

需求演化过程建模

张国生

(云南大学 信息学院, 云南 昆明 650091)

摘 要: 对需求任务、需求活动、需求工程过程以及需求工程过程系统进行形式化定义. 用信息熵对需求任务演化进行度量, 用联合熵对需求活动、需求工程过程以及需求工程过程系统演化进行度量. 从需求工程过程的角度, 用条件-事件网对需求工程过程的迭代、反馈进行微观演化建模. 从系统工程的角度, 用耗散结构理论对整个软件需求工程过程系统进行宏观演化建模.

关键词: 演化; 迭代; 反馈; 信息熵; 联合熵; 耗散结构理论

中图分类号: TP311

文献标识码: A

文章编号: 1000-7180(2012)05-0054-04

Requirements Evolution Processes Modeling

ZHANG Guo-sheng

(School of Information, Yunnan University, Kunming 650091, China)

Abstract: Requirements tasks, requirements activities, requirements engineering processes and requirements engineering processes system are formally defined. Requirements tasks are measured with information entropy. Requirements activities, requirements engineering processes and requirements engineering processes system are measured with joint entropy. From point of view of requirements engineering processes, microcosmic evolution of iteration and feedback of the requirements engineering processes are modeled with condition-event nets. From point of view of system engineering, macro evolution of the whole software requirements engineering processes system is modeled with dissipative structure theory.

Key words: evolution; iteration; feedback; information entropy; joint entropy; dissipative structure theory

1 引言

软件工程的基本目标是采用系统化的构建方法和工程化的管理手段去高效地开发出满足用户需求的软件系统. 软件需求工程的目的是为软件系统提供一个合理正确的开发文档, 因此, 需求工程无疑是软件工程之前的一个重要阶段, 它的成功实施是软件后续开发过程得以成功的先决条件. 近年来, 需求工程已经取得长足的进步, 但从其发展进程和未来的发展趋势来看, 需求工程不能仅仅局限在软件工程为需求阶段设定的目标上, 也就是说, 除了软件工程为它设定的目标外, 它更强调要用一种自然的方

式, 去系统地寻找和发现正确的和潜在的用户需求, 为这些需求建立适当的模型, 并在对模型进行正确性分析和有效性验证的基础上, 为软件开发的后续阶段提供合理、正确的软件规格说明^[1-2].

文献[3]将软件演化定义为软件系统在其生命周期中不断维护、不断完善的系统动力学行为. Lehman 认为软件演化是一个多层次、多循环、多代理的反馈系统. 来自于全局软件过程的可能约束, 包括技术、业务、市场、用户以及其它活动, 将形成一个多层次、多循环反馈系统, 为了改变系统的特征需要考虑、设计或适应、调整前向或反馈路径使外部可视行为达到所要求的变化^[4].

2 需求工程过程线性迭代模型

软件开发的螺旋式模型认为软件开发过程本质上是迭代的,以这种思想为基础,软件需求工程过程模型也是迭代的,如图1所示^[5-6].这是目前大多数需求工程过程采用的需求过程模型,它包含了软件需求工程的所有基本活动:需求抽取、需求分析和协商、需求文档化、需求有效性验证,并刻画了这些基本需求活动之间的关系.当然,这些活动并不是完全独立的,在具体的软件需求工程过程中它们常常交织在一起,与线性需求模型不同的是,线性迭代需求模型通过反馈实现软件需求工程过程及其中间结果的迭代、增量式开发.

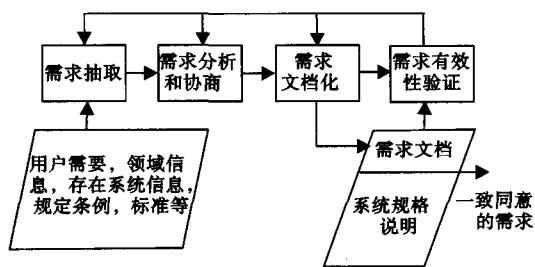


图1 需求工程过程线性迭代需求模型

需求工程过程线性迭代模型中各个基本活动的含义如下:

(1) 需求抽取:发现系统需求,一般通过咨询需求相关者获得,或从系统文档、领域知识、市场需求中获得;

(2) 需求分析和协商:详细分析需求,并与需求相关者协商决定哪些需求是可以接受的,协商的目的是要获得一致同意的需求集合;

(3) 需求文档化:记录一致同意的需求,并按某种规范和规定的详细程度撰写文档;

(4) 需求有效性验证:检查需求的一致性和完整性,其目的是要在需求文档被用来作为系统开发的基线之前尽可能完全地对需求文档进行系统化的检测.

3 熵与耗散结构理论

3.1 信息熵

1948年香农在《通信的数学理论》中给出了信息量的定量表述^[7].假定样本空间 X 有 n 个事件 X_i ($i=1, \dots, n$),每个事件出现概率为 p_i ($i=1, \dots, n$),且存在基本关系

$$\sum_{i=1}^n p_i = 1 \text{ 且 } p_i \geq 0 \text{ (} i = 1, \dots, n \text{)}$$

若有一函数 $H(P_1, \dots, P_i)$,要求它满足下述条件:

(1) 对于固定的 n 来说 H 是 $p_1, \dots, p_n$ 的连续函数;

(2) 若 $p_i = \frac{1}{n}, i = 1, \dots, n$,则对应的 $H\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$ 是 n 的单调递增函数;

(3) 若将试验分解成多个相继的试验,则原先的 H 值应为分解后多个试验相应各个 H 值的加权和;则将此函数 $H(p_1, \dots, p_i)$ 定义为熵

$$H(X) = -K \sum_{i=1}^n p_i \log p_i$$

3.2 联合熵

设 n 个随机变量 $X_1, X_2, \dots, X_n$ 均取有限值,它们的联合概率分布为 $p(x_1, x_2, \dots, x_n) = P\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$,则 $X_1, X_2, \dots, X_n$ 的联合熵定义为

$$H(X_1, X_2, \dots, X_n) = - \sum_{x_1, x_2, \dots, x_n} p(x_1, x_2, \dots, x_n) \log p(x_1, x_2, \dots, x_n)$$

3.3 耗散结构理论

一个系统要处于耗散结构,即动态有序,必须满足以下条件^[8]:(1)系统必须开放;(2)远离平衡态;(3)非线性相互作用;(4)涨落现象.这些条件是相互紧密联系的,根据这些条件可以将耗散结构概括为:在非平衡条件下产生的,依靠物质、能量、信息不断输入和输出条件来维持其内部非线性相互作用的有序系统.

一个开放系统的熵改变可以分成两部分 $dS = d_i S + d_e S$.其中, $d_i S$ 表示由于系统内部原因,使系统熵发生改变的部分,与外界无关; $d_e S$ 表示由于系统与外界有联系,形成系统熵改变的部分,若系统与外界无联系,则 $d_e S = 0$,否则 $d_e S \neq 0$.系统与外界的关系各式各样,其对系统熵的影响 $d_e S$ 可正、可负,系统总熵变化等于这两部分之和.因此,在开放系统中,只有 $d_e S < 0$,且 $|d_e S| > d_i S$,才有

$$dS = d_i S + d_e S < 0$$

整个系统的熵减少,系统变得有序,或者说只有在这时才有可能使系统从无序状态向有序状态转化,归纳起来可以认为,系统开放与外界有物质、能量交换时,由于交换使系统熵减少,且其变化数值大于系统内自发过程引起的熵产生时,整个系统的熵可能减少,系统向有序方向转化.

4 需求演化过程建模

4.1 需求任务、需求活动、需求过程系统及过程形式化

在软件需求工程过程中,需求任务完成特定的软件系统需求功能目标,多个需求任务构成需求活动,多个需求活动构成需求工程过程,多个需求工程过程构成需求工程过程系统。

定义 1 需求任务定义为一个 5 元组 $RT = (Inputs, Outputs, Pre, Post, IE_t)$, 其中

(1) Inputs 和 Outputs 分别表示需求任务的输入工作产品和输出工作产品;

(2) Pre 和 Post 为一阶谓词公式,分别表示执行需求任务所需的状态(前置条件)以及需求任务执行后的状态(后置条件);

(3) IE_t 为需求任务的信息熵。

定义 2 需求活动定义为一个 5 元组 $RA = (Inputs, Outputs, ra_body, ra_attributes, JE_a)$, 其中

(1) Inputs 和 Outputs 分别表示需求活动的输入工作产品和输出工作产品;

(2) ra_body 表示需求活动本身,既可以是一个需求过程,也可以是一组需求任务的集合,这些需求任务或需求过程作用于输入工作产品上;

(3) $ra_attributes$ 表示需求活动的自身属性,需求过程的不同需求活动所具有的属性以及所包含的需求任务各不相同,所产生的工作产品也不相同;

(4) JE_a 为需求活动所包含的多个任务的联合熵。

定义 3 满足下列条件的三元组 $N = (B, E; F)$ 称作一个条件-事件网

(1) $B \cup E \neq \emptyset$

(2) $B \cap E = \emptyset$

(3) $F \subseteq (B \times E) \cup (E \times B)$

(4) $\text{dom}(F) \cup \text{cod}(F) = B \cup E$, 其中

$\text{dom}(F) = \{x \in B \cup E \mid \exists y \in B \cup E: (x, y) \in F\}$

$\text{cod}(F) = \{x \in B \cup E \mid \exists y \in B \cup E: (y, x) \in F\}$

定义 4 设 $N = (B, E; F)$ 为一个条件-事件网

(1) 令 $x \in N$, 则

$\cdot x = \{y \mid (y, x) \in F\}$ 称为 x 的前集或输入集;

$x \cdot = \{y \mid (x, y) \in F\}$ 称为 x 的后集或输出集。

(2) 令 $x \subseteq N$, 则

$\cdot X = \{x \mid \exists y \in X: x \in \cdot y\} = \bigcup_{y \in X} \cdot y$ 称为 X 的前集;

$X \cdot = \{x \mid \exists y \in X: x \in y \cdot\} = \bigcup_{y \in X} y \cdot$ 称为 X 的后集。

定义 5 需求过程系统定义为一个 5 元组 $RPS = (B, E; F, JE_{ps}, c)$, 其中

(1) $(B, E; F)$ 是一个条件-事件网, B 称为条件集, E 称为事件集, JE_{ps} 为需求过程系统所包含的多个需求过程的联合熵, $c \subseteq B$ 称为 RPS 的一个情态。

(2) 事件 $e \in E$ 在情态 c 有发生权, 记作 $c[e]$, 当且仅当

$$\cdot e \subseteq c \wedge e \cdot \cap c = \emptyset$$

(3) 若事件 e 在情态 c 发生, 产生新的情态 c' , 记作 $c[e] \rightarrow c'$, 则

$$c' = (c - \cdot e) \cup e \cdot$$

定义 6 设 $RPS = (B, E; F, JE_{ps}, c)$ 为一个需求过程系统, $c_0 \in C$ 为 RPS 的一个实例, 则 $RP = (B, E; F, JE_p, c_0)$ 为一个需求过程, 其中, JE_p 为需求过程所包含的多个需求活动的联合熵。

4.2 需求演化过程的迭代、反馈及其动力学机制

需求演化是指软件需求从不明确到明确的变化过程。在问题域极其复杂或者需求不稳定时, 可以采用演化模型, 演化模型采用增量迭代和逐步展开的方式进行需求开发, 因此, 能够处理复杂的问题域和业务活动, 迭代是在明确了部分需求后就开始进行下一步的需求开发活动, 待完成后, 再针对新的需求进行迭代, 如此反复迭代, 直到系统所有的需求通过验证、确认为止, 而且不断迭代的方式还可以使它更好地应对需求的改变, 从而适应需求的不稳定性, 需求演化模型将需求开发看作是一个长期迭代的过程, 因此, 可以延长用户在项目中的作用时期, 提高用户的有效参与度。

一般来说, 需求工程过程开始于最初的用户需求和更高一级的面向系统的需求, 这些需求是逐渐地被分解并完善, 直到满足整个系统的需求为止。由于软件的需求演化过程是一个随时间而变化的过程, 因此, 可以把它看成是一个动力系统, 同时, 它也是一个离散的过程。

一个包含有 N 个子系统的系统, 从子系统层次来看, 当每个子系统存在概率测度, 它可以看成是子系统的不确定性, 那么在系统整体层次也必然存在不确定性, 这个不确定性是由子系统的不确定性造成的。

当需求任务的信息熵达到一定的阈值, 需求任务必须进行相应的演化; 当多个需求任务的联合熵达到一定的阈值, 需求活动必须进行相应的演化; 当

多个需求活动的联合熵达到一定的阈值,需求过程必须进行相应的演化;当多个需求过程的联合熵达到一定的阈值,整个需求过程系统必须进行相应的演化,而且,整个软件系统需求的演化遵循耗散结构理论开放系统的熵变原则。

4.3 需求演化过程模型

从系统工程的角度,外部环境和用户需求的变化必然会反馈到需求工程过程中,要求需求工程过程中的需求任务做相应的变化,从而导致需求活动、需求过程以及整个需求工程过程系统的演化,用耗散结构理论、联合熵、信息熵对整个需求工程过程系统、需求工程过程以及所包含的需求活动、需求任务进行度量,用扩展的条件-事件网为软件系统的需求演化过程建立形式化模型,如图2所示。其中, e_1 为需求抽取事件、 e_2 为需求分析和协商事件、 e_3 为需求文档化事件、 e_4 为需求有效性验证事件、 e_5 为需求演化反馈事件。在该模型中既模拟了需求的迭代,又模拟了需求的反馈,从而实现整个软件工程过程的增量、迭代、反馈和演化。

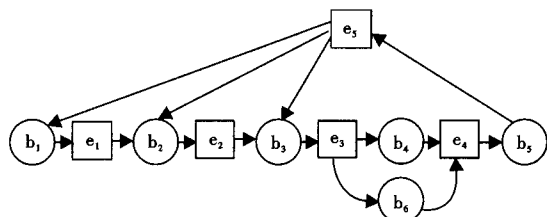


图2 需求演化过程形式化模型

5 结束语

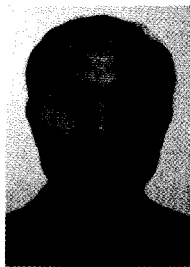
演化性是软件的基本属性,由于用户需求不断变化,导致软件的演化性和复杂性进一步增强,由于外部环境和用户需求的不断变化,软件系统必须不断演化以适应新的需求。需求演化是系统需求不断完善、发展的必然趋势,从需求工程过程的角度,对需求演化过程建模,突出了需求演化中的迭代和反馈。在需求演化过程中,用信息熵对需求任务进行演

化度量,用联合熵对需求活动、需求过程、需求过程系统进行演化度量,用条件-事件网模拟了需求工程过程的微观演化;用耗散结构理论,将需求工程过程系统看作一个开放系统,模拟需求工程过程系统的宏观演化。

参考文献:

- [1] 葛君伟,葛智,方义秋.非线性动力系统在软件需求分析中的研究[J].微电子学与计算机,2010,27(6):85-88.
- [2] Beecham S, Hall T, Rainer A. Defining a requirements process improvement model[J]. Software Quality Control, 2005(13):247-279.
- [3] Lehman M M. On understanding laws, evolution, and conservation in the large-program life cycle[J]. Journal of Systems and Software, 1980(1):213-221.
- [4] Belady L A, Lehman M M. A model of large program development[J]. IBM Systems Journal, 1976, 15(1):225-252.
- [5] 罗钧曼,武钰云,吴彬.模糊本体及其演化研究[J].微电子学与计算机,2011,28(5):140-143.
- [6] Kotonya G, Sommerville I. Requirements engineering: processes and techniques[M]. New York: John Wiley & Sons. 1998.
- [7] 香农 E. 通信的数学理论[M]. 沈永朝,译.北京:求实出版社. 1989.
- [8] 尼科利斯 G,普利高津 I. 非平衡系统的自组织[M]. 徐锡申,译.北京:科学出版社,1986.

作者简介:



张国生 男,(1968—),硕士,副教授。研究方向为软件工程及其形式化。