

中华人民共和国国家标准

GB/T 18905.4—2002/ISO/IEC 14598-4:1999

软件工程 产品评价 第4部分：需方用的过程

Software engineering—Product evaluation—
Part 4: Process for acquirers

(ISO/IEC 14598-4:1999, Information technology—
Software product evaluation—
Part 4: Process for acquirers, 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

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 4 部分，等同采用 ISO/IEC 14598-4:1999《软件工程 产品评价 第 4 部分：需方用的过程》(英文版)。

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：北京信息工程学院、中国电子技术标准化研究所。

本部分主要起草人：王凌、冯惠、罗锋盈、陈莹。

引 言

软件已经变得日益普及。由于更多的过程能自动利用计算机的能力来进行,因而增加软件的功能、完善软件产品的要求也在不断增长。今天现代化系统如此复杂,没有软件就无法实现其功能。随着可利用的软件产品种类的不断增长,加速了现货软件产品的商业应用,同时软件工程的迅速发展也减少了对定制软件的依赖。面向对象的开发方法是基于对现存的自包含单元程序库的扩展来进行应用系统的开发,也减少了对用户定制软件的需求。这就使重点转向了软件产品质量或自包含软件单元的质量上。

定制软件的开发倾向于重新编写程序,其结果不能满足用户的需求。使用定制软件会在调配、实施、培训和维护支持活动方面比预期要付出更多的人力和物力。获取商业现货软件产品或者复用内部现有的软件产品也不是没有风险。因为现货软件产品可能也需要客户化(即按用户要求进行修改);测试和分析的需求量可能很大;当产品过时或需要修正时,产品维护和支持会有问题;将软件产品集成到大型系统中可能比较困难;产品的质量可能与目标系统的质量需求不一致等等,因此问题可能会很多。

商业现货软件产品的种类繁多。包括:

- a) 独立使用的产品(即工资单、会计软件,消费软件,或压缩软件[即字处理软件,报表软件]);
- b) 作为部件集成到由其他软件和硬件部件构成的大型系统中(即操作系统,关系数据库管理系统,图形用户界面[GUI]);
- c) 嵌入到硬件中(即通信数据链接,可编程阵列逻辑[PAL]);
- d) 作为用于开发特定应用的可配置的软/硬件系统的嵌入部分(即分布式控制系统);
- e) 用作支持软件开发和维护过程的CASE工具(即编译程序,配置管理工具)。

独立使用的软件产品中的差错能影响生产率,造成资金流失,或导致不必要的重复工作。软件部件会难于集成,会影响整个系统的可靠性,或不符合系统的目标。CASE工具可能会把错误引入到开发的产品中或是难于使用。

因此关键是在获取软件产品或决定复用一个现存的软件产品或部件时,能评价软件产品的质量。评价可用来接受或拒绝一个单独的产品,或者从众多可选产品中选择一个满足目标应用所确立的质量需求的产品。评价过程的严格级别必须与产品完整性需求相对应。对用于关键性业务的软件产品进行评价时,其严格程度应是最高。

软件工程 产品评价

第4部分:需方用的过程

1 范围

GB/T 18905 的本部分包含了在获取现货软件产品、定制的软件产品或修改现有的软件产品时,对软件产品质量进行系统地测量、评估和评价的需求、建议和指南。它使用了 ISO/IEC 9126-1 中描述的软件质量模型,扩展了 GB/T 18905.1 中定义的软件质量评价的通用过程,以及使用了 GB/T 8566 中定义的获取过程。它能与 GB/T 17544、GB/T 18905.2、GB/T 18905.3 和 GB/T 18905.6 一起使用。本部分与 GB/T 18905.5 中的评价过程的步骤相类似,但两者使用的环境完全不同。对于委托第三方或第三方进行评价的情况,需采用 GB/T 18905.5。对于需要按照包的质量需求对软件包进行第三方测试的情况,可以应用 GB/T 17544。

本部分所描述的评价过程还有助于满足决定是接受一种单一的产品,还是从备选产品中选择一项产品的目标。评价过程可以根据应用的性质和完整性级别加以剪裁。这样也能以节约成本的方式充分地适应各种形式和用途的软件产品。

本部分旨在为计划获取软件产品的项目管理者、系统工程师、开发及维护软件的工程人员和最终用户,以及提供这类产品的供方所用,但不只限于他们。

在本部分中,评价过程的目标软件产品可以作为部件集成到较大的系统中,也可以单独使用。这些软件产品分为下面几类:

- a) 商业现货软件产品;
- b) 为其他应用或范围更广的普通应用而开发或获取的现有软件产品;
- c) 定制的软件产品或对现有软件产品的修改。

本部分中定义的评价过程也适用于 CASE 工具。由于在 GB/T 18234 中专门对 CASE 工具的评价进行了阐述,所以本部分不讨论 CASE 工具。

本部分旨在与其他标准一起使用。对具有较高完整性需求的系统,可以在本部分所描述的评价过程中包含附加需求,这些标准源自部门专用的标准,例如:IEC 880、DOA-167A、MOD-55 等等。

2 一致性

由于建议的通用性给用户提供了选择的自由,所以依从于本部分的简单声明是无效的。任何强制使用本部分作为贸易条件的组织负责规定和公开满足 6.1.1 中规定的强制性目标要求。这个规定的评价过程为本部分的给定应用制定依从性条款。第 6 章和第 7 章的全部活动都应视为具有适用性。对评价过程的需求也能在执行获取过程期间以合同的形式建立,因而,建立依从本部分所描述的评价过程也比较容易。

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 18905.4—2002/ISO/IEC 14598-4:1999

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

GB/T 18492—2001 信息技术 系统和软件完整性级别 (idt ISO/IEC 15026:1998)

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 18905.5—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

本部分采用下列定义。为便于引用,本部分所用到的其他标准的关键定义在附录 A 中重新给出。

4.1

商业现货软件 commercial-off-shelf software (COTS)

根据市场驱动的需要定义的、通过商业方式提供的、其适用性已经得到范围广泛的商业用户证实的软件。

注:可参见 IEEE Std 1062—1993 中的定义。

4.2

定制软件 custom software

根据用户的需求规格说明,为某个特定应用开发的软件。

4.3

现有软件 existing software

已经开发并且可以得到的软件;可以“原样”使用,也可以经修改后使用;由供方、需方或第三方提供。

注:亦可参见 GB/T 8566 对现货产品的定义。

5 软件产品评价的一般考虑

5.1 评价过程和获取过程间的相互关系

下面概括的获取过程的活动(在 GB/T 8566 中定义)与第 6 章和第 7 章中的一般评价过程的活动(在 GB/T 18905.1 中定义)相互结合起来。第 6 章的重点在获取 COTS 产品的过程中评价最终产品质量的应用方面。第 7 章的重点是获取定制的软件或修改现有软件过程中评价过程的应用。

- a) 启动——确定要获取的产品的软件需求、获取计划、验收策略和验收准则;
- b) 招标(标书)的准备——获取需求的规格说明和编制文档;
- c) 合同的准备和更新——选择供方、合同的准备和谈判,以及合同变化的控制;
- d) 监督供方——在合同执行期间实施评价活动,以期达到软件产品的验收和交付使用;
- e) 验收和完成——在验收和交付最终软件产品的过程中实施的活动。

注意:GB/T 18905.1 定义的一般评价过程未作为 GB/T 8566 中的一个过程来定义,但作为等同于策划—实施—检查—改进(PDCA)循环中的“检查”部分的一项基本功能,它在每一个生存周期过程中都将被执行。不过,一般评价过程也可以在任何 GB/T 8566 的过程(即开发、维护、获取、确认)中执行,因此它与 GB/T 8566 中所用的“过程”含义的抽象程度不同。

在实施一般的评价过程时,了解这种区别是很重要的。获取者需要定义在评价期间要遵循的以达到评价需求的评价和获取过程。在大型系统的开发环境中,要遵循的获取和评价活动需要与其他开发和集成活动结合起来,并在如 GB/T 18905.2 规定的项目测量计划中加以确定;也就是说,为评价所考虑的具体获取实施的内容包括以下几方面:

- 评价所要求的软件需求规格说明能构成用于招标(标书)所要求的获取需求的基础;
- 需要独立的初步评价活动以便预选软件产品和供方;

- 用于评价的供方及产品信息需求需要在获取需求中或在合同准备期间予以规定；
- 评价活动可在监督合同的执行期间作为部分要求的评价工作、部分产品的开发工作、或部分产品的正式验收工作来执行，或者在产品交付后执行。

5.2 评价过程的输入

5.2.1 系统需求

对目标软件确定评价需求开始于对整个系统的需求。系统需求确定了用户、用户目标、任务及特性，包括产品的使用环境，以及产品或系统的功能需求和其他需求。它们构成了随后的系统结构设计、软件需求规格说明和软件结构设计的基础。由于相关的法律和法规对获取过程和评价过程的严谨性和形式化程度有影响，因此，在此阶段还需要说明法律和法规的需求。

在对系统需求进行分解和设计期间，这些需求被分配为硬件配置项和软件配置项，还分配为包含系统规程的用户操作。系统开发生存周期内的设计活动会产生后来的获取或复用“现货”软件产品的决定。由于某些评价活动在形成决定的过程中起作用，因而，实际上，它们是这些设计活动的一部分。对要获取的软件产品的评价是单独进行的，在对最终产品进行系统集成和测试期间，软件配置项要与其他软件和硬件配置项(参照 GB/T 8566)集成起来。图 1 示出了用于评价和获取的大型系统工程环境。

本部分使用和获取的候选产品是能作为部件与大型系统进行集成，或是能独立使用的软件产品。其分类如下：

- a) 商业现货软件产品；
- b) 为其他应用或范围更广的普通应用而开发或获取的现有软件产品；
- c) 定制的软件产品或对现有软件产品的修改版本。

对于要集成到大型系统的软件配置项，需要对每一配置项定义软件需求。而在其它情况下，系统与软件配置项是一致的，并且可认为是等效的。

要获取的硬件配置项可能包含软件，如，驻存在固件中的操作系统(即 ROM、PROM)。当现有软件以这种方式构成硬件的一个完整部分时，通常需要与硬件配置项一起评价。

5.2.2 完整性级别需求

如果在可以接受的限制内，在包含系统的安全性、保密安全性、金融风险、环境风险以及社会风险等很重要的方面，软件是关键性的，就必须在获取和评价之前建立软件所要求的完整性级别，并形成文件。GB/T 18492—2001 中给出了完整性级别确定过程的指南。所确定的完整性级别决定了如何在评价过程中处置软件。

5.2.3 软件需求规格说明

应使用一个适当的定义良好的质量模型来规定软件需求。为此，除非有特殊原因需要使用另一个模型外，宜使用 ISO/IEC 9126-1 中的质量模型和定义。该模型定义了应用中软件的六种特性：功能性、可靠性、易用性、效率、可维护性和可移植性。这些特性能进一步分解为具有可测量或可评估属性的子特性。

宜按照直接与用户需要相关的外部度量(外部度量在 ISO/IEC 9126-2 中定义)来定义需求，并编制成需求规格说明文档。将用户的需要编成文档包括制定非正式的功能需求和性能需求的清单，以便为产品(或系统，如果产品是嵌入式的)准备一个完整的产品需求规格说明。那么，需求规格说明可以构成在获取过程的招标步骤中所用的获取需求的基础，也可以构成随后执行的产品评价的基础。

5.2.4 由其他方实施的评价

只要评价结果是值得信赖的，那么，通过使用第二方或第三方实施的评价活动的结果可以缩小当前评价过程的范围。这种评价活动可以由预先存在的认证活动、产品评价活动和(或)过程评估来组成。例如：

- 对用于产品开发的软件工程过程可实施标准化，以满足 GB/T 8566、GB/T 19000.3 或其他具体行业标准的需求；

- 软件开发所遵循的供方的质量体系可由第三方根据 GB/T 19001 的要求进行认证；
- 可由第二方或第三方对照 GB/T 18905.5 或 GB/T 17544 中的要求对软件产品进行评价；
- 开发可接受产品的供方的软件过程能力可由第三方根据 ISO 15504-8(在制定中)予以评估；
- 软件可作为大型系统开发阶段的一部分进行功能性评价；
- 软件产品可能原先已经被具有不同完整性需求的另一个应用评价过；
- 可能已经在组织内被其他方通过非正式或正式的评价活动实施过产品评价。

为获得和解释用于目标应用的外部评价结果，要求的附加成本和时间可能影响该方法的可行性。为了充分信任其它方的评价结果，仍有必要与评价者或供方进行协商。

注：对供方的软件工程过程、供方的质量体系或供方的单独能力等方面的评价结果还不足以证明软件产品包含了所需的质量特性，还需要实施其他产品评价方法(例如那些专门测量与最终用户需求对应的质量要素和质量属性的评价方法)。

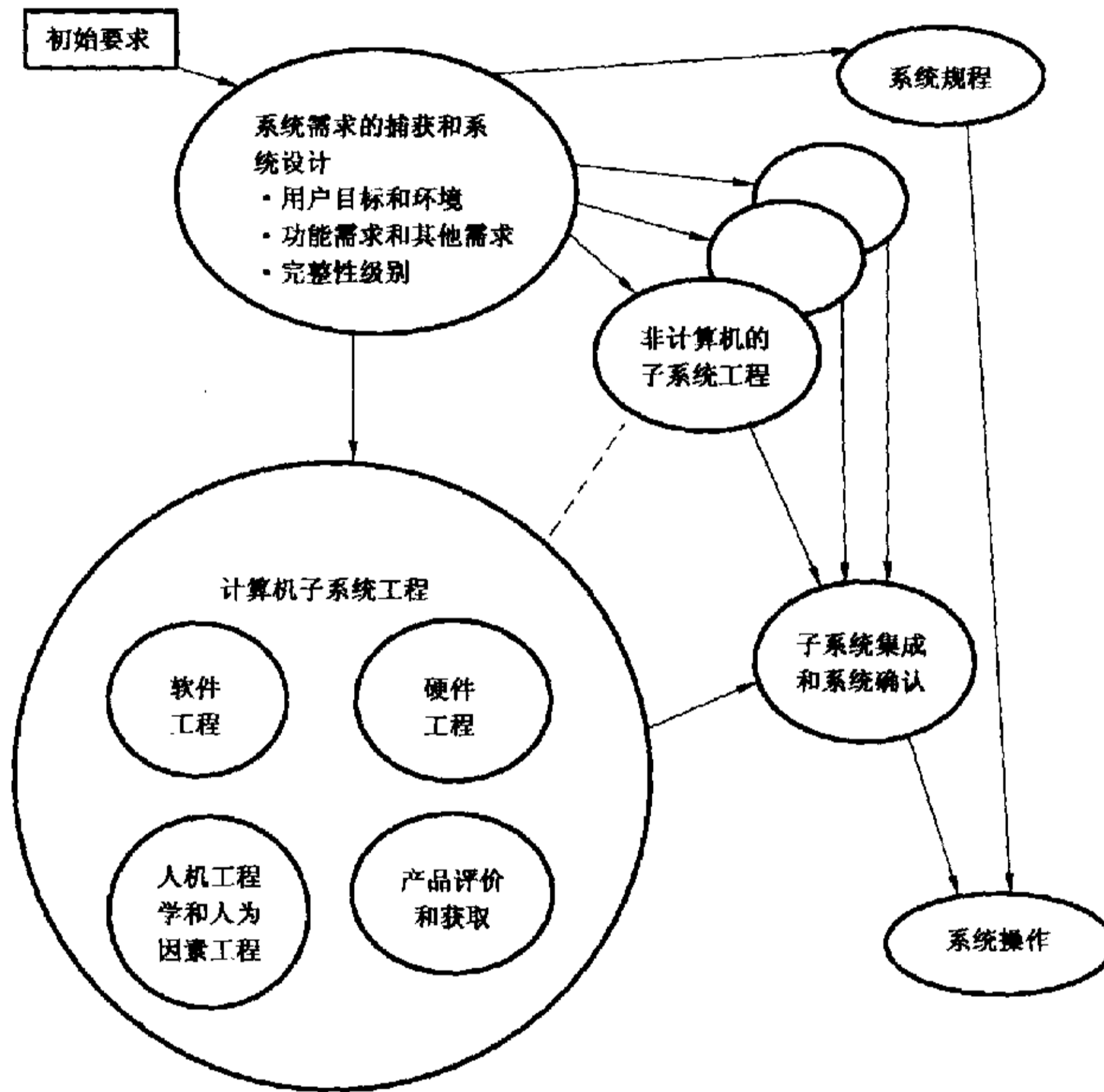


图 1 评价和获取软件产品的系统工程环境

5.3 剪裁

评价过程能用于范围广泛的获取需求、完整性需求和评价者目标。例如：

- 软件包的获取者可能希望只使用 GB/T 17544 来评价软件包；
- 软件产品的获取者可利用 GB/T 18905.5 来实施独立评价；
- 一个小的或单独的获取者可能要求带有最少文档的、不是很正式的评价过程；
- 对于消费类软件，评价过程的目标可以仅仅是从大量类似的产品中选择、测试和获取一种产品。正式的获取过程则简化到直接购买，而不需包括准备合同。

评价过程宜具有灵活性以适应每种应用的独特性，避免不必要的工作或无价值的工作，这是树立软件的必要信心的一个实际手段。软件所要求的完整性级别极大地确立了评价过程的严格性和正规性。

利用 GB/T 8566 中给出的剪裁指南和要获取的特定软件产品所要求的完整性级别，按照评价过程

来剪裁获取过程。对具有较高完整性需求的完整软件系统,通常会调用一整套获取活动和任务,包括 GB/T 8566 中规定的供应过程对应的活动和任务。通常,随着完整性级别的提高,获取过程的精确度、与获取过程相关的活动和任务的数目也随之增加。

附录 B 中的表 B.1 给出了一个按照目标软件完整性需求剪裁的集成获取过程和评价过程活动的例子。

6 获取现货软件产品期间的评价

GB/T 18905.1 中定义的通用软件产品的评价过程由四个主要步骤组成。在本部分对现货产品的获取过程中,这四个步骤特别体现和细化在强调对最终产品质量的评价上。然而,这并不排除针对特定的质量特性进行的对中间产品的评价。因此,这些步骤的具体实现细节与 GB/T 18905.1 中所描述的略有不同,但并非不一致。表 1 将按照评价过程的步骤和关键任务,以及输入和输出四个方面对评价过程加以概括。

表 1 获取现货产品期间的评价过程

输入	评价步骤	关键任务	输出
系统/软件需求	确立评价需求 (6.1)	规定目标、用途和范围。规定评价的精确度。确定评价的输入。确定要实施的评价,或由其他方实施的评价。确定要遵循的获取过程及如何将评价输入需求传达给供方。	评价需求 规格说明
评价需求规格 说明	规定评价 (6.2)	选择与软件产品特性相关的度量。确立评定等级。选择一组最有效的评价方法。确立对不同质量特性的评价结果以及有助于在特殊环境下对软件产品进行质量评估的其他方面加以概括的规程。	评价 规格说明
评价规格说明	设计评价 (6.3)	准备一份描述评价方法及评价进度的评价计划。确定评价活动和获取活动的结合点。	评价计划
评价计划	执行评价 (6.4)	执行选择的评价活动,分析和记录确定软件产品适合度的结果。分析已发现的缺陷的影响和调整产品用途的选项。对产品是否可接受以及最终决定是否购买做出结论。	评价记录 和结果

6.1 第 1 步 确立评价需求

6.1.1 确立评价的目的和范围

评价过程应:

- 按照能否客观评价软件产品的适用性,利用 ISO/IEC 9126-1 中的质量模型和定义来确立一组软件质量需求;
- 对软件质量特性确立适当的优先级;
- 确立适用的完整性级别的评价体系基础,包括确立评价活动中的精确度级别或细节需求,以及评价过程的输入和输出;

注:图 2 给出了一个软件产品评价过程的概貌。它给出了从获取者的角度看到的,评价过程的输入和输出的不同看法。

- 确立要遵循的获取过程,以及如何将评价输入需求传达给供方;

注:见附录 C 图 C.1 中组合的评价和获取过程的实例。

- 考虑下列内容后确立评价范围、用途和目标:

- 软件产品是否用于某个特定的应用、一组特定的应用或某一通用范围内的应用;
- 第二方或第三方是否曾做过评价,或是否计划将来实施某种评价活动;

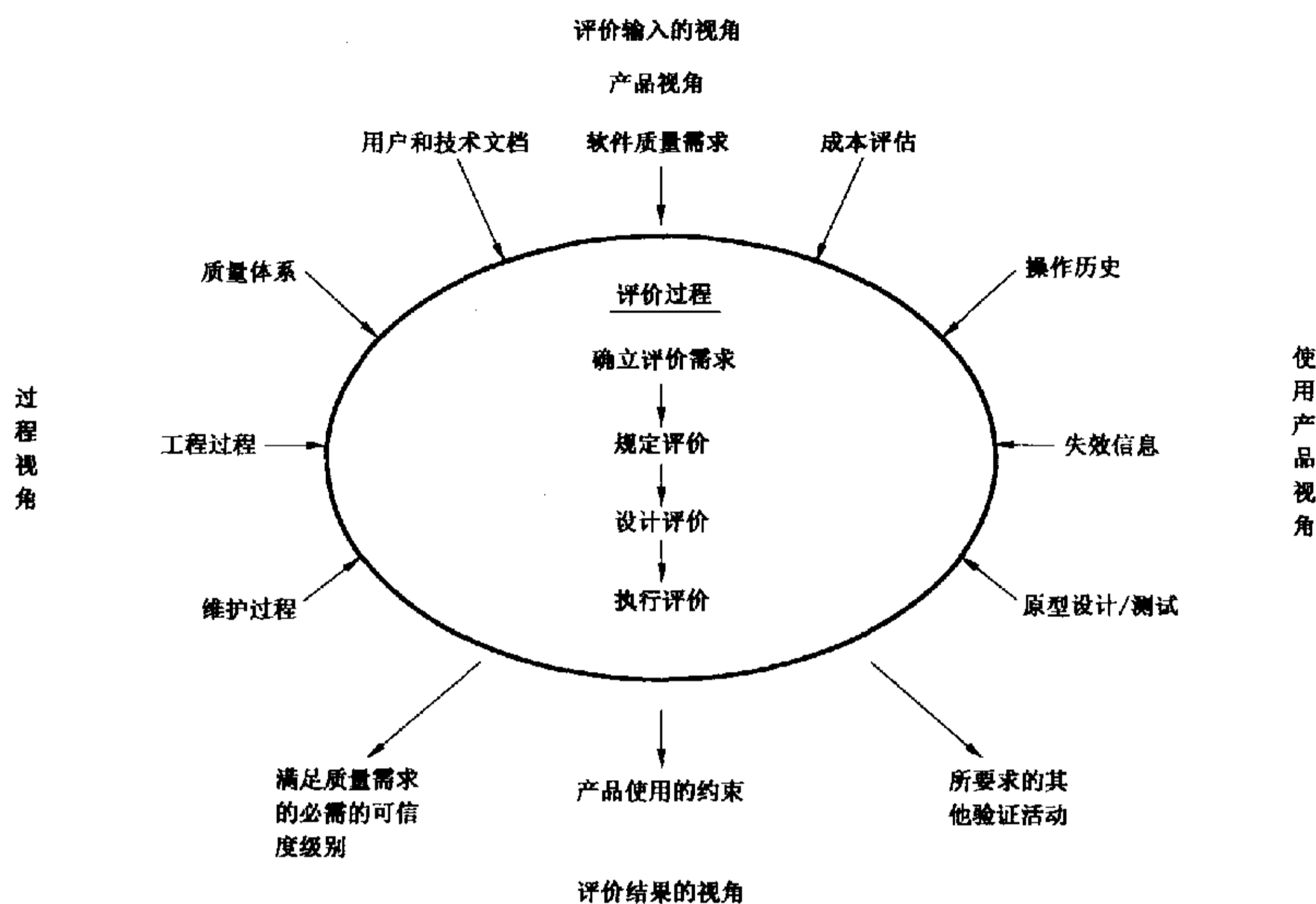


图 2 从获取者的视角看到的软件产品评价过程的概貌

6.1.2 规定评价需求

评价的需求规格说明宜：

- a) 确定用户及其目标、任务、特性及产品的用户环境；
- b) 确定软件应用的完整性级别(软件出错引起的风险),以及评价过程所需的精确度级别；
- c) 确定规章制度方面的需求[需要给立法方(或任何需要的各方)提供产品质量保证方面的文档](见 ISO 9126-1 中的一致性)；
- d) 确定产品的边界和接口,包括对软件产品的接口需求(即指通过接口传递的数据类型、数据格式、接口访问机制、失效/差错处理、计时问题、接口行为问题和接口状态的相关性和转换)(见 ISO 9126-1 中的互操作性)；
- e) 如果产品是要求与其他部件或产品进行集成的较大系统的一部分,要确定集成需求；
- f) 确定软件的质量需求,包括：
 - 1) 强制性需求和可选的需求之间的区别；
 - 2) 解释或理解需求所需的假设、例外、限制、排除情况或未解决的问题；
 - 3) 所有重要的质量特性以及优先级方面的用户需求(即,如果认为可维护性重要,就规定特定的可维护性需求)；
 - 4) 所有设计约束和环境约束;即,软件产品的使用造成的功能和性能限制,以及软件产品与用户应用中其他现有软件、定制软件和硬件集成的级别和复杂性引起的功能和性能限制；
 - 5) 所有项目管理约束;即,实施评价活动的资源和专业知识和、进度安排和预算限额、可能有的依赖性 or 风险、关键性的假设,或有关评价工作本身的一些假设；
 - 6) 使用不同于 ISO/IEC9126-1 中定义的质量模型的理由；
- g) 确定要评估的供方服务;即支持能力、应用开发能力和培训能力；
- h) 确定要评估的特定需求;即特定的技术可行性问题或设计实现问题；
- i) 确定彼此一致(即无冲突需求)的、并与应用的完整性级别一致的评价需求；
- j) 确定在未来应用中是否要复用产品,文档是否必须支持任何未来的产品评价；

- k) 确定获取过程,以及招标期间供方应提供的信息;
- l) 确定能减少产品评价工作的由第二方或第三方实施的评价。

注:评价的需求规格说明的详细和完整程度直接影响着评价的完整性级别,即,尽管评价的结果可能被用来指导其它阶段的下一步工作,预见问题,或者排除某些软件产品或供方,但仅仅基于初步需求的评价不能被认为是完整的评价。作为适当的设计或计划活动的一部分,这些工作通常会在主要的评价活动之前进行。在最终确定需求之前可能也需要一些评价工作。

6.2 第2步 规定评价

评价规格说明宜编成文档,从而使具有相当资格的人员能重复评价过程并得到相应的结果。

6.2.1 选择度量

评价规格说明宜确定:

- a) 要评价的产品的特性;
- b) 使用软件时,与质量的可测量方面相关的外部度量(附录B的表B.2表示一个外部度量的例子和可能的验收准则。注意:由于没有方便有效的规则可用,因此,可由用户根据经验来规定实际的验收数量阈值);
- c) “使用质量”的度量与用户对含有软件的系统的质量看法有关(见附录B表B.3的使用度量的示例);
- d) 对可接受范围的度量准则的有效描述(例如,必须有多少操作历史为给定的质量特性和给定的完整性级别提供一个合理的保证程度(见附录D.4和D.5中操作历史的具体细节));
- e) 任何打包的评价模块;
- f) 与评价需求相关的覆盖级别,在对原先由其他方完成的评价进行评审之后该评价需求是必须的;
- g) 评价要回答的检查表;
- h) 有助于回答问题的示例清单;
- i) 要使用的测试用例;
- j) 要收集和分析的数据及格式;
- k) 要采用的特定评价方法,包括评审或评估下列内容的一项或多项:
 - 1) 软件产品的用户和技术文档(包括在线文档)(见附录D.1);
 - 2) 基于供方的课程和培训的软件产品评价(见附录D.2);
 - 3) 软件工程过程,包括中间软件产品(见附录D.3);
 - 4) 供方的产品操作历史(见附录D.4);
 - 5) 用户的产品操作历史(见附录D.5);
 - 6) 供方的能力、支持和质量体系(见附录D.6);
 - 7) 原型或其他评价方法(见附录D.7);
 - 8) 产品缺陷清单和相关信息(通常在万维网上找到);
- l) 评估这些评价结果的方法;
- m) 对评估结果进行评级的合适的方法,以便从相似的产品中选择一项产品;
- n) 用于比较多个软件产品的评级方案。评级方案可以按质量特性的优先级来设定权重。

6.2.2 选择评价方法

宜规定评价方法的某种组合,以允许选择产品或确立产品的适用性。要评价的领域包括:

- a) 某些考虑的事项是否可能相互冲突(例如,“所选择的[评估]方法的成本是否在预算之内?”可能和“收集的方法是否表达了所有评价需求?”不相兼容)。在这种情况下,由评价者根据评价需求的优先顺序进行必要的权衡;

注:附录B中的表B.4给出了从成本和有效性两方面对软件质量特性的评价方法进行评级的例子。

- b) 评价是否在组合所选的方法方面提供足够的覆盖区域或范围,这要考虑如下问题:
 - 1) 怎样证明软件满足了它的规格说明;
 - 2) 各方法覆盖范围的重叠,以提供额外的信任度;
 - 3) 这组活动作为一个整体,是否提供了一种可接受的保证级别,保证它完全覆盖了人们所关心的那些软件质量特性;
 - 4) 这些方法的互补程度;
 - 5) 说明针对各种特性时,每种方法的有效性和客观性;
 - 6) 评价方法之间各种不同的途径(如,基于各种方法的评审、分析和测试);
 - 7) 因对应用所进行的评价活动取得信任,这些评价活动最终是整个系统开发周期的一部分活动;
 - 8) 相信由其他方完成的评价。
- c) 为了缩小功能上认为适合进一步评价的产品的选择范围,使用“非正式”的初步评价活动,如评审、调查或同行/用户体验、刊物对产品的评述、可查阅的产品用户文档、或产品评审的数据存储库等。

6.2.3 由其他方实施的评价

在相信由其他方实施的评价之前,宜考虑下列情况:

- a) 用精确度等级来说明评价需求的评价是否与应用的完整性级别一致?
- b) 评价报告是否指出了要评估的软件产品的版本、评价的范围、所用的判定准则和评价结论?
- c) 评价报告是否指出了软件产品或软件工程过程中的缺陷?是否提出了对那些缺陷的纠正措施和建议?以及这些纠正措施是否落实?
- d) 评价者是否有适当的专业知识,包括:
 - 1) 实施评价和分析的经验;
 - 2) 与国际公认的标准有关的软件质量方面的经验;
 - 3) 软件工程方面的专业知识;
 - 4) 与供方的完全独立性。

6.3 第3步 设计评价

产品评价计划宜确定:

- a) 供方或第三方是否愿意或能够提供对所需文档、设备、工具、软件、课程和(或)培训以及与此相关的成本等内容的访问;
- b) 与访问各种机密或专有信息相关的条件;
- c) 供方和第三方是否愿意或能够提供对具有相当专业知识的人员的访问,以便回答问题,与此相关的成本是什么,包括差旅成本;
- d) 根据评价需求,评价者开展评价所需的专业知识,以及获得这类特殊专业知识的成本;
- e) 为使产品适于全面测试所需要的预先测试;
- f) 为实施评价而提供测试环境(如测试硬件、支持设备和工具、专业人员)所需的成本;
- g) 评价的职责和所要求的进度安排;
- h) 在提供质量保证的评价方法中的任何限制和缺陷,以及这些限制和缺陷是否在计划中的其它地方已经提到;例如,评价方法无法囊括一个特定质量特性的所有子特性;
- i) 所用的各种评价方法之间的任何相互依赖性;即,建立评价方法的最佳顺序的顺序相关性(从一项测试中获取的信息可能在另一个测试中 useful);
- j) 必要的资源、总的评价成本以及每一种评价方法的成本;
- k) 评价活动和获取活动之间的连接点(见附录 C 中图 C.1 评价过程和获取过程的组合示例);
- l) 评价过程的决策点,它确定何时、因何原因认为评价是完成了(即接受或拒绝准则),并宜停止

评价；

- m) 下列是计划的各项评价活动：
 - 1) 宜遵循的规程和技术；
 - 2) 所要求的信息输入/输出和文档；
 - 3) 对生成的文档的任何格式和内容需求；
- n) 定义评价计划时做出任何例外或异常决定的背后所遵循的依据、理由及假设要形成文档；
- o) 评价工具；
- p) 开发和验证度量的规程，以及评价过程、度量和测量的标准化规程。

注1：GB/T 18905.6 定义了评价模块的概念，该评价模块系统地收集应用某一特定评价技术或方法对质量特性的指定方面进行评价所必需的信息。

注2：在安排评价方法的进度时，重要的是要认识到在各种不同的评价方法之间存在着高度的相互依赖性。也就是说，从一种方法获取的信息可能影响对另一种方法的理解。因为评价的本质是可重复的，在信息被获取时，可以再次考察这些问题。因此在评价时，评价计划可能会改变。举例来说，一旦进行评价，通常认为不必进行更详细程度的评价，或只作为一项附加需求。

注3：软件产品的评价可以在一个开发生存期中不同时间点分阶段进行，或在生存周期的某一时间点同时进行。不同的团体或个人可以负责评价的不同部分。当评价分阶段进行时，在每个阶段都重复评价活动的步骤，直到不需进行进一步的工作为止（见附录 E 中分阶段的评价过程的示例）。

6.4 第4步 执行评价

6.4.1 执行评价的方法

- a) 宜执行评价、编写评价文档及分析评价，以便：
 - 1) 建立能使该软件产品满足评价需求的适当的可信度等级；
 - 2) 确定各种关于评价需求的具体缺陷和任何确定这些缺陷范围所需的附加评价；
 - 3) 确定应用软件产品的各种特殊限制或条件；
 - 4) 确定评价自身的弱点或遗漏以及要求的附加评价；
 - 5) 确定任何未被评价覆盖的软件产品的应用选项；
- b) 评价的执行记录宜标识：
 - 1) 按照评价计划中规定的规程执行评价；
 - 2) 评价规程的执行步骤（包括所用的数据、安装过程和任何状态信息）、评价结果（对所有问题的解答和对答案来源的引用），以及软件产品的版本号；
 - 3) 某个评价活动的任何限制、约束、缺陷或例外，包括它们对过期的软件产品在使用、配置、修改或日常维护中的影响；
 - 4) 评价者和他们的资格；
 - 5) 要评估的产品版本和相应评价输入的不同，即文档编制或课程间的任何区别；
 - 6) 缺陷事件中的解决方案或“变通的工作”。

6.4.2 分析评价结果

分析评价活动的记录宜标识：

- a) 每个缺陷、任何相关分析以及每个缺陷是如何解决的。缺陷的解决情况可以包括：
 - 1) 任何一种其他评价方法已保证该缺陷不是主要缺陷；例如，大量的操作历史可能弥补有缺陷的软件工程过程；
 - 2) 能找到满意的“变通办法”以减轻缺陷的影响；如，对产品的修改，禁用或删除不需要的功能，用逆向工程的方式重新生成遗漏的设计需求；
 - 3) 初始的需求不是强制性的，且缺陷是能够接受的；
 - 4) 只要软件产品的使用受到特定条件或限制的控制，则缺陷是可接受的；
 - 5) 需要附加的评价工作以解决评价缺陷或差距。

- b) 为解决已确定的缺陷而执行的附加评价,以便:
 - 1) 确定缺陷的范围或影响;
 - 2) 建立不存在缺陷的信任度;
 - 3) 验证变通办法在技术上是可行的和(或)是合适的和可接受的;
 - 4) 当为纠正缺陷已做出一项或多项设计变更后,要验证软件的性能是否正确和可接受;
- c) 在必须限制或控制使用软件产品的情况下,限制是否会:
 - 1) 与满足应用的强制性需求的软件产品相冲突;
 - 2) 影响应用的设计、预算和进度安排;
 - 3) 要求额外的评价工作;
 - 4) 导致在应用中有失效的可能;
- d) 评价范围的任何例外和(或)每种评价结果的约束,例如:
 - 1) “该评价不包括对产品功能的详细评审;”或
 - 2) “如果成功地完成了对软件产品所需功能的整体评价,那么,相信该产品对所需的完整性级别是合格的。”
- e) 所有评价活动的综合结果允许对软件产品的评价做出一个全面结论。

6.4.3 做出评价结论

结论宜针对应用的完整性级别和实际的评价需求,说明软件产品在实际应用中是适合的和适用的。如果因发现了一些缺陷或者因缺少评价信息而不能“原样”地使用软件产品,那么有必要建议执行进一步的评价或者控制或限制软件产品在其目标应用中的使用。

结论可以用“需求一致性声明”的格式,即针对每一项特定需求,说明软件产品的特征、功能或服务已满足各项需求,并说明所用的评价方法已为这种符合性提供了足够的可信度。潜在的设计策略,如设计多样性的实现、配置的冗余度、接口的完整性检测和恢复技术,可以弥补软件产品的缺陷或潜在的失效。

评价可能会做出不接受该软件产品的决定,或者做出不主张遵照评价需求的决定,而建议重新评价另外的产品。最终的决定是买或者不买。

购买决定接着会产生一个购买软件产品的合同,这个合同还会带有产品验收测试形式的附加评价。

不买的决定会产生包括修改软件产品、开发一个定制的软件产品或改变需求在内的几种可能的选择方案。

7 定制软件的获取和现有软件的修改期间的评价

本章主要是在定制软件的获取或现有软件的修改期间评价过程的应用。附录 C 的图 C.2 中给出了一个为定制软件的获取和现有软件的修改,将获取过程和评价过程结合起来的例子。

7.1 第 1 步 确立评价需求

在 6.1 中定义的确立评价需求的过程也适用于本条。评价需求构成了获取需求的基础,而获取需求可作为招标书的一部分发送给预选供方。对现有软件的修改,评价必须主要集中在软件产品及接口的被更改的部分。

在发出招标书之前,宜从供方的能力、质量计划和软件工程过程等方面对预选的供方进行初步评价。

7.2 第 2 步 规定评价

6.2 中规定了对现货软件的评价,它也适用于评价定制软件和评价现有软件的修改。然而,附加测量需要作为供方开发过程的一部分,在测量中间产品质量的基础上预测最终的产品质量。GB/T 18905.3 提供了在开发期间测量中间产品质量的需求和指南。

7.3 第 3 步 设计评价

在考虑以下附加因素的条件下,6.3适用于定制软件的获取和对现有软件的修改。

在招标阶段,选择一个供方可能需要供方升级软件工程过程和维护过程,以求在软件开发或修改之前达到目标的完整性需求。

然后,所需的评价活动成为供方执行过程的一部分,即执行供方的开发、验证、联合评审和审核、测试和确认等活动的过程。这些需求将在供方必须遵循的质量或开发计划中规定。需方监督供方对计划的执行情况,并以供方和需方之间达成合同协议的形式确立计划需求。

7.4 第4步 执行评价

除了依照质量计划通过供方的执行活动和需方的监督活动来实施实际的评价之外,6.4中的评价执行需求也适用于这里。在产品交付需方之前,一项必要的准则是进行一次成功的软件产品验收测试。在使用产品之前,为处理在评价中发现的缺陷,需方可能决定对软件产品进行附加的修改。

附录 A
(资料性附录)
源自其他标准的定义

除非另外说明,以下定义均源自 GB/T 18905.1 所规定的定义。

A.1

需方 acquirer

从供方获得或采购系统、软件产品或软件服务的组织。

注:需方可以是下列人员之一:购买者、顾客、拥有者、用户、买方。

[GB/T 8566—2001]

A.2

获取 acquisition

获得系统、软件产品或软件服务的过程。

[GB/T 8566—2001]

A.3

属性 attribute

实体的可测量的物理或抽象的性质。

注:属性可以是内部或是外部的。

A.4

审核 audit

由授权人员对软件产品或过程执行的独立评估,以便评定是否符合需求。

[GB/T 8566—2001]

A.5

基线 baseline

在配置项的生存周期内的某一个特定时刻,已正式设计并固定了的,经正式批准的配置项的一个版本,而不管媒体是什么。

[GB/T 8566—2001]

A.6

CASE 工具 CASE tool

一种协助软件工程师的软件产品,已对 GB/T 8566 中定义的软件生命期活动提供自动支持。

注1:CASE 工具可以只在所选的功能领域提供支持,也可以在广泛的功能领域提供支持。

注2:CASE 工具可以用在以下几种模式:

- 作为一种单独的工具;在这种情况下,仅考虑它与环境要素间的兼容性。
- 在彼此直接联络的小组中;可以假定集成是预先定义的,也可能是专有的。
- 在较大的 SEE 框架中;在这种情况下,宜考虑工具使用框架中相关服务的能力。

[GB/T 18234—2000]

A.7

配置项 configuration item

一个配置中的实体,它满足一项最终使用功能,并能在给定的基准点上单独标出。

[GB/T 8566—2001]

A.8

合同 contract

通过法律约束当事双方的一个协议,或者一个组织内类似的内部协议,以保证软件服务的提供,或软件产品的供应、开发、生产、操作或维护。

[GB/T 8566—2001]

A. 9

开发者 developer

在软件生存周期过程中执行开发活动(包括需求分析、设计、测试直到验收)的一个组织。

[GB/T 8566—2001]

A. 10

直接量测 direct measure

不依赖于任何其他属性度量的一种属性的度量。

A. 11

评价模块 evaluation module

针对特定软件质量特性或子特性的评价技术包,包括:

- 评价方法和技术;
- 要评价的输入;
- 要测量和收集的数据;
- 验收准则;
- 支持的规程和工具。

[GB/T 18905.6—2002]

A. 12

外部测量 external measure

对系统行为的度量而得到的对产品的一种间接量测,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(定制的软件或现货的软件)和用户。

注2:在测试中发现的故障数量是对程序中故障数量的一种外部度量,因为故障数是在计算机系统运行程序的过程中计算的。

注3:外部度量可以用来评价更接近于最终设计目标的质量属性。

A. 13

外部质量 external quality

产品在特定条件下使用时,满足明确或隐含的要求的程度。

A. 14

失效 failure

产品完成所需功能的能力的终止,或在原先规定的限制内没有能力完成。

A. 15

故障 fault

计算机程序中的一个不正确的步骤、过程或数据定义。

注:该定义是取自 IEEE 610.12—1990。

A. 16

固件 firmware

硬件装置和驻留在硬件装置的只读软件中的计算机指令或计算机数据的组合,其软件不能在程序控制下方便地修改。

[GB/T 8566—2001]

A. 17

隐含要求 implied needs

当实体用于特定条件下时,尚未说明但又是实际需要的要求。

注:隐含的要求是未形成文档的真实要求。

A. 18

指标 indicator

能用来估计或预计另一度量的一种度量。

注1:预计的度量可以有相同或不同的软件质量特性。

注2: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量测。

A. 19

间接量测 indirect measure

从一个或一个以上的其他属性的度量得出的一种对属性的度量。

注:对计算机系统属性(例如对用户输入的响应时间)的外部度量就是对软件属性的一种间接量测,因这种量测受计算环境的属性和软件属性的影响。

A. 20

完整性级别 integrity level

项的某个特性的取值范围的一种表示,该特性取值范围对将风险保持在可接受的限度内是必需的。对于执行缓减功能的项,此特性是指项必须执行缓减功能的可靠性。对于因其失效能导致一个威胁的项,此特性是指对该失效的频率的限制。

[GB/T 18492—2001]

A. 21

中间软件产品 intermediate software product

软件开发过程的产品,用作对软件开发过程另一个阶段的输入。

注:在某些情况下,中间产品也可以是最终产品。

A. 22

内部量测 internal measure

对产品本身的一种度量,是直接的或是间接的。

注:代码行数、复杂度度量、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度量。

A. 23

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:在 GB/T 18905 中,与术语“外部质量”对应的“内部质量”基本上与 GB/T 6583 中“质量”的含义相同。

注2:术语“属性”与 4.1.1 中术语“特性”的含义相同,而术语“特性”在 ISO/IEC 9126 中有更具体的含义。

A. 24

量测 measure(verb)

执行一次测量。

A. 25

度量 measure(noun)

通过执行一次量测赋予实体属性的数字或类别。

A. 26

测量 measurement

使用一种度量,把标度值(可以是数字或类别)赋予实体的某个属性。

注:使用类别时,测量可以是定性的。如,软件产品的一些重要属性,例如源程序的语言(ADA,C,COBOL 等)就是定性的类别。

A. 27

度量(体制) metric

定义的测量方法和测量标度。

注1: 度量(体制)可以是内部的或外部的,可以是直接的或间接的。

注2: 度量(体制)包括把定性数据进行分类的方法。

A. 28

现货产品 “off-the-shelf” product

已经开发出来的、可得到的、可使用的、现成的或需要加以修改的产品。

[GB/T 8566—2001]

A. 29

操作者 operator

运行系统的组织。

[GB/T 8566—2001]

A. 30

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1: 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(GB/T 6583,注1)。

注2: 在GB/T 18905中有关的实体是指软件产品。

[GB/T 6583—1994]

A. 31

质量评价 quality evaluation

对实体能满足规定需求的程度的系统检测。

注: 当按照合同为某个特定用户开发产品时,其需求是正式规定的;当产品是为非特定用户开发时,如消费软件,其需求由开发组织来规定;当用户为比较和选择的目的评价产品时,需求可以是更一般的。

[GB/T 6583—1994]

A. 32

使用质量 quality in use

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

注1: 在含有软件的环境中,“使用质量”是关于质量的用户观点。它通过在环境中使用软件所产生的结果来测量,而不是根据软件本身的性质。

注2: 在GB/T 18905中,“使用质量”的定义目前不包括新的“安全性”特性,即特定用户使用产品的程度能满足在规定的使用环境中达到有效性、生产率和满意度等特定目标的需求。

注3: 在ISO 9241-11中,“易用性”的定义与ISO/IEC 9126-1中“使用质量”的定义相类似。使用质量可能会受所有质量特性的影响,因而它比“易用性”含义更广泛,在ISO/IEC 9126-1中,它定义为可理解性、可学习性、可操作性、有吸引力和一致性。

[ISO/IEC 9126-1]

A. 33

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。

A. 34

评级 rating

把测量值映射到相应的评定等级的活动,用于确定与软件某一质量特性相关的等级。

A. 35

评定等级 rating level

在有序标尺上的某个刻度,用于分类某一测量的标度。

注1: 评定等级能使软件按照明确或隐含的要求进行分类(评定)。

注2: 相应的评定等级与质量的不同视角有关,如“用户”、“管理者”或“开发者”的角度。

A. 36

招标[标书] request for proposal[tender]

需方使用的一种文件,用来向潜在的投标人表示它要获得特定系统、软件产品或软件服务的意图。

[GB/T 8566—2001]

A. 37

标度 scale

具有特性定义的一组值。

注: 标度类型的例子有: 与一组类别对应的标称标度、与一组有序刻度对应的序数标度、与一组等距的有序刻度对应的间隔标度,以及既有等距刻度,也具有绝对零度的比例标度。使用标称标度或序数标度的度量产生定性的数据,而使用间隔标度和比例标度的度量产生定量的数据。

A. 38

软件 software

信息处理系统的部分或全部程序、规程、规则及相关的文档。

注: 软件是独立于所记录媒体的智力创作。

[GB/T 5271.1—2000]

A. 39

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。

注: 产品包括中间产品和打算由开发者和维护者等用户使用的产品。

[GB/T 8566—2001]

A. 40

供方 supplier

同需方签订合同,并按合同的规定提供系统、软件产品或软件服务的组织。

注1: 术语“供方”是承制方、生产方、销售方或卖方的同义词。

注2: 需方可以指定本组织的一部分为供方。

[GB/T 8566—2001]

A. 41

系统 system

由一个或多个过程、硬件、软件、设施和人员组成的集合体,提供满足规定要求或目标的能力。

[GB/T 8566—2001]

A. 42

用户 user

使用软件产品执行特定功能的个人。

注: 用户可以包括操作者、软件结果的接受者或软件开发者或维护者。

A. 43

确认 validation

通过检查和提供客观证据证实某一特定预期用途的特殊需求已经满足。

注1: 在设计和开发中,确认关系到检查产品是否符合用户要求的过程;

注2: 确认一般是在规定的操作条件下对最终产品进行的。在早期阶段,这样做是必要的;

注3: “确认过的”一词用来表示相应的状况;

注4: 如果有几种不同的预期用途,可进行多种确认。

[GB/T 6583—1994]

A.44

验证 verification

通过检查和提供客观证据证实规定的需求已经满足。

注1：在设计和开发中，验证是指对某项指定活动的结果进行检查的过程，以确定该活动是否符合明确的需求。

注2：“验证过的”一词用来表示相应的状况。

[GB/T 6583-1994]

附录 B
(资料性附录)
表 格

表 B.1 每个目标软件完整性的评估和获取活动的裁剪实例

评价/获取活动	目标软件完整性		
	低	中	高
制定软件需求规格说明	可能不正式	是	是
制定获取需求	否	是	是
预选供方及准备招标	可能	可能	可能
评估供方能力	否	否	是
评价供方软件过程	否	否	是
评价可执行的产品	是	是	是
与供方一起评价产品操作历史	可能不正式	是	是
与客户一起评价产品操作历史	可能不正式	是	是
起草合同	否	是	是
执行正式的产品验收测试	否	是	是
执行附加的评估	否	可能	是

注1：低级别的完整性软件的例子是“压缩包装的”字处理软件或报表软件。但是，如果用报表软件来为化工厂计算一个关键的安全参数，那么，除非有缓和因素，即，有其他替代的计算方法使软件的完整性级别降低，否则必须要增加软件的完整性级别。

注2：中等级别的完整性软件的例子是用在不具备控制能力的工厂工艺控制系统的用户图形界面(GUI)软件。

注3：高级别的完整性软件的例子是用于空中交通管制系统的 UNIX 操作系统软件。

表 B.2 规定软件质量特性、子特性、外部度量的示例

目标软件的完整性	优先的质量特性	选择的子特性	选择的外部度量	可能的验收准则
低	1. 功能性	准确性	与预先计算的期望结果对照的准确结果数	95%
	2. 易用性	可操作性	与总评审信息数对照的清晰接收的评价信息数	80%
	3. 可移植性	可安装性	转移到新的平台上时，与总模块数对照的要被重新编译的模块数	<6个模块
	4. 效率	时间特性	从系统操作开始到接收系统响应之间所用的时间	<5秒钟
	5. 可靠性	容错度	与录入的总错误输入数对照的错误输入检测失效的次数	25%
	6. 可维护性 (不要求)	不要求	不要求	不要求

表 B.2(续)

目标软件的完整性	优先的质量特性	选择的子特性	选择的外部度量	可能的验收准则
高	1. 可靠性	可用性	指定操作周期中两次操作失效之间的平均时间	>6 个月
	2. 功能性	适用性	与强制性需求总数对照的在软件需求规格说明中满足强制性需求的数量	100%
	3. 可维护性	可变性	为确定的类似改变而需要改动的模块数	1
	4. 效率	资源利用	在最差的操作条件下 CPU 加载超过规定的操作周期的百分比	80%
	5. 易用性	可理解性	特定用户学会如何使用软件生成指定结果所需要的时间	<10 分钟
	6. 可移植性 (不要求)	不要求	不要求	--

表 B.3 规定软件使用中的质量的范例

用户	外部质量目标特性	使用质量的属性	使用质量测量的例子	验收准则
操作员 最终用户	功能性 可靠性 易用性 效率	有效性(把使用系统或产品以达到准确性和完整性的目标与能够达到这些目标联系起来)	——在解释显示结果时产生的错误数 ——在转录文档时偏离规定格式的数目	——越接近零越好 ——越接近零越好
		效率(与智力或体力消耗相关的有效性程度)	——任务时间和用户时间的劳动成本 ——所用设备和资源的成本 ——用户要求的培训成本	——越低越好 ——越低越好 ——越低越好
		满意度(使用的舒适性和可接受性)	——用户交互式满意度问卷调查表 ——在使用中肯定评论和否定评论的比例 ——所知的工作量问卷调查表	——越高越好 ——越高越好 ——每种用户类型的平均人数
维护者 安装者	可维护性 可移植性	有效性(把使用系统/产品以达到准确性和完整性的目标与能够达到这些目标联系起来)	——每一个改动请求所影响的模块数 ——在所有安装尝试中成功安装的次数	——越低越好 ——越接近 1.0 越好
		效率(与智力或体力消耗有关的所达到的有效性程度)	——任务时间和用户时间的劳动成本 ——所用设备和资源的成本 ——用户要求的培训成本	——越低越好 ——越低越好 ——越低越好
		满意度(使用的舒适性和可接受性)	——用户交互式满意度问卷调查表 ——在使用中肯定评论和否定评论的比例 ——所知的工作量问卷调查表	——越高越好 ——越高越好 ——每种用户类型的平均人数

表 B.4 评价方法的成本效益等级的例子

评价方法	成本等级	每种软件质量特性的评价方法的效益等级					
		功能性	可靠性	易用性	效率	可维护性	可移植性
最终产品的文档、课程和培训	低	高	低	高	中	低	高
中间产品	高	低	中	低	中	高	高
操作历史—供方	中	中	高	低	低	中	中
操作历史—客户方	中	高	中	高	中	高	高

注1: 此表表示了与特定产品质量特性相关的评价方法的相关总成本和相关有效性的定性评估等级。这种有效性评级假定评价成功地执行并达到适当的精确程度。此表可用于选择目标输入,这些目标输入需要对照那些在软件需求规格说明中规定的产品质量特性来进行评价,以全面评估产品质量特性的充足性。所需的评价能用来估算产品评价的总成本。

注2: 这里示出的成本和有效性等级是相关联的,并且取决于所进行的评价的范围。

附录 C
(资料性附录)
流程图

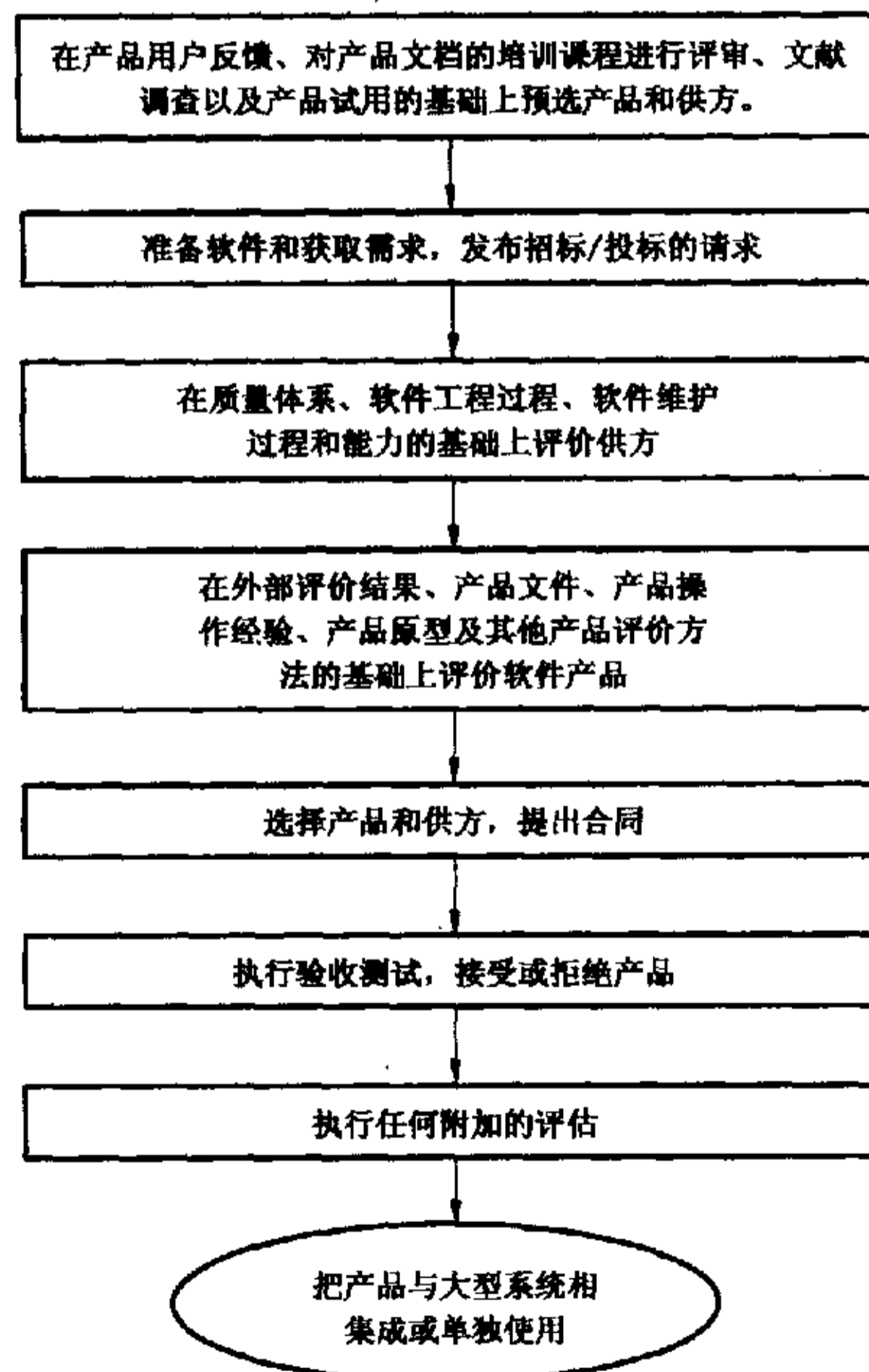


图 C.1 现货产品的评价/获取过程的例子

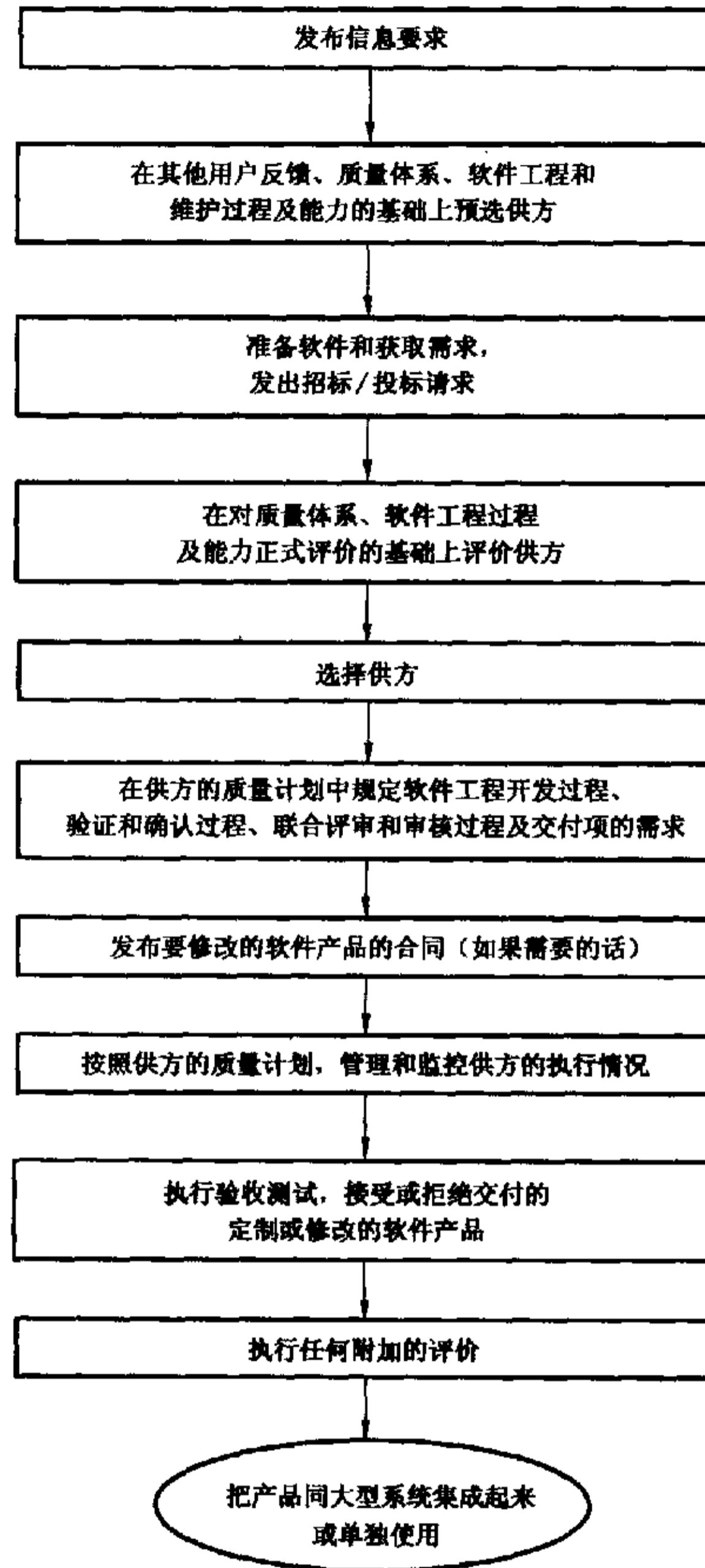


图 C.2 定制软件或对现有软件进行修改的评价/获取过程的例子

附录 D
(资料性附录)
评价方法

D.1 对用户和产品技术文档(包括联机文档)的评审

产品文档可以在对功能性和易用性需求以及可移植性和可维护性等方面进行评价时提供所有必要的信息。通过借阅文档或购买文档集能得到对相关软件产品文档的访问,并不需要实际购买该软件产品。虽然阅读软件产品文档可能不如参加课程或进行培训更有效,但可以证明是最经济的方法,尤其是对于评价者具有相关专业知识的情况。

D.2 基于供方课程和培训的评价

可以通过供方或第三方为众多的软件产品提供产品课程。在没有软件产品课程的情况下,可能会安排有经验的用户或软件产品的开发者开展专门的培训。提供产品课程或培训的好处在于能使评价者把精力集中在特定范围,并且在较短时间内获得有关产品功能和易用性方面的特定信息。通过阅读软件产品的文档也可能得到相同的信息,但要花费更多的时间。课程或培训的附加成本需要根据获取信息的效率和课程资料的普遍性两方面来衡量。

D.3 对软件工程过程的评估

对软件工程过程的评估是通过检查过程的中间产品,如:产品质量计划、需求规格说明、体系结构描述、详细设计描述、代码清单、验证和确认记录、代码检查和测试记录等来决定软件产品的质量。要达到这一目标,必须定义什么是构成一个可接受的软件工程过程的文档编制基线,该过程将为最终得到的软件产品提供充足的质量保证。

针对目标完整性级别剪裁 GB/T 8566 的需求可以定义一个可接受的基线,以便规定必要的开发和相关的支持活动。该基线包括确定下列内容:

- a) 必需的过程;
- b) 必需的过程输出文档(注:GB/T 18905.5 也在附录中列出了能用于评价软件工程过程的产品文件集);
- c) 对过程和过程输出文档的需求。

评估可以与 ISO/IEC 15504-8 中定义的供方过程能力级别的确定联系起来。例如,对有中级到高级完整性需求的产品可能要求表 D.1 中定义的软件过程文档编制基线:

表 D.1 软件过程文档编制基线

GB/T 8566—2001 过程	输出	GB/T 8566—2001 要求(条目)
计划编制	软件质量/开发计划和支持过程	7.1.2
系统需求分析	系统需求规格说明	5.3.2
系统需求验证	验证记录	6.4.2.3
系统结构设计	系统设计描述	5.3.3
系统设计验证	验证记录	6.4.2.4
软件需求分析	软件需求规格说明	5.3.4
软件需求验证	验证记录	6.4.2.3

表 D.1(续)

GB/T 8566—2001 过程	输 出	GB/T 8566—2001 要求(条目)
软件结构设计	软件结构描述	5.3.5
软件详细设计	软件设计描述	5.3.6
软件设计验证	验证记录	6.4.2.4
软件编码	代码	5.3.7
代码验证	验证记录	6.4.2.5
软件单元测试	测试记录	5.3.7
软件集成	测试记录	5.3.8,6.4.2.6
软件确认	测试记录	5.3.9,6.5
系统集成	测试记录	5.3.10
系统确认	测试记录	5.3.11,6.5
配置管理	计划、状态报告、发布、变更请求	6.2
培训	培训记录	7.4

对于有高完整性需求的系统,可以通过部门标准,如:IEC 880, Draft IEC 1508, DOA-167A 和 MOD-55 等来提出附加过程和产品需求。

供方的质量/开发计划及相关的方法学规程能用于评估供方对目标基线的符合情况。那么,一致性级别可以通过标识主要缺陷和评估其对软件产品质量的潜在影响来确定。附加的评价或变通的工作能说明缺陷的影响。

可以认为,有多种软件工程过程对特殊组织或不同种类的软件产品是有效的。评价过程应具有灵活性,以便适应多种多样合理的软件工程过程和方法。

建议软件工程评审按附录 E 所示的例子分阶段进行。当认为软件的完整性级别不要求对软件工程过程进行全面评价时,评审可以在第 I 阶段或第 II 阶段之后停止。

D.4 对供方操作历史的评审

对供方操作历史的评审能够提供一种非常有效的表明软件产品质量的方法。这可以通过评审软件产品的销售图表和使用该产品的行业及应用程序的详细资料来实现。这种评审也说明了软件版本的修订历史、维护修订的方法、解决顾客反映的缺陷的方法,以及已知缺陷的详细内容。执行评审最便利的方法是与供方的工程人员、销售人员和顾客支持人员进行沟通,以及检查所有的支持记录。

D.4.1 操作历史的需求

- 销售图表宜至少有六个月的时间;即用在评价中的销售数字应只包括在评价前六个月以上的销售额。这一准则基于的事实是:软件产品从交付、安装、委托代理直至服务到位可能需要花费多达六个月的时间;
- 软件产品宜至少经历一个主要的修订,并且宜有该修订中可实行的操作历史数据。这是基于这样一个假设:软件产品的质量依赖于它所经历的改进的数量;
- 软件产品的用户有向供方反馈缺陷报告的手段,而且也存在缺陷发生和处理缺陷的证据。

D.4.2 操作历史的评审

对产品操作历史的评审宜确定:

- 软件是否是通过修改另一产品而产生的,能否使用该产品的操作历史。在对该软件产品的改动次数和改动程度方面这可能是附带的;
- 软件产品运行的单元年数。由下列步骤算出:

- 1) 计算年销售额=(初始销售额[第一年的总销售额]+从现在起六个月累计的最终总额[假设一个单元投入运行前通常要延迟6个月])×(软件产品投放市场的年数)/2;
 - 2) 计算运行单元年数=(年销售额×责任周期[软件投入运行的时间百分比]×单元实际运行的百分比[这与固件有关,因为固件中的若干主机硬件单元可留作备件]);
- c) 操作经验是否提供了与应用所要求的功能有关的证据。例如,它是否被其他顾客用在类似的应用中?
 - d) 操作经验是否提供了已经充分运用了软件产品的功能的证据,例如,它是否已被用在广泛的应用和行业中?
 - e) 软件的每个修订版和软件产品的每个特定选项的运行单元年数;
 - f) 每种修订之间的区别,改动的程度以及每次改动是否是孤立的;
 - g) 是否存在文档化的证据支持操作历史数据;
 - h) 如何控制和跟踪软件产品的修订和相关硬件的修订;
 - i) 是否有可能定制软件产品的某个特定修订版,并暗示定制不是当前的修订版;
 - j) 接受和处理顾客提出的缺陷报告的供方过程;
 - k) 顾客是否保持最新的缺陷报告;
 - l) 软件的任何明显缺陷及其影响。

D.5 对顾客操作历史的评审

对那些正在使用或曾在某个应用中使用该软件产品的实际顾客的操作历史的评审,能使评价者在实际的工作条件下不带偏见地回答一些具体的问题。根据顾客的应用和所要求的应用之间的相似程度,从整体质量或特定功能中得到的质量保证可能与从原型或者甚至从扩展的实验用途中所得到的保证一样有用。进行评审最方便的办法是通过在工作现场与顾客进行沟通,甚至可能通过观看演示或查阅支持记录的方式。

评价者进行对顾客操作历史的评审宜:

- a) 建立与要求的应用的相似的等级;
- b) 试图观察工作中的软件产品或取得其他支持的证据;
- c) 询问顾客有关供方提供的支持形式和支持质量的问题;
- d) 确定免错操作的总数。

D.6 对供方的能力、支持和质量体系评审

对供方的支持能力级别和维护软件的能力的评价宜说明:

- a) 财政的稳定性、经验和能力;
- b) 产品开发环境的支持,包括集成工具、所用的操作系统、对其它部件/库对象的维护和使用;
- c) 产品与其他产品或产品组的接口,包括接口标准;
- d) 涉及第三方的可能性;
- e) 充分的提供产品支持的资源承诺;
- f) 充分的证明持续支持的顾客基础;
- g) 充分的用于纠正错误及环境支持的维护服务;
- h) 充足的对已经安装的平台引用;
- i) 形式化的和充分的发行、修订控制规程和实践证据;
- j) 形式化的回归测试和设计改动的正式评价;
- k) 文档化和形成惯例的问题报告和解决规程;
- l) 质量体系已就位;

- m) 用于软件和硬件的标准;
- n) 未来的开发计划;即与当前市场定位相关的策略已就位;
- o) 产品质量保证书。

D.7 原型和其他评价方法

D.7.1 原型

原型是一种评价方法,用于分析或微调需求,以确定使用软件产品的技术可行性,或排除与特定的功能性或易用性需求及与实现有关的未知风险或技术风险。原型可能或可能不利用全部软件产品功能,或者说明整个应用的需求。

应注意:原型常常要求供方提供对特别指明的设备、人员和文档的访问。对这些和其他因素,比如特定环境条件或服务所隐含的成本和进度问题,宜在确定软件产品原型的可行性时考虑。

除了通常的评价需求之外,原型宜:

- a) 使用足以表明需求已被评估的例子,并实际提供对关键操作参数的重新构造和模拟;
- b) 尽量整理文档以便能被第三方重复使用;
- c) 利用与提议的应用相关的历史数据。

D.7.2 其他评价方法

GB/T 18905.5 在附录中列出了与评价等级和软件质量特性相连的评价技术的清单。评价等级应该与评价的完整性级别相对应。其它的评价方法包括:

- a) 对软件结构设计的分析(可维护性);
- b) 软件的故障树分析(安全性、可靠性);
- c) 基于软件产品测试的随机统计应用(可靠性);
- d) 检测语法和语义的正确性的代码动态分析(可靠性);
- e) 软件设计的危险分析(安全性、可靠性);
- f) 对软件需求规格说明的评审(功能性);
- g) 代码检查(功能性);
- h) 软件的黑盒测试(功能性);
- i) 基准测试(效率);
- j) 需求的可追踪性分析(可维护性);
- k) 在部件间接口的模拟故障(坚固性)。

对具有高完整性级别的复杂软件,故障树分析或软件设计的危险分析能用来为进行评价而孤立“关键的”软件模块。这可以排除对不影响应用完整性的软件进行严格评价的必要。

附录 E
(资料性附录)
分阶段评价过程的例子

下列描述的是分阶段评价的例子,展示了在各个阶段对三个目标领域所要求的相关工作。

E.1 阶段 1:计划-需求阶段

E.1.1 产品的文档编制、进程和培训

本领域的某些评审最初几乎都是为了证实和确定软件产品的基本功能。这种初始的评审可由其他方来进行。根据完整性级别,宜考虑将这种非正式或初步的评价编成文档。

E.1.2 软件工程过程

在此开始对软件产品的软件工程过程进行评价是非常有利的,通常通过评审软件产品的质量/开发计划来进行。这一评价工作将指明该软件工程过程是否有可能产生满足文档集所要求的一般特性的文档和规程,而文档集是针对相关完整性级别的。适当的时候,对其他过程评价的评审可以认为是对这项工作的补充。

在此宜考虑一般特性和非特定需求,同时也要注意下列几点:

- a) 将供方的文档与规定的文档集进行准确匹配是不可能或不必要的;
- b) 在评估软件工程过程时,注意在整个过程中该过程的前后关系、导致该过程的前面几步的重要性,以及该过程的输入质量。

E.1.3 操作历史

对供方和适当顾客的操作历史的初始评审能够较早地提供是否需要进一步评价的指示。例如,操作历史较短和/或适当的顾客较少可能导致要考虑原型、对外部评价的评审,或提高软件工程过程评价的精确度等。

E.2 阶段 2:设计-获取阶段

E.2.1 产品的文档编制、进程和培训

在本阶段,需要对软件产品所需的功能进行更全面的评审。如果产品的进程或文档都不能证实所需的功能,那么,宜执行比适当的顾客操作历史更严格的评审。

E.2.2 软件工程过程

在本阶段评估质量计划的实际实施情况。这将涉及对实际产生的中间产品的评审,以确保它们与那些已确定的产品的一般特性保持一致。如果本阶段与投标过程重合,时间要素可能限制只对关键文档检查的评审。标书宜总是包含一个允许在合同裁定之前对在软件产品质量计划中引用的任何文档进行检测的阶段。

由于软件工程过程评审的成本较高,一般不考虑为较低完整性级别的软件产品进行软件工程过程的评审,除非觉得它的操作历史不足以提供产品的可靠性保证。

E.2.3 操作历史

本阶段宜集中在对现有软件产品的特殊版本和所需功能进行更广泛的操作历史的评审。如果对产品进程、文档及对适当的顾客的操作历史评价的评审不能为产品的可靠性和功能性提供足够的信任度,宜考虑使用原型。

E.3 阶段3:全面评价阶段

E.3.1 产品的文档编制、进程和培训

如果前面的阶段没有完成产品的功能性评价,那么在本阶段宜通过评审产品的文档或进程的参与者来完成产品的功能性评价。

E.3.2 软件工程过程

本阶段需要在把每个过程/产品的特定需求与供方的特定需求相比较的地方,对软件工程过程进行全面的评价。要考虑的补充点包括:

- a) 宜说明每一项需求,以确定供方的软件工程过程是否遵守需求的目标。通过与供方的软件工程师人员进行交流来做到这一点往往最容易;因此,在投标过程期间来确定它是不可行的或不够有效的;
- b) 也可要求供方成员展示其产品有充分质量保证的证据。这可能为没有专门被基线软件过程覆盖的领域提供了前景。它也有助于促成基线软件过程中的特定领域与部分不能对评价者直接展示证据的过程之间建立联系;
- c) 一些需求有主观因素,需要评价者的解释。因此,可能存在各种各样的满足需求目标的方法。
- d) 评审宜建立软件产品的修订历史,同时也宜确定每个修订版之间的改动程度。这包括对报告的缺陷的检查。宜考虑下列情况:
 - 1) 所有已知缺陷的状态;
 - 2) 对修复缺陷的所有改动所做的评审、评价和测试的等级;
 - 3) 当前设计修订的稳定性;
 - 4) 从最近发布之日起所经过的时间;
 - 5) 改动的次数和这些改动的规模或幅度;

如存在缺陷或不可以进行详细评审的地方,宜考虑如原型或黑盒测试这样的附加评价方法。

E.3.3 操作历史

如果在前面的阶段没有实施对操作历史的评价,那么,本阶段宜完成此项工作。

参 考 文 献

标准

- [1] ISO/IEC 指南 2:1991 关于标准化和相关活动的通用术语和定义
- [2] GB/T 5271.1—2000 信息技术 词汇 第1部分:基本术语
- [3] GB/T 5271.20—1994 信息技术 词汇 20部分:系统开发
- [4] GB/T 6583—1994 质量管理和质量保证 词汇
- [5] GB/T 19001—1994 质量体系 设计、开发、生产、安装和服务的质量保证模式
- [6] GB/T 17544—1998 信息技术 软件包 质量要求和测试
- [7] GB/T 18234—2000 信息技术 CASE工具的评价和选择指南
- [8] Radio Technical Commission for Aeronautics (RTCA) DO-178B/ED-12B, 航空系统和设备认证中的软件考虑
- [9] Ministry of Defence (UK) (MOD), Interim Defence Standard, 00-55(Parts 1,2)/Issue 1, The Procurement of Safety Critical Software in Defence Equipment.
- [10] IEC 880—1986 核电站安全系统中的计算机软件
- [11] IEEE Std 1062—1993 软件获取的建议实践

其他参考文献

- [12] Tremaine, DR. and de Grosbois, JFP. Guideline for the Qualification of Predeveloped Software. 907-C-H-6999002-0201, Ontario Hydro/AECL, April 22, 1993.
- [13] Ferguson, JR. and DeRiso, ME. Software Acquisition: A Comparison of DoD and Commercial Practices. CSU/SEI-94-SR-9 Software Engineering Institute, October, 1994.
- [14] Baker, ER., Cooper, L., Corson, BA. and Stevens, AE. Software Acquisition Management Maturity Model (SAM³). Program Manager. July-August, 1994, pp. 43-49.
- [15] Scott, JA., Preckshot, GG. and Gallagher, JM. Using Commercial-Off-the Shelf (COTS) Software in High Consequence Safety Systems (Draft Preprint). Lawrence Livermore National Laboratory, November, 1995.
- [16] Cochrane, Gail. Use of COTS/NDI in Safety-Critical Systems. (based on original Report prepared for the Federal Aviation Authority), TRW, 1996.
- [17] Brown, AW. and Wallnau, KC. Engineering of Component-Based Systems. Proceedings Second IEEE International Conference on Engineering of Complex Computer Systems. October, 1996.
- [18] Bevan, N. and Azuma, M. Quality in Use: Incorporating human factors into the software engineering lifecycle. Proceedings of the Third International Symposium on Software Engineering Standards. May, 1997.
- [19] Voas, J. and Miller, K. Interface Robustness for COTS-Based Systems. IEE Symposium on COTS and Safety-Critical Systems. Savoy Place, London, January, 1997.