

学生软件项目质量的模糊评价研究

曹起武

(辽宁机电职业技术学院 信息工程系, 辽宁 丹东 118009)

摘要: 学生学习成果的评价是教学过程中十分重要的环节之一。传统的试卷等评价方式可以比较准确地得出分数,而对于一些类似作品、项目等成果则很难给出客观的分数。针对计算机软件类专业学生软件项目的质量评价问题,文中展开了基于多属性模糊评价方法的评价研究。根据不同属性的特性进行设计,并采用相应的评语集合来评价各学生软件项目。通过建立基准评语集合把来自不同评语集合的语言属性评价价值转化为基准评语集合的模糊集,实现了不同评语集合之间的可比性,并进一步将它们有效集成,从而得出各学生软件项目的综合评价价值以及它们的排序。最后,通过一个实际算例来说明根据模糊集方式设计的评价方法,可以对学生的软件项目给出一个比较客观、准确的评价。

关键词: 学生软件项目;多属性决策;语言评价;模糊集;排序

中图分类号: TP311.1

文献标识码: A

文章编号: 1673-629X(2016)02-0043-04

doi: 10.3969/j.issn.1673-629X.2016.02.010

Fuzzy Evaluation Research on Quality of Student Software Projects

CAO Qi-wu

(Department of Information Engineering, Liaoning Jidian Polytechnic, Dandong 118009, China)

Abstract: The evaluation of students' learning achievement is one of the most important links in the teaching process. It is accurate to get the score by means of traditional evaluation methods like paper tests. But for some similar works, projects and results, it is difficult to give an objective score. In order to evaluate the quality of the student software projects for computer software professional, propose the fuzzy evaluation research based on multiple attributes. According to the respective characteristics of the different attributes, the appropriate linguistic evaluation sets are designed for assessing the projects. The basic linguistic evaluation set is designed, based on which various linguistic evaluation terms from different sets are all transformed into the fuzzy sets over it. Thus, various linguistic evaluation terms from different sets are comparable, and further integrated for calculating the overall values of the projects. In addition, the rankings of the student software projects can be obtained. Finally, an example is used for illustrating that an objective and accurate evaluation can be given to the students according to the evaluation of fuzzy set design.

Key words: student software projects; multiple attribute decision making; linguistic evaluation; fuzzy set; ranking

1 概述

学生学习成果的评价是高等教育教学过程中最后一个环节,而且是十分重要的一个环节。与中小学等较低教育层次对学生评价多数通过具有固定答案的试卷方式不同,高等教育过程所生成的学习成果多数是一些没有唯一固定标准的项目或成果,这就需要有一种方法可以科学、合理、公平地评价学生的学习效果及成果^[1-3]。评价方法的科学性与合理性直接决定了学生成果评价的准确性,而评价的准确性直接影响到学生在后续学习中的学习积极性和态度,所以随着信息化教学体系的不断完善,一个科学合理的评价体系是

成熟教学体系不可或缺的一部分。

随着信息技术的快速发展和在生产及生活中的广泛应用,各高校都开设了软件设计等专业,为信息产业提供必要的人才支撑。而作为衡量相关专业学生运用所学知识与技术,分析并解决实际问题能力的重要标准,其设计并实施的软件工程项目是最重要的学习成果。由于通常软件工程项目具有较大的开放性和一定的模糊性,计算机软件类专业课程教学过程中学生软件项目质量的评价问题,具有一定的挑战性^[4-5]。学生软件项目的评价问题也属于软件评价问题,相应的研究并不多见,目前主要通过神经网络法^[6]、灰色关联

收稿日期:2015-05-04

修回日期:2015-08-05

网络出版时间:2016-01-26

基金项目:辽宁省人文社会科学研究项目2014年立项课题(W2014351);辽宁省职教学会2015-2016年度科研规划立项课题(LZY15551)

作者简介:曹起武(1979-),男,硕士,副教授,研究方向为软件工程和数据库应用。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20160126.1517.010.html>

度法^[7]和模糊评价法^[8-9]等来解决对软件项目质量的评价问题。灰色关联度法与神经网络法的局限性在于要求各指标的评语集合具有相同个数的评语(元素),即各指标的评语集合具有相同粒度。同时,神经网络法还要有一定的训练样本,有时在实际决策过程中无法满足。学生软件项目质量具有难以度量和可见性差等特点,而模糊评价法可以很好地适应以上特点。针对以往采用模糊评价法所设计的算法,多数存在评价信息容易丢失的缺点^[8-9]。虽然很多设计者采用了多粒度的评语集方式进行解决,并把模糊矩阵的合成运算方式改为普通的矩阵运算方式。但是,上述算法具有随意性和粗糙性等缺点,基本没有解决评语集中元素按照各自实现的难易程度和所处的优劣程度分别赋值所引起的上述缺点。

需要指出,计算机软件类专业学生软件项目的评价问题具有一定的特殊性,主要体现在:针对一些评价准则(例如软件的特征等准则)无法给出精确的数值评价,只能给出定性的主观语言评价。值得注意的是,由于各评价准则具有不同的特性,因此它们相应的评价标准(如语言评价)并不相同,采用的语言评语集合不会相同,具有不同的粒度(各语言评语集合的元素个数不同)。例如,评价准则“软件项目前台界面设计的友好度,是否具有吸引力”的评价标准是:

- (1) 软件的界面简单、友好和有吸引力;
- (2) 软件的界面类似于一般的 Windows 桌面;
- (3) 软件的界面是文本性的,没有明显的功能提示。

评价准则“软件如需求分析那样工作吗?”的评价标准是:(1)是;(2)否。评价准则“软件的反应时间快吗?”的评价标准是:(1)非常快;(2)快;(3)一般;(4)慢;(5)非常慢。

目前,针对具有多粒度语言评价的计算机软件类专业学生软件项目的评价问题的研究并不多见。这种评价方法属于多属性决策,其特点是具有主观不确定性。文中认为,在计算机软件类专业课程教学过程中,如何评价学生软件项目质量的问题上,集结来自不同语言评语集合的主观语言评价,通过模糊评价法得到一个科学合理的评价结果,具有十分重要的意义,值得研究。

2 问题描述与语言评价基础

2.1 问题描述

在算法设计过程中,文中采用一系列符号来描述计算机软件类专业学生软件项目的多属性评价决策问题:

记 $S = \{S_1, S_2, \dots, S_m\}$ 为 m 个方案(即学生软件项

目)的集合, $m \geq 2$ 。

记 $C = \{C_1, C_2, \dots, C_n\}$ 为 n 个属性集合, $n \geq 2$ 。

$W = (w_1, w_2, \dots, w_n)$ 为属性的权重向量,其中, w_j 为属性 C_j 的权重, $0 \leq w_j \leq 1$ 且 $\sum_{j=1}^n w_j = 1$ 。

记 $\tilde{A} = [\tilde{a}_{ij}]_{m \times n}$ 为语言型决策矩阵。其中,将设计方案 S_i 中的属性 C_j , 用 \tilde{a}_{ij} 表示其短语形式的属性值。而在 $i = 1, 2, \dots, m, j = 1, 2, \dots, n$ 部分, \tilde{a}_{ij} 表示模糊语言类属性的评价值。在设计的过程中需要注意,应该在方案中采用不同语言评价集合的组合方式来应对不同属性所具有的不同特征。所以在该设计方案中, $\tilde{A} = [\tilde{a}_{ij}]_{m \times n}$ 作为语言型决策矩阵,其中不同属性上对各方案所做的评价值基本来自于不同粒度的语言评价集。采用该设计方法,需要解决的一个问题就是基准评语集合问题,也就是将来自各个语言评价集合中不同的语言属性评价值,利用一定标准规范为统一的语言评价集合,然后利用算法得出不同方案的综合评价数值并对其进行排序。

2.2 语言评价的有关概念

2.2.1 语言短语

语言短语的自身特征,可以很好地在模糊不定或是比较复杂的经济决策环境中,比较准确地表达出人们对于事物的主观判断或感觉^[10]。

定义 1:语言短语 \tilde{T} 在实体集合 R 中被定义为一组三角模糊数,表示为 (μ, α, β) , 将函数 $\mu_{\tilde{T}}: R^+ \rightarrow [0, 1]$ 作为语言短语 \tilde{T} 的隶属度函数。

$$\mu_{\tilde{T}}(x) = \begin{cases} \frac{x - \alpha}{u - \alpha}, & x \in [\alpha, u] \\ \frac{x - \beta}{u - \beta}, & x \in [u, \beta] \\ 0, & \text{其他} \end{cases} \quad (1)$$

其中, μ 为模值; β 与 α 分别代表语言短语 \tilde{T} 的上界值和下界值,并且 $\alpha \leq \mu \leq \beta$ 。

2.2.2 语言评价集合

各方案针对各属性评价之前,首先要对各个属性上采用的语言评价集合方案进行确定,根据属性间各自的特点,采用相适应的语言评价集合。

简单来说,方案中假设 $\text{REMARKSET} = \{t_0, t_1, \dots, t_g\}$ 为一个语言评价集合,并且这个集合为有序集合,由奇数个 $(g+1)$ 元素构成。例如, REMARKSET 是一个由 7 条评价语言短语构成的评价集合,即 $\text{REMARKSET} = \{t_0 = \text{“极差”}, t_1 = \text{“很差”}, t_2 = \text{“差”}, t_3 = \text{“一般”}, t_4 = \text{“好”}, t_5 = \text{“很好”}, t_6 = \text{“极好”}\}$ 。语言评价需要集合 REMARKSET 满足如下性质^[11-12]:

(1)具有有序性。当 $i \geq j$ 时,有 $t_i \geq t_j$,这里符号“ \geq ”表示“好于或等于”。

(2)存在逆运算算子“Neg”。当 $j = g - i$ 时,有 $\text{Neg}(t_i) = t_j$,其中的元素表示REMARKSET语言评价集合的元素个数。同时,语言评价集合REMARKSET中的最大短语为 t_g 。

(3)进行最大化 and 最小化运算。当 $t_i \geq t_j$ 时,有 $\text{Max}\{t_i, t_j\} = t_i$; $\text{Min}\{t_i, t_j\} = t_j$ 。

2.2.3 基准评语集合

在语言型的决策矩阵中,不同方案对于不同属性的语言评价可能来自不同的语言评价集合,也就是存在不同的粒度。这些评价不具备直接的可比性,必须通过某种规则或公式对其进行规范和统一。

文中采用的基准评语集合,或称为基准评价集合,内容为:“很棒”,“比较好”,“好”,“一般”,“较差”,“差”,“非常差”。该集合最后所生成的基准评价集合模糊集是由不同粒度的语言评价价值转化而来。为了在算法中方便表述,记基准评语集合为REMARKSET^B = { remark₀^B, remark₁^B, ..., remark_g^B }。

3 提出的方法

3.1 基准评语集合模糊集的转换

对于各个决策方案的评价,不同的属性可以采用不同的评价语言集合。解决这一问题的方法就是对其进行规范化。文中采用的方式是,将不同属性所对应的不同评语集合表达的各方案评价价值进行转换,形成基准评语集合模糊集。转换时首先给定关于属性 C_j 的评价语言集合REMARKSET^j,然后针对方案 S_i 的语言短语评价价值 $\text{remark}_i^j, \text{remark}_i^j \in \text{REMARKSET}^j, i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, K^j$ (K^j 为语言短语集合REMARKSET^j中语言短语的个数),采用下面的函数 τ 将 remark_i^j 转换为REMARKSET^B的模糊集^[13]。

$$\tau: \text{remark}_i^j \rightarrow F_i^j(\text{REMARKSET}^B) \tag{2}$$

其中, $F_i^j(\text{REMARKSET}^B)$ 为REMARKSET^B上的模糊集,并且

$$\tau(\text{remark}_i^j) = \{ (\text{remark}_0^B, \gamma_0^{ij}), (\text{remark}_1^B, \gamma_1^{ij}), \dots, (\text{remark}_g^B, \gamma_g^{ij}) \} \tag{3}$$

$$\gamma_l^{ij} = \max_{x \in \text{REMARKSET}^j} \min \{ \gamma(\text{remark}_i^j), \gamma_l(\text{remark}_x^B) \}, l = 0, 1, \dots, g \tag{4}$$

其中, $\gamma(\cdot)$ 与 $\gamma_l(\cdot)$ 分别为语言短语 remark_i^j 和 remark_l^B 的隶属度函数。

通过该方法,将语言混杂型决策矩阵 $\tilde{A} = [\tilde{a}_{ij}]_{m \times n}$ 中,基于不同语言评价集合的属性评价价值 \tilde{a}_{ij} (即

remark_i^j) 转化为基准评语集合REMARKSET^B的模糊集 $F_i^j(\text{REMARKSET}^B)$ 。

3.2 综合评价值的计算

学生软件质量的综合评价值决策方案通过加权和法计算得来,计算需要用到由3.1中得到的决策矩阵,该矩阵以基准评语集合REMARKSET^B的模糊集形式进行表达。

$$\text{Overall}_i = \sum_{j=1}^n \tau(\text{remark}_i^j) w_j, i = 1, 2, \dots, m \tag{5}$$

其中, w_j 为属性 C_j 的权重, $j = 1, 2, \dots, n$ 。

在质量评价方案中,各属性的权重可以邀请专家根据层次分析法给出,或者由决策者根据经验确定。

3.3 方案的排序与等级分配

评价所需要的基准等级集合REMARKSET^B的模糊集,记为

$$\text{Overall}_i = \{ (\text{remark}_0^B, o_0^{ij}), (\text{remark}_1^B, o_1^{ij}), \dots, (\text{remark}_g^B, o_g^{ij}) \}, i = 1, 2, \dots, m \tag{6}$$

该模糊集与由3.2计算得到的决策方案 S_i 的综合评价值 $\text{Overall}_i (i = 1, 2, \dots, m)$ 相同。根据上述结果,通过计算可以得到决策方案 S_i 的单点值型综合评价值 d_i 。

$$d_i = \frac{\sum_{k=0}^g ko_k^{ij}}{\sum_{k=0}^g o_k^{ij}}, i = 1, 2, \dots, m \tag{7}$$

决策方案可以按照 d_i 值的降序排列。

4 算例分析

文中算例针对5个计算机软件类专业学生课程学习过程中信息系统团队作业软件项目(记为方案 S_1, S_2, S_3, S_4 和 S_5)展开多属性评价,这些属性分别为:“软件的界面友好而有吸引力吗?”(C_1),“软件的反应时间快吗?”(C_2),“安全性”(C_3)和“软件如需求分析那样工作吗?”(C_4)。

针对属性“软件的界面友好而有吸引力吗?”采用语言集合{“软件的界面简单、友好和有吸引力”(简记为“友好”),“软件的界面类似于一般的Windows桌面”(简记为“一般”)和“软件的界面是文本性的,没有明显的功能提示”(简记为“差”)}来评价。针对属性“软件的反应时间快吗?”采用语言集合{“非常快”,“快”,“一般”,“慢”,“非常慢”}来评价。针对属性“软件的安全性”采用语言集合{“非常高”,“较高”,“高”,“一般”,“低”,“较低”,“非常低”}来评价。而针对属性“软件如需求分析那样工作吗?”的评价,采用语言集合{“是”,“否”}。

具体数据如表1所示。

表 1 学生软件项目的多属性评价数据

| 属性 | 软件的界面友好而有吸引力吗 | 软件的反应时间快吗 | 安全性 | 软件如需求分析那样工作吗 |
|-------|---------------|-----------|-----|--------------|
| S_1 | 友好 | 非常快 | 高 | 是 |
| S_2 | 一般 | 一般 | 高 | 否 |
| S_3 | 一般 | 快 | 较高 | 是 |
| S_4 | 一般 | 非常慢 | 一般 | 是 |
| S_5 | 差 | 快 | 一般 | 否 |

使用文中算法,以 7 个语言评语的集合为基准评语集合,将表 1 中的各列数据进行规范,可以得到各方案的综合评价值分别为:

$$\text{Overall}_1 = \{ (\text{remark}_0^B, 0), (\text{remark}_1^B, 0), (\text{remark}_2^B, 0), (\text{remark}_3^B, 0.126\ 9), (\text{remark}_4^B, 0.553\ 6), (\text{remark}_5^B, 0.521\ 9), (\text{remark}_6^B, 0.75) \}$$

$$\text{Overall}_2 = \{ (\text{remark}_0^B, 0.313\ 4), (\text{remark}_1^B, 0.376\ 0), (\text{remark}_2^B, 0.458\ 5), (\text{remark}_3^B, 0.559\ 7), (\text{remark}_4^B, 0.585\ 4), (\text{remark}_5^B, 0.174\ 5), (\text{remark}_6^B, 0.063\ 4) \}$$

$$\text{Overall}_3 = \{ (\text{remark}_0^B, 0.063\ 4), (\text{remark}_1^B, 0.126\ 9), (\text{remark}_2^B, 0.186\ 6), (\text{remark}_3^B, 0.414\ 6), (\text{remark}_4^B, 0.512\ 1), (\text{remark}_5^B, 0.759\ 9), (\text{remark}_6^B, 0.414\ 6) \}$$

$$\text{Overall}_4 = \{ (\text{remark}_0^B, 0.160\ 9), (\text{remark}_1^B, 0.339\ 4), (\text{remark}_2^B, 0.399\ 1), (\text{remark}_3^B, 0.660\ 9), (\text{remark}_4^B, 0.309\ 7), (\text{remark}_5^B, 0.313\ 4), (\text{remark}_6^B, 0.313\ 4) \}$$

$$\text{Overall}_4 = \{ (\text{remark}_0^B, 0.5), (\text{remark}_1^B, 0.373\ 1), (\text{remark}_2^B, 0.246\ 3), (\text{remark}_3^B, 0.470\ 6), (\text{remark}_4^B, 0.202\ 4), (\text{remark}_5^B, 0.196\ 4), (\text{remark}_6^B, 0.101\ 2) \}$$

进一步进行计算,根据式(7)得出各决策方案的单点值型综合评价值分别为: $d_{1i} = 4.970\ 6$, $d_{2i} = 2.600\ 2$, $d_{3i} = 4.067\ 4$, $d_{4i} = 3.126\ 7$, $d_{5i} = 2.237\ 4$ 。这样得出各决策方案的排序为: $S_1 > S_3 > S_4 > S_2 > S_5$ 。

5 结束语

为了克服当前针对计算机软件类专业学生软件项目的质量评价问题存在的不足,文中进行了模糊多属性评价研究。该算法针对不同属性特征制订了有针对性的评语集合,以此对不同的项目进行评价,解决了学生软件项目的开放性和模糊性特征。在完成基准评语集合的基础上,通过使用模糊运算法将不同评语集合的属性评价值转换为相对统一的基准评语集合模糊集。该方法实现了对不同粒度属性评价值的规范,并

在规范的基础上得到了各软件项目质量的综合评价值及质量排序。同时,该评价方式可以对学生软件项目的多方面进行评价,使得考核者有的放矢地解决学生项目中出现的普遍性问题^[14]。

可以说,文中提出的方法具有普遍的适用性和推广性,也克服了当前针对一般软件评价方法存在的不足,可以根据实际情况允许不同的属性采用不同的评语集合。此外,该算法易于操作,并且容易在计算机系统上实现,具有一定的可应用性。同时需要指出的是,由于使用该评价方法对项目进行评价的过程中有多个评委专家参加,所以此类问题的群决策支持研究将是该研究今后的主要发展方向,其解决实际问题的效果值得期待。

参考文献:

[1] 赵川平. 重视学生学习成果研究提升高等工程教育质量[J]. 中国高教研究, 2009(7): 90-91.

[2] 黄海涛. 美国高等教育中的“学生学习成果评估”: 内涵与特征[J]. 高等教育研究, 2010(7): 97-104.

[3] 黄海涛. 美国“学生学习成果评估”研究评析[J]. 清华大学教育研究, 2011(2): 119-124.

[4] Kwok R C W, Zhou D N, Zhang Q, et al. A fuzzy multi-criteria decision making model for IS student group project assessment[J]. Group Decision and Negotiation, 2007, 16(1): 25-42.

[5] 刘翠红. 《管理信息系统》课程考试内容与考试方式研究[J]. 中国信息界, 2011(3): 64-66.

[6] 杨根兴, 高大启, 宋国新. 基于神经网络的软件质量综合评价方法[J]. 华东理工大学学报: 自然科学版, 2004, 30(3): 292-295.

[7] 冯建湘, 唐 嵘, 王双维, 等. 软件质量灰色定量评价模式研究[J]. 哈尔滨工业大学学报, 2005, 37(5): 639-641.

[8] 董剑利. 软件质量的模糊综合评判研究与应用[J]. 计算机工程与设计, 2006, 27(12): 2196-2199.

[9] 陆 鑫, 廖建明. 基于模糊集理论的软件质量评估研究[J]. 电子科技大学学报, 2007, 36(3): 652-655.

[10] Zadeh L A. A computational approach to fuzzy quantifiers in natural languages[J]. Computing Mathematics Applications, 1983, 9(9): 149-184.


[11] Herrera F, Herrera-Viedma E. Linguistic decision analysis: steps for solving decision problems under linguistic information[J]. Fuzzy Sets and Systems, 2000, 115(1): 67-82.

[12] 徐泽水. 基于语言标度中术语指标的多属性群决策法[J]. 系统工程学报, 2005, 20(1): 84-88.

[13] Herrera F, Martinez L, Sanchez P J. Managing non-homogeneous information in group decision making[J]. European Journal of Operational Research, 2005, 166: 115-132.

[14] 胡恩华, 单红梅, 陈 燕. 企业核心竞争力的识别及综合模糊评价[J]. 系统工程, 2004, 22(1): 48-51.

学生软件项目质量的模糊评价研究

作者: [曹起武](#), [CAO Qi-wu](#)
作者单位: [辽宁机电职业技术学院 信息工程系, 辽宁 丹东, 118009](#)
刊名: [计算机技术与发展](#) 
英文刊名:
年, 卷(期): 2016 (2)

引用本文格式: [曹起武](#), [CAO Qi-wu](#) [学生软件项目质量的模糊评价研究](#) [期刊论文] - [计算机技术与发展](#) 2016 (2)