失败的问题,以弥补上述手动选择知识点的不足,从而达到正确生成知识网之目的。

文中在基于知识网最简约生成的大气环境分析系统基础上,对最简约生成算法中知识点的选择方法进行改进,提出基于知识网约束的合理性检验方法。针对知识网中存在的父子功能约束、功能信息流约束特点,采用自动选择知识点的方法,实现添加知识点到已选知识点中,构成完整的知识网。

1 知识网理论模型

众所周知,各种先进制造模式的研究若希望能在实践中出效果,那么,应该从系统的实用性出发,建立与企业制造模式相适应的、由功能模块或子系统组成的先进制造系统。如果把先进制造系统中的每一个功能模块都看作是一个 Agent,那么每个 Agent 都具有独立的功能,即它具有知识、目标和能力^[15]。此外,对于组成先进制造系统中的各个功能模块而言,它们又是由其他更小的子功能模块组成,以此类推,这些功能模块之间还存在着各种各样的联系和约束,功能模块、子功能模块,子子功能模块……,联系以及约束之间就形成了一张网,这个网称为 Agent 网。知识点及其联系也可以构成相应的知识网(Knowledge Mesh),它的结构与 Agent 网(Agent Mesh)结构相同。Agent 网就是实际的制造系统,所以它不能进行逻辑运算,因此可以

将 Agent 之间的联系映射为知识点之间的联系,每个 Agent 映射为与它对等的知识点。在定义知识网完备性理论之前,首先给出用到的基本概念及其相应的定义,详细内容见文献[16],图 1 是对其形象化的表示。

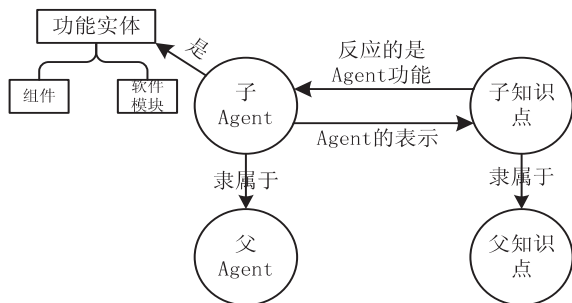


图 1 知识点和 Agent

图 1 清晰地表示了 Agent 的定义、知识点和 Agent 之间的对应关系、父 Agent 和子 Agent 的定义、父知识点和子知识点的定义。从图中可以看出,Agent 和知识点是一一对应的关系。

定义 1 定义知识网的根部的唯一的一个知识点为知识网的根知识点,根知识点存在且唯一^[16]。

定义 2 知识网(Knowledge Mesh) KM 可定义为如下的一个 6 元组:

$$KM = \{P, R, M, F_p, F_R, F_M\} \quad [16]$$

其中:

$P = \{p_1, p_2, \dots, p_m\}$ 是由 m 个知识点 p_1, p_2, \dots, p_m 所组成的有限知识点集。

$R \subseteq P \times P, R = \{r_1, r_2, \dots, r_n\}$ 是由 n 个复合联系 r_1, r_2, \dots, r_n 组成的有限复合联系集。

$M \subseteq P \times P, M = \{m_1, m_2, \dots, m_k\}$ 是由除了父知识点和子知识点以外,其他知识点间 k 个信息联系 m_1, m_2, \dots, m_k 组成的有限信息联系集合, $R \cap M = \emptyset$, 其中 \emptyset 是空集。

$F_p = \{F_{p_1}, F_{p_2}, \dots, F_{p_m}\}$ 是定义在集合 P 上的功能 $F_{p_1}, F_{p_2}, \dots, F_{p_m}$ 的有限集,其中 $\forall F_{p_i} \subset F_p$ 以及 $i = 1, \dots, m, F_{p_i}$ 是知识点 p_i 的所有的功能的有限集合,它是 F_p 的一个子集。

$F_R = \{F_{r_1}, F_{r_2}, \dots, F_{r_n}\}$ 是定义在集合 R 上的继承流及信息流的有限集合, $\forall F_{r_j} \subset F_R$ 和 $j = 1, \dots, n$, 其中 F_{r_j} 是复合联系 r_j 的继承流与信息流的有限集合,它是 F_R 的一个子集。

$F_M = \{F_{m_1}, F_{m_2}, \dots, F_{m_k}\}$ 是定义在集合 M 上的信息流 $F_{m_1}, F_{m_2}, \dots, F_{m_k}$ 的有限集, $\forall F_{m_l} \subset F_M$ 和 $l = 1, 2, \dots, k, F_{m_l}$ 是信息联系 m_l 的所有信息流有限集合,它是 F_M 的一个子集。

下面通过图 2 对知识网和 Agent 网的概念进行详细说明。Agent 网如图 2 右面粗圆圈表示,知识网如图 2 左面细圆圈表示,两者之间的映射关系如图 2 所

- 示。
1. 知识网中有且仅有一个根知识点, 见图中知识点“a”, 根知识点存在且唯一;
 2. 细圆圈表示知识点、加粗圆圈表示 Agent, 两者是一一对应关系;
 3. 小写字母 a 到 i 表示知识点名, 大写字母 A 到 E 表示 Agent 名;
 4. 粗实曲线——复合联系,
细实曲线——信息联系,
虚曲线——知识点和 Agent 之间的映射;
 5. 知识点“b”是根知识点“a”的子知识点, 同时又是知识点“e”的父知识点, 根知识点无父知识点, 最底层知识点无子知识点;
 6. Agent“B”是 Agent“A”的子 Agent, 同时也是 Agent“E”的父 Agent。

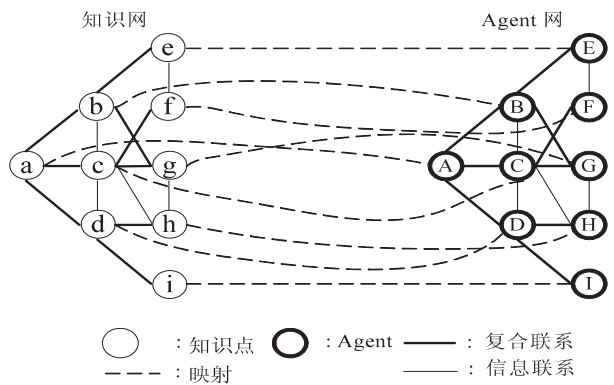


图2 Agent 网和知识网的映射关系

2 知识网完备性理论

由文献[16], 知识网的完备性的定义可表述为知识网中任意两个知识点间的 P 、 M 、 R 以及继承流和信息流是与该知识点完全相联的。

由上面的知识网完备性定义可知, 知识网中的继承流和信息流是影响知识网完备性的主要因素, 在知识网数据库^[9]中, 建立了表示知识点间继承流和信息流的基表、表示父子知识点间约束关系的基表、表示功能信息流关系的基表等。这些基表从数据的角度阐释了知识点之间的关系。

为保证本系统中各个功能模块之间进行信息交流和数据交换的正常工作, 可以利用关系数据库模型, 建立功能信息流约束关系基表以存储知识点功能所需要的功能信息流, 以约束动态知识网的生成, 确保重构生成的知识网的合理性。

表 1 表示的是功能信息流约束关系基表, 它包括三个元组:

1. 知识点功能(kwgn)。
2. 功能版本号(gnversion)。

3. 功能信息流(gninf)。

表 1 功能信息流约束关系基表结构

字段	类型(长度)	主键	非空
kwgn	Char(20)	Yes	Yes
gnversion	Char(10)	Yes	Yes
gninf	Char(100)		

表 2 表示的是知识点继承和信息流的联系基表, 它包括 5 个元组:

1. 根知识点号(rootkw)。
2. 根知识点版本号(rootversion)。
3. 输入知识点号(inkw)。
4. 输出知识点号(outkw)。
5. 父子知识点间(fzdj)。

该表是保存各个节点之间的关系, 其中表 2 中“rootkw(根知识点号)”表示该知识网的根知识点, “version(版本号)”对应于生成的相应知识网, 表中 inkw(输入知识点号)和 outkw(输出知识点号)分别代表继承流和信息流的输入起点和输出终点, 通过判断“fzdj(父子知识点间)”来判断输入知识点和输出知识点号是父子关系还是非父子的信息联系。“fzdj(父子知识点间)”为“否”的保存的是知识网中 $M = \{m_1, m_2, \dots, m_k\}$ 及其上的有限集 $F_M = \{F_{m_1}, F_{m_2}, \dots, F_{m_k}\}$ 的关系数据库基表实现。“fzdj(父子知识点间)”为“是”的保存的是知识网中集合 R 上的继承流和信息流的有限集 $F_R = \{F_{r_1}, F_{r_2}, \dots, F_{r_n}\}$ 及相应关系数据库基表实现。

表 2 知识点间继承流和信息流基表结构

字段	类型(长度)	主键	非空
rootkw	Char(20)	Yes	Yes
rootversion	Char(10)	Yes	Yes
inkw	Char(20)	Yes	Yes
outkw	Char(20)	Yes	Yes
fzdj	Char(3)	Yes	Yes

表 3 是父知识点功能与子知识点功能之间的约束关系表, 它包括四个元组:

1. 父功能(ffkw)。
2. 父功能版本号(ffkwversion)。
3. 子功能(cfkw)。
4. 子功能版本号(cfkwversion)。

父子功能约束关系基表中“ffkw(父功能)”和“ffkwversion(父功能版本号)”对应某一知识网中的父知识点, “cfkw(子功能)”和“cfkwversion(子功能版本号)”对应相应的知识网中父知识点的子知识点, 对于每一个父知识点必须包含其所有的子知识点, 这是由知识网完备性理论所决定的。这里的父知识点可能

还是其他知识点的子知识点,那么对于整个知识网数据库,必须保证所有的子知识点都包含在父知识点中,而父知识点必须完全包含其所有的子知识点。所以,通过建立父子功能约束关系基表来描述知识网的完备性,实现合理性检验功能。

表 3 父知识点功能和子知识点功能之间的约束关系基表结构

字段	类型(长度)	主键	非空
ffkw	Char(50)	Yes	Yes
ffkwversion	Char(50)	Yes	Yes
cfkw	Char(50)	Yes	Yes
cfkwversion	Char(50)	Yes	Yes

3 基于知识网约束合理性检验步骤

基于知识网约束的合理性检验是在知识网最简约生成算法实现的基础上完成的。通过功能信息流约束关系基表、知识点间继承流和信息流基表和父子知识点功能之间的约束关系基表中的约束和联系信息,采用自动选择知识点的方法,将实现添加知识点到已选知识点中,构成完整的知识网。详细步骤如图 3 所示。

4 实例分析

以气象环境保障平台知识化重构开发应用系统为例说明,本系统设计是在 Microsoft.NET 平台下,使用 Microsoft Visual Studio 2008 开发系统,以 Microsoft SQL Server 2000 为后台数据库^[17],使用 C#开发语言^[18]编写程序实现本系统功能。

登录系统进入最简约生成功能模块,选择无约束知识点选项,进入无约束知识点选择界面,进入该界面后,可以根据用户的功能需求手动选择底层功能模块信息,例如,文中选择“处理数据”模块下的污染数据 2D、泰州地图功能,“显示模块”下的显示风场、污染物浓度、显示温度、显示湿度、显示气压功能,点击“已选功能查看”,进入已选功能查看界面。

进入已选知识点查看界面后,点击合理性检验按钮,打开合理性检验界面。点击父子功能约束检验,可以查看满足知识网完备性约束的所缺少的父子功能约束信息,点击功能信息流检验,可以查看满足知识网完备性约束的所缺少的功能信息流信息,点击智能添加知识点按钮,即可自动添加生成知识网所必须的知识点信息。选择生成系统,弹出生成系统对话框,填入根知识点号和版本号,即可生成系统,生成动态知识网视图。

由此实例可见,文中基于知识网约束的合理性检验方法是可行的。

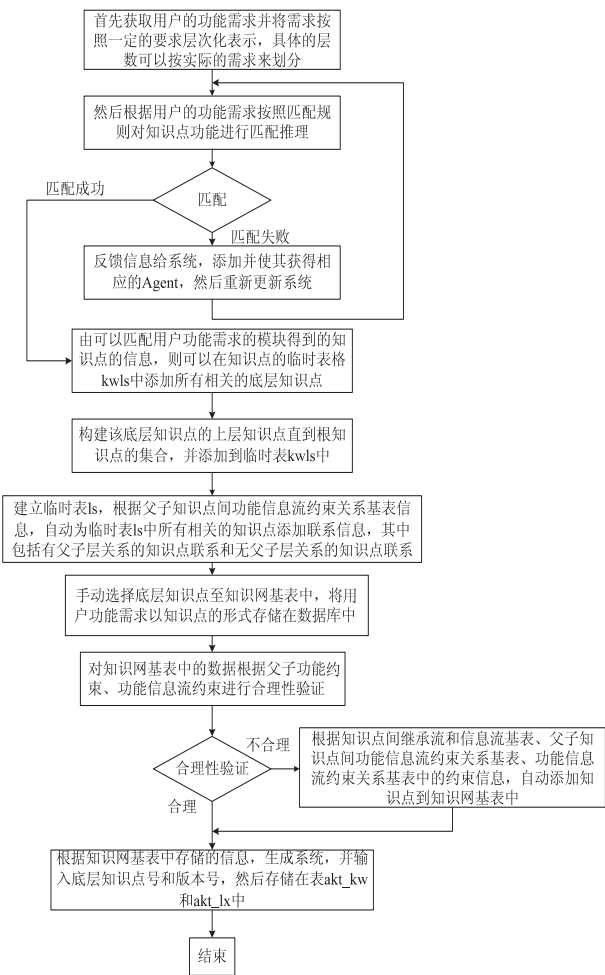


图 3 基于知识网约束合理性检验步骤

5 结束语

最简约生成理论与算法是知识网重构理论的重要核心理论,将最简约生成理论成功地应用到实际应用系统中是知识网重构理论的一大进步。对于最简约生成实现时可能遇到生成知识网失败的问题,文中根据知识网完备性理论^[15],同时利用父子知识点间的功能约束和父子知识点间功能信息流约束成功地实现了基于知识网的合理性检验方法,从而解决了这一问题,满足了知识化制造系统中最简约生成系统的需要。

参考文献:

[1] 严洪森,刘 飞. 知识化制造系统-新一代先进制造系统[J]. 计算机集成制造系统,2001,7(8):7-11.

[2] 薛朝改,严洪森. 基于组件技术的知识化制造系统自重构的实现[J]. 计算机集成制造系统,2004,10(专刊):39-45.

[3] 薛朝改,严洪森. 基于用户功能需求的知识网的自动生成研究[J]. 控制与决策,2005,20(9):996-1001.

[4] 薛朝改,严洪森. 基于 Agent 网的知识网的自重构研究[J]. 计算机集成制造系统,2003,9(11):995-1000.

数据挖掘系统									
DataGrid1									
fscndaccoun	frcvaccount	fsubmittime	fdelivertime	forqaddr	fdestaddr	fareacode	fcpid	fmsgtype	fsmtypr
3	5	2	0	4	0	1	0	3	5
1	5	2	0	1	0	0	0	14	3
3	5	2	0	4	0	1	0	3	5
6	3	4	0	0	0	3	0	5	0
3	5	2	0	4	0	1	0	3	5
6	3	2	0	0	0	3	0	5	0
2	5	2	0	3	0	0	0	14	3
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	0	0	3	0	7	2
3	5	2	0	4	0	1	0	3	5
3	0	1	0	0	0	0	0	0	2
3	5	12	0	0	0	3	0	7	2
4	0	5	0	0	0	0	0	0	5
4	0	5	0	0	0	0	0	0	5
4	0	5	0	0	0	0	0	0	5
6	3	12	0	5	0	0	0	5	0
2	3	12	0	3	0	0	0	14	3
3	0	12	0	0	0	0	0	0	3
1	5	12	0	2	0	0	0	14	3
3	5	12	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
3	5	2	0	4	0	1	0	3	5
4	5	5	0	0	0	3	0	1	2
4	0	5	0	0	0	0	0	0	5
4	0	5	0	0	0	0	0	0	5

属性约简剩余行数 50返回首页

图 3 属性约简数据

参考文献:

[1] Pawlak Z. Rough set; theoretical aspects of reasoning about data[M]. Dordrecht: Kluwer Academic Publishers, 1991.

[2] 王立宏, 吴耿锋. 基于并行协同进化的属性约简[J]. 模式识别与人工智能, 2003, 26(5): 630-635.

[3] 沈 玮, 赵佳宝. 一种新的启发式粗集决策表属性约简算法[J]. 计算机技术与发展, 2010, 20(10): 16-20.

[4] 王国胤, 于 洪, 杨大春. 基于条件信息熵的决策表约简[J]. 计算机学报, 2002, 25(7): 59-66.

[5] 叶东毅. Jelonek 属性约简算法的一个改进[J]. 电子学报, 2000, 28(12): 81-81.

[6] 苗夺谦. Rough Set 理论及其在机器学习中的应用研究[D]. 北京: 中国科学院自动化研究所, 1997.

[7] 胡可云. 基于概念格和粗糙集的数据挖掘方法研究[D]. 北京: 清华大学, 2001.

[8] Pawlak Z. Rough sets[J]. International Journal of Computer and Information Science, 1982, 11(5): 341-356.

[9] 刘 清. Rough 集及 Rough 推理[M]. 北京: 科学出版社, 2001.

[10] 王国胤. Rough 集理论与属性获取[M]. 西安: 西安交通大学出版社, 2001.

[11] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2005.

[12] 李智玲. 基于区分矩阵的粗糙集属性约简算法在数据挖掘中的应用研究[D]. 太原: 太原理工大学, 2007.

[13] Murphy K P. The Bayes net toolbox for matlab[J]. Computing Science and Statistics, 2001, 33(2): 1024-1034.

(上接第 46 页)

[5] 张 昱, 严洪森, 张 平. 基于 B/S 结构的知识化制造自重构子系统的实现[J]. 计算机技术与发展, 2008, 18(1): 25-29.

[6] 杨人子, 严洪森. 知识化制造系统中知识网的结构研究[J]. 计算机集成制造系统, 2008, 14(3): 595-601.

[7] 张 平, 严洪森, 余晓光. 基于混合算法的知识网运算表达式优化[J]. 计算机技术与发展, 2009, 19(3): 32-35.

[8] 杨人子, 严洪森. 基于知识网的知识表达度量方法及其应用[J]. 系统工程理论与实践, 2010, 30(6): 1067-1076.

[9] 王艳斌, 严洪森, 马力伟, 等. 知识化制造系统的知识网数据库设计[J]. 东南大学学报(自然科学版), 2004, 34(增刊): 24-29.

[10] 李金坚, 严洪森, 胡建悦. 基于知识网最简约生成的面向组件软件开发系统[J]. 计算机技术与发展, 2011, 21(1): 125-128.

[11] 殷乾坤, 严洪森, 王方顺. 知识化制造软件系统自动生成的实现[J]. 计算机技术与发展, 2011, 21(1): 16-18.

[12] 倪 兴, 张韵华. 动态测量数据的合理性检验方法研究[J]. 运筹与管理, 2011, 20(4): 113-115.

[13] 何 秋, 桂寿平, 朱 强. 区域物流系统动态学模型的建立与合理性检验[J]. 交通与计算机, 2002, 20(3): 30-33.

[14] 黄家贵, 詹武平. 合理性检验方法的研究[J]. 装备指挥技术学院学报, 2002, 13(4): 93-97.

[15] 张 洁, 刘世平, 李培根. 基于多 Agent 的车间重构模型[J]. 中国机械工程, 2000, 11(4): 432-434.

[16] Yan Hongsen. A new complicated-knowledge representation approach based on knowledge meshes[J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(1): 47-62.

[17] 李晓喆, 张晓辉, 李祥胜. SQL Server 2000 管理及应用系统开发[M]. 北京: 人民邮电出版社, 2002.

[18] 麦中凡, 陆永宁. C#编程语言[M]. 北京: 北京航空航天大学出版社, 2001.

基于知识网约束的合理性检验方法

作者: [刘楠楠](#), [严洪森](#), [季海](#)
作者单位: [东南大学 复杂工程系统测量与控制教育部重点实验室, 江苏 南京 210096](#)
刊名: [计算机技术与发展](#)
英文刊名: [Computer Technology and Development](#)
年, 卷(期): 2012(10)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_wjtz201210013.aspx