

# 软件边界组合测试的典型案例分析

刘 畅,王轶辰,刘 斌,阮 镰

LIU Chang,WANG Yi-chen,LIU Bin,RUAN Lian

北京航空航天大学 工程系统工程系,北京 100191

Department of System Engineering, Beihang University, Beijing 100191, China

E-mail: liuchang@dse.buaa.edu.cn

**LIU Chang,WANG Yi-chen,LIU Bin,et al.** Analysis of representative case in software boundary combination testing. *Computer Engineering and Applications*, 2009, 45(20):74-77.

**Abstract:** The concept and technical demand of software system functional combination and boundary testing are introduced. The insufficiency in native military software test case design is analyzed according to the testing process and method. Based on a testing project case in which a representative software defect is not detected, the causes in test case design are analyzed and the necessary of integrating the combination testing with the boundary testing is explained. A boundary combination testing method is proposed based on cause-effect chart. The feasibility and availability are validated through system testing.

**Key words:** software testing; boundary testing; combination testing; function decomposition; cause-effect chart

**摘要:**介绍了软件系统测试中功能组合测试与边界测试的概念和技术要求。针对我国目前军用型号软件测试的主要过程与方法,总结了在测试用例设计方面存在的不足。结合一个型号软件系统测试中未暴露的典型软件缺陷案例,分析了测试用例设计上的原因,说明了将功能组合测试与边界测试结合实施的必要性。提出了一种基于因果图的边界组合测试方法。经系统测试验证了该方法的可行性和有效性。

**关键词:**软件测试;边界测试;组合测试;功能分解;因果图

DOI: 10.3778/j.issn.1002-8331.2009.20.023 文章编号:1002-8331(2009)20-0074-04 文献标识码:A 中图分类号:TP311

随着我国军用装备型号的快速发展,软件在装备中比重的不断增加,型号软件的质量成为制约装备战斗效能的重要因素。据统计,目前我国装备系统中约 2/3 的失效是由软件错误引起的<sup>[1]</sup>。自从总装备部颁布《测试实验室和校准实验室通用要求》GJB2725A-2001 和《军用软件测试指南》GJB141-2004 以来,软件测试,特别是第三方测试作为保证和提高软件质量的重要手段,已经在型号软件研制过程中得到广泛开展。随着第三方软件测试的快速发展,在测试过程和技术方面都出现了一些值得总结提高的新问题。结合一个在型号软件系统测试中未暴露的典型软件缺陷案例,分析了系统测试过程和用例设计上的原因,并提出了一种基于因果图的边界组合测试方法,可以有效保证边界组合测试用例设计的全面性,从而提高测试质量,对我国目前尚处于起步阶段的软件测试工作具有重要的工程借鉴意义。

## 1 软件系统测试技术要求

### 1.1 功能组合测试

功能测试是对软件需求规格说明中的功能需求逐项进行的测试,以验证其功能是否满足要求。功能测试主要使用的测

试方法是等价类划分法,包括有效等价类和无效等价类。但等价类划分法着重考虑输入条件,未考虑各功能输入条件之间的联系即相互组合的情况。在功能组合状态下,软件可能会产生不同于独立功能测试时的新失效,因此必须采用一种适合于描述多种输入条件下的功能组合测试。目前比较适用于工程实践的功能组合测试方法是因果图法<sup>[2-5]</sup>。

### 1.2 边界测试

长期的测试工作经验证明,大量的错误是发生在输入或输出范围的边界上,而不是发生在输入输出范围的内部<sup>[6-7]</sup>。因此针对各种边界情况设计测试用例,可以测试出更多的错误。

边界测试使用边界值分析法设计测试用例。首先应确定边界情况,通常输入和输出等价类的边界,就是应着重测试的边界情况。应当选取正好等于、刚刚大于或刚刚小于边界的值作为测试数据,而不是选取等价类中的典型值或任意值作为测试数据。

## 2 型号软件系统测试用例设计方法

### 2.1 基于需求规格说明的测试用例设计方法

文献[6]介绍了目前我国型号软件系统测试的过程和技术方法。系统测试的主要依据是软件的需求规格说明,目的是检

**作者简介:**刘畅(1979-),男,博士研究生,主要研究领域为软件可靠性工程,嵌入式软件测试;王轶辰(1977-),男,博士,讲师,主要研究领域为软件测试,嵌入式软件自动化测试环境;刘斌(1967-),男,博导,教授,主要研究领域为嵌入式软件测试,软件可靠性工程;阮镰(1938-),男,博导,教授,主要研究领域为软件测试,软件可靠性工程。

收稿日期:2008-10-07 修回日期:2009-01-04

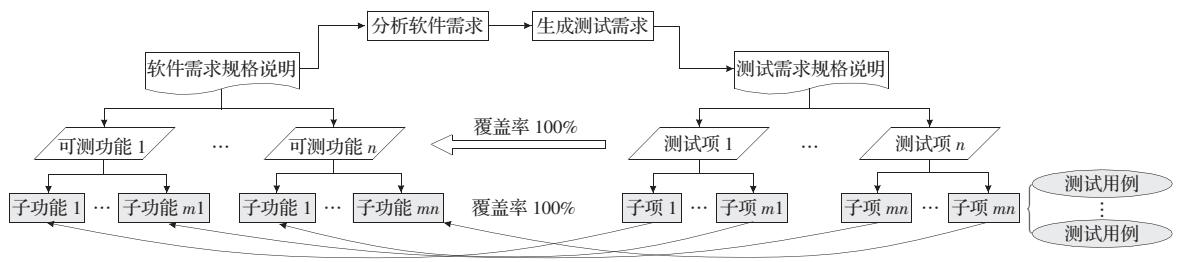


图 1 型号软件系统测试用例设计过程

验软件是否正确实现了需求中规定的各项功能。如图 1 所示,测试人员首先对软件需求文档进行分析,确定全部可测功能,对应每个可测功能设计相应的测试项。然后根据软件子功能的分解情况,将测试项相应分解设计为测试子项。测试项和测试子项的分解结果即形成测试需求规格说明,对测试需求规格说明设计的首要评价准则是测试项和子项要 100% 覆盖对应的软件可测功能和子功能。最后针对每一个测试子项设计具体的测试用例。测试项和子项的分解过程实际构成了测试设计的体系结构,将直接决定测试用例的设计范围。

## 2.2 测试用例设计存在的不足

从图 1 可以看出,目前型号软件测试用例设计过程采用的主要方法是功能分解法。这种层次结构的分解关系把被测软件划分成不同功能模块进行测试。这比较有利于满足和检验功能覆盖率的要求,但是对功能之间的组合关系就无从体现了。GJB2725A-2001 和 GJB141-2004 两份型号软件测试的纲领性文件中也未对功能组合测试做出规定和要求,这就导致了目前测试人员按照自己对软件的理解去设计功能组合测试用例,这样的弊端主要有:

- (1)如果软件需求规格说明包含不易被测试人员察觉的隐含需求,则难以考虑到功能组合的情况;
- (2)测试人员短期内难以深入理解系统,依靠各人对软件需求规格说明理解的不同,设计出的功能组合测试用例也不同,很容易遗漏测试用例,特别是容易忽略边界值组合的情况;
- (3)没有统一规定的功能和边界组合测试用例设计方法,设计驱动力不强,设计随意性比较大。若发生(1)和(2)中的设计缺陷,也缺乏依赖于设计技术的弥补机制。

## 3 基于典型测试案例的分析

在某型飞机综合显示系统软件(以下简称“综显软件”的第三方系统测试中,按照逐级分解软件需求的系统测试设计方式,进行了 100% 覆盖率的测试,各功能及边界测试用例设计充分,并且利用因果图进行了功能组合测试用例设计。但在测试工程结束后仍然出现了一个软件边界值设计缺陷引起的失效,造成了飞机试飞过程中综显软件死机。下面对这个软件失效进行介绍,并着重分析在系统测试中未能发现这一关键软件缺陷的原因。

### 3.1 软件失效介绍

软件“存储航路点”功能:综显软件具有存储航路点的功能。航路点表示一个地理位置点,由经、纬度两个数据元素组成。用户可以通过参数加载设备读入航路点信息,并存储在综显软件内供飞行使用。同时可以利用键盘修改航路点的经纬度数据。存储航路点的范围为[1, 20],单位为:个。

软件“显示航路点”功能:综显软件具有在显示屏上显

示航路点的功能,并在本次飞行线路中的各个航路点之间画出连接线和顺序号,以指示飞行员按计划飞行。可以按照不同量程改变显示的比例大小,量程范围为[30, 90, 120, 240, 360, 720, 960],单位为:km。

失效现象:当飞行员在试飞过程中,把 20 个航路点的经纬度编辑成地理位置相距比较近的 20 个点,然后在飞行时使用综显软件的显示航路点功能。当选择显示最大量程 960 km 时,满足在屏幕上显示全部 20 个航路点的条件。这时,软件在显示全部 20 个航路点功能时发生了内存溢出错误,造成死机。事后发现这个失效是由于显示航路点缓存区在“20”这个边界点上设计不足造成的。

### 3.2 测试设计原因分析

#### 3.2.1 测试用例设计缺陷

经过分析发现,之所以在系统测试中未能发现这个软件缺陷,是因为在测试用例设计中忽略了功能组合条件下边界值的测试,即边界组合测试。下面结合图 2 进行分析,图中:

(1)黑色圆点表示已设计的用例;

(2)白色圆点表示未设计的用例;

(3)带框的白色圆点表示在功能分解时已经设计但是在组合测试中未设计的用例。

纵坐标表示综显软件“存储航路点”功能中的航路点数量;横坐标表示“显示航路点”功能中的显示量程。

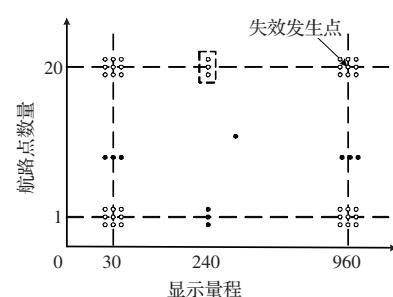


图 2 软件失效及测试用例设计分析

首先单独分析平行于横坐标,即纵坐标为 1 和为 20 的两条线。每条线上的 3 个黑色圆点表示在系统测试中,对综显软件“存储航路点”功能分别进行了下边界[0, 1, 2]个航路点和上边界[19, 20, 21]个航路点的测试,共计 6 组测试用例。类似地,平行于纵坐标的两条线上的黑色圆点表示对“显示航路点”功能也单独进行了显示量程的上、下边界测试。

当屏幕显示航路点时,实际是由上述两个变量进行输入组合。整体观察由这两组平行线相交组成的输入空间时发现,在 4 个相交点处产生了新的边界组合条件。图中白色圆点就表示着这些边界组合条件下的用例设计原则。由于功能组合的影响,在每个边界组合点处,应该新增 9 组边界组合测试用例。而

在实际测试中,测试人员忽略了将航路点为“1”和“20”的这两个边界条件加入到组合测试中,而只选取了在航路点数量为一般等价类值条件下,进行显示量程边界值的测试,导致测试用例未覆盖图2的失效发生点。可见这是由于在二维边界值组合情况下只进行了一维边界值测试,而忽略了另一维,导致在两个功能上边界组合点处的软件设计缺陷没有被测试出来。

综显软件的这个失效是显示航路点数量上边界值设计缺陷引起的,所以除了图2中的失效发生点,在全部显示量程范围内只要满足可以显示全部20个航路点的条件,都会引起相同的失效。

### 3.2.2 导致用例设计缺陷的原因分析

对应2.2节总结的3点不足,分析原因如下:

在综显软件的需求规格说明中对“存储航路点”和“显示航路点”功能的输入范围均做了明确规定,但是在“显示航路点”功能中只规定了量程的输入范围,而没有明确提及航路点显示个数范围上的要求。这就在“显示航路点”功能中形成了隐含边界需求,需要靠测试人员自己去发掘。

由于航路点的实际意义是地理坐标,在输入时是通过参数加载设备读取真实的航路点地理信息,正常情况下20个航路点的距离是非常远的,无法满足在量程为960 km的条件下全部显示,这也间接屏蔽了软件失效。而测试人员如果想构造{20个,960 km}这个测试条件,就需要人工修改模拟的航路点地理信息,这就需要有足够的测试用例设计技术作为驱动力。

由于没有强制规定的测试用例设计技术作为驱动,测试人员在利用因果图进行功能组合测试的过程中,只考虑了满足因果图中一般等价类的“TRUE”和“FALSE”的布尔关系,而忽略了功能组合中可能会产生的边界条件的组合,如图3所示。

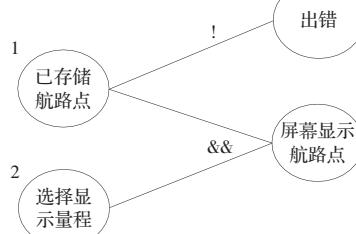


图3 显示航路点功能因果图

综上所述,在对这个功能组合的测试用例设计中,测试人员连续出现了2.2节中所列的3点错误。首先没有完全发掘该功能的隐含需求;然后未能充分理解该功能包含的2个基本功能组合之间的关系,造成一个边界值分析遗漏;最后由于缺乏设计技术上的驱动力,丧失了弥补这个缺陷的机会。

## 4 基于因果图的边界组合测试方法

针对这个未被测试发现的典型软件缺陷,结合上文分析的原因,从增加测试用例设计驱动力入手,提出了一种基于因果图的边界组合测试用例设计方法。

因果图是一种简明表示逻辑条件和相应动作的技术。该技术列出模块化的原因(输入条件)和结果(动作)事件,并为每个事件分配唯一标识号。将“原因”事件垂直列在左边,“结果”事件列在右边。再用线条将原因和结果连接起来,原因与结果的语义便直接或间接地显现出来。然后使用代表布尔表达式的符号标注图,将一个或多个原因与结果对应起来,最终按照判决

规则转换为测试用例。关于因果图的详细概念和使用方法可参阅文献[2-5]。

本方法在现有成熟因果图法的基础上进行扩展,应用边界值分析法对因果图中的“原因”事件进行再次分解。分解方法是在“原因”事件的输入域中,将事件再次分解为:

- (1)满足边界值的“TRUE”事件;
- (2)满足边界值的“FALSE”事件。

这样,一个具有边界值输入条件的“原因”事件就可以分解为:

- (1)满足边界值的“TRUE”子事件;
- (2)满足等价类的“TRUE”子事件;
- (3)满足边界值的“FALSE”子事件;
- (4)满足等价类的“FALSE”子事件。

以本案为例,图4所示即为经过分解后的“显示航路点”功能因果图,标识号1~10标明了分解后的各个子事件。与图3相比可以看出,经过对“原因”事件进行再次分解,对输入域的每个边界值进行了详细的规定,这样在因果图上,不同功能之间所有的边界值输入必须按照“&&”与关系进行组合,这样就100%覆盖了边界组合的情况。这种方法强制测试人员首先将“原因”事件的输入域进行边界值分析,然后再进行组合,是具有强制性的测试设计驱动力。在3.2.1节中忽略的4个边界组合点处的测试用例设计,在本方法的驱动下,也都可以得到保证。

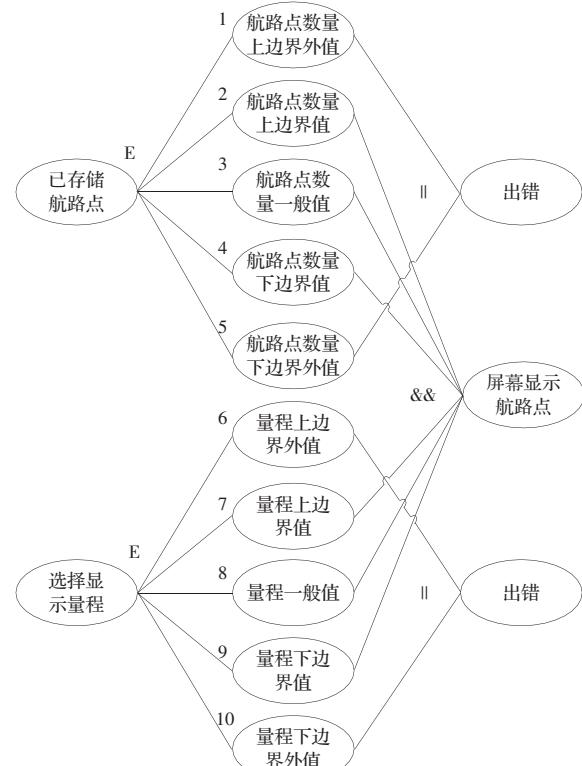


图4 扩展之后的显示航路点功能因果图

特别需要指出的是,对因果图中“原因”事件分解产生的各个子事件之间是“异”约束<sup>[2-4]</sup>,即图中的“E”约束符,子事件中一次只能有一个为“TRUE”。在本例中,生成的测试用例数量是:

$$\sum_{i=1}^5 \sum_{j=6}^{10} (i, j)$$

从因果图生成的测试用例包括了所有输入数据的取

“TRUE”与取“FALSE”的情况,构成的测试用例数目达到最少,且测试用例数目随输入数据数目的增加而线性地增加。

## 5 实例验证

在本型号软件测试项目的第二次系统测试中,采用了本文提出的基于因果图的边界组合测试方法,对同一版本软件的测试用例进行了重新设计与运行。表1列出了两次测试用例设计中对于边界测试用例数量和边界组合覆盖率的比较。在第一次测试中,缺乏足够的测试设计驱动,对于边界组合的情况,只有靠测试人员自己的理解去设计,所以边界组合覆盖率只有约35%。第二次测试采用本文提出的方法后,边界组合覆盖率达到100%。在第二次测试中不仅将该软件失效成功复现,并且对其他忽略的边界组合情况也全部做了测试,切实保证了测试质量。证实了本方法在软件测试工程中的可行性和有效性。

表1 两次测试的比较

	第一次测试	第二次测试
边界测试用例数量	42	96
边界组合覆盖率	约35%	100%

## 6 结论与展望

目前国内的软件测试研究和工程实践尚处于起步阶段,对测试类型进行针对性研究的文献还比较罕见。对于边界组合测试的研究,虽已有了一些成果<sup>[8-9]</sup>,但距普及至测试工程实践中尚需时日。本文提出的基于因果图的边界组合测试设计方法,充分利用了因果图这一成熟工具,具有很强的工程实践意义。并且目前已有商业软件工具支持基于因果图的测试用例设计<sup>[5]</sup>,

用户可以根据本文的方法定义“原因”事件,使这一过程变得更加快捷。

本文提出的方法目前只是从增强测试用例设计驱动力方面来保证测试质量。针对2.2节中的前两点不足,由于涉及到软件开发过程,特别是软件需求规格说明的质量,从而影响因素也更复杂,对软件测试工程是值得进一步研究的课题。

## 参考文献:

- [1] 阮镰,刘斌,陈雪松.软件可靠性测试及其测试环境[J].测控技术,2000,19(2):9-16.
- [2] Dustin E,Rashka J,Paul J.软件自动化测试:引入、管理与实施[M].于秀山,胡兢玉,译.北京:电子工业出版社,2003:157.
- [3] 朱少民.软件测试方法和技术[M].北京:清华大学出版社,2005:115,293.
- [4] 张福源,杨云.基于因果图的软件测试工具研究与设计[J].计算机学报,1993,16(3):219-224.
- [5] 北京旋极信息技术有限公司.Bender RBT:自动化的测试用例设计工具[EB/OL].[2008-09].http://www.watertek.com/product/cekkong-ceshi-bender%20rbt.htm.
- [6] 王轶辰,殷永峰,刘斌.航电设备软件系统测试过程研究[J].测控技术,2004,8:53-56.
- [7] Jorgensen P C.软件测试[M].韩柯,杜旭涛,译.北京:机械工业出版社,2003:70-75.
- [8] 赵瑞莲,董红霞.一种有效的边界测试点选取策略[J].计算机辅助设计与图形学学报,2007,19(2):251-256.
- [9] 董红霞.一种新的边界测试数据自动生成方法研究[D].北京:北京化工大学,2006.
- [10] 马志强,季振洲,胡铭曾.基于分类访问的低功耗联合式cache方案[J].哈尔滨工程大学学报,2007(1).
- [11] Smith J E,Dhadapkar A S.Dynamic microarchitecture adaptation via co-designed virtual machines[C]//2002 IEEE Solid-State Circuits Conference,Digest of Technical Papers,ISSCC,2002:198-199.
- [12] Albonesi D.Selective cache ways:On-demand cache resource allocation[C]//The 32nd Annual ACM/IEEE Int'l Symp on Micro-architecture.Haifa,Israel:IEEE,1999:248-259.
- [13] Ashutosh S.Managing multi-configuration hardware via dynamic working set analysis [C]//Proceedings of 29th Annual International Symposium on 2002.Anchorage, Alaska, USA; IEEE Computer Society,2002:233-244.
- [14] Kim N,Austin T,Mudge T.Challenges for architectural level power modeling(power aware computing)[M].Boston:Kluwer Academic Publishers,2001.
- [15] Guthaus M R,Ringenberg J S.Mibench:A free,commercially representative embedded benchmark suite[C]//Proc of the 4th IEEE Int'l Workshop on Workload Characterization,Dec 2001:3-14.
- [16] Chemical Physics,2005,123(4):1-4.
- [17] Tsai D-M,Lin B-T.Defect detection of gold-plated surfaces on PCBs using entropy measures[J].Int J Adv Manuf Technol,2002(20):420-428.
- [18] 范为福,牛正光,王东云.基于局部熵的织物疵点检测[J].中原工学院学报,2007(5):16-19.
- [19] 徐蝶,刘书明.图像中局部熵描述的合理性及其应用[J].信息技术,2005(11):59-75.

(上接70页)

## 参考文献:

- [1] Intel.Intel XScale microarchitecture.2001.
- [2] Zhang C,Vahid F,Najjar W.A highly-configurable cache architecture for embedded systems[J].ACM SIGARCH Computer Architecture News,2003,31(2):136-146.
- [3] Albonesi D H.Selective cache ways:On-demand cache resource allocation[J].Journal of Instruction Level Parallelism,2000:248-259.
- [4] Zhang C,Vahid F,Najjar W.Energy benefits of a configurable line size cache for embedded systems[C]//IEEE International Symposium on VLSI Design,February 2003.
- [5] Malik A,Moyer B,Cermak D A low power unified cache architecture providing power and performance flexibility[C]//Proceedings of the 2000 International Symposium on Low Power Electronics and Design,2000:241-243.
- [6] 肖斌,方亮.低功耗的可重构数据 Cache 设计[J].计算机工程与设计,2007,28(7):1508-1510,1707.

(上接73页)

- [5] 吴军,李勇.机器视觉技术在疵点检测中的应用[J].上海毛麻科技,2007(3):42-43.
- [6] 祝双武,郝重阳.基于离散小波变换的织物疵点检测方法[J].计算机工程与应用,2007,43(25):111-113.
- [7] Biroli G,Bouchaud J -P,Tarjus G.Are defect models consistent with the entropy and specific heat of glass-formers[J].Journal of