



中华人民共和国国家标准

GB/T 16260.3—2006/ISO/IEC TR 9126-3:2003

软件工程 产品质量 第3部分：内部度量

Software engineering—
Product quality—
Part 3: Internal metrics

(ISO/IEC TR 9126-3:2003, IDT)

2006-03-14 发布

2006-07-01 实施

中华人民共和国国家质量监督检验检疫总局 发布
中国国家标准化管理委员会

目 次

前言	I
引言	II
1 范围	1
2 符合性	1
3 规范性引用文件	1
4 术语和定义	2
5 符号和缩略语	2
6 软件质量度量的使用	2
7 度量表的阅读和使用	3
8 度量表	4
8.1 功能性度量	4
8.2 可靠性度量	4
8.3 易用性度量	5
8.4 效率度量	5
8.5 维护性度量	6
8.6 可移植性度量	6
附录 A (资料性附录) 使用度量时的考虑	27
附录 B (资料性附录) 使用质量的度量、外部度量和内部度量的用法(框架实例)	31
附录 C (资料性附录) 度量标度类型和测度类型的详细解释	37
附录 D (资料性附录) 术语	42
附录 E (资料性附录) 单纯的内部度量	44

前 言

GB/T 16260《软件工程 产品质量》分为如下几部分：

- 第 1 部分(即 GB/T 16260.1):质量模型；
- 第 2 部分(即 GB/T 16260.2):外部度量；
- 第 3 部分(即 GB/T 16260.3):内部度量；
- 第 4 部分(即 GB/T 16260.4):使用质量的度量。

本部分为 GB/T 16260 的第 3 部分。

本部分等同采用 ISO/IEC TR 9126-3:2003《软件工程 产品质量 第 3 部分:内部度量》。

为便于使用,本部分做了下列编辑性修改:

- a) “ISO/IEC 9126 的本部分”改为“本部分”；
- b) 删除了国际标准的前言,修改了国际标准的引言；
- c) 纠正了国际标准中的一些错误。9126-3 的第 E.1 节中,“……that will influence the same or all of the overall software……”中的“same”应改为“some”,即“……影响软件的一些或全部特性和子特性”。

本部分的附录 A、附录 B、附录 C、附录 D 和附录 E 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。

本部分起草单位:中国电子技术标准化研究所、上海计算机软件技术开发中心、杭州波导软件有限公司。

本部分主要起草人:韩红强、杨根兴、王欣、王凌、冯惠、丁志刚、吴黎辉。

引 言

本部分提供了内部度量以测量 GB/T 16260.1—2006 中定义的 6 个内部质量特性的属性。本部分所列的度量并非一个完备集。开发者、评价者、质量管理者和需方可以从本部分中选择合适的度量,用来定义质量需求、评价软件产品、测量质量情况或做其他用途;亦可以修改度量或使用本部分未包括的其他度量。本部分适用于各种软件产品,但并非每种度量适用于各种软件产品。

GB/T 16260.1—2006 定义了软件质量特性,及这些特性如何被分解为子特性。但在该部分中并没有描述这些子特性是如何被测量的。对于这些特性和子特性的测量,GB/T 16260.2—2006 定义了外部度量,GB/T 16260.3—2006 定义了内部度量,而 GB/T 16260.4—2006 定义了使用中的质量的度量。内部度量测量软件本身,外部度量测量包含该软件的计算机系统的行为,使用中的质量的度量则测量在指定条件中使用该软件的效果。

本部分旨在与 GB/T 16260.1—2006 一起使用。极力推荐读者在使用本部分之前阅读 GB/T 16260.1—2006 和 GB/T 18905.1—2002,尤其是在读者不熟悉使用软件度量进行产品规格说明和产品评价的情况下。

第 1 章至第 7 章以及附录 A 到附录 D 为 GB/T 16260 的第 2 部分、第 3 部分和第 4 部分共有。附录 E 仅供本部分使用。

软件工程 产品质量

第3部分：内部度量

1 范围

GB/T 16260 的本部分定义了依据 GB/T 16260.1—2006 定义的特性和子特性来定量测量软件内部质量的内部度量。本部分旨在与 GB/T 16260.1—2006 一起使用。

本部分包括以下内容：

- a) 如何使用软件质量度量的解释；
- b) 每个子特性的基本度量集；
- c) 在软件产品生存周期内如何使用这些度量的实例。

本部分没有为某个评定级别或依从性等级而设置这些度量值的范围，因为这些值是依据每个软件产品或软件产品的一部分的自身特性而定的，也即依赖于软件分类、完整性级别和用户需求等因素。一些属性可能会有期望的取值范围，但不依赖于特定用户的需求，而范围的确定往往依赖于一般因素，例如人类认知因素。

本部分可用于各种应用软件。用户可以选择、修改及应用本部分中的度量和测度，也可以针对独特的应用领域定义特定应用的度量。例如，对于安全性和安全保密性等质量特性的具体测量可参见有关国家标准和国际标准。

本部分旨在针对以下使用者：

- a) 需方(从供方获得或采购系统、软件产品或软件服务的个体或组织)；
- b) 评价者(实施评价的个体或组织。例如评价者可以是测试实验室、软件开发组织的质量部门、政府组织或用户)；
- c) 开发者(执行开发活动的个体或组织，开发活动包括软件生存周期过程中的需求分析、设计、测试直至验收等活动)；
- d) 维护者(执行维护活动的个体或组织)；
- e) 供方(按所签合同向需方提供系统、软件产品或软件服务的个体或组织)，其在合格性测试中确认软件质量时使用；
- f) 用户(使用软件产品执行具体功能的个体或组织)，其在验收测试中评价软件产品质量时使用；
- g) 质量管理者(执行软件产品或软件服务的系统性检查的个体或组织)，作为质量保证和质量控制的组成部分评价软件质量时使用。

2 符合性

符合性不作要求。

注：在 GB/T 16260.1—2006 质量模型中有关于度量的一般符合性要求。

3 规范性引用文件

下列文件中的条款通过 GB/T 16260 的本部分的引用而成为本部分的条款。凡是注明日期的引用文件，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 16260.3—2006/ISO/IEC TR 9126-3:2003

- GB/T 5271.20—1994 信息技术 词汇 20 部分 系统开发(eqv ISO/IEC 2382-20:1990)
- GB/T 8566—2001 信息技术 软件生存周期过程(idt ISO/IEC 12207:1995)
- GB/T 16260.1—2006 软件工程 产品质量 第1部分:质量模型(ISO/IEC 9126-1:2001,IDT)
- GB/T 16260.2—2006 软件工程 产品质量 第2部分:外部度量(ISO/IEC TR 9126-2:2003, IDT)
- GB/T 16260.4—2006 软件工程 产品质量 第4部分:使用质量的度量(ISO/IEC TR 9126-4:2004,IDT)
- GB/T 18491.1—2001 信息技术 软件测量 功能规模测量 第1部分:概念定义(ISO/IEC 14143-1:1998,IDT)
- 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)
- GB/T 18905.6—2002 软件工程 产品评价 第6部分:评价模块的文档编制(ISO/IEC 14598-6:2001,IDT)
- ISO 9241-10:1996 使用视觉显示终端(VDT)办公的人类工效学要求 第10部分:对话原则

4 术语和定义

在GB/T 18905.1—2002和GB/T 16260.1—2006中定义的术语适用于本部分,并且这些术语列在了附录D中。

5 符号和缩略语

SQA:软件质量保证(组)

SLCP:软件生存周期过程

6 软件质量度量的使用

GB/T 16260的第2、第3和第4部分提出了与第1部分“质量模型”一起使用的一组软件质量度量(外部质量、内部质量和使用质量的度量)的建议。这些部分的用户可以修改已定义的度量,和/或也可以使用未列出的度量。当使用一个已修改的或一个未在各部分中定义的新度量时,用户宜说明这些度量与第1部分中的质量模型或任何其他所用的替代质量模型之间的关系。

GB/T 16260的用户宜从第1部分中选择用于评价的质量特性和子特性,确定要采用的适当的直接测度和间接测度,确定相关的度量,并以客观的方式解释测量结果。GB/T 16260的用户也可以从GB/T 18905系列标准中选择软件生存周期中的产品质量评价过程。上述这些标准给出了测量、评估和评价软件产品质量的方法,旨在供开发者、需方和独立的评价者使用,特别是那些负责软件产品评价的人员(见图1)。

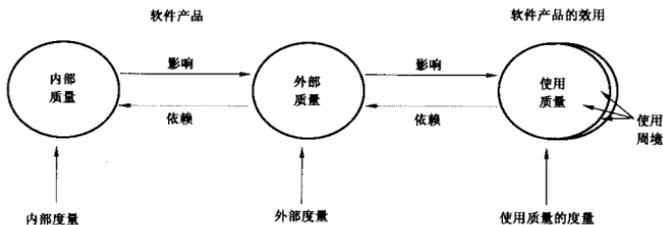


图1 度量类型之间的关系

内部度量可用于开发阶段的非执行软件产品(例如标书、需求定义、设计规格说明或源代码等)。内部度量为用户提供了测量中间可交付项的质量的能力,从而可以预测最终产品的质量。这样就可以使用户尽可能在开发生存周期的早期察觉质量问题,并采取纠正措施。

外部度量可以通过测量该软件产品作为其一部分的系统行为来测量软件产品的质量。外部度量只能在生存周期过程中的测试阶段和任何运行阶段使用。在所属系统环境下运行该软件产品即可获得这样的测量。

使用质量的度量是测量产品在特定的使用周境下,满足特定用户达到特定目标所要求的有效性、生产率、安全性和满意度的程度。这只能在真实的系统环境下获得。

用户的质量要求可用使用质量的度量、外部度量甚至是内部度量的质量需求来规定。这些由度量规定的需求宜作为产品评价时的准则。

建议尽可能采用与目标外部度量有密切关系的内部度量,以便能用这些内部度量来预测外部度量的值。然而,往往很难设计出一个能够在内部和外部度量间提供密切关系的严格的理论模型。因此,假设模型可能是模糊的,所以在使用度量时,外部度量和内部度量关系密切程度模型应该使用统计建模的方法。

GB/T 16260 的第1部分附录A中的A.4列出了与有效性和可信性相关的建议和需求。另外,本部分的附录A列出了使用度量时的一些考虑细节。

7 度量表的阅读和使用

第8章列出的度量依据GB/T 16260.1—2006中的特性和子特性进行分类。下面是表中的每个度量应给出的信息:

a) 度量名称

在内部度量表和外部度量表中的相应度量的类似名称。

b) 度量目的

在度量应用中以回答问题的形式进行描述。

c) 应用的方法

提供一个应用的大纲。

d) 测量、公式和数据元素计算

给出测量公式,并解释所用的数据元素的意义。

注:在某些情况下一个度量对应多个公式。

e) 测量值解释

给出范围和最佳值。

f) 度量标度类型

度量中使用的标度类型。包括：标称标度、顺序标度、间隔标度、比率标度和绝对标度。

注：附录 C 有详细的解释。

g) 测度类型

所用的类型是：规模类型（例如功能规模、源代码规模）、时间类型（例如经时时间、用户时间）、计数类型（变更数、失效数）。

注：附录 C 有详细的解释。

h) 测量输入

测量中使用的数据来源。

i) 在 GB/T 8566 中的应用

标识出应用度量的软件生存周期过程。

j) 目标用户

标识测量结果的使用者。

8 度量表

本章列出的度量并不是一个完备集，而且可能还有未经确认的。这些度量按照 GB/T 16260.1—2006 中介绍的软件质量特性和子特性的顺序列出。可用的度量并不局限于这里所列的。其他特定用途的具体度量将由其他相关文档提出，例如：功能规模的测量、精确时间效率的测量。

注：推荐使用有关标准、技术报告或指南中规定的度量或测量。功能规模测量的定义见 GB/T 18491.1—2001。精确时间效率测量的实例见 ISO/IEC 14756。

在一个特定环境中应用度量之前，应确认该度量（见附录 A）。

8.1 功能性度量

内部功能性度量用来预测该软件产品是否将满足规定的功能需求和隐含的用户要求。

8.1.1 适用性度量

内部适用性度量（见表 8.1.1）指明一组属性，这组属性用来明确评估规定任务的功能，以及确定执行任务的充分性。

8.1.2 准确性度量

内部准确性度量（见表 8.1.2）指明一组属性，这组属性用来评估软件产品达到正确结果或商定结果的能力。

8.1.3 互操作性度量

内部互操作性度量（见表 8.1.3）指明一组属性，这组属性用来评估软件产品与指定系统之间交互的能力。

8.1.4 安全保密性度量

内部安全保密性度量（见表 8.1.4）指明一组属性，这组属性用来评估软件产品抵御非法访问系统和/或数据的能力。

8.1.5 功能性的依从性度量

内部功能性的依从性度量（见表 8.1.5）指明一组属性，这组属性用来评估软件产品遵循与功能性有关的用户组织的标准、约定或法规的能力。

8.2 可靠性度量

在软件产品的开发期间，内部可靠性度量用来预测该软件产品是否满足规定的可靠性要求。

8.2.1 成熟性度量

内部成熟性度量（见表 8.2.1）指明用于评估软件成熟度的一组属性。

8.2.2 容错性度量

内部容错性度量（见表 8.2.2）指明一组属性，这组属性用于评估软件产品在发生操作故障或违反

其规定的接口时维持期望的性能水平的能力。

8.2.3 易恢复性度量

内部易恢复性度量(见表 8.2.3)指明一组属性,这组属性用来评估软件产品在失效时能重新建立一个有足够的性能水平,并恢复直接受影响的数据的能力。

8.2.4 可靠性的依从性度量

内部可靠性的依从性度量(见表 8.2.4)指明一组属性,这组属性用来评估软件产品遵循与可靠性有关的用户组织的标准、约定和法规的能力。

8.3 易用性度量

内部易用性度量用来预测该软件被理解、学习和操作,被吸引以及遵循易用性法规和指南的程度。所实施的测度应尽可能用于建立验收准则或进行产品之间的比较。这意味着测度应是对已知值的项进行计数。结果应报告均值和均值的标准差。

8.3.1 易理解性度量

用户能够选择一个适合他们期望的用法的软件产品。内部易理解性度量(见表 8.3.1)用来评估新用户能否理解下列内容:

- 软件是否合适;
- 如何用于特定任务。

8.3.2 易学性度量

内部易学性度量(见表 8.3.2)用来评估用户要用多长时间才能学会使用各项具体的功能,以及评估帮助系统和文档的效用。

易学性与易理解性有很密切的关系。易理解性的测量可作为软件潜在的易学性指标。

8.3.3 易操作性度量

内部易操作性度量(见表 8.3.3)评估用户是否能够操作和控制软件。易操作性度量可按 ISO 9241-10:1998中的对话原则进行分类:

- 软件对任务的适合性;
- 软件的自描述性;
- 软件的可控制性;
- 软件与用户期望的符合性;
- 软件的容错性;
- 软件对个性化的适应性;

要测试的功能选择与功能使用的预期频率、功能的关键性程度以及预期的易用性问题有关。

8.3.4 吸引力度量

内部吸引力度量(见表 8.3.4)用来评估软件的外观,并受屏幕设计、颜色等因素的影响。这一点对于消费者产品特别重要。

8.3.5 易用性的依从性度量

内部易用性的依从性度量(见表 8.3.5)评估与易用性相关的标准、约定、风格指南或法规的遵循性。

8.4 效率度量

内部效率度量用来在测试或运行期间预测软件产品所表现出的效率。为测量效率,应定义明确的条件,即应定义一个参照环境(必须在软件规格说明中定义)中的硬件配置和软件配置。在引用测量的时间特性值时,应说明参照环境。

8.4.1 时间特性度量

内部时间特性度量(见表 8.4.1)指明一组属性,这组属性用来预测在测试或运行中包含软件产品的计算机系统的时间特性。

8.4.2 资源利用性度量

内部资源利用性度量(见表 8.4.2)指明一组属性,这组属性用来预测在测试或运行中包含软件产品的计算机系统对硬件资源的利用情况。

8.4.3 效率的依从性度量

内部效率的依从性度量(见表 8.4.3)指明一组属性,这组属性用来评估该软件产品遵循与效率相关的用户组织的标准、约定或法规的能力。

8.5 维护性度量

内部维护性度量用来预测为修改软件产品需要付出的工作量。

8.5.1 易分析性度量

内部易分析性度量(见表 8.5.1)指明一组属性,这组属性用来预测维护者或用户在试图诊断软件产品的缺陷或失效的原因时、或标识需要修改的部分时所耗费的工作量或耗费的资源。

8.5.2 易改变性度量

内部易改变性度量(见表 8.5.2)指明一组属性,这组属性用来预测维护者或用户在对软件产品进行某项特定的修改时所需耗费的工作量。

8.5.3 稳定性度量

内部稳定性度量(见表 8.5.3)指明一组属性,这组属性用来预测对软件产品进行任何修改后的稳定程度。

8.5.4 易测试性度量

内部易测试性度量(见表 8.5.4)指明一组属性,这组属性用来预测软件产品中已设计并实现的自动测试辅助功能的总量。

8.5.5 维护性的依从性度量

内部维护性的依从性度量(见表 8.5.5)指明一组属性,这组属性用来评估软件产品遵循与维护性有关的用户组织的标准、约定或法规的能力。

8.6 可移植性度量

内部可移植性度量用来预测在移植活动期间,软件产品对实现者或系统的行为所产生的影响。

8.6.1 适应性度量

内部适应性度量(见表 8.6.1)指明一组属性,当用户试图使软件产品适用于不同的特定环境时,这组属性用来预测软件产品对用户需付出努力程度的影响。

8.6.2 易安装性度量

内部易安装性度量(见表 8.6.2)指明一组属性,当用户试图在一个指定用户环境中安装软件时,这组属性用来预测软件产品对用户需付出努力程度的影响。

8.6.3 共存性度量

内部共存性度量(见表 8.6.3)指明一组属性,这组属性用来预测当软件产品与其他软件产品共享相同的可运行硬件资源时对其他软件产品产生的影响。

8.6.4 易替换性度量

内部易替换性度量(见表 8.6.4)指明一组属性,当用户试图在规定环境及使用条件中用该软件代替其他指定软件时,这组属性用来预测软件产品对用户需付出努力程度的影响。

8.6.5 可移植性的依从性度量

可移植性的依从性度量(见表 8.6.5)指明一组属性,这组属性用来评估软件产品遵循与可移植性有关的用户组织的标准、约定或法规的能力。

表 8.1.1 适合性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	度量类型	测量输入	在 GB/T8566 中的应用	目标用户
功能的充分程度如何？	对已实现的适合于执行特定任务的功能进行计数,然后测量它与实现的功能的比率。 可测量下列指标: —全部或部分设计规格说明; —完成的模块/部分软件产品。	$X=1-A/B$ A=在评价中检测到有问题的功能数 B=检查的功能数	$0 \leq X \leq 1$ 越接近 1,越充分	绝对标准	X = 计数 A = 计数 B = 计数	需求规格说明 设计文档 源代码 评审报告	6.5 确认 6.6 联合评审	需求方 开发者	
功能实现的完整性	对评价中检测到遗漏的功能进行计数,并与需求规格说明中描述的功能数相比较。 注:测量过程的输入是更新的需求规格说明。在生存周期中标识的任何变更必须在测量过程使用之前对需求规格说明进行更新。	$X=1-A/B$ A=评价中检测到遗漏的功能数 B=在需求规格说明中描述的功能数	$0 \leq X \leq 1$ 越接近 1,越完整	绝对标准	X = 计数 A = 计数 B = 计数	需求规格说明 设计文档 源代码 评审报告	6.5 确认 6.6 联合评审	需求方 开发者	
功能实现的覆盖率	对不能正确实现或遗漏的功能进行计数,并与需求规格说明中描述的功能数相比较。 注:用功能项评审。	$X=1-A/B$ A=检测到的不能正确实现或遗漏的功能数 B=需求规格说明中描述的功能数	$0 \leq X \leq 1$ 越接近 1,越正确	绝对标准	X = 计数 A = 计数 B = 计数	需求规格说明 设计文档 源代码 评审报告	6.5 确认 6.6 联合评审	需求方 开发者	
功能规格说明的稳定性(易变性)	对在生存周期开发阶段中变更(增加、修改、删除)的功能进行计数,并与需求规格说明中描述的功能数相比较。	$X=1-A/B$ A=在生存周期开发阶段变更的功能数 B=在需求规格说明中描述的功能数	$0 \leq X \leq 1$ 越接近 1,越稳定	绝对标准	A = 计数 B = 计数 X = 计数 数	需求规格说明 评审报告	6.5 确认 6.3 质量保证 5.3 鉴定测试 6.8 问题解决 5.4 运作	开发者 维护者	

表 8.1.2 准确性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
计算的准确性	准确性需求已完全实现的程度如何?	对已经实现准确性需求的功能进行计数,并与有特定准确性需求的功能数相比较。	$X=A/B$ A=在评价中已证实的实现准确性需求的功能数 B=需要实现特定准确性需求的功能数	$0 \leq X \leq 1$ 越接近 1,越完全	绝对标准	X=计数/计数 A=计数 B=计数	需求规格说明 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
精度	完全实现数据项的特定级别精度的程度如何?	对满足特定级别精度需求的数据项进行计数,并与有特定精度需求的数据项总数相比较。	$X=A/B$ A=在评价中已证实的满足特定级别精度要求的数据项数 B=需要特定级别精度要求的数据项数	$0 \leq X \leq 1$ 越接近 1,越完全	绝对标准	X=计数/计数 A=计数 B=计数	需求规格说明 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

表 8.1.3 互换性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
数据的可交换性(依据数据格式)	正确实现接口数据格式的程度如何?	对按规格说明已经正确实现的接口数据格式进行计数,并与规格说明中要交换的数据格式相比较。	$X=A/B$ A=按规格说明正确实现接口数据格式数 B=规格说明中要交换的数据格式数	$0 \leq X \leq 1$ 越接近 1,越正确	绝对标准	X=计数/计数 A=计数 B=计数	需求规格说明 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
接口的一致性(协议)	正确实现接口协议的协议如何实现程度如何?	对按规格说明已经正确实现的接口协议进行计数,并与规格说明中要实现接口协议的接口数相比较。	$X=A/B$ A=在评审中已证实的按规格说明正确实现的接口协议数 B=规格说明中要实现接口协议的接口数	$0 \leq X \leq 1$ 越接近 1,越一致	绝对标准	X=计数/计数 A=计数 B=计数	需求规格说明 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

表 8.1.4 安全保密性度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标度类型	测量类型	测量输入	在 GB/T 8566 中的应用	目标用户
访问的可靠性	访问记录可审核能力如何?	对照规格说明被正确记录的访问类型进行计数,并与规格说明中需要记录的访问类型计数相比较。	$X=A/B$ A=按规格说明被记录的访问类型数 B=规格说明中要求记录的访问类型数	$0 \leq X \leq 1$ 越接近 1,审核能力越强	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 源代码 评审报告	6.5 确认 6.6 联合评审	需求方 开发者
访问的可控性	对系统访问的可控程度如何?	对照规格说明已被正确实现的访问需求,并与规格说明中的可控访问需求相比较。	$X=A/B$ A=按规格说明被正确实现的访问需求数 B=规格说明中的可控访问需求数	$0 \leq X \leq 1$ 越接近 1,越可以控制	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 源代码 评审报告	6.5 确认 6.6 联合评审	需求方 开发者
数据的抗流篡性	数据的抗流篡性完全实现的程度如何?	对照规定实现的实例进行计数,并与需求中规定为能扰乱/破坏数据的操作/访问实例相比较。	$X=A/B$ A=在评审中已证实的实现的数据抗流篡的实例数 B=需求中确定为能扰乱/破坏数据的操作/访问实例数	$0 \leq X \leq 1$ 越接近 1,越完全	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 源代码 评审报告	6.5 确认 6.6 联合评审	开发者
数据加密	实现数据加密的完备程度如何?	对照规定已实现的实例进行计数,并与规格说明中要求加密/解密功能的数据项实例数相比较。	$X=A/B$ A=在评审中已证实的实现的数据项实例数 B=规格说明中要求加密/解密功能的数据项实例数	$0 \leq X \leq 1$ 越接近 1,越完全	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 源代码 评审报告	6.5 确认	开发者

注:使用这一度量时要考虑安全保密性级别。

注:数据加密,例如,开放数据库中的数据,公用通信设备中的数据。

表 8.1.5 功能性的依从性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
功能性的依从性	遵循与产品有关的功能性有关的法规、标准和约定的程度如何?	对已满足的需要依从的项进行计数,并与规格说明要求的依从性项数相比较。	$X=A/B$ A=在评价中已证实的正确实现的与功能性的依从性有关的项数 B=依从性项目总数	$0 \leq X \leq 1$ 越接近 1, 依从性越强	绝对标准	X=计数 A=计数 B=计数	依从性的规格说明和相关的标准、约定或法规设计源代码评审报告	6.4 验证 6.6 联合评审	需方 开发者
系统回标准的依从性	遵循与接口有关的法规、标准和约定的程度如何?	对已满足的需要依从的接口进行计数,并与规格说明要求的依从的接口数相比较。	$X=A/B$ A=在评审中已证实的规格正确实现的接口数 B=需要依从的接口总数	$0 \leq X \leq 1$ 越接近 1, 依从性越强	绝对标准	X=计数 A=计数 B=计数	需求规格说明设计源代码评审报告	6.4 验证 6.6 联合评审	开发者 需方

注: 必须检查标准规定的所有属性。

表 8.2.1 成熟性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
故障检测	已评审的产品中检测到多少故障?	对评审中检测到故障的绝对数,并与此阶段估计会检测出的故障数相比较。	$X=A/B$ A=在评审中检测到的故障数 B=在评审中估计会检测出的故障数(根据过去的历史或引用的模型)	$0 \leq X$ X 值越大, 产品质量越好, 但 $A=0$ 并不表明被评审项中没有隐含的故障。	绝对标准	X=计数 A=计数 B=计数	值 A 取自评审报告 值 B 取自组织数据库	6.4 验证 6.6 联合评审	需方 开发者

注 1: 本度量仅在开发过程中用于预测。

注 2: 在阐述特性时有必要将 X 归到 [0, 1] 区间。

表 8.2.1 (续)

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测试类型	测量输入	在 GB/T 8566 中的应用	目标用户
故障排除	有多少故障已被纠正? 故障被排除的比例是多少?	对设计/编码期间被排除的故障进行计数,并与设计/编码期间的评审中检测出的故障数相比较。	$X=A$ A=设计/编码中已纠正的故障数 $Y=A/B$ A=设计/编码中已纠正的故障数 B=评审中检测到的故障数	$0 \leq X$ X 越大,残存的故障越少 $0 < Y \leq 1$ 越接近 1,越好 (越多故障被纠正)	绝对标准 比率标准	X=计数 A=计数 Y=计数/计数 B=计数	值 A 取自故障排除报告 值 B 取自评审报告	6.4 验证 6.6 联合评审	需方 开发者
测试充分性	测试计划覆盖了多少要求的测试用例?	对计划的测试用例进行计数,并与为获得充分测试覆盖率而要求的测试用例数相比较。	$X=A/B$ A=在测试计划中设计并在评审中证实的测试用例数 B=要求的测试用例数	$0 < X$ X 越大越充分	绝对标准	X=计数/计数 A=计数 B=计数	值 A 取自测试计划 值 B 取自需求	6.3 质量保证 6.8 问题解决 6.4 验证	开发者 维护者

注:在概述特性时有必要将 X 归到[0..1]区间。

表 8.2.2 容错性度量

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测试类型	测量输入	在 GB/T 8566 中的应用	目标用户
避免失效	为避免关键的和严重的失效,可以控制多少种故障模式?	对可避免的故障模式进行计数,并与要考察的故障模式数相比较。	$X=A/B$ A=在设计/编码过程中可以避免的故障模式个数 B=要考察的故障模式个数	$0 < X$ X 值越大,避免失效越好。	绝对标准	X=计数/计数 A=计数 B=计数	值 A 取自评审报告 值 B 取自需求规格说明文档	6.4 验证 6.5 确认 6.6 联合评审 6.8 问题解决	开发者 需方 维护者
		故障模式数注册、故障模式举例、数据出界、死锁							
		注册、故障树分析技术可被用来检测故障模式。							

表 8.2.2 (续)

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
抵御误操作	有多少具有抵御误操作能力的功能被实现?	对抵御因误操作引起严重的功能失效而实现的功能进行比较, 并与要考虑的误操作模式数相比较。 注解: 除系统失效以外, 还有数据损坏。	$X=A/B$ A=为抵御误操作模式而实现的功能数 B=要考虑的误操作模式数 注解: 误操作模式、错误的参数数据类型、错误的数据输入序列、错误的操作序列 注解: 故障树分析技术用来检测误操作模式。	$0 \leq X$ X 越大, 抵御抗误操作的能力越强。	绝对标准	X = 计数/计数 A=计数 B=计数	值 A 取自评审报告 值 B 取自需求规格说明文档	6.4 验证 6.5 确认 6.6 联合评审 6.8 问题解决	开发者 需求方 维护者

表 8.2.3 易恢复性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
易恢复性	在异常事件或在需要时, 产品本身的复原能力如何?	对已实现的复原需求进行计数, 并与规格说明中的复原需求数相比较。 复原需求的例子: 数据库检查点; 事务检查点; 重复功能; 还原功能。	$X=A/B$ A=在评审中证实的已实现的复原需求数 B=规格说明中的复原需求数	$0 \leq X \leq 1$ X 值越大易复原性越好	绝对标准	X = 计数/计数 A=计数 B=计数	A 取自评审文档 B 取自需求或设计文档	6.4 验证 6.6 联合评审	开发者 维护者

表 8.2.3 (续)

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
复原的有 效性	复原能力的有 效性如何?	对已实现的满足目 标修复时间(通过 计算或模拟)的修 复需求进行计数, 并与规定的目标时 间的修复需求数相 比较。	$X=A/B$ A=已实现的满足目标修复 时间的修复需求数 B=有规定目标时间要求的 修复需求数	$0 \leq X \leq 1$ X 越大,有效 性越好	绝对标准	X=计数 A=计数 B=计数	A取自评审文 档 B取自需求或 设计文档	6.4 验证 6.6 联合评审	开发者 维护者

表 8.2.4 可靠性的依从性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
可靠性的依 从性	遵循与产品的 可靠性有关的 法规、标准和 约定的程度如 何?	对已满足的 需要依从的 项目进行计 数,并与规 格说明要求 的依从性项 数相比较。	$X=A/B$ A=在评价中证实的已正确 实现与可靠性依从相关的项 数 B=依从项总数	$0 \leq X \leq 1$ 越接近 1,依从 性越好	绝对标准	X=计数 A=计数 B=计数	依从性规格说 明和相关标 准、约定或法 规 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需方 开发者

表 8.3.1 易理解性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
描述的完整 性	在产品描述中 被描述的功能 (或功能类型) 的比例是多 少?	对被充分描 述的功能进 行计数,并与 产品的总功 能数相比较。	$X=A/B$ A=在产品描述中被描述的 功能(或功能类型)数 B=功能(或功能类型)的总 数	$0 \leq X \leq 1$ 越接近 1,越完 整	绝对标准	X=计数 A=计数 B=计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者

注 1: 指出潜在用户在阅读了产品描述后能否理解产品的能力。

注 2: 见 ISO/IEC 9127 消费者软件包。

表 8.3.1 (续)

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标度类型	测量输入	在 GB/T 8566 中的应用	目标用户
演示能力	需要演示的功能中具有演示能力的比例是多少?	对具有充分演示能力的功能进行计数,并与要求有演示能力的总功能数相比较。	$X = A/B$ A=在评审中已证实和演示的功能数 B=要求有演示能力的功能总数	$0 \leq X \leq 1$ 越接近 1, 越有能力	X = 计数/总数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
功能的可见性	产品功能对用户是明显可见的比例是多少?	对用户是显而易见的功能(或功能类型)数进行比较。	$X = A/B$ A=对用户是显而易见的功能(或功能类型)数 B=功能(或功能类型)的总数	$0 \leq X \leq 1$ 越接近 1, 越好	X = 计数/总数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
功能的易懂性	产品功能能被用户正确理解的比例是多少?	对能被用户理解的界面功能数,并与用户界面功能的总数相比较。	$X = A/B$ A=能被用户理解的界面功能数 B=用户界面功能总数	$0 \leq X \leq 1$ 越接近 1, 越好	X = 计数/总数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

注: 演示产品如何使用过程的演示步骤, 包括“向导”。

注: 说明用户能否通过界面(例如通过检查菜单)来定位这些功能。

表 8.3.2 易学性度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标度类型	测量输入	在 GB/T 8566 中的应用	目标用户
用户文档和/或帮助机制的完整性	用户文档和/或帮助机制中描述的功能的比例是多少?	对帮助机制和/或文档中已实现的功能, 并与产品功能总数相比较。	$X = A/B$ A=描述过的功能数 B=提供的功能总数	$0 \leq X \leq 1$ 越接近 1, 越完整	X = 计数/总数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

注: 可能有三项度量: 文档的完整性、帮助机制的完整性或帮助及文档组合使用的完整性。

表 8.3.3 易用操作性度量

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	质量标度类型	测量类型	测量输入	在 GB/T 8566 中的应用	目标用户
输入的有 效性检查	输入项提供了 对有效数据进 行检查的比例 是多少?	对检查有效数据的 输入项进行计数, 并与在输入数据项 中可能进行有效性 检查的总项数相比 较。	$X=A/B$ A =对有效数据 输入项数 B =对有效数据可能进行检 查的输入项数	绝对标度 $0 \leq X \leq 1$ 越接近 1, 越好	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者
用户操作 的易取消 性	功能在完成之 前能被取消的 比例是多少?	对已实现的能在完 成之前被用户取消 的功能进行计数, 并与要求具备预取 能力的功能总数 相比较。	$X=A/B$ A =已实现的能被 用户取消的功能数 B =要求具备预取 能力的功能总数	绝对标度 $0 \leq X \leq 1$ 越接近 1, 易取 消性越强	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者
用户操作 的易还原 性	功能被还原 的比例是多 少?	对已实现的能在动 作完成之后被用户 还原的功能进行计 数, 并与功能总数 相比较。	$X=A/B$ A =已实现的功能 被用户还原的功能 数 B =功能总数	绝对标度 $0 \leq X \leq 1$ 越接近 1, 越有 还原能力	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者
易定制性	能在运行中被 定制的功能的 比例是多少?	对已实现的能在操 作中被用户定制 功能进行计数, 并 与要求具有定制能 力的功能总数相比 较。	$X=A/B$ A =能在运行中 被用户定制的功能 数 B =要求具有定制能 力的功能总数	绝对标度 $0 \leq X \leq 1$ 越接近 1, 易定 制性越强	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者
物理易访 问性	能为有生理缺 陷的用户定制 访问的功能的 比例是多少?	对已实现的能为用 户定制的功能进行 计数, 并与功能的 总数相比。	$X=A/B$ A =能被定制的功能 数 B =功能的总数	绝对标度 $0 \leq X \leq 1$ 越接近 1, 物理 易访问性越强	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需方 开发者

注: 物理易访问性的例子如无能使用鼠标的人或盲人。

表 8.3.3 (续)

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
运行状态的易监控性	具有运行状态的监控能力的功能的比例是多少？ 注：状态也包括进程的监控。	对已实现的功能监控的功能进行计数，并与要求有监控能力的功能总数相比较。	$X=A/B$ A=具有状态监控能力的功能数 B=要求有监控能力的功能总数	$0 \leq X \leq 1$ 越接近 1, 易监控性越好	绝对标准	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
操作的一致性	操作行为与系统其他部分类似操作相同的比例是多少？ 注：操作行为不一致的操作实例数	对操作行为不一致的操作实例数，并与操作的总数相比较。	$X=1-A/B$ A=操作行为不一致的操作实例数 B=操作的总数	$0 \leq X \leq 1$ 越接近 1, 越一致	绝对标准	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
消息的明确性	消息能够自解释的比例是多少？ 注：清晰的出错消息向用户解释采取什么动作能从出错中恢复。	对已实现的有明确解释的消息进行计数，并与要实现的消息总数相比较。	$X=A/B$ A=已实现的带有明确解释的消息数 B=要实现的消息总数	$0 \leq X \leq 1$ 越接近 1, 越清晰	绝对标准	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
界面元素的准确性	界面元素能自解释的比例是多少？ 注：当元素使用简洁文字或提供了“帮助框”或“工具条”时就称为带有自解释。	对带自解释的界面元素进行计数，并与界面元素的总数相比较。	$X=A/B$ A=带自解释的界面元素数 B=界面元素的总数	$0 \leq X \leq 1$ 越接近 1, 越清晰	绝对标准	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
运行差错性的易恢复性	能够容忍用户的差错的功能的比例是多少？ 注：运行差错性的易恢复性是指当用户输入了错误的命令或数据时，系统能够识别并纠正错误的功能。	对已实现的、能容忍用户差错的功能进行计数，并与要实现有容错能力的功能总数相比较。	$X=A/B$ A=已实现的有用户容错能力的功能数 B=要求有容错能力的功能总数	$0 \leq X \leq 1$ 越接近 1, 越易恢复	绝对标准	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

表 8.3.4 吸引力度量

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
交互吸引力	界面用户对用户的吸引程度如何?	对用户的问卷调查。	用问卷的方式向用户调查界面用户对用户的吸引程度,要考虑颜色与图像设计等属性。 注解:吸引力的潜在因素包括:项的排列(水平和垂直)、分组、颜色的使用,适当及合理的图像大小、空白/间隔/边界的使用、动画、排印及 3D 界面等	分类评估	顺序标度	$X = \text{计数(计数表得分)}$	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者
外观的易用性	有多大比例的界面元素能从外观上被定制?	由专家进行。	$X = A/B$ A=能被定制的界面元素的数量 B=界面元素类型的总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	$X = \text{计数/计数}$ A=计数 B=计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

注:这与屏幕布局或实物模型有关。

表 8.3.5 易用性的依从性度量

度量名称	度量的目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
易用性的依从性	遵循与产品的易用性有关的要求的程度如何?	对已经满足的要求的依从性项进行计数,并求规格说明要求依从的总项数相比。	$X = A/B$ A=在评价中已证实的正确实现与易用性的依从性相关的项数 B=依从性的总项数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	$X = \text{计数/计数}$ A=计数 B=计数	依从性的规格说明和相关的标准、约定或法规设计源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

表 8.4.1 时间特性度量

度量名称	度量目的	应用的方法	测量公式及 数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
响应时间	估计完成一项特定任务要多长时间?	评价操作系统和应用系统调用的效率。据此估计响应时间。 可在下列事项中测量： ——部分或全部的设计规格说明； ——测试完成事务的路径； ——测试完成的模块或部分软件产品； ——在测试阶段完成的软件产品。	$X = \text{时间}$ (计算的或模拟的)	越短越好	比率标准	$X = \text{时间}$	已知的操作系统在系统调用中的估计用时	6.4 验证 6.6 联合评审	需求 开发者
吞吐量时间	估计在单位时间内能执行的任务数?	评价在系统中处理资源的效率。根据在处理资源时对整个的应用调用,产生一个因子	$X = \text{每单位时间完成的任务个数}$	越大越好	比率标准	$X = \text{计数}$	已知的操作系统在系统调用中的估计用时	6.4 验证 6.6 联合评审	需求 开发者
周转时间	估计整批完成一组相关任务需要多长时间?	评价操作系统和应用系统调用的效率。据此估计完成一组相关任务的响应时间。 可在下列事项中测量： ——部分或全部的设计规格说明； ——测试完成事务的路径； ——测试完成模块或部分软件产品； ——在测试阶段完成软件产品。	$X = \text{时间}$ (计算的或模拟的)	越短越好	比率标准	$X = \text{时间}$	已知的操作系统在系统调用中的估计用时	6.4 验证 6.6 联合评审	需求 开发者

表 8.4.2 资源利用性度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	度量标准类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
I/O 利用率	估计完成特定任务的 I/O 利用率是多大?	估计此项应用对 I/O 利用率的需求。	X=缓冲区域数目 (计算的或模拟的)	比率标准	X=大小	源代码	6.4 验证	开发者
I/O 利用的消息密度	在负责实现系统调用的代码行中涉及 I/O 利用率的消息的密度是少?	对属于 I/O 失败的错误及警告消息行计数, 并与估计的承担系统调用的代码行数相比较。	X=A/B A=与 I/O 相关的错误消息数 B=直接与系统调用相关的代码行数	绝对标准	X=计数/计数 A=计数 B=计数	源代码	6.4 验证	开发者
内存利用率	估计产品为完成特定任务将占用多大的内存?	估计内存的需求。	X=以字节为单位的内存大小 (计算的或模拟的)	比率标准	X=大小	内存利用的估计大小	6.4 验证	开发者
内存利用的消息密度	在负责实现系统调用的代码行中, 涉及内存利用率的消息的密度是少?	对属于内存失败的错误及警告消息行计数, 并与估计的承担系统调用的代码行数相比较。	X=A/B A=与内存相关的错误消息数 B=直接与系统调用相关的代码行数	比率标准	X=计数/计数 A=计数 B=计数	源代码	6.4 验证	开发者
传输利用率	估计传输资源利用率的总数是多少?	通过估计传输容量来估计传输资源利用率的需求。	X=位/时间 (计算的或模拟的)	比率标准	X=时间	已知的操作系统在系统调用中的估计用时	6.4 验证	开发者

表 8.4.3 效率的依从性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测量类型	测量输入	在 GB/T 8566 中的应用	目标用户
效率的依从性	遵循与产品的要求有关的法规、标准及约定的程度如何？	对已经满足的要求的依从性进行计数，并与规格说明要求依从的项数相比较。	$X = A/B$ A=在评价中已证实有效的项数 B=依从性项总数	$0 \leq X \leq 1$ 越接近 1, 依从性越好	绝对标准	X = 计数/计数 A = 计数 B = 计数	依从性的规格说明、相关的标准、约定和法规设计源代码 评审报告	6.4 验证 6.6 联合评审	需求方 开发者

表 8.5.1 易分析性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标准类型	测量类型	测量输入	在 GB/T 8566 中的应用	目标用户
活动的记录	系统状态记录的完整程度如何？	对照规定在日志中记录的行计数，并与需要记录的项数比较。	$X = A/B$ A=在评审中已证实按照规范定义的项数 B=在规格说明中定义的项数	$0 \leq X \leq 1$ 越接近 1, 记录系统状态的项数越提供得越完整 注解：在概述特性时有必要将 X 归到 [0, 1] 区间。	绝对标准	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 取自需求规格说明。	6.4 验证 6.6 联合评审	维护者 用户
诊断功能准备的情况	诊断功能准备得全面程度如何？	对照规定已实现诊断功能的项进行计数，并与规格说明中要求的诊断功能数相比较。 注解：此度量也可用来测量失效分析能力及因果分析能力。	$X = A/B$ A=在评审中已证实按照规范定义的项数 B=要求的诊断功能数	$0 \leq X$ 越接近 1, 诊断功能实现得越好。 注解：在概述特性时有必要将 X 归到 [0, 1] 区间。	绝对标准	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 取自规定的要求。	6.4 验证 6.6 联合评审	维护者 用户

表 8.5.2 易改变性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	度量值解释	度量标度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
变更的记录	规格说明和程序模块的变更是否在代码的注释行中被恰当地记录下来?	模块变更信息记录比率。	$X = A/B$ A=在评审中已证实有变更注释内容的功能/模块的变更数 B=按源代码变更的功能/模块的总数	$0 \leq X < 1$ 越接近 1, 可记录的变更越多。	绝对标度	X = 计数/计数 A = 计数 B = 计数	配置控制系统 版本记录 规格说明	6.4 验证 6.6 联合评审	开发者 维护者 需方

表 8.5.3 稳定性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	度量值解释	度量标度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
变更的影响	修改后发生不利影响的频率是多少?	对检测到的在修改之后发生不利影响的次数进行计数, 并与已执行的修改次数相比较。	$X = 1 - A/B$ A = 检测到修改之后发生不利影响的次数 B = 实施修改的次数	$0 \leq X \leq 1$ 越接近 1, 越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 也取自评审报告。	6.4 验证 6.5 联合评审	开发者 维护者 需方
修改影响的局部化	软件产品修改后, 受到的影响有多大?	对一次修改中受影响的变量进行计数, 并把它与产品的变量总数进行对比。	$X = A/B$ A = 在评审中证实受修改影响的变量数 B = 变量的总数	$0 \leq X \leq 1$ 越接近 0 修改的影响越小。	绝对标度	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 也取自评审报告。	6.4 验证 6.6 联合评审	开发者 维护者 需方

表 8.5.4 易测试性度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
内置测试功能的完整性	内置测试能力的完整性如何？	对按照规定已实现的内置测试的功能进行计数，并与需求规格说明中内置测试的功能数相比较。	$X = A/B$ A = 在评审中证实按照规定实现的内置测试的功能数 B = 要求内置测试的功能数	$0 \leq X \leq 1$ 越接近 1, 越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 取自需求规格说明或设计文件。	6.4 验证 6.6 联合评审	开发者 维护者
易测试性的自主性	软件测试的独立性如何？	对依赖于其他系统的已经用桩模块模拟的测试项进行计数，并与其他系统的测试项相比较。	$X = A/B$ A = 依赖于其他系统的已经用桩模块模拟的测试项总数 B = 依赖于其他系统的测试项总数	$0 \leq X \leq 1$ 越接近 1, 越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 取自需求规格说明或设计文件。	6.4 验证 6.6 联合评审	开发者 维护者 需求方
测试进程的可见性	在测试期间内对测试结果展示的完整程度如何？	对按规定已实现的检查点进行计数，并与设计要求规定的检查点数目相比较。	$X = A/B$ A = 在评审中证实按规定已实现的检查点总数 B = 设计中的检查点总数	$0 \leq X \leq 1$ 越接近 1, 越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	A 取自评审报告, B 取自设计文件。		开发者 维护者 需求方

表 8.5.5 维护性的依从性度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
维护性的依从性	遵循与产品的维护性有关的法规、标准及约定的程度如何？	对要求的依从性已满足的项进行计数，并与规格说明中要求的依从性项数相比较。	$X = A/B$ A = 在评价中证实已正确实现的与维护性的依从性有关的项数 B = 依从项总数	$0 \leq X \leq 1$ 越接近 1, 依从性越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	依从性的规格说明相关的标准、约定和法规设计源代码 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需求方

表 8.6.1 适应性度量

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
数据结构 的适应性	产品对数据结构的变更的适应程度如何?	对适应后能够操作且没有限制的数据结构进行计数,并与要求适应能力的数据结构结构的总数相比较。	$X=A/B$ A=经评审证实能适应后能够操作且不含限制的数据结构的计数 B=要求有适应能力的数据结构结构的总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计数 A=计数 B=计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需求方
硬件环境 (对硬件 设备和网 络设施的 适应性)	产品对与 H/W 相关的环境变更的适应程度如何?	对已实现的能在指定的多种 H/W 环境中按规定取得所要求的结果的功能进行计数,并具有 H/W 环境适应能力要求的功能总数相比较。	$X=A/B$ A=经评审证实已实现的能在指定的多种 H/W 环境中取得所要求结果的功能数 B=具有 H/W 环境适应能力要求的功能总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计数 A=计数 B=计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需求方
组织环境 的适应性 (对组织 的基础设 施的组织 适应性)	产品对组织变更的适应程度如何?	对已实现的能在指定的多种组织及业务环境中按规定取得所要求的結果的功能进行计数,并与具有组织环境适应能力要求的功能总数相比较。	$X=A/B$ A=经评审证实已实现的能在指定的多种组织及业务环境中按规定取得所要求结果的功能数 B=具有组织环境适应能力要求的功能总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计数 A=计数 B=计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需求方

表 8.6.1 (续)

度量名称	度量目的	应用的方法	测量、公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
用户对移植的友好性	对产品移植操作的容易程度如何?	对已实现的能按规定支持易被用户掌握的功能的计数,并与要求易于掌握的功能总数相比较。	$X=A/B$ A=经评审证实按规定支持易被用户掌握的功能的计数 B=具有易于掌握要求的功能总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需方
系统软件环境的适应性(对 OS、网络软件和合用软件的适应性)	产品对与系统软件相关的环境变更的适应程度如何?	对已实现的能在指定的环境中按规定的功能所要求结果的功能数进行计数,并与具有系统软件环境适应性要求的功能总数相比较。	$X=A/B$ A=经评审证实已实现的能在指定的多种系统软件环境中获取所要求结果的功能数 B=具有系统软件环境适应性要求的功能总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计数 A = 计数 B = 计数	需求规格说明 设计 评审报告	6.4 验证 6.6 联合评审	开发者 维护者 需方

表 8.6.2 易安装性度量

度量名称	度量目的	所采用的方式	测量、公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
易于重新安装	重复安装操作是否容易?	对已实现的重新安装操作的操作数进行计数,并与要求的重新安装操作的总计数相比较。	$X=A/B$ A=经评审证实已实现的重新安装操作的操作数 B=要求重新安装操作的总计数	$0 \leq X \leq 1$ 越接近 1,越容易	绝对标度	X = 计数/计数 A = 计数 B = 计数	评审报告	6.5 确认	开发者

表 8.6.2 (续)

度量名称	度量目的	所采用的方式	测量、公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
安装的工 作量	安装需要付出 多少工作量?	对已实现的自动安 装的步骤进行计 数,并与规定要实 现的安装步骤总数 进行比较。	$X=A/B$ A=经评审证实自动的安装 步骤数 B=要求的安装步骤总数 注解:规定的,如,视图/命令/ 手动操作的数目,以达到目标 操作。	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计 数 A = 计数 B = 计数	评审报告	6.5 确认	开发者
安装的灵 活性	安装时的灵活 性和可定制性 如何?	对已实现的按预定 可定制的安装操作 进行计数,并与有 定制能力要求的安 装操作总数进行比 较。	$X=A/B$ A=经评审证实已实现的按 规定可定制的安装操作的计 数 B=要求可定制安装的操作 总数 注解:可定制的,如,嵌套深 度、操作视图个数	$0 \leq X \leq 1$ 越接近 1,越灵 活	绝对标度	X = 计数/计 数 A = 计数 B = 计数	需求规格说明 评审报告	6.5 确认	开发者

表 8.6.3 共存性度量

度量名称	度量目的	所采用的方式	测量、公式及数据元素计算	测量值解释	度量标度类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
可行的共 存性	产品在对其他 软件不产生不 良影响的前提 下与其他软件 灵活地共享环 境的程度如 何?	对能按规定共存的 产品的实体进行计 数,并需要在产 品环境中共存的 实体总数相比较。	$X=A/B$ A=能按预定共存的产品实 体数 B=需要在产品环境中共存 的实体总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标度	X = 计数/计 数 A = 计数 B = 计数	要求 规定 评审报告 测试报告	6.4 验证 6.6 联合评审	需求方 开发者 维护者

表 8.6.4 易替换性度量

度量名称	度量目的	所采用的方式	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
数据的连续使用	在产品被替代后所有数据保持不变的总数有多少?	对产品被替代后规定继续使用的数据项进行计数,并与软件被替代后要求使用的旧数据总数比较。	$X = A/B$ A=经评价证实按规定在替代之后继续使用的数据个数 B=从原软件中要求使用的旧数据总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标准	X = 计数/计数 A=计数 B=计数	设计源代码 评审报告 测试报告	6.4 验证 6.6 联合评审	需求开发者 维护者
功能的内含性	有多少功能保持不变?	对被产生类似结果的新软件覆盖的功能数进行计数,并与旧软件的功能总数比较。	$X = A/B$ A=经评审证实被产生类似结果的新软件覆盖的功能数 B=旧软件的功能总数	$0 \leq X \leq 1$ 越接近 1,越好	绝对标准	X = 计数/计数 A=计数 B=计数	设计源代码 评审报告 测试报告	6.4 验证 6.6 联合评审	需求开发者 维护者

表 8.6.5 可移植性的依从性度量

度量名称	度量目的	所采用的方式	测量、公式及数据元素计算	测量值解释	度量标准类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
可移植性的依从性	遵循与产品的可移植性有关的法规、标准及约定的程度如何?	对已满足要求的依从性的项进行计数,并与规格说明中要求的依从性项数相比较。	$X = A/B$ A=在评价中证实已正确的与可移植性有关的项数 B=依从项总数	$0 \leq X \leq 1$ 越接近 1,依从性越好	绝对标准	X = 计数/计数 A=计数 B=计数	依从性的规格说明 相关的标准、约定和法规 设计 源代码 评审报告	6.4 验证 6.6 联合评审	需求开发者

附录 A
(资料性附录)
使用度量时的考虑

A.1 测度的解释

A.1.1 测试使用环境与运行使用环境之间的潜在差异

在策划使用度量或解释测度时,理解清楚软件所要的使用环境,以及测试使用环境与运行使用环境之间的潜在差异是很重要的。例如:在类似的软件系统中,“学会操作所需的时间”的测度,对技术熟练的操作者与非熟练的操作者常常不一样。下面给出一些潜在差异的例子:

a) 测试环境与运行环境之间的差异

测试环境与运行环境之间是否有明显差异?

下面是一些实例:

- 具有较高/相当/较低的运行计算机 CPU 性能的测试环境;
- 具有较高/相当/较低的运行网络和通信性能的测试环境;
- 具有较高/相当/较低的运行操作系统性能的测试环境;
- 具有较高/相当/较低的运行用户界面性能的测试环境。

b) 测试的执行与实际运作的执行之间的差异

测试的执行与用户环境中运行的执行之间是否有明显差异?

下面是一些实例:

- 测试环境中功能的覆盖率;
- 测试用例的抽样率;
- 实时事务的自动测试;
- 压力负载;
- 每周 7 h×24 h(不间断)运行;
- 用来测试异常和差错的合适数据;
- 周期性处理;
- 资源利用率;
- 中断级别;
- 生产强度;
- 干扰。

c) 观察用户的特征

测试时用户的特征与运行时用户的特征是否有明显差异?

如下列实例:

- 混合类型的用户;
- 用户的技能水平;
- 专业用户或一般用户;
- 受限用户组或公共用户。

A.1.2 影响结果有效性的问题

下列问题可能会影响所收集的数据的有效性。

a) 收集评价结果的规程

- 借助工具或设施自动收集/手工收集/问卷调查或面谈。

- b) 评价结果的来源
 - 开发者的自述报告/评审者的报告/评价者的报告。
- c) 结果数据的确认
 - 开发者自查/由独立评价者检查。

A.1.3 测量资源的平衡

在每个阶段使用的测度的平衡是否适合于评价的目的?

在为内部测量、外部测量和使用质量的测度应用适当的度量范围时,平衡所用的工作量是很重要的。

A.1.4 规格说明的正确性

软件规格说明与实际操作要求之间是否有明显差异?

在不同阶段评价软件产品时,所采取的测量就是与产品的规格说明进行比对。因此,通过确认和验证来确保用于评价的产品规格说明能反映出运行中真实和实际的需要是非常重要的。

A.2 度量的确认

A.2.1 度量的理想性质

为了从质量评价中获得有效结果,度量应具有下列性质。若某种度量不具备这些性质,那么,度量描述应说明对其有效性的约束,并尽可能解释如何处理这类情况。

- a) (度量的)可靠性:可靠性与随机误差有关。如果随机变量不影响度量的结果,则度量是没有随机差错的。
- b) (度量的)可重复性:由相同的评价者使用相同的评价规格说明(包括在相同的环境中)和相同的用户类型及环境,对相同产品重复进行的度量宜在适当的容差范围内得出相同的结果。这里所谓适当的容差应包括诸如疲劳、学习效应等因素。
- c) (度量的)可再现性:由不同的评价者使用相同的评价规格说明(包括在相同的环境中)和相同的用户类型及环境,对相同产品进行的度量宜在适当的容差范围内得出相同的结果。

注1:建议对测量结果的可变性进行统计分析。

- d) (度量的)可用性:度量应明确指出其约束使用条件(如特定因素存在的条件)。
- e) (度量的)指示性:度量标识软件应改进的部份或改进的项,并给出与期望值进行比较的测量结果的能力。

注2:与只检查所需的项目不同,对选定或建议的度量宜提供使用度量可用性的书面证据。

- f) (测度的)正确性:度量应具备下列性质:
 - 1) (测度的)客观性:量度的结果与其数据输入应是有据可查的,即不受评价者、测试用户的感觉或观点的影响(除非满意度或吸引力度量,因为用户的感觉与观点也是测量的对象)。
 - 2) (测度的)公正性:度量不应偏向任何特殊的结果。
 - 3) (测度的)充分精确性:精确性由度量的设计,特别是作为度量基础的材料的选择来确定。度量的用户将描述度量的精确性和灵敏性。
- g) (测度的)意义:测量应产生有关软件行为或质量特性的有意义的结果。度量也应具有成本效益:即成本越高的度量,提供的结果应越具价值。

A.2.2 度量有效性的证实

度量的用户应标识一些证实度量的有效性的方法,例如:

a) 相关性

质量特性值(对运行使用中的主要度量的测度)中的变化可以用度量值中的变化来解释,用线性系数的平方表示。

利用相关性度量,评价者不用直接测量就可以预测质量特性的值。

b) 跟踪

若度量值 M 和质量特性值 Q (对运行使用中的主要度量的测度) 直接相关, 给定一个产品或过程, 当值 $Q(T_1)$ 变为 $Q(T_2)$ 时, 度量值也以相同的趋势, 从 $M(T_1)$ 变为 $M(T_2)$ (若 Q 值增加, 则 M 值也增加)。

评价者不必直接测量而是通过使用那些具有跟踪能力的度量就可以检测质量特性随时间周期的变化。

c) 一致性

若质量特性值 (对运行使用中的主要度量的测度) Q_1, Q_2, \dots, Q_n 对应于产品或过程 $1, 2, \dots, n$, 当有关系 $Q_1 > Q_2 > \dots > Q_n$ 时, 则对应的度量值也有关系 $M_1 > M_2 > \dots > M_n$ 。

评价者可以使用具有一致性能力的度量来关注软件的异常部件和易出差错的部件。

d) 可预测性

若使用时间 T_1 的度量来预测时间 T_2 质量特性值 Q (对运行使用中的主要度量的测度), 则预测误差值应在预测允许的范围。预测误差 = (预测值 $Q(T_2)$ - 实际值 $Q(T_2)$) / 实际值 $Q(T_2)$ 。

评价者可以通过可预测性的度量预测质量特性今后的变化趋势。

e) 可判别性

度量能够判别出软件质量的高低。

评价者可以使用具有判别能力的度量对软件部件进行分类和对质量特性值定级。

A.3 使用度量进行估计(判断)与预测(展望)

用如下两种方法在早期阶段估计和预测软件产品的质量特性是最具价值的度量。

A.3.1 利用当前的数据预测质量特性

a) 利用回归分析来预测

当通过使用特性(属性)的当前值(数据)来预测同一特性(属性)的未来值(测度)时, 根据一个足够长的时间内观察到的一组数据进行回归分析是有用的。

例如在测试阶段(活动)中获得的 MTBF(平均失效间隔时间)的值可用于估计在运行阶段的 MTBF。

b) 利用相关性分析来预测

当用不同属性的当前测量值来预测特性(属性)的未来值(测度)时, 使用一个确认的表明相关性的函数进行相关性分析是有用的。

例如在编码阶段, 模块的复杂性可用来预测在维护过程中程序修改和测试所花费的时间与工作量。

A.3.2 根据当前的事实估计当前的质量特性

a) 利用相关性分析来估计

在估计不可直接测量的属性的当前值时, 若任何其他的测度与目标测度密切相关时, 相关性分析方法是有效的。

例如软件产品中遗留的故障数是不可测的, 但它可以用检测出的故障数及故障趋势进行估计。

对于不能直接测量的属性进行预测的那些度量应用下述解释来估计:

- 使用模型来预测属性;
- 使用公式来预测属性;
- 基于经验来预测属性;
- 使用合理判断来预测属性。

对于不能直接测量的属性进行预测的那些度量可以用下述解释来确认:

- 标识要预测的属性的测度;
- 标识要用来预测的度量;

- 进行基于确认的统计分析；
- 将结果归档；
- 定期地重复上述工作。

A.4 检测易发生质量问题的部件中的偏差或异常

下列质量控制工具可用来分析在软件产品部件中的偏差和异常情况：

- a) 流程图(软件的功能模块)
- b) 排列分析和排列图
- c) 直方图和散点图
- d) 运行图、相关图和层次图
- e) 鱼骨图
- f) 统计过程控制(软件功能模块)
- g) 检查单

上述工具可用于标识源于数据的质量问题,这些数据是通过应用度量来获得的。

A.5 显示测量结果

- a) 显示质量特性评价的结果

对于每个质量特性和子特性可用下列图示法显示质量评价的结果：

雷达图、条形图、数字化的直方图、多变量图、重要性能矩阵图等。

- b) 显示测度

可利用一些有用的图形表示,如排列图、趋势图、直方图、相关图等。

附录 B

(资料性附录)

使用质量的度量、外部度量和内部度量的用法(框架实例)

B.1 引言

本框架实例是一个高层描述,它描述的是关于如何在软件开发和实现过程中使用 GB/T 16260.1—2006 中的质量模型和相关的度量来获得满足用户要求的质量产品。本实例所示的概念可用不同的定制形式来实现,以适应个体、组织或者项目。本实例使用的 GB/T 8566—2001 中的生存周期过程可作为传统软件开发生存周期的基准,使用的 GB/T 18905.3—2002 中的质量评价过程的步骤可作为传统软件产品质量评价过程的基准。只要能理解基本概念,如果用户愿意,也可以把这些概念映射为其他的软件生存周期模型。

B.2 开发及质量过程的概述

为了测量可交付项(即使用质量、外部质量和内部质量)的质量,表 B.1 描述了一个示例模型,它把软件开发生存周期过程的活动(从活动 1 到活动 8)与其关键的可交付项及相关的引用模型联系起来。

第一行描述软件开发的生存周期过程的活动。(可为适应独特的要求来定制)。第二行描述可能作为测量类别(如使用质量、外部质量或内部质量)的一种实际的测度或者预测。第三行描述可以测量质量的关键可交付项,第四行描述在每个过程活动中可适用于每个可交付项的度量。

表 B.1 质量测量模型

	活动 1	活动 2	活动 3	活动 4	活动 5	活动 6	活动 7	活动 8
阶段	需求分析 (软件与系统)	体系结构设计 (软件与系统)	软件的详细设计	软件编码与测试	软件集成及软件的合格性测试	系统集成及系统合格性测试	软件的安装	软件的验收支持
模型的引用	所需的用户质量 所需的内部质量 所需的外部质量	预测的使用质量, 预测的外部质量, 测量的内部质量	预测的使用质量, 预测的外部质量, 测量的内部质量	预测的使用质量, 测量的外部质量, 预测的外部质量, 测量的内部质量	预测的使用质量, 测量的外部质量, 预测的外部质量, 测量的内部质量	预测的使用质量, 测量的外部质量, 测量的内部质量	预测的使用质量, 测量的外部质量, 测量的内部质量	测量的使用质量, 测量的外部质量, 测量的内部质量
活动的关键可交付项	用户的质量需求(规定的), 外部的质量需求(规定的), 内部的质量需求(规定的)	软件/系统体系结构设计	软件详细设计	软件代码,测试结果	软件产品,测试结果	集成的系统,测试结果	安装的系统	交付的软件产品
用于测量的度量	内部度量(外部度量可用于确认需求规格说明)	内部度量	内部度量	内部度量 外部度量	内部度量 外部度量	内部度量 外部度量	内部度量 外部度量	使用质量 度量 内部度量 外部度量

B.3 质量途径步骤

B.3.1 概述

开发周期中对质量的评价分为下列步骤。步骤 1 必须在需求分析活动中完成,步骤 2 到步骤 5 必须在上述定义的过程活动中重复进行。

B.3.2 步骤 1:质量需求的确定

对质量模型中定义的质量特性和子特性,用表 B.2 中的两个例子对每类测量(使用质量、外部和内部质量)确定用户要求的权重。根据分配的相对权重,允许评价者集中精力于最重要的子特性上。

表 B.2 用户要求的特性与权重 a)

使用质量		
	特性	权重(高/中/低)
	有效性	高
	生产率	高
	安全性	低
	满意度	中

表 B.2 用户要求的特性与权重 b)

外部与内部质量		
特性	子特性	权重(高/中/低)
功能性	适合性	高
	准确性	高
	互操作性	低
	安全保密性	低
	功能性的依从性	中
可靠性	成熟性(硬件、软件、数据)	低
	容错性	低
	易恢复性(数据、过程、技术)	高
	可靠性的依从性	高
易用性	易理解性	中
	易学性	低
	易操作性	高
	吸引力	中
	易用性的依从性	高
效率	时间特性	高
	资源利用性	高
	效率的依从性	高

表 B.2 b) (续)

外部与内部质量		
特性	子特性	权重(高/中/低)
维护性	易分析性	高
	易改变性	中
	稳定性	低
	易测试性	中
	维护性的依从性	高
可移植性	适应性	高
	易安装性	低
	共存性	高
	易替换性	中
	可移植性的依从性	高

注：权重可用高/中/低的方式表示，也可在1~9的范围内用顺序标度来表示（例如1~3=低、4~6=中、7~9=高）。

B.3.3 步骤2：评价的规格说明

每个开发过程活动都要实施本步骤。

质量模型中定义的质量子特性均标识要应用的度量 and 要求的级别，以便达到在第1步骤中设定的用户要求，并按表B.3的例子加以记录。

对内容阐述的基本输入及用法说明可在表B.1的例子中得到，其中解释了在开发周期的这一阶段中能测量什么。

注：在开发周期的特定活动中，表中的某些行可以是空的，因为在开发过程的早期，不可能测量所有子特性。

表 B.3 质量测量表 a)

使用质量测量类别				
	特性	度量	要求的级别	评估实际结果
	效率			
	生产率			
	安全性			
	满意度			

表 B.3 质量测量表 b)

外部质量测量类别				
特性	子特性	度量	要求的级别	评估实际结果
功能性	适合性			
	准确性			
	互操作性			
	安全保密性			
	功能性的依从性			

表 B.3 b) (续)

外部质量测量类别				
特性	子特性	度量	要求的级别	评估实际结果
可靠性	成熟性(硬件/软件/数据)			
	容错性			
	易恢复性(数据、过程、技术)			
	可靠性的依从性			
易用性	易理解性			
	易学性			
	易操作性			
	吸引力			
	易用性的依从性			
效率	时间特性			
	资源利用性			
	效率的依从性			
维护性	易分析性			
	易改变性			
	稳定性			
	易测试性			
	维护性的依从性			
可移植性	适应性			
	易安装性			
	共存性			
	易替换性			
	可移植性的依从性			

表 B.3 质量测量表 c)

内部质量测量类别				
特性	子特性	度量	要求的级别	评估实际结果
功能性	适合性			
	准确性			
	互操作性			
	安全保密性			
	功能性的依从性			
可靠性	成熟性(硬件/软件/数据)			
	容错性			
	易恢复性(数据、过程、技术)			
	可靠性的依从性			

表 B.3 c) (续)

内部质量测量类别				
特性	子特性	度量	要求的级别	评估实际结果
易用性	易理解性			
	易学性			
	易操作性			
	吸引力			
	易用性的依从性			
效率	时间特性			
	资源利用性			
	效率的依从性			
维护性	易分析性			
	易改变性			
	稳定性			
	易测试性			
	维护性的依从性			
可移植性	适应性			
	易安装性			
	共存性			
	易替换性			
	可移植性的依从性			

B.3.4 步骤3:评价的设计

每个开发过程活动都要实施本步骤。

制定一个包括可交付项的测量计划(类似于表 B.4 中的例子),这些交付项用作要实施的度量和测量过程的输入。

表 B.4 测量计划

子特性	要评价的可交付项	应用的内部度量	应用的外部度量	应用的使用质量度量
1. 适合性	1. 2. 3.	1. 2. 3.	1. 2. 3.	(不适用)
2. 满意度	1. 2. 3.	(不适用)	(不适用)	1. 2. 3.
3.				
4.				
5.				
6.				

B.3.5 步骤4:评价的执行

每个开发过程活动都要实施本步骤。

执行评价计划,填写表 B.3 实例中的每一列。GB/T 18905 系列标准可用作编制计划和执行测量过程的指南。

B.3.6 步骤 5:反馈给组织

每个开发过程活动都要实施本步骤。

一旦所有测量均已完成,要把结果映射到表 B.1 中并以报告的形式将结论写成文件。同时标识产品质量需要改进的特定区域以使其满足用户的需要。

附录 C (资料性附录)

度量标度类型和测度类型的详细解释

C.1 度量标度的类型

当度量的用户得到了测量结果并使用该测度进行计算和比较时,宜对每种测度标识下列度量标度类型。对某些测度,平均值、比率或差值可能没有意义。度量标度类型有:标称标度、顺序标度、间隔标度、比率标度和绝对标度。 $M' = F(M)$,这里 F 是一个容许函数。每个测量标度类型的描述包含容许函数的描述(若 M 是一个度量,则 $M' = F(M)$ 也是一个度量)。

a) 标称标度

$M' = F(M)$,这里 F 是一一对一的映射。

标称标度包括分类,例如软件的故障类型(数据、控制、其他)。只有用相同类型的频率计算时,平均值才有意义。只有用经过映射的每种类型的频率计算时,比率也才有意义。因此,平均值和比率可以用来代表相同类型频率的早期和后来情况之间或两个类似情况之间的差。否则,它们可以用来相互比较每种类型各自的频率。

注 1: 例如:城市交通线标识号、编译器出错消息标识号。

注 2: 含义说明:只是不同类别的数。

b) 顺序标度

$M' = F(M)$,这里 F 是一个单调递增的映射,即:若 $M(x) \geq M(y)$,则 $M'(x) \geq M'(y)$ 。

顺序标度包括排序,例如:软件的失效按严重程度排序(忽略不计的、轻微的、严重的、灾难性的)。只有用经过映射的相同顺序的频率计算时,平均值才有意义。只有用经过映射的每种顺序的频率计算时,比率也才有意义。因此,平均值和平均值可以用来代表相同顺序频率的早期和后来情况之间或两个类似情况之间的差。否则,它们可以用来相互比较每种顺序的频率。

注 1: 例如:学校的考试成绩:优、良、及格和不及格。

注 2: 含义说明:每个量取决于它们在顺序中的位置,如中值。

c) 间隔标度

$M' = aM + b$ ($a > 0$)

当两次测度之间的差值有经验意义时,间隔标度包括排序的等级标度。但间隔标度中两次测度的比率可能没有相同的经验意义。

注 1: 例如:温度(摄氏、华氏、开氏),实际计算时间与预测的时间的差。

注 2: 含义说明:算术平均值和任何依赖排序的值。

d) 比率标度

$M' = aM$ ($a > 0$)

当两次测度之间的差值及两次测度的比例有相同的经验意义时,比率标度包括排序的等级标度。平均值和比率有各自的含义,它们给出了值的实际含义。

注 1: 例如:长度、重量、时间、规模、计数。

注 2: 含义说明:几何平均、百分比。

e) 绝对标度

$M' = M$,它们只能按一种方式测量。

任何与测度有关的说明都是有意义的。例如,当测量的单位相同时,一个比率标度类型的测度除以另一个比率标度类型的测度,结果是一个绝对值。一个绝对标度类型的测量值事实上不带任何单位。

注 1: 例如:带注释的代码行数除以代码的总行数。

注 2: 含义描述:一切事情。

C.2 测度类型

C.2.0 概述

为了设计一个收集数据、正确解释其含义并且把测度规范化以便进行比较的过程,度量的用户宜标识并考虑度量所使用的测度类型。

C.2.1 规模测度类型

C.2.1.0 导引

按其定义中所声称的测度内容,本类型的测度代表软件的一种特殊规模。

注:软件可以有多种表示规模的方法(就像任何一个实体可以进行多维测量——质量、体积、表面积等等。)

用一种规模测度来使其他的测度规范化,可以根据规模单位给出可比值。下列描述的规模测度类型可用于软件质量的测量。

C.2.1.1 功能规模类型

功能规模是软件可能有的的一种规模类型(一维)的例子。任何一个软件实例可能会有多个功能规模,例如取决于:

- a) 测量软件规模的目的(它影响到在测量中包含的软件范围);
- b) 所用的特定功能规模测量方法(它将改变其单位和标度)。

GB/T 18491.1—2001 提供了概念定义和使用功能规模的测量方法(FSM 方法)的过程。

为了规范化地使用功能规模,必须确保采用相同的功能规模方法,基于同样的目的,还要确保要比较的不同软件已经过测量,因而具有可比较的范围。

尽管下列内容经常声称代表了功能规模,但不能保证它们等同于应用 FSM 方法所获得的功能规模,也不能保证它们依从于 GB/T 18491.1—2001 标准。不过,在软件开发中,如下的方法仍被广泛使用。

- 1) 电子表格数;
- 2) 屏幕数;
- 3) 要处理的文件或数据集合数;
- 4) 用户需求规格说明描述的逐条列举的功能需求数。

C.2.1.2 程序规模类型

本条中,术语“程序设计”代表当执行时导致一些动作的表达式,术语“语言”代表所用的表达式类型。

1. 源程序规模

应解释程序设计语言,它应提供如何处理诸如注释行这样的不可执行语句。经常使用下列测度:

非注释性源语句(NCSS)包括可执行语句和带有逻辑性源语句的数据声明语句。

注1:新程序规模

开发者可能使用新开发的程序规模来代表开发与维护工作产品的规模;

注2:变更的程序规模

开发者可能使用变更的程序规模来代表包含修改过的部件的软件规模;

注3:计算的程序规模

计算程序规模的公式的例子:新代码行 + 0.2 * 修改过的部件中的代码行(NASA Goddard)

可能有必要更详细地区分下列源代码语句的类型:

I. 语句的类型

逻辑性源语句(LSS):LSS 测量软件指令的数量。这些语句不考虑与行的关系,独立于表现它们的物理格式。

物理源语句(PSS):PSS 测量软件的源代码行数。

II. 语句的属性

可执行语句;

数据声明语句;

编译程序命令语句;

注释性源语句;

III. 源

修改的源语句;

增加的源语句;

删除的源语句;

- 新开发的源语句(= 增加的源语句 + 修改的源语句);
- 重用的源语句(= 原来的源语句 - 修改的源语句 - 删除的源语句)

2. 程序字数规模

可采用下列 Halstead 测度方法计算测量值:

程序的词汇数 = $n_1 + n_2$; 观察到的程序长度 = $N_1 + N_2$, 其中:

- n_1 : 程序源代码中被程序语言预留的不同操作符的字数;
- n_2 : 程序源代码中由编程人员定义的不同操作数的字数;
- N_1 : 程序源代码中不同操作符出现的次数;
- N_2 : 程序源代码中不同操作数出现的次数。

3. 模块数

本测量计算可独立执行的对象个数, 例如程序中的模块个数。

C.2.1.3 利用的资源规模测度类型

本测度类型标识要评价的软件在运行中所用的资源。例如:

- a) 存储器的数量, 例如在软件执行过程中, 临时和永久占用的磁盘或内存的数量;
- b) I/O 负载, 例如通信数据流量总数(对网络中的备份工具有意义);
- c) CPU 负载, 例如每秒钟 CPU 指令集占用的百分比(本测度类型对测量 CPU 的利用率或在并发/并行系统中软件的多线程运行时测量进程分配的效率时有意义);
- d) 文件与数据记录, 例如文件或记录的位长度;
- e) 文档, 例如文档的页数。

注意峰值(最大值)、最小值和平均值, 以及时间周期及观察的次数等数据可能很重要。

C.2.1.4 特定的操作规程步骤类型

本测度类型标识在人工界面的设计规格说明或用户手册中规定的规程的静态步骤。本测量值可能依测量所用的描述类型的不同而有所区别, 例如用户的操作规程可以用图形也可以用文字来表示。

C.2.2 时间测度类型

C.2.2.0 导引

时间测度类型度量的用户应记录时间周期、检查过多少站点及有多少用户参与了这一测量。有多种以时间为单位进行测量的方式, 例如:

a) 实时单位

这是物理时间单位, 如秒、分或小时。这种单位常用来描述实时软件的任务处理时间。

b) 计算机器时间单位

这是计算机处理器的时钟时间, 即 CPU 时间的秒、分或小时。

c) 正式的日程表上的时间单位

包括工作小时、日历(日、月或年)。

d) 部件的时间单位

在有多个站点时,部件时间单位标识各个站点,部件时间单位是每个站点单独时间的累计。这种单位通常用来描述部件的可靠性,如部件的失效率。

e) 系统时间单位

在有多个站点时,系统时间不标识单独的站点,而标识整个系统中所有运行的站点。这种单位常用来描述系统的可靠性,如系统的失效率。

C.2.2.1 系统运行时间类型

系统运行时间类型为测量软件的可用性提供了基础。主要用于评价可靠性。应确定软件是间断运行还是连续地运行。如果软件是间断运行的,应确保在软件运行期间对时间进行测量(这显然可以扩展到连续运行的情况)。

a) 经时时间

当软件在不变的情况下使用时,如系统每周运行时间长度相同。

b) 机器加电时间

用于实时的、嵌入的或操作系统软件,它在系统运行的全部时间内都得到充分使用。

c) 规格化的机器时间

类似于机器加电时间,但把多台机器上不同的加电时间数据汇集起来并用一个修正因子进行调整。

C.2.2.2 执行时间类型

执行时间类型是指为完成特定任务所需要执行软件的时间。应分析几种尝试的分布,应计算均值、方差和最大值。应检查在特定条件下,特别是在过载条件下的执行时间。执行时间类型主要用于评价效率。

C.2.2.3 用户时间类型

用户时间类型测量单个用户在使用软件完成任务时所花费的时间。例如:

a) 会话时间

会话开始和结束的时间。如一个家庭银行系统的用户提取钱的行为。对于交互程序来说,只研究交互的易用性问题,不研究空闲时期。

b) 任务时间

单个用户每次试图运行软件完成任务所花费的时间。应定义好测量的起点和终点。

c) 用户时间

从开始到某个时间点,用户使用软件所花费的时间(从开始时起,用户使用软件大约有多少小时的时间或天数)。

C.2.2.4 工作量类型

工作量类型是指与某特定项目任务有关的生产时间。

a) 个人工作量

开发者、维护者或操作者为完成特定任务进行工作所需要的生产时间。个人的工作量只是每天一定数量的生产小时数。

b) 任务工作量

任务工作量是指所有单个的项目人员(开发者、维护者、操作者、用户或其他)为完成特定任务进行工作的人员工作量的累计值。

C.2.2.5 事件的时间间隔类型

本测度类型是指在观察期间,一个事件与下一个事件之间的时间间隔。可用观察时段的频率代替本测度。本测度可以典型地用来描述相继发生的失效之间的时间。

C.2.3 计数测度类型

C.2.3.0 索引

若对软件产品的文档属性进行计数,则为静态计数类型。若对事件或人的动作进行计数,则为动态

计数类型。

C.2.3.1 检测的故障数类型

本测量对在评审、测试、纠正、运行或维护期间检测到的故障个数进行计数。按照故障所造成的影响,可为这些故障的严重程度进行分类。

C.2.3.2 程序结构的复杂度类型

本测量对程序结构的复杂度进行计数。例如不同路径的数目或 McCabe 圈复杂度。

C.2.3.3 检测不一致的个数类型

本测量对调查所准备的不一致项数进行计数。

a) 不符合的项数

例如:

- 与需求规格说明的规定项不相符;
- 与法律、法规或标准不相符;
- 与协议、数据格式、介质格式、字符编码不相符;

b) 用户期望的不能实现实例数

本测量对所列举的满意或不满意的项数进行计数,这些项描述用户合理的期望与软件产品性能间的差别。

本测量可用问卷的方式向测试者、客户、操作者或最终用户就发现的缺陷进行调查。例如:

- 功能是否可用;
- 功能是否有效地执行;
- 功能是否可用于用户特定的预期使用;
- 功能是否是预期的、需要的或不需要的;

C.2.3.4 变更数类型

本类型标识检测出的已经变更的软件配置项。如在源代码中发生变更的行数。

C.2.3.5 检测到失效数类型

本测量对在产品开发、测试、运作或维护过程中检测出的失效个数进行计数。根据这些失效造成的影响,可以按严重性的等级进行分类。

C.2.3.6 尝试(试验)次数类型

本测量对与故障造成的缺陷相关的尝试次数进行计数。例如在评审、测试和维护中的尝试次数。

C.2.3.7 人工操作过程中的点击类型

当用户与软件在操作中发生互动时,本测量对用户作为的动态步骤活动所产生的点击个数进行计数。本测量量化了人类工效的易用性及使用的工作量。因此,本测量可用于易用性测量。如执行任务时的点击次数,眼睛活动的次数等。

C.2.3.8 记分类型

本类型标识算术计算的记分或结果。记分可包括计数或按检查表进行或不进行加权计算。例如检查表的记分;问卷调查的记分;Delphi 方法等。

附录 D
(资料性附录)
术 语

D.1 定义

除非特别指出,全部的定义都引自 GB/T 18905.1—2002 和 GB/T 16260.1—2006。

D.1.1 质量 quality

外部质量 external quality

产品在特定条件下使用时,满足明确或隐含要求的程度。

内部质量 internal quality

产品属性的总和,决定了产品在特定条件下使用时,满足明确和隐含要求的能力。

注 1: 当术语“特性”在本部分中用于更特定的意思时,使用术语“属性”(而不是 3.1.3 中使用的术语“特性”)。

质量 quality

实体特性的总和,表示实体满足明确或隐含要求的能力。

注 2: 在某种契约的环境或在某个受控的环境中,如核安全领域,要求是明确规定的,而在其他环境中,宜确定和定义隐含的要求。

使用质量 quality in use

特定用户使用产品满足其要求的程度,以达到在特定应用环境中的有效性、生产率和满意度等特定目标。

注 3: 使用质量是在包含软件的环境中质量的用户观点,它可以用在环境中使用软件的结果来测量,而不是根据软件本身的性质来测量。

注 4: 在 GB/T 18905.1—2002 中使用质量的定义目前不包括新的“安全性”特性。

质量模型 quality model

一组特性及特性之间的关系,它提供规定质量需求和评价质量的基础。

D.1.2 软件与用户 software and user

软件 software

信息处理系统的部分或全部程序、过程、规则及相关的文档。[GB/T 5271.1—1993]。

注 1: 软件是独立于所记录媒体的智力创作。

软件产品 software product

一组计算机程序、规程以及可能有的相关文档和数据。[GB/T 8566—2001]

注 2: 产品包括中间产品及专为开发者或维护者这样的用户所准备的产品。

用户 user

使用软件产品执行特定功能的个人。

注 3: 用户可以包括操作者、软件结果的接受者或软件的开发者或维护者。

D.1.3 测量 measurement

属性 attribute

实体的可以测量的物理或理论上的性质。

直接测度 direct measure

不依赖于任何其他属性测度的一种对属性的测度。

外部测度 external measure

从对系统行为的测度导出的对产品的一种间接测度,其中产品是系统的一部分。

注1:系统包括任何相关的硬件、软件(定制的软件或现货软件)和用户。

注2:在测试中发现的故障数量是对程序中的故障数量的外部测度,因为故障的数量是在计算机系统运行程序的过程中计数,以便标识代码中故障的数量。

注3:外部测度可以用来评价更接近于最终设计目标的质量属性。

指标 indicator

能用来估计或预测另一测度的一种测度。

注4:预测的测度可以是针对相同或不同的软件质量特性。

注5:指标可用来估计软件质量的属性和开发过程的属性,它们是对属性的不精确的间接测度。

间接测度 indirect measure

从一个或一个以上的其他属性的测度导出的一种对属性的测度。

注6:对计算机系统属性(例如对用户输入的响应时间)的外部测度就是对软件属性的一种间接测度,因这种测度要受计算环境的属性和软件属性的影响。

内部测度 internal measure

对产品本身的一种测度,或是直接的或是间接的。

注7:代码行数、复杂度、在走查和Fog索引中发现的故障数都是对产品本身进行的内部测度。

测度(名词) measure(noun)

通过进行一次测量赋予实体属性的数或类别。

测量(动词) measure(verb)

进行一次测量。

测量 measurement

使用一种度量,把标度值(可以是数或类别)赋予实体的某个属性。

注8:“类别”可用于指示属性的定性测度。如软件产品的一些重要属性,例如源程序语言(Ada,C,COBOL等)就是定性的类别。

度量 metric

定义的测量方法和测量标度。

注9:度量可以是内部的或外部的。

度量包括对定性数据进行分类的方法。

附 录 E
(资料性附录)
单纯的内部度量

E.1 单纯的内部度量

单纯的内部度量用来测量软件设计和软件产品的代码的某些属性,它们将影响软件的一些或全部特性和子特性。

表 E. 1.1 单纯的内部度量

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标准度类型	测度类型	测量输入	在 GB/T 8566 中的应用	目标用户
可追踪性	度量从需求到实现映射功能的过程中软件产品的文档、设计结构和代码的有效性。		$X = A/B$ A=在评审中已证实的可追踪的项数 B=被检测的项数	$0.0 \leq X \leq 1.0$ 越接近 1.0, 越好	绝对标度	X = 计数/计数 A = 计数 B = 计数		中的应用	目标用户
图复杂度	测量软件设计和编码结构的复杂性级别。		$e - n + 2p$ e : 边数 n : 结点数 p : 不连通部分的数目						
信息流复杂度	测量设计控制结构的复杂度。(参照 IEEE 982.1)		IFC(信息流复杂度) = (输入 × 输出) ²						
自描述性									
模块性	测量更新和概括基于程序功能/数据、执行序列和控制流层次的功能性知识的容易程度。		$X1 = A1/B1$ 其中 A1 = 功能上互相关联的模块数 B1 = 模块数 $X2 = A2/B2$ 其中 A2 = 数据结构上互相关联的模块数 B2 = 模块数						
自包含性									
程序规模	测量程序的规模。		$(N_1 + N_2) \log_2 (n_1 + n_2)$ N_1 : 操作符出现次数 N_2 : 操作数出现次数 n_1 : 操作符总数 n_2 : 操作数总数						

表 E.1.1.1 (续)

度量名称	度量目的	应用的方法	测量公式及数据元素计算	测量值解释	度量标准类型	度量类型	测量输入	在 GB/T 8566 中的应用	目标用户
条件语句的复杂性级别。	测量编码模块的复杂性级别。		$X=A$ $A=$ 条件语句数	$0.0 \leq X$		$X=$ 大小 $A=$ 大小			
统一数据基准	测量数据统一性。		$X=A/B$ $A=$ 评审中已证实的具有统一名字的数据基准数 $B=$ 数据基准的总数	$0.0 \leq X \leq 1.0$ 越接近 1, 越好	绝对标准	$X=$ 计数/计数 $A=$ 计数 $B=$ 计数			
变量的适当性	测量变量的适当性。		$X=A/B$ $A=$ 评审中已证实的具有适当名字变量的变量数 $B=$ 变量的总数	$0.0 \leq X \leq 1.0$ 越接近 1, 越好	绝对标准	$X=$ 计数/计数 $A=$ 计数 $B=$ 计数			
数据耦合的模块率	测量数据耦合的模块率。		$X=A/B$ $A=$ 评审中已证实的数据耦合的模块数 $B=$ 模块总数	$0.0 \leq X \leq 1.0$ 越接近 1, 越好	绝对标准	$X=$ 计数/计数 $A=$ 计数 $B=$ 计数			
程序语句	测量程序源语句。		$X=A$ $A=$ 程序语句的总数	$0.0 \leq X$		$X=$ 大小 $A=$ 大小			
平均模块大小	测量平均模块规模。		$X=A/B$ $A=$ 所有模块中源语句的总行数 $B=$ 模块总数	$0.0 \leq X$	绝对标准	$X=$ 大小 $A=$ 大小			
功能耦合的模块率	测量功能耦合的模块率。		$X=A/B$ $A=$ 评审中已证实的功能耦合的模块数 $B=$ 模块总数	$0.0 \leq X \leq 1.0$ 越接近 1, 越好	绝对标准	$X=$ 计数/计数 $A=$ 计数 $B=$ 计数			