

文章编号:1001-9081(2005)02-0283-03

专家系统规则库覆盖度量的研究及实践

李茹¹,赵小东¹,吕国英¹,褚诚缘²

(1. 山西大学 计算机与信息技术学院, 山西 太原 030006;

2. 中国科学院 计算技术研究所, 北京 100080)

(liru@sxu.edu.cn)

摘要:针对专家系统规则库的软件测试和评价是保证专家系统质量与可靠性的重要环节。基于数据流的结构化测试技术,提出了一种规则库执行流图 RBEF,用于描述规则库中不同规则之间相互连结、制约关系,进而给出了一组规则库覆盖度量准则和测试用例的设计方法,最后讨论了一个已实现并付诸于应用的基于 RBEF 的软件测试工具 ESRTS。

关键词:软件测试;专家系统;规则库;覆盖度量准则;软件测试工具

中图分类号: TP311.56 **文献标识码:** A

Study and practice of coverage metrics for rule-base of expert system

LI Ru¹, ZHAO Xiao-dong¹, LU Guo-ying¹, CHU Cheng-yuan²

(1. school of Computer Science & Information Technology, Shanxi University, Taiyuan Shaanxi 030006, China;

2. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: Software testing and evaluation for rule-base of expert system is an important part for ensuring quality and reliability of expert system. Motivated by structural testing techniques based on data flow analysis, a RBEF was proposed, which showed the coupling and restriction relationship of rules. Furthermore, rule-base coverage metrics criteria and test case design methods were put forward. And finally an implemented software testing tool ESRTS based on RBEF was analyzed.

Key words: software testing; expert system; rule-base; coverage metrics criteria; software testing tool

0 引言

专家系统是一种能在某特定领域以人类专家水平解决该领域中疑难问题的计算机软件。如同其他软件一样,专家系统在投入实际运行之前应经过充分的测试以保证其质量,而其质量很大程度上取决于存储在系统中知识的正确性,因此对知识库的测试就成为专家系统软件测试的重要部分。

在基于规则的专家系统中,知识是以 if-then 形式的规则存储的。对规则库的测试,一种通常的方法是从软件功能规约角度出发,通过大量测试用例的运行及之后运行结果和预期结果的比较,得出系统的质量评价^[1]。这种测试方法存在下述缺陷:

1) 测试用例不能保证规则库充分覆盖;

2) 规则库的许多故障可能来源于不可预见的规则间交互作用的结果,而测试用例可能从来都没有激发产生这些规则间的交互;

3) 由于规则库覆盖的不充分性,使得专家系统在质量评价方面存在片面性。

针对上述问题,本文提出一种面向专家系统规则库的确认、验证和测试(Verification, Validation and Testing, VV&T)的方法。这种方法以一种图模型来表示规则库,运用数据流测

试技术^[2],提出一组规则库覆盖度量准则,并在此基础上形成测试用例的设计原则,使规则库得以充分的覆盖,在最大程度上为专家系统的质量评价提供依据。

1 一种规则库的图模型

使用基于图的方法进行规则库结构分析的早期成果包括 CASNET^[3]、KB-Reducer^[4]、COVER^[5]等。这些研究在不同程度上对规则库的 VV&T 提供了方法支持,但侧重各不相同,并不适用于对规则库进行覆盖度量分析。由此,提出了用规则库执行流图 RBEF (Rule-Base Execution Flow Chart) 来描述规则库。

定义 1 基于规则库的图模型是一个有向无环图 $G = (N, E)$, N 是节点的有限集, $E \subseteq N \times N$ 是有向边的有限集。图中的节点包括一个对应于推理起点的入口节点 S , 一个对应于推理终点的出口节点 G , 以及若干中间节点。为了区分图中出现的各类中间节点,下面给出定义:

1) 类节点:规则库作为一种分类系统,可以被划分为若干类,每一类有其最终结论,将这些类结论反映在 RBEF 图中称为类节点,所有类节点对应于出口节点 G ;

2) 子类节点:RBEF 图中从入口节点到出口节点的连续推理链中表示那些中间结论的节点;

收稿日期:2004-07-17;修订日期:2004-10-19 基金项目:国家 863 计划项目(2004AA115460)

作者简介:李茹(1963-),女,山西浑源人,副教授,主要研究方向:多媒体技术、人工智能、智能软件工程; 赵小东(1976-),男,山西晋中人,硕士研究生,主要研究方向:人工智能、软件工程; 吕国英(1964-),女,河北沧州人,副教授,主要研究方向:人工智能;褚诚缘(1962-),男,天津武清人,高级工程师,主要研究方向:智能信息管理。

3) 事实节点:表示那些规则前件中已知输入值的节点,在 RBEF 图中,入口节点 S 对应于所有事实节点;

4) 操作符节点:在 RBEF 图中出现的操作符节点包括 AND, OR, NOT 和 NOFM (NOFM 表示“如果规则前件中 M 个条件中有 N 个条件为真就激发规则”),这些操作符节点作为规则前件中各个子条件的连接符。

规则库的 RBEF 图的构建是由所有规则形成的子图按照推理链顺序组合连接而成,图 1 给出了一个具有 10 条规则的示例规则库的 RBEF 图表示:

- 1) if P_1 and P_2 then R_1
- 2) if P_6 then R_1
- 3) if P_1 and P_2 then R_2
- 4) if P_2 and P_4 then R_2
- 5) if nofm(P_4, P_5) and P_7 then R_3
- 6) if P_3
- 7) if P_8 and R_1 then G_1
- 8) if P_8 or R_2 then G_2
- 9) if P_8 or R_3 then G_4
- 10) if P_8 then G_3

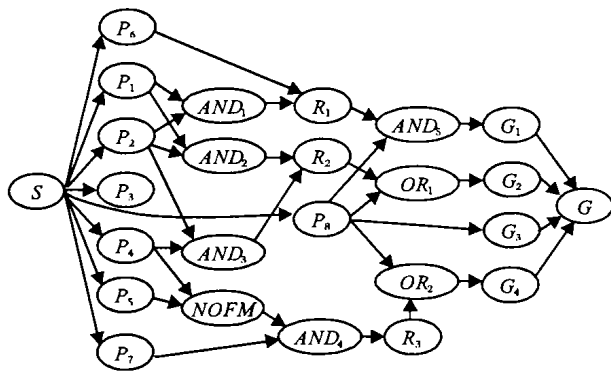


图 1 示例规则库的 RBEF 图表示

其中包括入口节点 S , 出口节点 G , 事实节点 $P_i (i = 1, 2, 3, \dots, 8)$, 子类节点 $R_j (j = 1, 2, 3)$, 类节点 $G_k (k = 1, 2, 3, 4)$ 和若干操作符节点。

定义 2 基于规则库的 RBEF 图的执行路径 (Execution Path) 可以表示为一个 RBEF 子图,它是根据一组具体的事实沿着特定的推理链激发的所有规则集合形成的 RBEF 图。它包括一个入口节点,一个出口节点,一组事实节点及其在推理链上涉及到的所有操作符节点、子类节点和与结论相关的类节点。如图 2 即为图 1 中示例规则库中的一个 RBEF 子图。

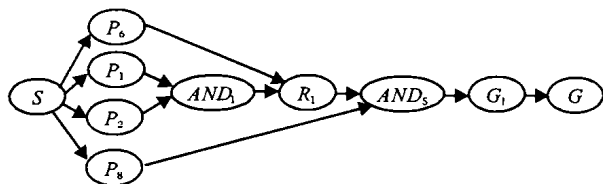


图 2 示例规则库的一个 RBEF 子图

2 基于 RBEF 图的测试准则

参照结构测试的数据流测试以及 Rapps 和 Weyuker 制定的一组数据流测试覆盖准则^[6],我们给出基于规则库 RBEF

图的测试准则。

设 T 为基于规则库 RBEF 图的一个测试用例集, P_T 为与测试用例集 T 相对应的 RBEF 图中执行路径的集合。

准则 1 (类覆盖准则) 测试用例集 T 称为类节点覆盖充分的,当且仅当 P_T 覆盖了 RBEF 中所有的类节点。

准则 2 (子类覆盖准则) 测试用例集 T 称为子类节点覆盖充分的,当且仅当 P_T 覆盖了 RBEF 中所有的子类节点,满足该准则的测试用例集是类覆盖准则测试用例集的超集。

准则 3 (类-子类组合覆盖准则) RBEF 图中,每一个子类节点与每一个类节点之间可能存在若干执行路径(也可能不存在这样的执行路径),测试用例集 T 称为类-子类组合覆盖充分的,当且仅当 P_T 在每一个子类节点与每一个类节点之间的若干执行路径中至少执行一次,满足该准则的测试用例集是类和子类覆盖准则测试用例集的超集。

准则 4 (类-事实组合覆盖准则) RBEF 图中,每一个事实节点与每一个类节点之间可能存在若干执行路径(也可能不存在这样的执行路径),测试用例集 T 称为类-事实组合覆盖充分的,当且仅当 P_T 在每一个事实节点与每一个类节点之间的若干执行路径中至少执行一次。满足该准则的测试用例集是类覆盖准则测试用例集的超集,但不是子类覆盖的超集。

准则 5 (有向边覆盖准则) 测试用例集 T 称为有向边覆盖充分的,当且仅当 P_T 覆盖了 RBEF 图中所有的推理关系(所有有向边)。该准则保证规则库中每条规则在执行路径中以任意可能的方式被覆盖。

显然,在一般意义上,前三个准则的测试强度是依次递增的。准则 1 和准则 2 是最基本的,准则 3 是一个实用的测试覆盖准则,它完全包含准则 1 和准则 2。准则 3 和准则 4 的结合是更强的实用准则。准则 5 覆盖程度最强,由于测试量上的组合爆炸问题限制了其实用性,但对于高可靠性要求的系统,仍要对关键路径采用此准则。

性质 类-子类组合覆盖准则和类-事实组合覆盖准则不存在互包含性,其中符合一个覆盖准则的测试用例集不能保证另一个覆盖准则的成立。

3 测试用例的设计

在实际中,规则的形式是多样的,可以是任意个数的事实节点组合归结为子类节点,或事实节点和子类节点的组合归结为类节点等。如果将其看作一个分类系统,使用分支逻辑,所有事实节点、子类节点和类节点可取值为 True 或 False,那么一个规则库测试用例集可以表示为一个分类系统测试数据集的子集。因此在设计测试用例时,应根据规则库规模大小,可靠性要求高低等,按照一定的测试覆盖准则,从总体测试数据集中选取适当的测试用例,设计原则如下:

- 1) 对每一个类节点,设计一个测试用例使之唯一归结为该节点;
- 2) 增加一组测试用例,使其遍历所有未使用的子类节点;
- 3) 增加一组测试用例,使其覆盖规则库中所有的类-子类组合关系及其由事实节点直接归结为类节点的组合关系;

4) 增加一组测试用例,使其覆盖规则库中所有的类-事实组合关系;

根据上述测试用例的设计原则生成一组测试用例集,其完备性由它对规则库的覆盖程度决定。一个不完备的测试用例集可根据覆盖分析信息从总体测试数据集中不断补充相应的测试用例,以完成一定覆盖准则基础上的规则库质量评价,图3体现了测试用例的设计原则在规则库质量评价模型中应用。

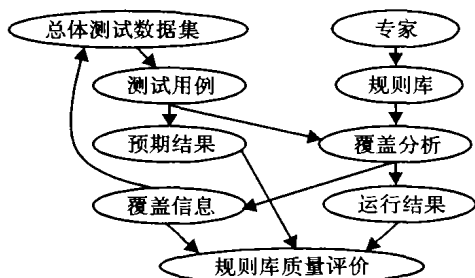


图3 规则库质量评价模型

一般来说,如果设计出的测试用例集满足类-子类组合覆盖准则,我们可以确信规则库中大部分的推理关系已被测试。如果设计出的测试用例集不能满足这一层次的覆盖则意味着规则库的测试是不充分的。

4 工具支持

ESRTS(Expert System Rule-base Testing System)是结合国家863项目“农业专家系统开发平台的测试与评价”开发的专家系统规则库软件测试工具。该工具基于上述理论,以结构测试为主,服务的对象是规则库的开发人员。ESRTS提供了规则库RBEF图的自动生成和规则推理链分析自动化、静态分析的错误检测、测试用例的覆盖分析和质量度量的自动化等,图4给出了其系统结构。

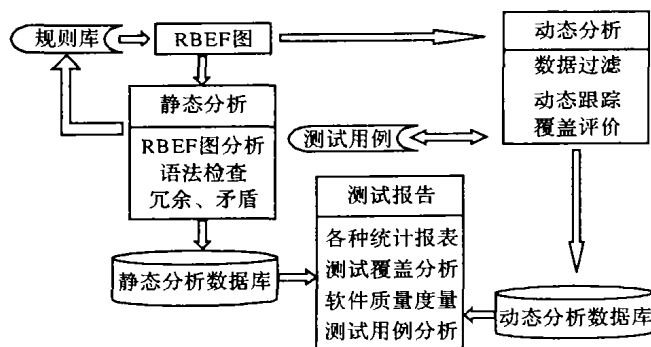


图4 ESRTS的系统结构

4.1 静态分析

规则库RBEF图的构建过程伴随着冗余规则、冲突规则和含糊规则的检测,实现了ESRTS的静态分析。静态分析的目的是鉴别规则库中的循环、虚悬条件、无用结论和孤立规则,它可以通过对RBEF图的深度优先遍历得出信息:

- 1) 如果在遍历时有回溯边,那么该图实际上不是一个RBEF图,在规则库中存在循环;
- 2) 如果在遍历时产生了一个深度优先森林而不是一棵树,那么规则库中存在虚悬条件;

3) 如果存在一个规则结论,而且它所代表的节点并未连接到任何规则的前件节点,那么它是一个无用结论;

4) 如果遍历时产生了一个深度优先森林,并且包括无用结论,那么规则库中存在孤立规则。

4.2 动态分析

ESRTS的动态分析是通过运行测试用例,动态跟踪规则的推理链路来收集规则库的执行信息,如规则库RBEF图中节点的遍历次数,测试执行时的不同取值(True/False),测试用例的覆盖情况以及跟踪程序的异常中止等。

ESRTS通过提供数据过滤和覆盖评价功能,实现了综合测试用例的自动生成。当某一测试覆盖准则的测试用例集不完备时,ESRTS可根据反馈的覆盖信息生成综合测试用例,当然,该测试用例必须通过人类专家审查以决定其在规则库所处理的问题领域中是否有意义。

4.3 测试报告

ESRTS的测试报告利用静态分析和动态分析获取的数据,向用户提供了功能强大的统计报告和详尽的软件质量评估。根据不同的覆盖准则要求,测试报告从规则库的复杂性、正确性、可理解性、完整性等各种角度给出了质量评价和规则库的可靠性估计,为规则库的测试和测试用例的生成及其相关问题领域专家系统规则库的比较提供了重要依据。

5 结语

将软件需求规约用于对专家系统进行功能测试时,具有形式化程度低、规范性差、缺乏客观的测试评判标准和测试充分性不足的缺点。基于RBEF图的测试覆盖度量方法以RBEF图为标准设计测试用例,测试过程也受到了RBEF图的规范,克服了上述不足。通过实践,将测试工具ESRTS应用到“北京农业专家系统开发平台PAID3.0”和“哈尔滨工业大学农业专家系统开发平台ESDK”规则库的测试工作中,效果良好,现已生成了较有价值的测试报告。

参考文献:

- [1] KEEFE O, D E O, Leary. Expert system verification and validation: a survey and tutorial[J]. Artificial Intelligence Review, 1993, 7: 3-42.
- [2] 郑人杰. 计算机软件测试技术[M]. 北京: 清华大学出版社, 1992.
- [3] WEISSM, KULIKOWSKI CA, AMARELS, et al. A model-based method for computer-aided medical decision-making[J]. Artificial Intelligence, 1978, 11: 145-172.
- [4] GINSBERG A. A new approach to checking knowledge bases for inconsistency and redundancy[A]. In Proceedings of Third Annual Expert Systems in Government Conference[C], 1988. 102-111.
- [5] PREECE AD, SHINGHAL R. Foundation and application of knowledge base verification[J]. International Journal of Intelligent Systems, 1994, 9: 683-701.
- [6] RAPPS S, WEYUKER E. Selecting software test data using data flow information[J]. IEEE Transactions on Software Engineering, 1985, 11(4): 367-375.