

组合测试

在针对不同软件开展测试设计的过程中,可以发现,大量软件的输入、配置等呈现组合性特征。如果对所有可选组合都进行测试,极易造成组合爆炸问题,可能产生过多的测试用例,影响测试效率,甚至导致测试方案不可行。测试者有必要掌握基本的组合测试方法,以应对软件输入数据、环境等呈现组合性特征的情况。

一、实验目标

掌握应用组合测试技术进行测试用例设计的方法。了解组合测试技术所适用的问题,能够对给定测试问题进行抽象,将待测试内容抽象为多种因素的组合,能够分析组合因素之间的约束等,并利用工具完成测试用例的构造。目标知识与能力如表 3-1 所示。

表 3-1 目标知识与能力

知 识	能 力
(1) 组合关系、组合约束、组合生成算法 (2) 组合测试方法	(1) 问题分析: 能够将软件测试问题表达成组合数学问题 (2) 设计/开发解决方案: 软件需求规格分析; 测试用例设计 (3) 使用现代工具: 能够使用组合测试工具来辅助开展测试设计

二、实验内容与要求

许多在线购物系统支持组合检索策略,允许用户选定某个组合配置,然后在该配置下浏览、检索相应的商品,例如,如图 3-1 所示的计算机选购组合检索。然而,系统的组合检索并不总是能够正确地实现,时常有不符合预期的情况发生。

本实验要求用组合测试方法开展在线购物系统组合选购功能的测试。具体实验内容如下。

(1) 分析所选在线购物系统的组合选购功能有哪些输入控制因素。

(2) 以某类典型商品的选购为例,分析可选的组合因素、每个组合因素的可选取值水平,以及组合之间存在的限制约束(对于可选取值过多的组合因素,设计策略选择其中代表性的取值来参与测试)。

品牌:	戴尔 无极 攀升 雷豹 京天 华硕 宁美 HUAWEI Authorized Reseller 授权经销商 清华同方 硕扬												
电脑整机:	台式机	笔记本	一体机	游戏本	平板电脑	服务器/工作站	笔记本配件	平板电脑配件					
电脑组件:	组装机	主板CPU套装	CPU	显示器	显卡	机箱	内存	主板					
二手电脑整机:	二手台式机	二手笔记本	二手一体机	二手平板电脑	二手笔记本配件	二手平板电脑配件	二手游戏本						
电脑形态:	主机+显示器	单主机	一体机	迷你机	工作站	其他							
高级选项:	用途 v	处理器 ^	内存容量 v	特性 v	显卡 v	显示器尺寸 v	屏幕尺寸 v	分辨率 v	优选服务 v	系列 v	硬盘容量 v	系统 v	机箱大小 v
	Intel i3	Intel i5	Intel i7	AMD Ryzen7	Intel i9	AMD Ryzen9	Intel奔腾 赛扬	AMD	Ryzen锐龙	Intel至强	Intel 其他	其他	
	Intel i5低功耗版	Intel i7标准电压版	Intel i7低功耗版	AMD系列	Intel CoreM	Intel赛扬	Intel赛扬	Intel奔腾					

图 3-1 计算机选购组合检索

(3) 基于微软的 PICT 工具生成因素取值水平组合,作为对组合选购功能开展测试的测试输入。

(4) 将 PICT 生成的输入水平组合具体化为可实际执行的测试用例,执行测试,思考正确测试结果的特征,该如何校验测试结果,并在部分测试用例上对该思路加以验证(不要求编程验证,人工实验验证亦可)。

(5) 讨论如果没有 PICT 工具,可以采用哪些算法思路来自动构造组合测试用例。

三、实验环境

- (1) Word 与 Excel 办公软件。
- (2) 可访问的在线购物系统。
- (3) PICT 工具: <https://github.com/Microsoft/pict>。

四、评价要素

评价要素如表 3-2 所示。

表 3-2 评价要素

要素	实验要求
组合因素分析	应合理抽取、抽象组合因素;应识别其中的限制约束
组合构造	有效基于 PICT 工具构造组合选项
测试执行与结果分析	实施测试,并给出检验测试结果正确性的思路
算法设计	给出基本有效的组合输入构造算法思路

问题 分析

1. 组合测试技术

若一个软件的输入数据或配置可以被抽象为 N 个因素的组合,每个因素的可取值隶属一个有限的枚举空间,则可以用组合测试方法^[1,2,3]加以测试。组合测试构造出一组关于 N 个输入因素的组合实例,能够对输入空间的两因素(Pairwise)或更多因素(N -way, $N > 2$)

组合形成覆盖。

在软件测试实践中,许多经验表明,大多数缺陷由少量几个因素的共同作用而引发,例如,未妥善处理某两个输入因素的组合而造成。因此,组合测试实践中一般寻求对两到三个因素的组合形成覆盖。两因素组合测试是目前常用的组合测试方法,这种测试方法在生成的组合规模和检错能力方面取得了相对较好的平衡。

两因素组合测试可以基于正交表实现,正交表具有清晰的数学结构,可以实现两因素组合取值等概率出现的均匀覆盖,对于软件测试教学中理解组合测试较有帮助。但其一方面是不易获取,另一方面是一味追求等概率出现可能要求一个组合取值被测试多次,带来一定冗余,增加了测试代价。

另一种选项是用微软提供的 PICT 组合测试工具来生成取值组合。PICT 不仅支持描述各个因素和候选取值水平,还能够表达多因素取值之间的约束,具有较强的组合生成能力。本实验基于 PICT 来实施组合测试。

2. 实验问题的解决思路与注意事项

1) 识别输入因素

Web 系统中的输入主要是本文框、复选框等。

2) 组合因素及其取值水平、取值约束获取

并非每一输入都适合直接作为组合因素,真实应用中参与组合的输入项可能很多,每个输入项也可能有不少候选取值。因此,重点是如何对输入项及其取值进行抽象。一种抽象思路是从系统实现和风险角度进行抽象。例如,某些输入项的应用内部处理机制可能相同,测试一个,一般也就排除了同类输入的风险,对于这种情况,多个输入可能测试一个即足够。同样,一个文本框的输入可能有非常多种,甚至近乎无限多种,可以将其抽象为填、不填,或者单个单词、多个单词等多种不同填充状态。在此基础上,可以识别组合因素和水平之间的约束,典型的约束包括取值互斥、取值恒定相同等。

3) 基于 PICT 的测试构造

PICT 工具提供了丰富的功能,包括描述组合间的约束,可以查阅官方文档来了解其使用。

4) 测试执行与结果确认

关于如何确定测试执行的结果是否符合预期,基本的方法就是由人工分析组合语义,对照软件功能,设定预期结果。预期结果一般不一定是一个固定结果,有可能软件输出符合特定模式即表明其在组合输入下能够正常工作。

另外,一些软件对于同样一个检索提供不同的访问途径,即有多种软件用法都可以实现相同目的的检索任务。在这种情况下,印证同一功能在不同软件使用方式下的结果是否一致,可以帮助确定组合输入是否得到了正确处理。

3. 难点与挑战

(1) 组合因素和水平抽取:该过程需要结合软件实现和相关质量风险对原始输入项及其取值进行抽象。如何合理抽象是一个难点,不当的抽象一方面很可能导致组合爆炸,产生过量测试用例;另一方面,缺乏依据的抽象难以使人对测试结果产生信心。

(2) 测试用例构造与测试结果确认:测试用例构造是一个从组合因素和水平到具体测试用例的过程,但其中因素多、从抽象因素和水平到真实输入项和数据的映射关系复杂,如

何理顺关系、高效构造用例是一个难点。真实系统如何确认软件在组合输入下表现正常是另一个难点,是否能找到途径实现某种程度的测试结果自动判定具有相当大的挑战。

(3) 组合生成算法: 如何从组合因素、水平、约束构造输入组合,这一问题涉及复杂数据结构 and 算法,需要具有扎实的基础,且广泛查阅资料才能解决。

实验方案

本实验以京东商城 Web 版的组合选购功能为测试对象,研究应用组合测试的思想来设计测试用例的方法,尝试对组合因素展开高充分度的测试。

1. 输入控制因素分析

通过关键词搜索或者类别浏览功能,进入一个如图 3-2 所示的商品选购页面。在输入关键词为“测试 Web”时,可以发现有以下输入项可以帮助选择商品。

- 搜索框关键词: 可以输入单个或多个关键词。
- “品牌”: 多种品牌,可以多选。
- “计算机与互联网”: 图书内容类别,单选。
- “其他图书”: 多个选项,单选。
- “出版社”: 十个以上选项,可以多选。
- “类别”: 多选类别选项。
- “高级选项”: 包括“包装”“是否套装”等,单选。
- 工具栏选项: “京东物流”“货到付款”“仅显示有货”“新品”“PLUS 专享”等。

在 Web 页面上单击某组合选项(如“中小学教辅”)后可能呈现其他动态输入项,这些输入项在图 3-2 的界面中隐藏,可以在组合选择过程中出现。具体输入项难以全部枚举,但这些输入项基本特征类似。



图 3-2 商品选购页面

以上用于选择商品的输入项基本为互不相干的枚举类型,各个选择因素之间及各个可

选条目之间推测不存在直接联系,功能相对独立。即使选择一个输入项后,导致另一个选项不可选,或者从隐藏变成出现,这种联系也是与商品检索结果相关,而不是直接由输入项间的逻辑关联造成的。另外,选项部分单选、部分多选,多个选项共同决定最终的检索结果,因此存在在组合选择下系统出现异常的风险。输入为枚举类型、独立性强、共同发生作用,该实验对象较适合采用组合测试的方法来开展测试。

当然,组合测试仅用于检验组合搜索相关的缺陷,如中文网站无法搜索英文商品之类的缺陷风险暂不在本次测试的考虑范围内。

2. 组合因素和水平分析

本实验对象中,参与组合的因素包括搜索框关键词、“品牌”“计算机与互联网”“其他图书”“出版社”“类别”“高级选项”“京东物流”“货到付款”“仅显示有货”“新品”“PLUS 专享”等,总计组合因素超过 20 个。诸如品牌等组合因素,其可选水平值达 10 个以上。如果直接对所有输入因素及其原始取值进行组合,则存在组合爆炸的问题,产生的测试用例过多,测试低效而难以开展。

如何精简组合因素和水平?可以尝试对软件的缺陷风险加以分析,识别可能导致缺陷的主要因素和水平,对此加以组合,提高测试效率。对所有输入因素而言,可以将其分为以下几个类别,其取值水平也可加以抽象。

(1) 搜索框关键词:关键词检索和其他勾选等方式的检索输入对应的处理机制不同,缺陷风险有差别,有必要作为独立组合因素来测试。其取值可以抽象为单个关键词、多个关键词两种,而多个关键词中两个、三个关键词,其差别明显小于单个和多个关键词,只需测试一种往往就能对检索机制的执行路径加以有效覆盖。

(2) 单选检索输入:“计算机与互联网”“其他图书”“系列”等组合选项为单选输入项,这些输入项对各类商品不同,是根据内容动态生成的输入因素。对这些因素,测试其中一个往往即能触发背后的程序执行路径。因此,对动态单选输入项,可以先取其中一个加以测试,如果测试资源充足,再考虑进行补充。诸如“计算机与互联网”的输入下设有多个选项,对这些选项,软件的处理逻辑是类似的,没有必要一一加以区分。可以考虑测试“选择”和“不选”两种取值水平,这两种水平检测了明显不同的两个内部程序处理分支。

(3) 多选检索输入:对于如“品牌”“出版社”等动态多选输入因素,测试一个即具有代表性,可统一将这些输入因素抽象为多选检索输入。多选检索输入有“不选”“单选”“多选”三种典型输入。没有必要将每个实际输入情况都当作组合水平来加以测试,因为许多取值水平所能检测的软件内部实现是基本相同的。

(4) 高级选项:高级选项在界面上单独列出,UI 风格不同,有理由怀疑其实现与其他输入选项存在差异,因此需要单独作为输入因素加以测试。同时,“包装”“是否套装”等各高级选项推测也是动态生成的,测试一个即具有代表性。选项的取值水平也可抽象为“选择”和“不选”两种触发执行代码不同的情况。

(5) 工具栏选项:“京东物流”“货到付款”“仅显示有货”“新品”“PLUS 专享”等,这些选项与商品特征无关,属于偏固定的选项,触发的内在执行代码可能有所不同,因此有必要作为独立输入因素加以测试,其中每个因素有“勾选”和“不勾选”两种取值水平。

在以上输入项组合因素中,通过实际尝试,发现“京东物流”和“货到付款”与“新品”“PLUS 专享”存在冲突,前者被勾选时,后者自动消失,无法勾选。“新品”和“PLUS 专享”

也存在冲突,勾选一个,则另一个不允许选择。假如以上情况是软件正常功能设置,则存在输入因素之间的互斥约束。

对于单选输入,在 Web 页面上进行选择后可能引发选项动态变化,隐藏的组合选项被列出,或目前呈现的组合选项被隐藏。例如,单击输入因素“出版社”中的“清华大学出版社”这一取值水平后,会出现“供应链管理类型”选项,而选择另一些输入因素水平,则不会出现。即“出版社”和“供应链管理类型”这两个组合因素存在关联。因素及其取值水平间存在由于各自覆盖图书范围差异而引发的出现约束,一些选项组合下无书可选,因此选项互斥。然而,上述约束动态性强,不易列举;在“单选检索输入”“多选检索输入”“高级选项”这样的因素及其水平抽象下,对于输入生成影响不大,并不会总是导致因素和水平的组合无效;且任意一次组合输入下的测试过程都能够测试到因检索结果图书范围引起的输入项动态变化,因此,也没有必要将相关约束逐一列出。

总体来看,精简后的测试输入组合因素、水平及其约束如表 3-3 所示。

表 3-3 组合因素、水平及其约束

组合因素	取值水平	约束
搜索框关键词	单关键词、多关键词	
单选检索输入	选择、不选	
多选检索输入	不选、单选、多选	
高级选项	选择、不选	
京东物流	勾选、不勾选	勾选后“新品”“PLUS 专享”不可选
货到付款	勾选、不勾选	勾选后“新品”“PLUS 专享”不可选
仅显示有货	勾选、不勾选	
新品	勾选、不勾选	勾选后“PLUS 专享”不可选
PLUS 专享	勾选、不勾选	勾选后“新品”不可选

3. 测试生成

借助 PICT 工具,可以获得商品组合检索功能的测试输入组合,构造组合测试下的测试用例集。

首先,下载 PICT 工具。从 PICT 的 Github 主页可以下载其可执行程序 pict.exe,该程序无须安装即可运行。工具文档也可在 Github 查询。

在测试项目的目录(任一目录)下,新建一个“input_space.txt”文件,将组合因素及其取值水平填入文件中,内容如图 3-3 所示。注意标点符号要使用英文标点,文件路径尽量为英文,.txt 文件编码尽量为 ANSI 格式,以免出现编码错误。

在命令行下执行命令“pict.exe input_space.txt”,可以生成测试输入组合,并在屏幕上显示。执行“pict.exe input_space.txt > tests.xls”可以将生成结

搜索框关键词: 单关键词, 多关键词
 单选检索输入: 选择, 不选
 多选检索输入: 不选, 单选, 多选
 高级选项: 选择, 不选
 京东物流: 勾选, 不勾选
 货到付款: 勾选, 不勾选
 仅显示有货: 勾选, 不勾选
 新品: 勾选, 不勾选
 PLUS 专享: 勾选, 不勾选

图 3-3 未考虑因素的组合空间描述文件

果重定向到文件,并用 Excel 打开查看。生成的 9 个输入组合如图 3-4 所示。

搜索框关键词	单选检索输入	多选检索输入	高级选项	京东物流	货到付款	仅显示有货	新品	PLUS专享
单关键词	选择	多选	选择	勾选	勾选	勾选	勾选	不勾选
多关键词	不选	单选	不选	不勾选	不勾选	不勾选	勾选	勾选
单关键词	不选	多选	不选	勾选	不勾选	勾选	不勾选	勾选
多关键词	选择	多选	选择	不勾选	勾选	不勾选	不勾选	不勾选
多关键词	选择	不选	不选	勾选	不勾选	勾选	不勾选	不勾选
单关键词	不选	单选	选择	不勾选	勾选	勾选	不勾选	不勾选
单关键词	不选	不选	选择	勾选	勾选	不勾选	勾选	勾选
多关键词	选择	不选	不选	不勾选	勾选	勾选	勾选	勾选
多关键词	选择	单选	选择	勾选	不勾选	勾选	勾选	不勾选

图 3-4 不考虑约束生成的输入组合

然后,在输入空间描述文件“input_space.txt”中加入如图 3-5 所示约束描述语句,表达“京东物流”“货到付款”“新品”“PLUS 专享”之间的限制。为简化约束,这里假定“新品”“PLUS 专享”因冲突而不可选情况即对应其取值水平“不勾选”。若要更准确地建模,可以引入一个取值水平“不可用”,并增加关于该取值的约束,表示该选项对用户不可见,因而而不可选。

```
IF [京东物流] = "勾选" THEN [新品] = "不勾选" AND [PLUS 专享] = "不勾选";
IF [货到付款] = "勾选" THEN [新品] = "不勾选" AND [PLUS 专享] = "不勾选";
IF [新品] = "勾选" THEN [PLUS 专享] = "不勾选";
IF [PLUS 专享] = "勾选" THEN [新品] = "不勾选";
```

图 3-5 组合因素间限制的约束描述

考虑约束后,利用 PICT 工具生成的测试用例如图 3-6 所示,共需测试 11 种组合。该组合数量意外地较之不考虑约束还要多,这是因为考虑约束后,组合限制多,算法求解策略发生变化,优化空间也未必如不考虑约束大。

搜索框关键词	单选检索输入	多选检索输入	高级选项	京东物流	货到付款	仅显示有货	新品	PLUS专享
单关键词	选择	多选	不选	不勾选	勾选	不勾选	不勾选	不勾选
多关键词	不选	不选	选择	不勾选	不勾选	勾选	不勾选	勾选
单关键词	不选	多选	不选	勾选	不勾选	勾选	不勾选	不勾选
单关键词	选择	单选	选择	不勾选	不勾选	不勾选	不勾选	勾选
多关键词	选择	不选	不选	不勾选	不勾选	不勾选	勾选	不勾选
单关键词	不选	不选	选择	不勾选	不勾选	勾选	勾选	不勾选
多关键词	选择	不选	选择	勾选	勾选	勾选	不勾选	不勾选
多关键词	不选	单选	不选	不勾选	不勾选	勾选	勾选	不勾选
多关键词	不选	多选	不选	不勾选	不勾选	不勾选	不勾选	勾选
多关键词	选择	多选	选择	不勾选	不勾选	不勾选	勾选	不勾选
多关键词	不选	单选	不选	勾选	勾选	不勾选	不勾选	不勾选

图 3-6 考虑约束生成的输入组合

4. 测试用例构造和执行

步骤 3 生成的测试输入仍是较为抽象的测试输入,在执行测试用例前,还需要进一步将测试输入组合具体化,并设定待测系统预期输出或行为表现,得到具体待执行的测试用例。

图 3-7 列出了本实验采用的一个测试输入具体化方案,在该方法中将单关键词检索具体化为搜索“测试”,多关键词检索具体化为检索“测试 Web”;将单选检索输入具体化为“计算机与互联网”和“其他图书”等单选检索项的输入,其输入水平选择也具体化为选择“操作

系统”等,其他以此类推。并且,输入构造也考虑了选项单击后引起的组合选择因素变化。如此,测试输入中所有待填入、设定、勾选的内容全部具体化,测试用例可以确定地进行复用,如果发现缺陷,也便于软件调试者复现问题。

京东商城为内容密集型网站,其页面呈现元素和系统上线商品高度相关。图 3-7 中,“五金工具”“测量类”是搜索框中仅检索“测试”单一关键词时出现的组合选项,在检索“测试 Web”时不存在。测试过程中也可以发现很多预期可行的输入组合在实际测试中由于商品内容变化,变得不再可行。为此,图 3-7 中划去了部分不可选择的输入选项。“新品”和“PLUS 专享”选项大部分情况下不可选。实验根据仅有的几组可选择场景,重新设计了勾选选择(加粗未划掉部分为新的输入设定,如“不勾选”划去“不”变为“勾选”),以确保选项能够被勾选,对这两项输入因素的影响加以检测。

	搜索框关键词	单选检索输入	多选检索输入	高级选项	京东物流	货到付款	仅显示有货	新品	PLUS 专享	预期结果
1	测试	“五金工具”选择 “测量工具”	“品牌”多选	不选	不勾选	勾选	不勾选			商品列表 与检索设 定相符
2	测试 Web	不选	不选	“包装”选择	不勾选	不勾选	勾选	不勾选	勾选	
3	测试	不选	“出版社”多选	不选	勾选	不勾选	勾选			
4	测试	“五金工具”选择 “手动工具”	“品牌”单选	“测量类”选择	不勾选	不勾选	不勾选		勾选	
5	测试 Web	“计算机与互联网 “选择”操作系统”	不选	不选	不勾选	不勾选	不勾选	勾选		
6	测试	不选	不选	“国产/进口” 选择	不勾选	不勾选	勾选	勾选	不勾选	
7	测试 Web	“其他图书”选择 “学者”	不选	“形式”选择	勾选	勾选	勾选			
8	测试 Web	不选	“出版社”单选	不选	不勾选	不勾选	勾选	勾选		
9	测试 Web	不选	“类别”多选	不选	不勾选	不勾选	不勾选		勾选	
10	测试 Web	“其他图书”选择 “标准与规范”	“品牌”多选	“客户评分”选 择	不勾选	不勾选	不勾选	勾选		
11	测试 Web	不选	“类别”单选	不选	勾选	勾选	不勾选			

图 3-7 实际待执行的测试用例

关于测试用例的运行结果,可以通过检索所得商品的标题、商品介绍等信息加以判断。正确的检索结果一方面所得每一商品应与检索设定相符,另一方面不应遗漏匹配的商品。

检索所得是否与设定相符可以通过抽查商品实现,即抽查若干检索所得商品的标题、商品介绍,确认其是否与检索设定相符。检索是否遗漏可以通过逆向分析判断。首先找一个已知商品,如一本测试书籍,已知其出版社、类别等,然后按已知信息检索该书,检查该书是否在所得结果中,若不在,则说明结果可能有问题。

除此以外,还可通过多检索对照分析来评判测试结果。在商品组合搜索中,可以通过多种方式搜索同样的目标内容。例如,在检索输入项“品牌”和“出版社”以及页面顶部的检索输入框中均可以搜索“清华大学出版社”,若结果存在较大差异,则可能预示错误。另外,一般而言,增加更多的检索设定,其结果应该减少,如果更强的条件设定下得到的反而是更多的检索结果,则说明商品检索也可能存在问题。对多选检索选项,例如“出版社”,选择多个出版社的检索结果应该是单选情况下的合集。

本实验按上述思路执行测试,发现京东商城绝大部分情况下的检索结果符合预期。个别情况下检索出现意外,如表 3-4 所示。这些意外未必是系统缺陷,也可能是特殊的功能设置,或者是卖家输入商品信息有误造成的结果。

表 3-4 测试结果中的意外情况

序号	意外问题	描述
1	检索结果不一致	在检索输入框中输入“测试 Web”，检索图书。执行 3 次测试，分别在检索输入项“品牌”“出版社”，以及页面顶部的检索输入框中各自增加“清华大学出版社”限制，商品检索结果存在数量上的较大差异，按“品牌”增加限制后检索所得结果明显过少（仅 1 页结果，其他为 10 页）
2	增加输入限制结果变多	测试用例 7，在某些检索选项下，勾选“京东物流”，此时检索结果为空，但进一步勾选“仅显示有货”，商品内容又突然变得非空，显示软件内在逻辑不清晰，可能处理有误
3	检索输入项的动态变化异常	测试用例 8 实施限制出版社的检索，选择一个检索结果较少的出版社后，“计算机与互联网”这一组合因素候选选项变化，但变化后的选项与限制出版社后所得的图书列表不符，出现了图书列表上不该有的类别“操作系统”。该组合因素是否是动态更新、更新的逻辑是什么，似乎并不很清晰，与预期要么不变、要么根据检索结果变化的设想不相符，也不太符合习惯逻辑

5. 算法构造测试用例

组合测试的生成一方面希望所得测试用例对各种因素及其取值水平的组合形成覆盖；另一方面，希望形成覆盖所需的测试用例数量尽量少，测试代价低；此外，在同等规模、同样覆盖的情况下，也希望各因素和水平的覆盖能够均衡，加强测试抽样的散布性，增加命中缺陷的潜在可能。

最优测试的生成需要较为高深的数学理论。一种相对简单的替代方案是采用演化计算的思路来进行组合生成。大致步骤如下。

(1) 生成初代组合测试集种群。采用随机生成的方法，从 n 个因素的组合空间中选择若干组合配置，构造出一个能够覆盖必要因素水平组合的测试输入集 T 。重复该过程 m 次，构造 m 个测试输入集构成的种群 $G_1 = \{T_1, T_2, \dots, T_m\}$ 。该种群能够保证覆盖，但在测试集规模、均衡性方面未必最优，后续任务即优化测试集，使之更为接近目标结果。

(2) 交叉变异构造子代种群。定义综合评估测试覆盖、组合规模和因素水平均衡度的指标，从父代种群中选择指标高的测试用例集进行交叉，以及变异替换测试用例集中部分输入，得到子代测试输入集种群。这部分测试输入集一部分表现更佳，有一部分引入了新的变异，带来更多的测试用例发展方向。

(3) 从上一代种群和子代种群中选择指标更好的测试输入集，构造一个新的测试种群。重复步骤(2)~步骤(3)若干次，直到有测试输入集的综合指标达到较高的水平，可以作为组合测试的测试用例。

该方法不能保证生成的测试输入集绝对最优，但其所得结果同样可以优化组合空间，降低测试代价。

小 结

本实验展示了在一个真实案例上应用组合测试方法的过程，处理了其中的因素和取值水平抽象等问题，考虑了一些组合约束，初步尝试应对组合因素及其取值水平动态变化等问

题,为组合测试应用提供了参考。但实验中仍有许多考虑不全面、测试不深入之处,在组合测试之外,测试实践者往往还需要根据待测软件的需求规格和缺陷风险,增加测试用例,弥补组合测试技术不擅长之处。



附件资源

- (1) 组合描述文件。
- (2) 测试用例表。



参考文献

- [1] 史亮. 软件测试实战: 微软技术专家经验总结[M]. 北京: 人民邮电出版社, 2014.
- [2] 聂长海. 组合测试[M]. 北京: 科学出版社, 2015.
- [3] 严俊, 张健. 组合测试: 原理与方法[J]. 软件学报, 2009, 20(6): 1393-1405.