

扩展产生式规则的网络故障诊断专家系统

刘培奇, 李增智, 赵银亮

(西安交通大学计算机系统结构及网络研究所, 710049, 西安)

摘要: 根据自然语言理解和网络故障诊断专家系统的特性, 提出了一种扩展产生式规则的知识表示方法. 扩展产生式规则将规则的前提、结论和建议分别用概念图表示, 是一种混合知识表示方法. 在扩展产生式规则知识表示方法的基础上, 设计出了网络故障诊断专家系统的模型. 该系统具有自然语言接口, 可将网络中有关故障问题的汉语语句转换成概念图; 经过专家系统推理后, 可将推理结果和推理过程转换成汉语输出. 为了增强专家系统知识获取的灵活性, 在系统中还增加了 Apriori 关联规则挖掘算法, 通过对陷阱协议数据单元挖掘, 实现在线知识获取. 利用该系统模型已开发出网络故障诊断专家系统原型.

关键词: 自然语言理解; 网络管理; 故障诊断; 专家系统; 知识表示

中图分类号: TP182 **文献标识码:** A **文章编号:** 0253 - 987X(2004)08 - 0783 - 04

Network Fault Diagnosis Expert System Based on Extended Production Rules

Liu Peiqi, Li Zengzhi, Zhao Yinliang

(Institute of Computer Architecture and Network, Xi an Jiaotong University, Xi an 710049, China)

Abstract: According to the characteristics of natural language understanding and expert system of network fault diagnosis, the extended production rule was presented. This rule is a hybrid knowledge representation method that the premise, conclusion and advice of the rules are conceptual graphs. Based on this knowledge representation, a model of expert system with an interface of natural language understanding for network fault diagnosis was designed. It can translate Chinese sentences about faults into conceptual graphs, infer some results from extended production rules, and output Chinese sentences about results and the process of inference. In order to acquire the knowledge flexibility, Apriori algorithm mining association rules from traps protocol data unit of network was added to acquire knowledge online. The prototype of the expert system for faults diagnosis of networks had been implemented in Visual Prolog 5.0 by using this expert system model.

Key words: natural language understanding; network management; fault diagnosis; expert system; knowledge representation

自从 Internet 网络诞生以来, 计算机网络无论在技术上还是在规模和结构上都得到了飞速发展. 面对庞大、复杂的计算机网络, 仅以人工的方式很难处理网络中出现的各种故障. 如何高效、准确地进行网络故障的定位、排除和恢复, 是网络故障管理面临的重大课题. 人工智能的理论、方法和工具, 特别是专

家系统的最新成果, 为计算机网络故障诊断与分析提供了新的研究方向. 但是, 目前的专家系统还存在着接口不够友好、知识表示方法不完备和知识获取方法不灵活等问题. 本文提出一种具有在线知识获取和自然语言接口的网络故障诊断专家系统设计模型, 并依据该模型设计了网络故障诊断专家系

收稿日期: 2003 - 09 - 10. 作者简介: 刘培奇(1972~), 男, 博士生; 李增智(联系人), 男, 教授, 博士生导师. 基金项目: 国家自然科学基金资助项目(60173066).

统原型。

1 网络故障诊断专家系统模型

网络故障诊断专家系统结构如图1所示,下述是其主要模块。

(1) 预处理与词法分析模块. 它可对输入的汉语语句进行切词、同义词处理和词汇识别. 若输入语句中的词汇不在词典中,系统会报告出错信息. 另外,该模块还对输入语句中不重要的词汇进行适当的忽略,增强系统识别句子的灵活性。

(2) 句法分析模块. 利用词典和短文结构文法规则对输入语句进行句法分析,正确的语句会生成句法分析树,以此作为生成概念图的依据。

(3) 语义分析模块. 它由概念图生成器和匹配/推理器组成. 首先,利用概念图生成器将句法分析树转换成概念图的表示形式,再将概念图与图式库中的图式进行匹配/推理,得到符合语义的概念图,即工作图(Working Graphs, WG)。

(4) 专家系统模块. 它由推理机、解释器、知识获取、知识库维护以及一些知识库组成. 推理机利用WG和知识库中的知识进行推理,找出网络故障. 在

推理中,根据工作图是条件还是结论,可自动选择正向推理和反向推理. 为了处理网络故障诊断中的不确定性,专家系统可根据规则可信度(CF)值进行不确定推理,推理的结论为概念图,即答案图. 解释器根据记录的推理路径,对推理结论的合法性作出解释. 知识库由知识库模块负责维护,当专家系统中知识不足时,知识获取模块负责获取。

(5) 语言生成模块. 它由语言生成器和优化模块组成. 利用扩展短语结构文法(APSG)库,将专家系统输出的答案图转换成自然语言文本. 为了使生成的语言符合自然语言规范,可由优化处理器处理。

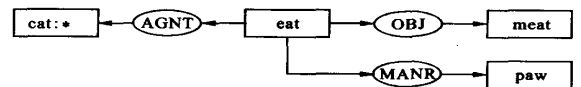
网络故障诊断专家系统主要由专家系统和自然语言接口两部分组成,其中的专家系统为整个系统的核心. 输出结果是以概念图的形式表示推理结论、推理结论合法性的解释和系统异常现象报告,即答案图. 然后,调用自然语言生成部分,对概念图进行搜索,将概念图转换成自然语言. 自然语言接口可实现自然语言理解和输入/输出领域内的汉语语句,为用户提供更加自然、友好的用户接口。

2 扩展产生式规则的表示方法及推理机制

为了便于专家系统推理和自然语言理解,采用了扩展产生式规则知识表示(EPR)方法. 它是基于概念图和产生式规则的混合知识表示。

2.1 概念图

概念图是由 Sowa^[1,2]提出的,直观上它是一个有限、连通、有向的二分图,图中用到了概念节点和关系节点. 概念节点表示研究领域中的一个概念、实体、属性等,关系节点揭示了概念之间的关系. 例如,语句“A cat eats meat with paw”的概念图如图2所示。

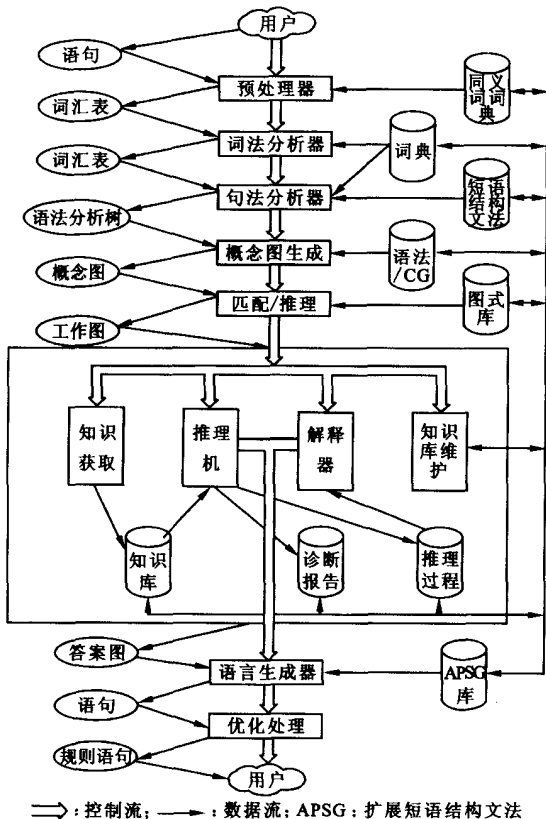


概念节点用矩形框中的概念标识符表示;关系节点用椭圆中的关系标识符表示;有向弧表示节点间的作用关系

图2 概念图表示法实例

2.2 EPR方法

EPR方法是在概念图和产生式规则基础上发展起来的知识表示方法. 产生式规则的前提和结论分别用概念图表示,并且在规则中引入了建议、规则强度 RC、前提重要度 IM 和规则可信度 CF 的概念. 一



⇒: 控制流; →: 数据流; APSG: 扩展短语结构文法

图1 网络故障诊断专家系统结构

条 EPR 知识的形式化表示可参见文献[3].

2.3 推理机制^[3]

网络故障诊断专家系统的推理机制包括不确定推理、完全匹配推理、投影匹配推理和最大连接匹配推理.

完全匹配推理同一般产生式规则的推理方式类似,是一种最理想的推理方式.只要规则的前提概念图与事实概念图完全相同,就认为匹配成功,得到结论概念图.

投影匹配推理是通过相关概念图之间的概念特化实现推理的.设 u, v 为概念图, $u \supset v$, 即 v 是 u 的概化, u 是 v 的特化, 并存在映射 $v \rightarrow u$, 则称子图 v 为 v 在 u 上的投影. 该推理可实现不完全推理.

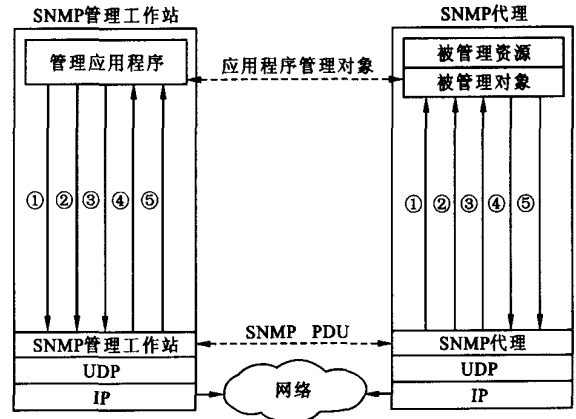
最大连接匹配推理又称相容匹配推理,是建立在概念图之间的概念概化基础上的推理.概念 c_1 和 c_2 是相容的,指 c_1 和 c_2 存在最大公共子类 c_3 ,即对任意的概念节点 c ,如果 $c \subset c_1, c \subset c_2$, 则 $c \subset c_3$. 最大连接就是建立在相容概念基础上的连接,它累加了原始图的相容概念,保留了概念图的不同部分.通过概念图中的概念概化实现最大连接匹配推理.

3 知识获取

故障诊断专家系统知识获取的结构如图3所示,它包括简单知识获取、交互式知识获取和基于算法 Apriori 的关联规则获取.

关联规则获取是利用经典的 Apriori 算法对简单网络管理协议 (SNMP) 的陷阱 (Trap) 数据进行知识挖掘,找出告警关联规则.网络的陷阱数据主要是依靠管理工作站与代理之间的通信获取的. SNMP 的体系结构如图4所示,其中的 SNMP 代理管理若

干个管理对象,而 SNMP 管理工作站管理 SNMP 代理.管理工作站与代理之间利用 GetRequest、GetNextRequest、GetResponse、SetRequest 和 Trap 原语进行通信、查询和改变管理信息库中的信息,其中的 Trap 是当代理发现某个紧急事件后,向管理工作站发送的协议数据单元(PDU).为了使管理工作站及时、有效地进行监控,该工作站用 GET 原语定时轮巡所有代理,以获取有关信息.



: GetRequest ; : GetNextRequest ; : SetRequest ;
 : GetResponse ; : Trap

图4 SNMP体系结构

在网络管理中,SNMP 陷阱的 PDU 格式比较复杂,可将网络告警的 PDU 定义为一个四元组

$$\text{Alarm} = (\text{Time}, \text{Source}, \text{Type}, \text{Severity})$$

其中:Time 是一个时间戳 (Timestamp),表示从系统重新初始化到网络出现故障的时间间隔;Source 为故障源,是网络设备标识,一般为网元 SNMP 代理的 IP 地址;Type 为网络告警类型,分为通用陷阱和特定陷阱,通用陷阱在 SNMP 协议中定义了 7 类,特定陷阱由制造商定义,其数目一般在 100 ~ 1 000 之间;Severity 为网络出现告警的严重程度,分为已清除 (Cleared)、不确定 (Indeterminate)、一般信息 (Information)、警告告警 (Warning)、重大告警 (Major) 和严重告警 (Critical). 若将网络中的每条告警视为一个交易,则所有告警可组成网络告警事务数据库,然后用 Apriori 算法^[4,5]挖掘网络中的故障诊断关联规则.通过知识挖掘得到的关联规则有以下 3 类^[6].

- (1) 告警间关联规则:表示当一个告警出现后,可引起其他告警的出现.
- (2) 告警故障关联规则:表示当一个告警出现后,可引发的计算机网络故障的现象.
- (3) 告警事务关联规则:表示当一个告警出现后,可引发的计算机网络的业务障碍.

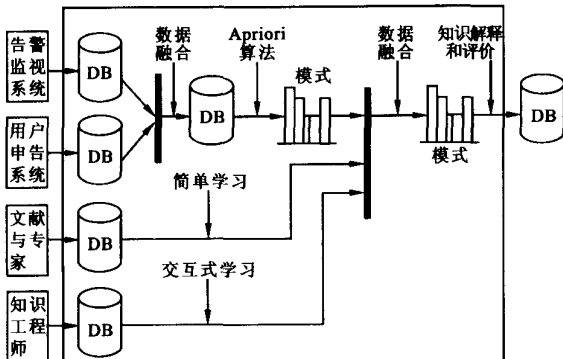


图3 知识获取的结构

4 网络故障诊断专家系统的设计

网络故障诊断专家系统原型是利用 Visual Prolog 语言实现的. 在实现过程中用到 100 多个谓词, 下面描述了几类重要的谓词.

(1) 知识库由 fault.kdb、language.kdb 和 apsg.kdb 等数据库组成, 在数据库中主要有表示语法属性、概念图和规则的谓词. 其中: 语法属性谓词, 如 modal、adjective、noun 等, 都是一元谓词, 表示自然语言中的语法属性; 概念图谓词由三元谓词 conception、四元谓词 relation 和三元谓词 graphs 组成, 分别用来描述概念节点、关系节点和概念图; EPR 规则用七元谓词 rules 描述.

(2) 用谓词 check 和 token 实现词法分析, 其中: token 是三元谓词, 负责提取句子中的词汇; check 是一元谓词, 负责检查词汇的合法性.

(3) 语法分析与生成概念图同时进行, 当分析的是合法语句时, 就生成相应的概念图. 在这一部分中主要使用了 3 类谓词, 即 Is-词法属性、S-语法属性和 Cg-语法属性谓词. 其中: Is-词法属性是一元谓词, 辨认词汇在句子中的语法作用, 如 Is-noun、Is-modal 等; S-语法属性是四元谓词, 对语句进行语法分析, 如 S-verbp、S-verb、S-nounp、S-wh 等; Cg-语法属性是语法属性对应的概念图, 如 Cg-nounp、Cg-verbp 等, 分别表示句子中名词和动词部分对应的概念图.

(4) 语言生成由 Generate 调用谓词 S-NP、S-VP 生成概念图对应的句子, 其中: S-NP 是六元谓词, 它通过对概念图的遍历生成句子的名词短语; S-VP 是七元谓词, 它通过对概念图的遍历生成句子的动词短语.

(5) 专家系统主要由 Gbes 分别调用三元谓词 goes0、goes1 和 goes2 实现完全匹配、投影匹配和最大连接匹配推理, 推理过程记录在 History 表中. 二元谓词 Show-Condition 和 Show-Rule 从推理的 History 表

中逐步取出规则号、规则和规则的条件, 来说明推理结果的真实性. 在推理知识不足时, 二元谓词 Apriori 从网络故障的 Traps 中挖掘故障诊断的关联规则.

5 结 论

本文针对网络管理中存在的不足之处, 提出了一种网络故障诊断专家系统的设计模型. 该模型以扩展产生式规则的表达方法为基础, 具有自然语言接口能力. 为了增强知识获取的灵活性, 在专家系统中应用 Apriori 算法对 Traps 进行关联规则挖掘. 目前利用 Visual Prolog5.0 语言开发了基于该模型的网络故障诊断专家系统原型, 系统对网络故障领域内汉语语句的理解能力可达 87%, 只要进一步完善知识库, 该系统就可以应用到实际网络故障诊断中. 在汉语切词算法的准确性方面还有待于进一步研究.

参考文献:

- [1] Sowa J F. Conceptual structure: information processing in mind and machine [M]. Boston: Addison-Wesley Publishing Co Inc, 1984.
- [2] Sowa J F. Conceptual graphs for database interface [J]. IBM J Res & Dev, 1997, 20(4): 336-357.
- [3] 刘培奇, 李增智, 赵银亮. 扩展产生式规则知识表示方法 [J]. 西安交通大学学报, 2004, 38(6): 587-590.
- [4] Agrawal R, Imielinski T, Swami A. Mining association rules between sets of items in large databases [A]. The ACM SIGMOD Conference on Management of Data, Washington DC, 1993.
- [5] Han Jiawei, Kamber M. Data mining: concepts and techniques [M]. San Francisco, USA: Morgan Kaufmann Publishers, 2001. 227-230.
- [6] 刘康平, 朱海萍. 告警关联与故障诊断专家系统研究与实现 [J]. 计算机工程, 2002, 28(6): 11-68.

(编辑 苗 凌)