智能科学技术著作丛书

基于进化优化的软件变异测试理论及应用

姚香娟　巩敦卫　著

科学出版社

智能科学技术著作丛书

基于进化优化的软件变异测试理论及应用

姚香娟　巩敦卫　著

科学出版社
北京

内 容 简 介

本书阐述基于进化优化的软件变异测试基本原理及应用，内容主要涉及等价变异体检测、基于相关性分析的变异体约简、变异测试数据进化生成及变异准则改进等。本书除了详细阐述所采用的原理与方法之外，还给出不同方法在基准和工业软件测试中的应用，以及全面的方法对比和结果分析。本书是作者多项国家和省部级科研项目系列研究成果的结晶。

本书可供计算机、自动化等专业的教师及研究生阅读，也可供自然科学、工程技术领域的研究人员及软件测试人员参考。

图书在版编目(CIP)数据

基于进化优化的软件变异测试理论及应用/姚香娟，巩敦卫著. —北京：科学出版社，2014

(智能科学技术著作丛书)

ISBN 978-7-03-042656-7

Ⅰ. ①基… Ⅱ. ①姚…②巩… Ⅲ. ①软件-测试 Ⅳ. ①TP311.5

中国版本图书馆 CIP 数据核字(2014)第 280558 号

责任编辑：孙 芳 王迎春 / 责任校对：桂伟利

责任印制：徐晓晨 / 封面设计：陈 敬

科学出版社出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

北京中石油彩色印刷有限责任公司 印制

科学出版社发行 各地新华书店经销

*

2015 年 3 月第 一 版 开本：720×1000 1/16

2015 年 3 月第一次印刷 印张：9

字数：160 000

定价：65.00 元

(如有印装质量问题，我社负责调换)

《智能科学技术著作丛书》编委会

《智能科学技术著作丛书》序

“智能”是“信息”的结晶，“智能科学技术”是“信息科学技术”的辉煌篇章，“智能化”是“信息化”发展的新动向、新阶段。

“智能科学技术”(intelligence science&technology，IST)是关于“广义智能”的理论方法和应用技术的综合性科学技术领域，其研究对象包括以下几项。

自然智能(natural intelligence，NI)，包括人的智能(human intelligence，HI)及其他生物智能(biological intelligence，BI)。

人工智能(artificial intelligence，AI)，包括机器智能(machine intelligence，MI)与智能机器(intelligent machine，IM)。

集成智能(integrated intelligence，II)，即人的智能与机器智能人机互补的集成智能。

协同智能(cooperative intelligence，CI)，指个体智能相互协调共生的群体协同智能。

分布智能(distributed intelligence，DI)，如广域信息网、分散大系统的分布式智能。

“人工智能”学科自1956年诞生五十余年来，在起伏、曲折的科学征途上不断前进、发展，从狭义人工智能走向广义人工智能，从个体人工智能到群体人工智能，从集中式人工智能到分布式人工智能，在理论方法研究和应用技术开发方面都取得了重大进展。如果说“人工智能”学科的诞生是生物科学技术与信息科学技术、系统科学技术的一次成功结合，那么可以认为，现在“智能科学技术”领域的兴起是在信息化、网络化时代又一次新的多学科交融。

1981年，中国人工智能学会(Chinese Association for Artificial Intelligence，CAAI)正式成立，几十年来，从艰苦创业到成长壮大，从学习跟踪到自主研发，团结我国广大学者，在人工智能的研究开发及应用方面取得了显著的进展，促进了智能科学技术的发展。在华夏文化与东方哲学影响下，我国智能科学技术的研究、开发及应用，在学术思想与科学方法上具有综合性、整体性、协调性的特色，在理论方法研究与应用技术开发方面取得了具有创新性、开拓性的成果。智能化已成为当前新技术、新产品的发展方向和显著标志。

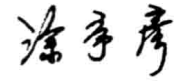
为了适时总结、交流、宣传我国学者在“智能科学技术”领域的研究开发及应用成果，中国人工智能学会与科学出版社合作编辑出版《智能科学技术著作丛书》。需要强调的是，这套丛书将优先出版有助于将科学技术转化为生产力以及对社会和国民经济建设有重大作用和应用前景的著作。

我们相信，有广大智能科学技术工作者的积极参与和大力支持，以及编委的共同努力，《智能科学技术著作丛书》将为繁荣我国智能科学技术事业、增强自主

创新能力、建设创新型国家做出应有的贡献。

为庆祝《智能科学技术著作丛书》出版，特赋贺诗一首：

智能科技领域广，
人机集成智能强。
群体智能协同好，
智能创新更辉煌。



中国人工智能学会荣誉理事长

2005 年 12 月 18 日

前　言

提高软件可靠性的重要途径之一是在软件投入使用之前进行大量测试，以发现软件中存在的缺陷甚至错误。变异测试作为一种面向缺陷检测的测试方法，具有排错能力强、方便灵活及自动化程度高等显著优点，目前，在单元测试、接口测试、并行软件测试及面向对象软件的测试等方面都得到了广泛应用。研究结果表明，与传统测试方法相比，变异测试生成的测试数据具有更好的错误检测能力。

尽管人们对变异测试的研究已有几十年时间，也取得了很多可喜的研究成果，但是从已有的研究成果来看，尚有如下核心问题没有得到有效解决：

(1) 等价变异体问题。事实证明，等价变异体是广泛存在的，而等价变异体的存在给变异测试带来诸多不便。但是，等价变异体在所有变异体中能够占到多大的比例，还没有比较系统的研究结论。另外，所有不能被现有测试数据杀死的变异体并不一定都是等价变异体，这类变异体称为顽固变异体。了解了顽固变异体的分布特性，同样会对变异测试提供很多帮助。此外，等价变异体和顽固变异体之间的数量关系也是一个值得研究的课题。

(2) 变异体数量庞大的问题。利用传统变异测试生成变异体的数量庞大，从而限制了变异测试在实际软件中的应用。虽然已经有多种变异体约简方法，但是已有方法适用的语句类型单一，作用范围窄，约简幅度小。另外，现有方法均没有考虑变异算子本身的优劣，使得选择过程具有很大的盲目性，这说明研究新的变异体约简理论与方法势在必行。

(3) 生成变异测试数据的效率低下。有些变异体和原程序的语义非常接近，利用传统方法很难生成杀死这些变异体的测试数据。虽然已有多种变异测试数据生成方法，但是已有方法基本上是基于传统的测试技术，而没有考虑变异测试的特殊问题，这就导致求解变异测试数据生成问题的效率受到很大限制，这说明研究高性能的变异测试数据生成方法迫在眉睫。

鉴于此，作者在主持的多项国家和省部级科研项目的资助下，一直从事基于进化优化的变异测试理论与方法的研究工作，给出了基于人工分析的等价变异体和顽固变异体检测方法，以期为等价变异体问题的研究提供帮助；研究了基于相关性分析的变异体约简方法，以解决变异体数量庞大的问题；建立了变异测试数据生成问题的多目标优化模型，并提出相应的进化方法进行求解，以提高测试数据生成的效率和质量；提出了一种基于语句占优关系的变异体检测方法，以降低变异测试的成本；最后，将所提方法应用于大量程序的测试中，并与现有方法进行

比较，以评价所提方法的性能。这些成果不但丰富了软件变异测试理论与方法，提高了软件测试效率，而且拓展了遗传算法的应用范围，因此，具有重要的理论意义和实用价值。

本书在各章首先引出问题的研究动机，然后详细阐述所提理论与方法的具体思想和步骤，最后给出所提方法在实际软件测试中的应用，通过不同方法的对比来验证所提方法的性能。在撰写本书的过程中，尽量做到思路清晰，语言表达准确、通俗易懂，各章节数学符号统一，公式规范，图表清楚。

在撰写本书的过程中，博士研究生张功杰提供了部分书稿内容。另外，本书的出版得到国家自然科学基金（项目编号：61203304，61375067）、江苏省自然科学基金（项目编号：BK2012566）、中央高校基本科研业务费专项基金（项目编号：2012QNA41）的联合资助，在此一并表示感谢。

尽管变异测试已经取得一定的研究成果，但是变异测试的理论及应用均有大量问题尚待进一步深入研究。期望本书能够为读者的进一步研究提供一定的启迪，对变异测试的发展发挥一定的作用。

限于作者水平，书中难免存在不妥之处，敬请读者批评指正。

作　者

2014 年 11 月于中国矿业大学

主要变量及表示方法

被测程序：G
语句：s
语句集：S
测试数据：$\boldsymbol{X}$
测试数据集：$\mathfrak{I}$
变异体：M
变异体构成的集合：$\wp$
变异算子：m
变异条件语句：ℓ
随机变量：Z
种群：Pop
种群规模：Pop_size
控制流图：D

目　录

第 1 章　进化变异测试入门

任何软件都不可避免地存在这样或者那样的错误。对软件进行测试可以有效降低软件错误的数量，从而提高软件的质量。因此，软件测试这一工作越来越受到人们的重视。

同时，软件测试也是一项非常烦琐的工作，需要投入大量的人力、物力和时间。研究结果表明，软件研发机构大约有 50%的工作量花在测试上。对一些要求高可靠性、高安全性的软件，用在测试方面的工作量更高。如果能实现测试工作自动化，将大大缩短测试所需的时间，从而缩短软件的开发周期，提高软件的质量和市场竞争力。而进行软件自动测试的核心，是采用有针对性的理论和方法，生成有效的测试数据，以满足既定的测试充分性准则[1]。

变异测试是一种面向缺陷检测的软件自动测试技术，具有排错能力强、方便灵活及自动化程度高等优点，既可以用来生成测试数据，又可以用来衡量测试数据集的检错能力。但是，变异测试需要消耗大量计算资源，从而很难在实际测试中得以应用。因此，如何提高变异测试的效率是值得深入研究的课题[2]。

对复杂软件的测试问题，采用诸如遗传算法等智能优化方法进行求解，以期取得更高的求解效率，是近年来软件工程界一种全新的研究方法，并且取得了很多可喜的研究成果[3]。

鉴于此，本书针对变异测试问题，研究基于进化优化的复杂软件变异测试理论及应用。该问题属于计算机、自动化与应用数学等多个学科有机交叉的研究方向，不仅有广泛而重要的应用背景，还具有鲜明的新颖性与挑战性。相应研究成果不但能够缩减软件测试成本，提高软件质量，而且能够扩大遗传算法的应用范围，因此，具有重要的理论意义和实用价值。

本章内容安排如下：1.1 节～1.3 节分别是软件测试简介、变异测试简介和软件进化测试简介；1.4 节给出研究现状分析，并指出当前研究存在的问题；在 1.5 节给出本书的主要内容及结构安排之后，1.6 节总结本章内容。

1.1　软件测试简介

软件测试是在软件投入市场使用之前，对软件的需求分析、规格说明及源代码等进行的最终复审，是保障软件质量的重要手段和必要依据。

软件测试主要通过技术、流程、工具、人员及管理手段，检测软件文档、软件中

间产品和最终产品，查找和报告软件缺陷、错误及隐患的方法，通过跟踪缺陷、错误及隐患的修正过程，确保软件产品、中间产品和文档符合软件工程过程需求和用户的最终需求。

当前，随着软件应用领域的不断深入，软件的复杂程度日益增大，开发周期不断缩短，对软件质量的要求也逐步提高。因此，如何提高软件测试水平和效率，从而有效保证软件质量，是值得业内和学术界深入研究的课题，也是软件企业谋求发展的必经之道。

1.1.1 软件测试基本方法

按照软件开发的不同阶段，软件测试可以分为单元测试、集成测试、确认测试及系统测试等。其中，单元测试主要针对程序的最小组成单元(模块)进行正确性检验。通过测试，保证各程序模块能够实现正确的功能；集成测试把已经通过单元测试的模块进行组装后，再对组装后模块的体系结构进行测试；确认测试主要检查成品软件是否能够完成需求规格说明中规定的各种功能和软件配置是否正确等；而系统测试则把已经完成确认测试的软件投入实际运行环境，与其他系统组合在一起进行测试。

按照是否执行被测程序，软件测试分为静态测试和动态测试两种。其中，静态测试不执行被测软件，仅通过分析或检查程序的语法、结构、过程及接口等判定程序有无错误；而动态测试需要以测试数据为输入运行被测软件，通过检查运行结果与预期结果的差异判定程序是否错误。

按照测试过程中是否需要程序代码，软件测试又可以分为黑盒、白盒及灰盒测试三种。其中，黑盒测试不需要程序代码，只需要知道程序应该完成的功能或约束。黑盒测试完全忽略被测程序的内部结构，只检查程序是否能够正确地接收输入数据，并产生正确的输出；白盒测试需要详细的程序代码，检查软件内部的状态是否正常。通常，测试人员依据程序的内部逻辑结构和相关信息设计所需的测试数据，并以该测试数据为输入运行程序，通过检查程序的内部状态，确定该程序是否发生异常；而灰盒测试则是黑盒和白盒测试的融合，它既关注程序的功能，也关注程序的内部状态，同时具备黑盒测试和白盒测试的优点。

本书主要考虑应用范围广泛的白盒测试。软件测试的重要性是毋庸置疑的，而软件测试的核心是测试数据。如何根据一定的测试充分性准则生成相应的测试数据是软件测试的关键问题。因此，下面简要介绍白盒测试中常用的测试数据生成方法。

1.1.2 测试数据生成

生成测试数据的过程是非常烦琐的，往往需要花费大量的时间。如果能实现

测试数据的自动生成，将有效减轻测试人员的劳动强度，提高软件测试的效率和软件质量，从而节省软件的开发成本。总体来讲，有四种测试数据生成方法，分别为随机法、静态法、动态法和试探法。

随机法通过对被测程序的输入空间随机采样，生成覆盖测试目标的数据。该方法简单易行，不受软件类型的限制，且能够快速生成大量的测试数据。但是该方法在生成测试数据时具有很大的盲目性，特别是对于一些复杂的被测程序，很难满足所有的测试要求。因此，人们提出各种自适应随机测试方法，目的是提高所生成测试数据的空间分布特性[4,5]。

静态法仅对程序进行静态的分析和转化，不需要实际执行被测程序，如 Botella等[6]提出的符号执行法、王志言和刘椿年[7]提出的区间算术法等。在生成测试数据策略方面，张德平等[8]提出一种资源受限的测试数据分配算法。静态法节省了执行程序的时间，但是该类方法通常需要执行大量的代数和(或)区间运算，特别是当输入变量个数很多时，计算量往往是非常大的。另外，该类方法无法处理输入空间为无穷区间的软件测试。

动态法基于程序的实际运行生成测试数据，且生成测试数据的过程是确定的，如 Miller 和 Spooner[9]提出的直线式程序法、Korel[10]提出的改变变量法、王雪莲等[11]提出的基于前向分析的动态程序切片方法、Callagher 和 Narasimhan[12]提出的罚函数法，以及 Gupta 等[13]提出的迭代松弛法等。该类方法节省了人工分析所需的工作量，但生成测试数据的时间往往很长，且最终结果和初始测试数据的好坏关系很大，因此，得到的测试数据往往不是最佳的。

与动态法不同的是，试探法虽然也基于被测程序的实际运行，但是生成测试数据的过程带有很大的随机性，这类方法包括禁忌搜索、微粒群优化、人工免疫、模拟退火和遗传算法等。本书在生成测试数据时，主要采用试探法对建立的模型进行求解。

1.2　变异测试简介

变异测试是一种面向缺陷检测的软件测试技术，最早由 DeMillo 等[14]和 Hamlet[15]提出。变异测试具有排错能力强、方便灵活及自动化程度高等优点，既可以揭示软件的缺陷，又可以衡量测试数据集的检错能力。因此，变异测试自诞生以来，得到了快速发展，应用范围也非常广泛。

1.2.1　变异测试基本原理

变异测试的基本原理是：首先，采用变异算子对被测程序做微小的合乎语法的变动，例如，将关系运算符“>”替换为“<”，从而产生大量的新程序，每个新程