



中华人民共和国国家标准

GB/T 18905.6—2002/ISO/IEC 14598-6:2001

软件工程 产品评价 第6部分:评价模块的文档编制

Software engineering—Product evaluation—
Part 6: Documentation of evaluation modules

(ISO/IEC 14598-6:2001, IDT)

2002-12-04 发布

2003-05-01 实施



中华人民共和国
国家质量监督检验检疫总局 发布

中 华 人 民 共 和 国

国 家 标 准

软件工程 产品评价

第 6 部分:评价模块的文档编制

GB/T 18905.6—2002/ISO/IEC 14598-6:2001

*

中国标准出版社出版

北京复兴门外三里河北街16号

邮政编码:100045

电话:68523946 68517548

中国标准出版社秦皇岛印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

*

开本 880×1230 1/16 印张 1 1/4 字数 46 千字

2003年4月第一版 2003年4月第一次印刷

印数 1—1 500

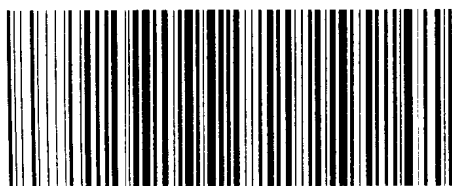
*

书号:155066·1 19326 定价 15.00 元

网址 www.bzchs.com

版权专有 侵权必究

举报电话:(010)68533533



GB/T 18905.6-2002

前 言

GB/T 18905—2002《软件工程 产品评价》分为六个部分：

- 第1部分：概述；
- 第2部分：策划和管理；
- 第3部分：开发者用的过程；
- 第4部分：需方用的过程；
- 第5部分：评价者用的过程；
- 第6部分：评价模块的文档编制。

本部分为 GB/T 18905—2002 的第 6 部分，等同采用 ISO/IEC 14598-6:2001《软件工程 产品评价 第 6 部分：评价模块的文档》(英文版)。

本部分附录 C 中 C.5.1 的表中第 3 级(良)有印刷勘误，故将原文的]0.80 .. 0.00]改为]0.80 .. 0.90]。

本部分的附录 A、附录 B、附录 C、附录 D 是资料性附录。

本部分由中华人民共和国信息产业部提出。

本部分由中国电子技术标准化研究所归口。

本部分由中国电子技术标准化研究所负责起草。

本部分主要起草人：罗锋盈、陈莹、王凌、冯惠。

引 言

软件产品评价依赖于—组评价技术和提供软件质量特性信息的度量。许多度量及相关的方法使用测量结果来实施特定的软件产品评价。ISO/IEC 9126-2 和 ISO/IEC 9126-3 提供了与子特性对应的度量实例。这些度量难以在一个组织中一致地使用。但它可能对特定应用开发新度量是必需的。因此,组织中的支持机构(见 GB/T 18905.2)为了在组织内正确和一致地使用度量,有必要对各个度量都分别加以规定。对度量、相关方法及使用指南的文档编制格式宜标准化。评价模块的内容提供这种要求的解决方案。

评价模块规定了应用于评价质量特性并标识其要求的证据的评价方法。它也定义了技术应用中的基本评价规程和测量结果的报告格式。

编制评价模块文档的一致性方法有以下优点:

- 在描述评价模块的理论基础中,提供了基准通用。
- 为编制评价模块文档或开发评价模块,确定了最低需求。
- 在收集归并大量参与的评价模块时,提供了必需工具。

评价模块提供了一种灵活的和结构化的方法,使度量适用于评价中间产品或最终产品。使用本部分产生的评价模块,有助于确保软件产品评价可重复性、可再现性和客观性。

评价模块文档编制的格式需考虑下列因素:

- 适用于软件产品的评价环境。
- 该格式支持根据技术发展水平来开发新度量的要求。
- 该格式提供度量及应用的精确定义。
- 它提供了评价人员进行评价所需的信息。

附录 A 提供了新评价模块开发过程的指南。

附录 B、附录 C 和附录 D 是评价模块的实例。

目 次

前言	I
引言	II
1 范围	1
2 一致性	1
3 规范性引用文件	1
4 术语和定义	1
5 评价模块的概念	2
6 评价模块文档编制格式	2
附录 A (资料性附录) 评价模块的开发	5
附录 B (资料性附录) 评价模块的实例——故障密度	6
附录 C (资料性附录) 评价模块的实例——功能性	9
附录 D (资料性附录) 评价模块的实例——易用性和使用质量	18
参考文献	22

软件工程 产品评价

第6部分:评价模块的文档编制

1 范围

本部分定义了用于描述评价模块的文档编制的结构和内容。评价模块旨在 ISO/IEC 9126 和本系列标准的环境中使用。

本部分旨在供测试实验室、研究机构等评价技术领域的专家在制定新的评价模块时使用。

2 一致性

如果评价模块的文档满足第6章(评价模块的文档格式)的要求,则认为它符合本部分。

3 规范性引用文件

下列文件中的条款通过本部分的引用而成为本部分的条款。凡是注日期的引用文件,其随后所有的修改单(不包括勘误的内容)或修订版均不适用于本部分,然而,鼓励根据本部分达成协议的各方研究是否可使用这些文件的最新版本。凡是不注日期的引用文件,其最新版本适用于本部分。

GB/T 8566:2001 信息技术 软件生存周期过程(idt ISO/IEC 12207:1995)

GB/T 18905.1—2002 软件工程 产品评价 第1部分:概述(ISO/IEC 14598-1:1999, IDT)

GB/T 18905.2—2002 软件工程 产品评价 第2部分:策划和管理(ISO/IEC 14598-2:2000, IDT)

GB/T 18905.3—2002 软件工程 产品评价 第3部分:开发者用的过程(ISO/IEC 14598-3:2000, IDT)

GB/T 18905.4—2002 软件工程 产品评价 第4部分:需方用的过程(ISO/IEC 14598-4:1999, IDT)

GB/T 18905.5—2002 软件工程 产品评价 第5部分:评价者用的过程(ISO/IEC 14598-5:1998, IDT)

ISO/IEC 9126-1 软件工程 产品质量 第1部分:质量模型

4 术语和定义

本部分采用下列术语和定义。

4.1

评价模块 evaluation module

用于测量软件质量特性、子特性或属性的评价技术包。

注:此包包括:

- 评价方法和技术;
- 对评价的输入;
- 待测量和待收集的数据;
- 支持规程和工具。

4.2

评价技术(用于评价的技术) evaluation technology

用于评价的技术、工具、度量、测量及其他技术信息。

5 评价模块的概念

软件产品的评价可说是一项全面性的工作。各个质量特性和子特性可能需要采用不同的评价技术,收集不同的数据。为了应付这种复杂性,评价活动应被构造成多个可管理的单元。每个单元能覆盖质量的一个或多个方面。但每个单元应集中在采用特定的评价技术对一个特定的质量方面做出评价。进行这些评价所需的信息应收集起来并整理成包,供将来使用。这种包就称为评价模块。

采用标准格式评价模块的好处在于:

- 它支持评价模块的开发,因为它提供了一个内容表,使得评价必需何种信息及如何处理这些信息(原则、度量、工具等)成为可视;
- 它支持评价模块的使用,因为信息以同构的方式可用;
- 它支持评价模块的重用,因为它便于建立和维护评价模块库;
- 它支持评价模块的标准化,因为其格式符合标准的要求。

评价模块收集了应用特定的评价技术进行质量特性特定方面的评价所需的所有信息。其中阐明了要测量哪个软件质量特性的特定方面。测量的进行过程以及先决条件和准确性都做出了规定。

评价模块提供了评价技术、度量和测量之间的联系。当 GB/T 18905.3、GB/T 18905.4 或 GB/T 18905.5 推荐一种评价技术的应用时,可从评价模块库中选出合适的评价模块(见 GB/T 18905.2)。

评价模块的文档有 EM(评价模块)0 至 EM5 共 6 个部分,它们各有不同的用途,并有可选的附录 EMA。

EM0 提供评价模块的正式信息,并介绍评价模块中描述的评价技术。

EM1 定义评价模块的适用范围。

EM2 提供相关引用标准。

EM3 包括评价模块所需的定义。

EM4 规定评价需要的输入产品,并定义待收集的数据和待计算的测量。

EM5 包括如何解释测量结果的信息。

可选的附录 EMA 包含了应用评价模块的详细过程。尽管 EMA 是可选的,但仍建议将它包括进来。

注:评价模块文档编制格式符合 GB/T 1.1—2000 描述的标准的正式要求。这有利于评价模块的标准化。

6 评价模块文档编制格式

评价模块的文档编制应按照 6.1、6.2、6.3、6.4、6.5、6.6 和 6.7 予以格式化。

6.1 EM0:前言和引言

本章涉及组织的和项目规定的需求。

6.1.1 前言

本节应提供如下信息:

——准备、核准、组成和变更;

——与其他标准和其他文档的关系。

6.1.2 引言

本节应介绍评价模块的原则、背景和技术原理。

注:评价方法的正式描述见 6.2.3。

6.2 EM1:范围

6.2.1 特性

本条应确定评价模块可评价的特性、子特性或属性。

注:评价模块可提供给一个或多个特性或子特性。

描述特性、子特性或属性的质量模型应予确定。除非有特殊情况,宜使用 ISO/IEC 9126-1 中的模型。

6.2.2 评价级别

本条描述由评价模块定义的评价级别。评价级别与被评价的特性、子特性或属性的重要性有关。建议在描述级别时考虑软件的用法和软件产品的环境(例如:安全条件、保密约束、经济风险和应用约束)。

级别根据采用的评价技术和要获得的评价结果来定义评价的深度或全面性。不同的评价级别给出不同的软件产品质量的可信度。

注:级别可公式化为 GB/T 18905.5 中描述的 A、B、C、D 四级。软件完整性级别在 GB/T 18492 中描述。

6.2.3 技术

本条应描述评价模块中应用的评价技术。应包含或适当引用评价所依据的有关理论、模型或试探法。

注:评价技术的例子有可靠性生长模型、基准测试和对代码的静态分析。

6.2.4 适用性

本条应确定评价模块的适用范围。

注1:例如,评价模块适用于特殊的程序设计语言,或命令语言类。

应描述在软件生存周期中评价模块在何处可以用。若拟用在特定的软件生存周期过程中,则应加以规定。

注2:软件生存周期过程在 GB/T 8566 中规定。

6.3 EM2:引用

本条应提供对标准的和技术文档的引用。若评价模块依赖于其他评价模块的结果,则应在这里提及。

6.4 EM3:术语和定义

本条应定义在评价模块中使用的技术术语。或者是,指出可找到定义的引用信息。

6.5 EM4:输入和度量

6.5.1 对评价的输入

本条应确定评价所要求的输入。输入应分类为产品部件、产品信息、支持信息和使用中的产品信息。

注1:产品部件分类信息包括:软件需求规格说明,软件设计描述,程序描述,源代码,可执行代码和用户文档。

注2:产品信息分类信息包括:软件需求评审报告,软件设计评审报告,程序评审报告,单元测试报告和用户文档评审报告。

注3:支持信息分类信息包括:质量保证计划,配置管理计划,程序测试计划,编程语言和编译程序的描述。支持信息不作评价,只用作进行评价所需的背景信息。

注4:使用中的产品信息分类信息包括:描述系统行为的测试报告和操作报告。系统包括相关的硬件、软件 and 用户。

6.5.2 数据元素

本条应规定哪些数据元素应从输入中提取。

注1:数据元素的例子:包含注释的代码行数;用户手册中语句长度的频率分布;每条帮助消息的字数;每小时操作中观察到的失效数;规定的词汇扫描在每个模块中发现的特定类型的标记数。

注2:通常,数据元素是计算测量的来源资料。但在某些情况下,原始数据本身可组成度量。

6.5.3 度量和测量

本条应描述如何用度量根据数据元素计算测量。

若将度量合并成“高阶”度量,则其依赖性应显式说明。

测量前应申明所有的假设和前提均已满足。

6.6 EM5:结果解释

6.6.1 测量映射

本条应规定测量的意义,也就是说,测量结果的解释。这包含评价尺度,由定义的度量获得的值与它相映射。若此映射是有价值的,那么需要映射的算法(功能)的细节应予以定义,或对其来源加以引用。如果单个特性、子特性或属性得到几种测量,那么本条应描述如何将这几种测量合并成特性、子特性或属性的等级评定。

应规定测量的精确度。

6.6.2 报告

本条应描述报告内容,提供应用评价模块的结果。在某些情况下,所得值的可视化工作是重要的,应鼓励使用。

6.7 EMA:应用规程

注:使用 EMA 是可选的,如果使用,应列出下列内容。

6.7.1 技术术语的定义

本条应定义那些在 6.4 中未定义,但在评价模块的 EMA 部分或引用的来源中使用的技术术语。

6.7.2 所需要的资源

本条应规定应用评价模块时需要的资源。宜包括:需要的软件工具(宜标识任何需要的软件工具,宜引用通用工具类型和专有工具);需要的硬件软件;测试装备或其他设备;技能和资格(宜标识评价者或评价组织需要的任何专业技能和资格(如,证书));应用工作(宜估算评价模块的典型应用所需的工作,并且,如果这类工作依赖产品的属性(如,代码行数),宜给出估计算法);以及其他资源。

6.7.3 评价指令

本条应描述所遵循的规程的所有细节。宜包括证据(例如,代码的抽样)的选择,原始数据的产生和记录、计数规则、从原始数据计算度量的算法、结果的记录,对工作记录的保留和最后文档的要求。特别是,宜强调采取措施确保结果的可追踪性和可重复性。描述的过程宜符合 GB/T 15481 的要求。

6.7.4 文档编制

本条应概述使用评价模块而产生的内部文档。概要报告宜符合 GB/T 15481。

附 录 A
(资料性附录)
评价模块的开发

本附录提供了开发新评价模块的过程指南。ISO/IEC 9126-2、ISO/IEC 9126-3 和 ISO/IEC 9126-4 可作为该过程的输入。评价模块的开发过程宜包括五个步骤：

A.1 评价模块需求的确定

当确定了新评价模块的需求,并决定开发评价模块时,第一步宜确定新评价模块的具体需求。包括确定质量模型和质量特性或子特性。还应决定评价的严格度。

A.2 评价模块的规格说明

在评价模块的需求已确定的基础上,下一步是规定针对评价的技术和评价输入(例如,源代码),以及度量集和基础数据元素集。ISO/IEC 9126-2、ISO/IEC 9126-3 和 ISO/IEC 9126-4 可能有用。

A.3 评价规程的开发

本步骤把正式的规格说明与前一步结合起来并增加一些程序性的内容。对度量和数据元素的解释宜在评价环境中予以阐明。宜估计所需的资源,制定详细的评价规程。评价规程的试用可能是必要的。

A.4 评价规程的描述

本步骤中,宜根据评价模块格式来描述前一步骤制定的评价规程,也就是说,该描述应符合本部分。

A.5 评价模块的验证和确认

评价模块宜对照规格说明进行评审。宜由掌握评价技术的专家来做这项工作。确认宜确保评价模块代表了技术最新发展状态和现有的技术对本组织是不太熟悉的。

评价模块宜在不同的环境中由不同的小组人员来进行测试(确认)。所得的经验宜作为一种改进的输入反馈给评价模块开发组。

附录 B

(资料性附录)

评价模块的实例——故障密度

信息技术 软件产品评价 评价模块:故障密度

B.0 引言

本评价模块用于确定程序的故障密度。在设计和测试期间检测出大量故障数能减少软件中潜在的故障,这些潜在的故障将引起操作期间的失效。在操作开始阶段中较高的故障数引起频繁的失效并降低产品的可靠性。因此,在程序运行前应确保故障密度在规定阈值以下。

总的来说,在软件产品中遗留的故障数是难以统计的。但是,通过使用模型和故障检测的历史数据,可以估计这个数。故障密度即利用这个估计数计算出来。一般估计规程如下:

- 1) 选择一个合适的可靠性增长模型(RGM),例如,采用指数型的 RGM 和 S 形的 RGM;
- 2) 在测试阶段的特定时刻,记录所测故障的累积数目。
- 3) 决定需要适合所记录的数据集曲线的 RGM 等式的参数个数。
- 4) 当 RGM 等式中的时间(t)趋向于无穷时,可计算出潜在故障的估计数。

B.1 范围

B.1.1 特性

可靠性 成熟度 故障密度

B.1.2 评价级别

GB/T 18905.5 中所定义的级别 B。

B.1.3 技术

使用可靠性增长建模技术。在最终软件产品中的故障数使用可靠性增长模型作出预测。

B.1.4 适用性

- 1) 这通常在系统测试期间被使用,并适用于所有类型的编程语言。
- 2) 当必须将故障密度值与用不同编程语言写的程序中的其他值进行比较时,规模值应标准化。

B.2 引用文件

.....

B.3 术语和定义

下列定义适用于评价模块:

故障 fault

软件中的一种缺陷。

失效 failure

一组已定义事件中某一事件的发生(或某些事件未发生)。

LOC(代码行数) (Lines of Code)

代码行数。

ELOC(代码差错行数) (Erroneous Lines Of Code)

检测到并加以修改的代码行数。

EELOC(估计的代码差错行数) (Estimates Erroneous Lines Of Code)

估计的代码差错行数。

NCLOC(无注释的代码行数) (Non-Commented Lines Of Code)

没有注释的代码行数。

FDV(故障密度值) (Fault Density Value)

指出每单位产品量故障数的值。

B.4 输入和度量

B.4.1 评价的输入

下列源用作评价的输入:

- 1) 产品部件:源代码;
- 2) 产品信息:程序测试报告、程序评审报告、程序验证报告。

B.4.2 数据元素

为了应用评价技术,必须收集下列数据元素:

- 1) 测得的故障数(ELOC)。
- 2) 每个失效的检出时间。时间测量的方法必须一致。例如,CPU 时间或日历时间。
- 3) 代码行数(NCLOC)。

B.4.3 度量和测量

故障密度值用下面公式计算:

$$FDV = (EELOC - ELOC) / NCLOC$$

B.5 结果的解释

B.5.1 测量的映射

FDV(在此没有度量映射)。

公司内部实验性尺度源:

$FDV \leq 10E-4$	优秀
$10E-4 < FDV \leq 10E-3$	良好
$10E-3 < FDV \leq 10E-2$	合格
$FDV > 10E-2$	差

注:必须依据应用阶段或软件领域来剪裁阈值。

B.5.2 报告

作为应用评价模块的结果报告下列信息:

- 1) 源代码的标识;
 - 2) 故障密度值:FDV;
- 相应等级(优秀、良好、合格、差)。

B.6 应用规程

B.6.1 所用技术术语的定义

RGIM

可靠性增长模型。可靠性增长模型用于估计出错的代码行数。

B.6.2 需要的资源

使用评价模块时,下列资源必须是可用的:

需要的软件工具

- 1) LOC 计数工具;

- 2) 修改行计数工具;
- 3) 可靠性数据收集工具和分析工具(可选)

硬软件平台

无特别需求。

测试装备或其他设备

无特别需求,但建议使用可靠性数据采集和分析工具。

技能和资格

需要有使用 RGM 方面的知识。

应用方面的人力

大部分工作与失效的测试和记录有关。若采用可靠性和数据采集工具,则 RGM 的应用和 FDV 计算只需要有限的工作。

其他特殊资源

无特殊需求。

B.6.3 评价指令

- 1) 样本的选择

选择样本源代码。采样率宜超过一半。

- 2) 原始数据的生成

失效数据从测试报告中提取。若测试报告不可用,则进行测试。为了采用可靠性增长模型 RGM,需要最小的测试时间为 $\times \times \times$ 。

对于每一个失效,记录其出现的位置和时间。

用修改时间计数修改的 ELOC 和每个样本的总的 NCLOC。

- 3) 算法

通过使用可靠性增长模型 RGM,估计潜在的故障 LOC 数。故障代码行的估计数通过应用 RGM 模型求出。

$$EELOC = RGM - ((\text{失效, 时间}))$$

计算故障密度值:

$$FDV = (EELOC - ELOC) / NCLOC$$

注:必须在这儿详细解释如何使用 RGM 模型或工具,或者,必须提供对用于估计的软件工具和手册所做的引用。

- 4) 工作和最终文档的保留需求

每周做一次测试,在测试阶段做趋势分析。

动作:若故障密度的计算值大于一些规定值(等级=差),则再次进行源代码评审、重测试和调试,并且再次计算故障密度。

B.6.4 文档编制(内部)

下列信息在内部文档中记录:

- 1) 源代码样本的标识和版本;
- 2) 测试文档的标识和版本;
- 3) 用于估计 EELOC 的失效集;
- 4) 应用日期和负责人。

附录 C
(资料性附录)
评价模块的实例——功能性

信息技术 软件产品评价 评价模块:功能性

C.0 引言

评价模块用于确定一个软件系统或部件的“功能性”,该功能性是指软件部件在规定的条件下满足明确的和隐含功能的程度。

如果有一个妥善归档的软件或系统需求的定义,并且明确的要求描述得当,正确进行评价是可能的。在任何情况下,评价都要考虑如何编制文档。

ISO/IEC 9126 按照下面段落中描述的子特性定义了“功能性”。而通过对更多子项的测量说明每个子特性是可测量的。

这里描述的每个子特性的基本项将是该子特性值定义的大致等级。此评价公式当前并不存在,并且显露标准的采用将有信心在其应用环境中达成共同协议。

与评价模块相关的质量模型对与每个子特性有关的元素项考虑了多种度量。并在“检查表”中描述了每个待测元素项的度量,对每个问题都提供了可能有的答案(y/n/d):y表示“是”,n表示“否”,d表示应不适用而被废除。

答案 y/n 涉及度量值与期望值的比较:如果测得值等于或优于期望值,则回答为“是”,否则为“否”。

“功能性”特性通过下列规则进行评价:

$$V_c = \sum V_{sc_i} / n_{sc}$$

$$|V_{sc_i}| = \sum m_i / (n - nd)$$

V_c 为特性的测量值;

V_{sc_i} 为 i 子特性的测量值;

n_{sc} 为子特性数目;

m_i 当第 i 个回答是肯定时为 1,否则为 0;

n 为测量的总数;

nd 为废除的问题数。

C.1 范围

C.1.1 特性

功能性度量指出软件部件是否满足规定的需求。软件部件也应满足隐含的用户要求;换句话说,软件部件隐含的类型需求在评价之列。

功能性的度量包括表示 5 个子特性指标:

- 适合性;
- 准确性;
- 互操作性;
- 依从性;
- 安全性。

适合性度量测量的是,在对所要求的功能进行测试以及用户操作期间满足功能的比率;换句话说,

找出执行任务的不符合文档中的需求的功能。

准确性度量测量的是精确度：

- 计算出错的范围；
- 执行任务的实际结果与期望结果之差；
- 实际操作规程与文档化规程(例如,手册)间的不一致。

互操作性度量测量的是软件部件同其他系统、其他软件产品和其他设备的通信级别：

- 数据传输能力；
- 命令交换能力。

依从性度量测量的是软件部件对照环境的规章或规则的标准化级别。

安全性度量测量的是针对非法访问与非法操作的防范性能级别。

C.1.2 评价级别

在 GB/T 18905.5 的附录 B(资料性附录)中描述了评价级别的选择准则。

在本模块中,针对不同的方面(如安全性、经济性,……),特性测量的临界性会影响评价准确性的定义,执行的测量数量的定义,及拟采用的技术。确定评价级别的问题是:若“功能性”不能满足需求,会存在什么问题?

下表指出每个级别的级别和条件,考虑如下方面:

- 安全性;
- 经济性;
- 保密性;
- 环境。

级别的选择方式是通过选用比分析每个方面得到的级别更高的级别。

级 别	安全性	经济性	保密性	环 境
A	有许多人伤亡	财政危机(公司不能生存)	对战略性数据和业务的保护	不可恢复的环境损害
B	威胁人的生命	巨大经济损失(公司被兼并)	对关键性数据和业务的保护	可恢复的环境损害
C	对财产有破坏,有少数人受伤	重大经济损失(公司受影响)	针对出错风险的保护	局部污染
D	对财产有小的损害,对人没有危险	轻微经济损失	没有具体的风险	无环境风险

C.1.3 技术

下面的表格说明了评价功能性所采用的评价技术,从选定级别所在行开始,直至末行为止;也就是说,若选定的评价级别是 B,则从 B 级到 D 级的技术均适用于此评价。

评价功能性的技术

A 级	形式化证明(目前在 A 级尚不存在评价功能性的相应技术)
B 级	部件测试(白盒测试)
C 级	评审、代码检查
D 级	功能性测试(黑盒测试)

下面是表中所示技术的初步描述。

形式化证明

最一般的程序证明的方法是“归纳断言”。目的是开发评价所用的有关软件部件的一组定理。该方法先写出有关软件部件的输入条件和正确结果的断言。独立证明是必须的,以显示程序总会终止。其他的证明方法有“谓词变换”,“子目标归纳”,“计算归纳”,“结构化归纳”,“中断断言”等。

将来该评价模块会扩大到覆盖该评价级别,但是目前,尚不能在该级别测量功能性。

部件测试

每个软件部件都针对需求加以测试。同时,对完整地集成软件部件加以测试。

白盒测试

该技术允许检验软件部件的内部结构:测试者从对程序逻辑的检验中导出测试数据。

白盒测试是关于测试用例执行或覆盖程序逻辑(源代码)的程度。有效的覆盖准则是“语句覆盖”:即要求程序中的语句要至少被执行一次。更严格的逻辑覆盖准则是“路径覆盖”、“分支覆盖”、“条件覆盖”、“分支/条件覆盖”和“多条件覆盖”。

评审

有关软件部件及所有的相关文档的可视检查/分析的过程。

代码审查

指阅读或可视性检查软件部件。目标是发现差错,而不是寻找差错的解决方法。但是此技术不能发现“高级”差错,如设计中的差错。代码审查被看作是基于一计算机审查如代码静态分析的辅助审查。

检查表

评审活动基于已定义的检查表,该检查表允许有少数主观的活动被重复。检查表中的问题必须尽可能简单;问题目标必须是基本信息。检查表的作用是基于应用经验进行修订、集成和删减。

代码的静态分析

通过分析源代码可减少软件中差错的存在(或可能)。静态分析的一种类型是“控制流”,其中源代码被细分成段,并且检查各段间的关系,以验证不存在不能执行的段,或者不存在不能到达“停止”状态的路径。

分析的另一种类型涉及“调用图”,或“软件系统的结构”,描述所有软件单元的嵌入情况。

功能测试

这是试图发现软件产品与它的外部规格说明之间的差距的过程。分析规格说明并导出一系列测试用例。妥善定义测试用例并采用特定的技术和方法(等价类划分、边界值、分析、因果图、差错猜测方法等)都是很重要的。

黑盒测试

把软件视为黑盒子,即不用关心程序的内部行为和结构。测试者仅对找到程序不根据其规格说明运转所处的环境感兴趣。如果不存在充分的细致测试,则有必要定义充分的测试用例。

C.1.4 适用性

本评价模块的范围是确定软件产品评价的“功能性测量”。本评价模块在两类条件出现时适用:

——有关评价过程需求的条件:

——当该特性在评价时评价过程的具体需求得到满足时适用。

——有关对评价的文档输入的条件:

——适用于评价过程的文档编制输入的可用性。

评价过程的需求:

功能性的软件需求

适合性

——应把所有功能需求编成文档;

——应在文档中描述产品的硬件结构;

——应在文档中描述产品的软件结构;

——应定义所有输入、处理和输出;

- 应明确标识交给实验室进行评价的软件部件为评价中的软件部件；
- 应标识所有软硬件的测试需求；
- 应把执行的测试结果编成文档；
- 应充分规定测试的规格说明；
- 对于功能性需求，所有设计部件应是可追踪的；
- 应描述每次测试的软硬件环境；
- 准确性；
- 程序和数据本身和在文档中都没有矛盾；
- 在软件文档中提及的所有功能应可完整而正确地执行。

安全性

- 应定义安全性需求；
- 应确定所有的安全性威胁、安全性目标和安全性实施功能；
- 安全性实施功能应达到安全性目标。

注：上述安全性需求需要与评价中的软件产品的特定需求结合起来。

对评价的文档输入：

适用于对评价过程的文档输入的可用性

下表描述了对每个评价级别的评价过程所需的文档或部件。同时，该表也指出了在通用软件生存周期内，评价模块何时可用。

在表中，所有对某一等级所规定的内容对其他较低等级也有效。

	需要的软件部件，文档	适用的软件生存周期阶段	评价的需求
等级 A	软件需求评审报告、软件需求验证报告、需求测量报告、软件设计评审报告、软件设计验证报告、设计测量报告、用户文档评审报告、软件规格说明方法和工具的说明、设计方法和工具的说明、编程语言和编译程序的说明	在所有软件生存周期阶段	不适用
等级 B	程序评审报告、程序验证报告、程序评审报告、程序测量报告、程序测试计划、程序测试报告、单元测试计划、单元测试报告、系统需求分析、系统规格说明和设计、系统测试计划、配置管理计划、配置管理报告、质量保证计划、质量保证报告	在整个软件生存周期阶段	在测试活动(7)中与“开发者”的合作，在机器码执行时“目标环境”的可用性，在“评价环境”和“目标环境”之间信息改变后的可用性
等级 C	源代码、编程语言和编译程序的说明、软件需求规格说明(5)、软件设计说明、系统评审报告、系统验证报告、系统测试计划、系统测试报告	软件开发阶段之后	C语言(6)编写的源代码，评审过程中和“开发者”的合作
等级 D	可执行的产品(1)、产品说明(2)、用户手册(3)、系统手册、测试用例(4)	产品交付之前	

(1) 指在评价环境中运行产品，或有权使用可能运行产品的目标环境。