

中创软件丛书

# 软件质量保障与测试

朱 鸿 金陵紫 著

科学出版社

中创软件丛书

# 软件质量保障与测试

朱 鸿 金陵紫 著

科学出版社

1 9 9 7

## 内 容 简 介

本书是系统地论述软件质量保障的原理、技术、方法和理论基础的专著。全书分为三个部分。第一部分介绍软件质量管理的基本思想、方法和技术；第二部分重点介绍软件质量保障中的关键技术——软件动态测试，书中内容较全面地覆盖了在软件研究和软件生产实践中常见的测试方法，反映了当前软件测试领域的国际最新动态和前沿；第三部分探讨软件测试的理论基础。

本书中许多结果来自作者近年来的研究工作，具有较强的系统性，并附有大量实例，内容深入浅出，力求理论与实践相结合。

本书可供计算机科研人员、软件工程管理、开发人员参考，也可作为高等学校计算机、自动化、通信等专业高年级学生、研究生的教材和参考书。

### 图书在版编目(CIP)数据

软件质量保障与测试/朱鸿, 金陵紫著. —北京: 科学出版社, 1997. 8

(中创软件丛书)

ISBN 7-03-006124-1

I. 软… I. ①朱…②金… III. ①软件质量-质量控制  
②软件质量-测试 IV. TP311.5

中国版本图书馆 CIP 数据核字(97)第 12437 号

科学出版社出版

北京东黄城根北街16号

邮政编码: 100717

北京双青印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

\*

1997年8月第 一 版

开本: 850×1168 1/32

1997年8月第一次印刷

印张: 8 5/8

印数: 1—2 200

字数: 215 000

定价: 20.00 元

## 中创软件丛书编辑委员会

主 编 徐家福

副主编 杨芙清 董韫美 景新海

编辑委员 (按姓氏笔画排列)

马绍汉 朱三元 李 未

施伯乐 程建平 董继润

魏道政 鞠丽娜

## 中创软件丛书序言

软件乃程序及其文档之总称. 细言之, 有个体含义、总体含义、学科含义之别. 软件与硬件犹如阴与阳, 天与地, 二者互为依存, 缺一不可, 同为计算机系统之基本成分. 就其学科含义而论, 二者之发展构成计算机科学技术之发展.

软件之作用有三: 一为用户与硬件之接口界面; 二为计算机系统之指挥机构; 三为计算机体系设计之重要依据.

软件一词出现于 60 年代初. 至 60 年代中期, 其规模与复杂度日益增大, 软件开发难以控制, 出现“软件危机”. 于 1968 年北大西洋公约一次学术会议上提出“软件工程”, 视软件开发为工程, 以工程方法开发软件, 研究软件开发范型、模型与方法. 80 年代以来, 结构构件化、开发自动化、表示形式化、接口自然化等逐渐成为研究热点. 研究软件开发环境与工具, 探讨软件体系结构, 企图从总体上解决软件问题. 然而, 鉴于软件开发之复杂度, 迄今工作虽多, 鲜有突破.

40 年来, 我国学者在系统软件、支撑软件、应用软件诸方面均有建树. 惜乎反映研究成果之著作, 问世者甚鲜. 山东中创软件工程有限公司高瞻远瞩, 深感软件人才与教育之重要, 高质量著作作为二者所必需. 因此, 设立了“中创软件丛书基金”, 奖励杰出软件人才, 出版“中创软件丛书”, 并成立“中创软件丛书编委会”, 负责选题、审议、编纂等事. 此举蒙国家科委、中国计算机学会软件专业委员会、各高等院校与科研机构以及广大软件工作者之大力支持. 愿本丛书遵“质量第一”之原则, 自 1997 年起每年均有一至二部著作问世, 冀收促进发展、嘉惠来学之效.

中创软件丛书编委会

1997 年 5 月

• i •

## 序

“质量”一词,观点不同,定义各异.《辞海》一书称质量为产品或工作的优劣程度,其重要性不言而喻.忽视软件质量以致酿成灾难之事例屡见不鲜.软件产品质量之评价标准众多,正确性则居其首.然而断定软件产品之正确,殊非易事.60年代以来,学者致力于“正确性验证”研究,著述虽多,鲜能实用.经验表明,测试乃软件质量保障之重要实用手段.

本书以质量管理和质量保障原理为指导,以作者多年从事软件研究与开发之经验(特别是,作者在英国参加英国贸工部的软件测试项目的经验)为基础,系统论述软件生产质量保障的原理、方法和技术.全书共十四章,除引言一章外,其余诸章可概括为三部分,即质量、测试方法以及测试理论.阅读全书,感其特点如下:

1. 系统性.作者从不同角度,详细阐述“质量”一词的含义,介绍并分析了各种有代表性的测试方法,进而论述了测试充分性的公理系统、充分性度量理论以及软件测试的逻辑基础.循序渐进,系统性强.

2. 科学性.作者寓理论于方法,具体与抽象处理得当.基于前人工作,又不囿于前人之见,对软件测试领域有代表性的文献进行了系统整理,其中包括作者本人的研究成果.立论有据,论述严谨,科学性强.

3. 实用性.书中配有大量实例,便于读者理解原理与方法,所述各种测试方法均经实践检验,富有成效.软件质量保障与测试工作者如能掌握本书内容,则庶几窥其全貌.全书理论联系实际,实用性强.

朱、金二君早岁师从于吾，天资颖慧，根基扎实，勤学敏思，成果突出，著述此书，造福于民。余乐而为之序。

徐家福

1996年12月于南京大学

## 前 言

当你乘坐飞机旅游的时候,你是否知道你所乘坐的飞机是计算机控制的?当你通过电话和远方的亲朋好友交谈的时候,你是否知道电话是通过计算机控制的交换台接通的?当你在收看电视节目或在灯光下阅读时,你是否知道电网是由计算机控制的?当今信息时代,计算机已深入到各行各业,深入到我们生活的方方面面.当我们憧憬着信息高速公路和多媒体将为人类所带来的新生活的时候,你是否也意识到人类越来越依赖于计算机的正确运行了?计算机可依赖吗?本书的中心问题就是计算机藉以完成各种控制和服务的软件的质量问题.

软件质量问题由来已久,早在 60 年代,人们就发现软件质量低下,可靠性差,维护费用高昂.面对软件规模的迅速扩大,复杂性的急剧增大,软件开发能力却远远跟不上应用需求的高速发展,计算机科学家曾惊呼软件面临着危机.时至今日,计算机在工业控制、医疗、通讯、交通、航天、航空、经济、金融等领域的应用对软件的质量提出了前所未有的新的挑战.计算机科学家们认为人类目前所掌握的软件开发技术尚不能达到在这些应用领域中所提出的极高的质量要求.因软件质量问题而造成的重大事故也时有发生.例如,今年夏天就曾发生欧洲用阿里雅娜火箭发射卫星时因软件故障而爆炸的事件.因此,软件质量问题不容乐观.

60 年代末 70 年代初,以 C. A. R. Hoare 和 E. W. Dijkstra 等人为代表的一批计算机科学家针对软件质量问题的严重性,提出了以用形式化功能规约描述用户需求,并证明程序与功能规约的一致性为特征的形式化途径,从而向人们勾绘出一幅大幅度提高软件质量、大幅度减少甚至完全取消软件测试与维护费用的软件开发理想境界.形式化途径为软件生产所勾画的美妙前景吸引了



•

包括本书作者在内的许多从事计算机科学的青年学生,大大地刺激了计算机科学工作者对软件及其开发方法的理论研究,其理论也从证明程序与功能规约的一致性逐步发展成为从形式化功能规约形式地推导与功能规约一致的程序,从由程序员人工地证明和推导发展成为在软件工具的帮助下自动和半自动地进行证明和推导。

然而,在过去的四分之一世纪里,软件生产并未沿着形式化途径所描绘的方向发展。相反,被形式化途径所抛弃的、以软件测试为中心的软件质量保障技术在软件生产实践中得到了迅速发展,软件测试已成为软件生产中必不可少的质量保障手段。近年来,人们开始将多年软件生产实践的经验教训系统地总结成软件生产质量保障的理论。本书的作者从1990年至1994年在英国工作期间,先后参加了由计算机软件工业界和学术界联合进行的研究项目,从而有机会了解到形式化途径的不足之处和两种途径相互补充的必要性。同时深感探讨以软件生产实践经验为基础的软件质量保障方法的理论基础的重要性。本书试图以质量管理和质量保障的一般理论原则为指导,介绍软件生产质量保障的原理、方法和技术。长期以来,软件测试的理论基础一直为理论计算机科学界所忽视。本书的内容包含了作者近年来在软件测试理论基础研究方面的部分研究心得,力图抛砖引玉,引起理论计算机科学界对软件测试理论基础的重视。

本书中所介绍的一部分研究结果曾与英国 Open 大学 Patrick Hall 教授和 John May 博士进行过多次讨论。与英国 Nuclear Electric, plc 公司的 Ian Andrews 博士, Gordon Hughes 博士, Rolls Royce 公司的 Trevor Cockram 先生, Lloyd's Register 公司的 Kevin Czachur 先生, BAeSEMA 公司的 Stewart Gardiner 博士和 Alison Machenzie 女士, Lucas Electrics 公司的 Andy Ashdown 博士, NEI Control Systems 公司的 Jonathan Parkinson 博士等人的合作与接触以及参加由英国贸工部资助的 CONTESSE 和 FASGEP 研究项目均对作者理解软件生产中的问题和本书中

所包含的一些思想的形成有颇多帮助. 本书中的部分内容曾在南京大学计算机科学系软件专业研究生课程中以及软件工业界有关软件质量保障技术的培训讲座上使用, 研究生课程和培训班的参与者们对这部分内容提供了大量的反馈信息. 山东中创软件工程有限公司为本书的出版提供了资助, 南京大学软件研究所徐家福教授对本书的出版给予了大力支持和帮助, 并欣然作序, 北京航空航天大学计算机科学与工程系高仲仪教授和南京大学徐家福教授仔细审阅了全书, 并提出了宝贵的意见. 还有许多专家学者对本书的出版给予了关心和支持, 在此一并致谢.

书中部分研究结果源于国家杰出青年基金、国家“863”计划306主题、国家教委优秀年轻教师基金资助的研究项目.

由于水平所限, 错误之处在所难免, 恳请批评指正.

朱 鸿 金凌紫

1996年9月于南京大学

# 目 录

## 中创软件丛书序言

### 序

### 前言

第一章 引言 .....	( 1 )
参考文献 .....	( 6 )
第二章 质量 .....	( 7 )
2.1 质量与质量观 .....	( 7 )
2.2 软件质量因素 .....	( 12 )
2.2.1 质量因素的层次结构 .....	( 12 )
2.2.2 质量要素间的关系 .....	( 18 )
参考文献 .....	( 21 )
第三章 质量度量 .....	( 22 )
3.1 度量的数学理论 .....	( 22 )
3.1.1 度量的数学结构 .....	( 23 )
3.1.2 标度与标度类型 .....	( 25 )
3.1.3 度量结果运算的有意义性 .....	( 27 )
3.1.4 间接度量 .....	( 29 )
3.2 软件质量度量 .....	( 31 )
3.2.1 McCall 的质量度量 .....	( 31 )
3.2.2 软件复杂性度量 .....	( 35 )
3.3 软件质量的综合评价 .....	( 38 )
参考文献 .....	( 40 )
第四章 质量管理 .....	( 41 )
4.1 质量管理一般原理 .....	( 41 )
4.1.1 概念与术语 .....	( 41 )

4.1.2	质量体系	(45)
4.1.3	质量体系的认证	(48)
4.2	软件生产的质量体系	(50)
4.2.1	质量链与质量环	(50)
4.2.2	软件质量体系要素	(54)
4.2.3	文档与质量记录	(54)
4.2.4	软件质量控制技术	(56)
4.3	配置管理与修改控制	(56)
	参考文献	(59)
<b>第五章</b>	<b>静态测试</b>	<b>(60)</b>
5.1	概念与术语	(60)
5.2	测试方法	(62)
5.2.1	结构化走通	(62)
5.2.2	Fagan 检查	(64)
5.3	测试内容	(65)
5.3.1	需求定义的静态测试	(65)
5.3.2	设计文档的静态测试	(68)
5.3.3	源代码的静态测试	(72)
	参考文献	(74)
<b>第六章</b>	<b>动态测试原理</b>	<b>(76)</b>
6.1	概念与术语	(76)
6.2	测试过程与方法	(78)
6.3	测试的充分性	(81)
	参考文献	(88)
<b>第七章</b>	<b>控制流测试</b>	<b>(89)</b>
7.1	程序结构的流图模型	(89)
7.2	语句测试、分支测试与路径测试	(93)
7.3	循环体测试	(99)
7.4	以复杂度为基础的测试	(102)
7.5	决策谓词测试	(103)
	参考文献	(107)

<b>第八章 数据流测试</b> .....	(108)
8.1 概念与术语 .....	(108)
8.2 数据的定义-引用关系测试 .....	(111)
8.3 数据间交互作用的测试 .....	(114)
8.4 计算环境的测试 .....	(116)
8.5 结构性数据与动态数据的测试 .....	(118)
8.6 过程间的数据流测试 .....	(121)
参考文献.....	(123)
<b>第九章 排错性测试</b> .....	(125)
9.1 植错法 .....	(125)
9.2 变异体测试 .....	(127)
9.2.1 测试方法与过程 .....	(127)
9.2.2 理论基础 .....	(129)
9.2.3 优缺点 .....	(129)
9.3 弱变异体测试 .....	(131)
9.4 小结 .....	(133)
参考文献.....	(134)
<b>第十章 分域测试与功能测试</b> .....	(137)
10.1 软件输入空间的分解.....	(137)
10.1.1 按照需求的分解 .....	(137)
10.1.2 按照实现的分解 .....	(141)
10.1.3 按照需求分解与按照实现分解的结合 .....	(143)
10.2 边界条件测试.....	(143)
10.3 功能函数测试.....	(148)
10.3.1 代数测试 .....	(149)
10.3.2 摄动测试 .....	(150)
参考文献.....	(151)
<b>第十一章 测试方法的揭错能力</b> .....	(152)
11.1 揭错能力比较.....	(152)
11.1.1 评判标准 .....	(152)
11.1.2 随机测试与选择性测试的比较 .....	(153)

11.1.3	静态测试与动态测试的比较 .....	(156)
11.1.4	动态测试方法之间的比较 .....	(158)
11.2	影响测试效率的因素 .....	(161)
11.2.1	人为因素 .....	(162)
11.2.2	软件类型 .....	(162)
11.2.3	错误类型 .....	(163)
11.2.4	测试充分度 .....	(164)
	参考文献 .....	(165)
<b>第十二章</b>	<b>测试充分性的公理系统</b> .....	<b>(167)</b>
12.1	Weyuker 公理系统 .....	(167)
12.1.1	公理系统 .....	(168)
12.1.2	公理之间的关系 .....	(174)
12.1.3	公理系统的推广 .....	(177)
12.1.4	与其它性质的关系 .....	(179)
12.1.5	对充分性准则的评价 .....	(182)
12.2	路径准则的公理系统 .....	(185)
12.2.1	Weyuker 公理与路径准则的性质 .....	(186)
12.2.2	BHB 公理 .....	(189)
12.2.3	公理系统 .....	(195)
12.2.4	对控制流准则的评价 .....	(196)
	参考文献 .....	(203)
<b>第十三章</b>	<b>充分性度量理论</b> .....	<b>(205)</b>
13.1	数值系统 .....	(205)
13.2	公理 .....	(207)
13.3	度量理论性质 .....	(212)
13.4	各种度量方法的特征 .....	(214)
13.4.1	选择性测试与随机测试 .....	(214)
13.4.2	限于程序定义域的和限于功能规约定义域的度量 .....	(220)
13.4.3	结构性测试与排错性测试 .....	(221)
	参考文献 .....	(222)
<b>第十四章</b>	<b>软件测试的逻辑基础——归纳推理</b> .....	<b>(223)</b>

14.1	Weyuker 公理系统的归纳推理解释 .....	(224)
14.1.1	极限识别归纳推理模型 .....	(224)
14.1.2	充分性公理的归纳推理解释 .....	(227)
14.1.3	测试充分性与软件正确性的关系 .....	(230)
14.2	充分性度量公理系统的归纳推理解释 .....	(235)
14.2.1	概率近似归纳推理模型 .....	(236)
14.2.2	充分性度量的归纳推理解释 .....	(239)
14.2.3	测试充分度与软件可靠性的关系 .....	(242)
	参考文献 .....	(246)
	索引 .....	(249)

# 第一章 引言

软件质量问题是本书的中心问题.我们将讨论什么是软件质量,如何提高我们所开发的软件的质量,以及如何测量一个软件产品的质量等问题.

为什么要研究软件质量问题呢?这几乎是一个不言而喻的问题.早在60年代软件危机初期,人们就认识到了软件复杂度高,开发周期长,可靠性差,开发和维护费用大等问题.其中可靠性差就是软件质量问题的集中表现.而软件质量差又是软件维护费用大的主要因素之一.由此可见,软件质量问题由来已久.

近年来,随着计算机应用领域的迅速扩大,计算机软、硬件新技术的不断涌现,人们对软件质量提出了新的更高的要求.目前,计算机已广泛地应用于航空、航天、工业控制、交通、银行、金融、医疗等领域.在这样的应用领域中,软件质量往往关系到人民生命财产和生态环境的安危.一旦软件发生故障,就可能造成人的生命和财产的巨大损失或生态环境的破坏.例如,

- 在西方国家,计算机已用于核电站的控制与安全保护,一旦其软件发生故障,将会造成人员伤亡和环境污染.

- 在伦敦希思罗国际机场,平均每三分钟就有一架飞机起落.如果调度飞机场飞机着陆次序的软件发生故障,将会造成飞机相撞的灾难.

- 波音747等大型客机都使用计算机自动导航系统,如果它发生故障,就可能造成机毁人亡的事故.

- 在西方国家,铁路信号系统已计算机化.英国已计划在不久的将来实现铁路信号和扳道的全部自动化.铁路交通的安全性将进一步依赖于计算机的正常运行.

- 在巴黎地铁系统中,由于使用计算机进行有效的调度和安



全控制,火车的行驶速度得以加快,火车之间的间隔得以缩短.

这些系统中的计算机及其软件具有一个共同的特征,这就是,它们的正确运行关系到人的生命安危.因此,对这类软件的质量往往具有极高的要求.这样的系统究竟质量如何呢?它们是否可以依赖呢?让我们再来看几个例子.

- 由于控制放射性治疗设备的软件错误,在加拿大已经造成了多起癌症病人因受到过量放射性辐射而死亡的事故.

- 在英国,由于放射性治疗设备中计算辐射量的软件错误,发生了癌症病人因放射性辐射剂量太低而得不到适当治疗的事故.

- 在伦敦,救护车调度软件刚投入使用几小时就发生故障,造成急诊病人延误达十几小时.

- 飞机自动控制软件的错误造成了飞机在特技飞行表演时坠毁.

这类因计算机软件错误而造成的事故引起了西方工业国的高度重视,把这类一旦发生故障就可能危及人的生命、财产和生态环境的软件称为安全性第一的软件(safety critical software).它们的根本特征是安全性具有至关重要的意义.

所谓软件安全性,简单地讲,就是计算机软件能够使其所控制的系统始终处于不危及人的生命、财产和自然环境的安全状态的性质.除此之外,安全性第一的软件还有如下重要特征:

### (1) 高质量需求

安全性第一的软件对软件质量往往具有很高的要求.例如,英国 Sizewell-B 核电站安全保护系统的可靠性要求失败率不高于  $10^{-4}$ . 美国联邦航空管理局用于空中交通控制的关键计算机系统的可靠性要求是系统总失效率低于  $10^{-7}$ ,即软件处于故障状态的时间每年不超过三秒钟.一个城市火车调度系统中,对安全性具有关键作用的处理器的失败概率必须低于  $10^{-12}$ /小时.民用航空系统对系统可靠性的需求是人们常常引用的数据,它要求在每小时操作中出现“灾难性事故”的概率低于  $10^{-9}$ .

### (2) 高复杂性