

多特征层次化答案质量评价方法研究

崔敏君¹ 段利国¹ 李爱萍^{1,2}

(太原理工大学计算机科学与技术学院 太原 030024)¹ (武汉大学软件工程国家重点实验室 武汉 430072)²

摘 要 社交媒体中的问答对可以为自动问答系统提供答案,但有些答案的质量不高,因此答案质量评价方法具有研究价值。已有的评价方法没有考虑问题类别特征,对不同类型的问题采用统一的评价方法。因此提出了一个层次分类模型。首先分析问题类型;然后提取文本、非文本、语言翻译性、答案中的链接数4类特征,依据特征分类影响力随问题类型不同而不同这一客观现象,采用逻辑回归算法对各类型问题的答案质量进行评价,取得了较好的实验效果;最后分析了影响各类问题答案质量的主要特征。

关键词 层次分类模型,问题类别,答案质量评价,特征分析

中图分类号 TP391.1

文献标识码 A

DOI 10.11896/j.issn.1002-137X.2016.1.022

Research on Multi-features Hierarchical Answer Quality Evaluation Method

CUI Min-jun¹ DUAN Li-guo¹ LI Ai-ping^{1,2}

(Department of Computer Science and Technology, Taiyuan University of Technology, Taiyuan 030024, China)¹

(State Key Laboratory of Software Engineering, Wuhan University, Wuhan 430072, China)²

Abstract Social media question-answer pairs can provide the answer to the automatic question and answering system, but the quality about some of the answers is not so high. So the evaluation method of answer quality has the research value. The existing evaluation methods without considering the problem of question types feature use the uniform evaluation method for different question types. This paper presented a hierarchical classification model. Firstly we analyzed the types of question, and then extracted four features of text, non-text, language translation, number of links in the answer. According to this objective phenomenon that the influence of feature classification varies with the types of different questions, we used logistic regression algorithm to evaluate various types of answer quality based on these features, achieving good results. Finally the main features that influence the answer quality of all kinds of questions were analyzed.

Keywords Hierarchical classification model, Question types, Answer quality evaluation, Feature analysis

1 引言

答案抽取是自动问答系统研究的难点。充分利用如 Yahoo Answer、Twitter、Facebook、百度知道等问答社区中的问答对为用户推荐答案是自动问答系统实现的有效方法。然而由于种种原因,社区问答中不少问答对质量堪忧,这些低质量的答案出现在问答系统中会严重影响问答系统的质量。如何评价问答对的质量是本文研究的目标。

在评价答案质量方面,Jeon等利用基于文本特征的最大熵模型^[1]和核密度估计方法^[2]评价答案的质量;Zhu等^[3]提出了一个基于13维特征的多维质量评估模型来分析和评价答案质量;刘高军等^[4]利用百度知道中“问答对”的文本、用户、统计信息、问题和答案的关联性4类特征进行分类,根据分类结果对问答对质量进行分析;田作辉、Li^[5,6]利用排序学习算法选出高质量答案;A. Figueroa、Wang^[7,8]等依据相似问

题提供答案的思想,提出了深度学习模型来选出高质量答案。

综上所述,现有方法大都依据文本或非文本特征对答案质量进行评价,均未注意到不同特征对评价答案质量的影响力随问题类别而异这一现象。例如,对于是非类问题,答案中大多会出现“是”或者“否”;对于事实类问题,答案中往往会有一些名词;对于建议类问题,答案中一般会出现“我认为”、“我觉得”等一系列词。所以,可以充分利用问题类别特征分析答案质量。

本文提出了一个层次分类模型进行答案质量评价。该模型利用文本特征、非文本特征、语言翻译性、答案中的链接数4类特征,采用逻辑回归算法,实现基于特征集的答案质量分类器,取得了较好的效果。

2 层次分类模型

层次分类模型如图1所示,包含分析层和评价层。分析

到稿日期:2015-03-21 返修日期:2015-07-01 本文受武汉大学软件工程国家重点实验室开放课题项目(SKLSSE2012-09-30),山西省自然科学基金项目(2013011015-2)资助。

崔敏君(1991—),女,硕士生,CCF学生会员,主要研究方向为自然语言处理,E-mail:279546713@qq.com;段利国(1970—),男,博士,副教授,CCF高级会员,主要研究方向为中文信息处理;李爱萍(1974—),女,博士,副教授,CCF会员,主要研究方向为软件形式化及智能信息处理。

层通过对问题分类提取问题类别特征,并提取问答对的文本特征、非文本特征、语言翻译性、答案中的链接数等特征。评价层利用上述特征,采用逻辑回归算法^[9]对答案质量进行评价。

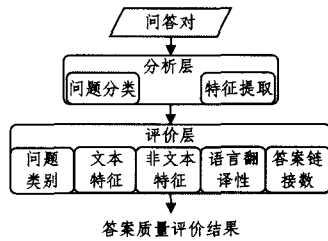


图1 层次分类模型设计

3 特征分析

3.1 问题类别特征分析

问题分类对答案质量评价至关重要。首先,问题分类可以大大缩小答案的范围;其次,问题分类可以决定答案评价策略。例如问题“保温瓶为什么可以保温?”为原因类,可以根据原因类的答案质量评价策略去评价答案的好坏。

疑问词作为问句的特征标识,在问题分类中起重要作用。通过对大量问题的分析和归纳,不同类别的问题其疑问词及句式如表1所列。本文主要针对事实类、过程类、建议类、是非类、原因类、描述类以及列举类问题进行研究。

表1 问题类别、对应疑问词及句式举例

问题类型	疑问词举例	对应句式	问题举例
原因类	为何 为什么	疑问词+原因	保温瓶为什么可以保温?
列举类	哪些	疑问词+名词	参加北京奥运会的国家都有哪些?
描述类	什么 怎样 怎么样	疑问词+是动词 疑问词+名词	什么是转述句? 最近天气怎么样?
过程类	怎 怎么样 如何	疑问词+方法/ 一般动词	世界大战是如何发生的?
事实类	哪里 哪儿 多少 什么	疑问词+一般名词 (时间/地点/人名等)	2014年世界杯在哪里举办?
建议类	有何 有什么 怎样 怎么办	疑问词+看法/建议?	你对学校或老师在教育孩子的教育方面有何意见和建议?
是非类	是否	疑问词+名词/从句	是不是马克思写了《资本论》?
其他类			

分析表1可以发现,仅凭疑问词对问题类型进行评价具有较大的局限性。例如依据“为什么”、“为何”等可以确定问题类型是原因类,依据“哪些”可以确定为列举型,而通过问题中的“怎样”、“如何”、“什么”很难判别问题类型,因此本文利用疑问词和依存句法特征对问题进行分类。

依存句法分析是指在给定文法下分析自然语言的层次结构^[10],通过分析语言单位内成分之间的依存关系揭示其句法结构。本文采用了哈尔滨工业大学的依存句法分析器,通过分析句法结构,使用基于疑问词-问句焦点特征的方法^[10]进行问题分类,该方法步骤如下。

- ①对问句分词并提取疑问词;
- ②对问句进行句法分析,得到依存弧和弧之间的关系类型;

③提取问句主干和疑问词附属成分作为问题分类的焦点特征。

表2为一个问题的词法、句法结构及其疑问词和焦点特征实例。

表2 一个问题的词法、句法结构及其疑问词和焦点特征的实例

分析	举例
问句	哪个国际人道主义机构对难民进行了食品援助?
分词与词性标注	哪个/r 国际/n 人道主义/n 机构/n 对/p 难民/n 进行/v 了/u 食品/n 援助 v? /wp
疑问词	哪个
句法依存结构分析	
问题焦点特征	机构 进行 援助
问题类型	事实类

实验表明,该方法可以去掉问句分类的噪音,提取特定类别的焦点特征,提高问题分类准确率。

3.2 4类特征分析

(1)文本特征

①表层语言特征

表层语言特征是指不需要经过复杂的分析就能从答案上下文中提取出来的特征。在社区问答系统相关研究中发现,表层语言特征对于答案质量分析有极其重要的作用^[11]。好的答案应包含合理的表层语言特征^[12]。表层语言特征包括问题的长度、答案的长度、问题中的名词数、答案中的名词数、问题中的动词数、答案中的动词数、答案中非停用词数7项特征。

②答案问题长度比

此特征给出了答案和问题长度的比值。在某些问题类别中好的答案应有相对大的问题答案长度比。例如对于过程类问题,其答案往往会对问题中事件的发展进行详细的描述,需要加一些解释去支持它的可信度,所以答案的长度要远大于问题的长度。

③相似特征

在问答对里答案和问句之间存在一定的相似性,相似度高的答案往往质量会比较高。本文利用基于向量空间模型的方法来计算问题和答案之间的相似度^[13]。计算公式如下:

$$\text{sim}(w, w') = \frac{\sum_{i=1}^n (w_i \times w'_i)}{\sqrt{\sum_{i=1}^n w_i^2 \times \sum_{i=1}^n w'^2_i}} \quad (1)$$

w 为一个问题句子,其对应的向量为 $w = \{w_1, w_2, \dots, w_n\}$;对应的待比较的答案句子为 w' ,其对应的向量为 $w' = \{w'_1, w'_2, \dots, w'_n\}$ 。

④标点符号密度

此特征用于表示问答对里一些特殊符号的密度。好的答案里应尽可能少地出现无用标点和特殊符号,例如:“(一)”。该特征包含问题中的特殊标点符号密度和答案中特殊标点符号密度两项内容。

(2)非文本特征

问答社区中用户的作用举足轻重,所以本文使用用户在问答社区中的属性作为非文本特征进行分析。非文本特征包含以下内容。

浏览次数:某用户在平台中提出的某个问题被查看的次数。

采纳数:某用户在平台中对某个问题的回答被采纳为正确答案的数目。

采纳率:某用户在平台中总回答被采纳为正确答案数与总回答数之间的比值。

用户等级:某用户在网站中的级别,越活跃等级越高。

经验值:根据用户的等级而定,等级越高经验值越高^[14]。

财富值:根据用户参与平台的活跃度而确定,例如用户在平台中的回答被采纳为正确答案或做任务即可获得相应财富值。

点赞数:某用户在平台中回答某个问题被其他用户点赞的数目。

被踩数:某用户在平台中回答某个问题被其他用户反对的数目。

点赞数所占比:点赞数与点赞数和被踩数总和的比值。

回答数:某用户在平台中回答其他用户提出问题的总数。

回答被赞同数:某用户在平台中回答的全部问题的总点赞数。

提问数:某用户在平台中总的提问数。

(3)语言翻译性特征

此特征表示问题词和答案词之间的语义关联的概率,对于某些特定类别,在问题和答案中期望有上下文的语义关联。例如原因类的问题,答案中期望包含因为等词语。

(4)答案中的链接数特征

此特征表示答案中含有链接数,加入该特征对分析特定问题类别的质量评价有重要影响。例如对于事实类或过程类问题,答案需要提供一些背景知识或证据来提高答案的可信度,而这些证据往往很长,所以在答案中加入链接来说明。

4 实验及结果分析

4.1 实验数据及评价指标

作者从百度知道采集了从2006年7月到2007年4月包括7种问题类别的6000多个问答对,平均每个问题对应3.7

个答案,如表3所列。同时提取这些问答对中的文本、非文本、语言翻译性、答案是否链接等特征作为数据,训练分类器,评价答案的质量。本文还对答案质量进行人工标注。

表3 各问题类型问答对数据统计

类别	原因类	列举类	描述类	过程类	事实类	建议类	是非类	总数
高质量	698	592	690	671	722	893	452	4718
低质量	145	178	210	211	208	200	143	1295
总数	843	770	900	882	930	1093	595	6013

本实验使用准确率、AUC作为答案质量评价指标。

4.2 分类器选择

实验选用了一些流行的机器学习算法,包括支持向量机、逻辑回归、序列最小优化、投票机制^[12],由于各算法原理不同,使用的数据类型也不同,为了提高实验的可信度和可对比度,实验用4种算法处理本文提取出来的特征数据集,采用十折交叉验证的方法,根据实验结果选择最合适的算法作为答案质量评价的分类器。表4呈现了各分类器的性能,实验结果显示,在所有机器学习算法中逻辑回归在各项评价指标上的表现都是最好的,所以选用逻辑回归作为本实验分类器。

表4 各分类器性能比较(分类准确度P,召回率R,F值和ROC曲线面积AUC)

分类器	P	R	F	AUC
SVM	0.658	0.811	0.727	0.500
逻辑回归	0.839	0.828	0.832	0.840
投票感知	0.656	0.800	0.721	0.491
最小序列优化	0.744	0.759	0.726	0.678

4.3 答案质量分析结果

本实验使用weka软件作为工具,采用十折交叉验证的方法,基于文本(T)、非文本(N)、语言翻译性(F)、答案是否有链接(L)等特征的不同组合,利用逻辑回归算法对各问题类别的答案质量进行评价。实验结果如表5所列。

表5 不同特征组合对各类问题答案质量的评价结果

问题类别	T		T+N		T+N+F		T+N+L		T+N+F+L	
	准确率	AUC	准确率	AUC	准确率	AUC	准确率	AUC	准确率	AUC
原因类	84.02%	0.826	84.27%	0.840	84.35%	0.853	84.27%	0.838	84.38%	0.856
列举类	77.03%	0.682	78.10%	0.717	78.09%	0.727	78.10%	0.732	78.01%	0.724
描述类	82.14%	0.840	83.32%	0.870	83.32%	0.883	84.21%	0.892	82.82%	0.850
过程类	80.23%	0.788	81.35%	0.802	81.35%	0.822	82.04%	0.814	82.10%	0.819
事实类	78.34%	0.767	78.38%	0.772	78.38%	0.782	80.32%	0.828	80.72%	0.892
建议类	83.23%	0.840	83.72%	0.867	84.04%	0.852	83.32%	0.847	84.30%	0.877
是非类	66.32%	0.668	75.23%	0.707	75.23%	0.712	75.23%	0.731	75.31%	0.762
平均值	78.75%	0.773	80.62%	0.796	80.68%	0.806	81.06%	0.811	81.09%	0.825

从实验结果可以看出,对于原因和建议类问题,用户回答此类问题时经常用一些特殊的短语,例如:“因为……”、“我认为……”、“我的意见是……”、“你应该……”等词。所以加入语言翻译性特征后,评价答案质量的准确率和AUC都比加入前有所提升。

对于事实类、过程类、描述类问题,在答案中需要一些证据和事实,所以很多用户回答该类问题经常会使用链接,因此对于此类问题,加入答案链接特征后,评价答案质量的准确率和AUC都比加入前有所提升。

对于是非类问题,由于该类问题的答案一般很短,只是回

答“是”或者“否”,因此文本特征对于该类问题的准确率不高,评价该类问题时可能需要用户更多的情感特征。而对于列举类问题,大部分答案还是高质量的,但有些答案并未将问题所需全部列举出来,需要综合几条答案才可以得到用户满意的答案,所以此类问题并不容易分析。

每个类别在文本特征的基础上加上非文本特征,准确率和AUC都有所提升,说明非文本特征对于质量评价是很有用的。有些类别的问题在加入全部特征后,其准确率有所下降,经过分析发现,对于少数问答对,加入所有特征反而会产生冗余信息进而影响其性能。但对7类问题的准确率和

AUC 求平均值发现,加入各特征组合后,总体性能都有所提高,其准确率均在 75%以上,说明本文的层次分类模型能够有效地评价答案质量。

综合分析上述 AUC 数据发现,融合多种特征确实能提高答案质量评价性能。因为篇幅有限,本文只列出是非类问题的 ROC 曲线。如图 2 所示,可以清晰看出依次加入文中特征后,其 ROC 曲线下方面积(也就是 AUC)的值是逐渐变大的,也就是说融合多特征确实能提高答案质量评价性能。

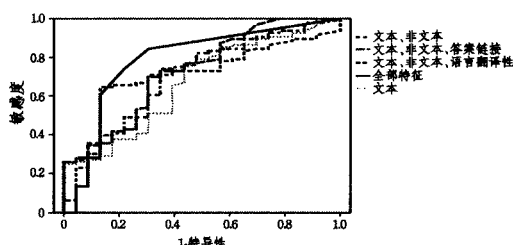


图 2 非类问题答案质量评价 ROC 曲线

4.4 影响质量评价的特征分析

按照回归系数(回归系数大于 10),表 6 列出了影响各类问题答案质量评价的最有影响力特征。

表 6 各问题类型答案质量评价的最有影响力特征

问题类别	特征	特征个数
原因类	点赞数所占比 采纳率 语言翻译性	10
	答案中非停用词数 相似性 答案问题长度比	
	答案中名词数 答案长度 用户等级 答案中动词数	
列举类	点赞数所占比 采纳率 答案中非停用词数	8
	答案中名词数 答案长度 用户等级	
	答案问题长度比 问题中名词数	
描述类	点赞数所占比 采纳率 答案中链接数	10
	答案中非停用词数 相似性 答案问题长度比	
	答案中名词数 答案长度 用户等级 答案中动词数	
过程类	点赞数所占比 采纳率 答案中链接数	11
	答案中非停用词数 相似性 答案问题长度比	
	答案中名词数 答案长度 用户等级 答案中动词数 问题中名词数	
事实类	点赞数所占比 采纳率 答案中链接数	9
	答案中非停用词 答案问题长度比 答案中名词数	
	答案长度 用户等级 答案中动词数	
建议类	点赞数所占比 采纳率 语言翻译性 答案中非停用词	8
	答案问题长度比 答案长度 用户等级 答案中动词数	
	答案问题长度比 答案长度 用户等级 答案中动词数	
是非类	点赞数所占比 采纳率 用户等级	6
	点赞数 答案长度 问题答案长度比	

依据表 6 可知,在答案质量预测中,非文本特征往往很重要。分析数据可得,在社区问答系统中,对于用户的回答,好的答案往往会获得大家的支持。而问答对中标点符号密度没有进入前列,经过分析发现用户在回答时大多使用规范的格式,很少使用特殊标点符号,所以该特征对质量评价影响不大。而其他特征没有进入前列说明这些特征可能并不适合该模型。

结束语 本文针对不同问题类别的答案质量评价准则不同的情况,提出了一个层次分类模型来评价答案质量。通过分析各类别问题,发现评价不同问题类别的答案质量需要包含不同特征。在社区问答系统中,两种最难评价的问题类别是列举类和是非类,需要更多证据和背景知识评价它们。对于其他类别的答案,例如描述、建议、过程、原因类等问题,使用本文的方法评价答案质量的结果较好。经过分析影响质量评价的特征,得出了影响各问题类别答案质量的最有影响力

特征。发现质量好的答案需要完整的事实和证据,被写在好的结构中,以及有好的用户反馈。综上所述,本文利用层次分类模型对答案质量的评价取得了很好的结果。

下一步工作计划将情感分析加入层次分类模型中,用于对各类别问题答案质量的分析。进一步计划将本文模型扩展到其他 UGC 内容,如微博、TWEET、论坛。

参考文献

- [1] Berger A, Della Pietra S, Della Pietra V J. A maximum entropy approach to natural language processing[J]. Computational Linguistics, 1996, 22(1): 39-71
- [2] Hwang J N, Lay S R, Lippman A. Nonparametric multivariate density estimation: A comparative study[J]. IEEE Transactions of Signal Processing, 1994, 42(10): 2795-2810
- [3] Zhu Z, Bernhard D, Gurevych I. A Multi-Dimensional model for assessing the quality of answers in social Q&A sites[R]. Technische Universität UKP Lab, 2009
- [4] Liu G J, Ma Y Z, Duan J Y. Assessment of Quality of “Questions and Answerw” in Community Q&A System[J]. Journal of North China University of Technology, 2012, 24(3): 31-35 (in Chinese)
- [5] 刘高军, 马祝忠, 段建勇. 社区问答系统中“问答对”的质量评价[J]. 北方工业大学学报, 2012, 24(3): 31-35
- [6] Tian Z H. Answer selection for non-factoid question[D]. Harbin; Harbin Institute of Technology, 2013 (in Chinese)
- [7] 田作辉. 非事实类问题的答案选取[D]. 哈尔滨: 哈尔滨工业大学, 2013
- [8] Li S G, Manandhar S. Improving question recommendation by exploiting information need[C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics. Portland, Oregon: ACL, 2011: 1425-1434
- [9] Figueroa A, Neumann G. Category-specific models for ranking effective paraphrase in community Question Answering[J]. Expert System With Applications, 2014, 41(10): 4730-4742
- [10] Wang B X, Liu B Q, Wang X L, et al. Deep learning approaches to semantic relevance modeling for Chinese question-answer pairs[J]. ACM Transactions on Asian Language Information Processing, 2011, 10(4): 1-16
- [11] Li C, Chao W H, Chen X M. Quality evaluation and prediction for question and answer in chinese community question answering[J]. Computer Science, 2011, 38(6): 230-236 (in Chinese)
- [12] 李晨, 巢文涵, 陈小明. 中文社区问答中问题答案质量评价和预测[J]. 计算机科学, 2011, 38(6): 230-236
- [13] Wen X, Zhang Y, Liu T, et al. Syntactic Structure Parsing Based Chinese Question Classification[J]. Journal of Chinese Information Processing, 2006, 20(2): 33-39 (in Chinese)
- [14] 文勋, 张宇, 刘挺, 等. 基于句法结构分析的中文问题分类[J]. 中文信息学报, 2006, 20(2): 33-39
- [15] Hu H F. Research on the method of the feature representation and fusion for evaluating the quality of user generated answers [D]. Harbin; Harbin Institute of Technology, 2013 (in Chinese)
- [16] 胡海峰. 用户生成答案质量评价中的特征表示及融合研究[D]. 哈尔滨: 哈尔滨工业大学, 2013

$LbnR_{0.3 \wedge V1}^{\geq} X = \{(x_1, 0.82), (x_7, 0.76), (x_9, 0.85), (x_{11}, 0.63), (x_{20}, 0.47)\}$

$bnR_{0.3 \wedge V1}^{\geq} X = \{(x_1, 0.82), (x_7, 0.76), (x_9, 0.85), (x_{11}, 0.63), (x_{13}, 0.85), (x_{14}, 0.85), (x_{19}, 0.85), (x_{20}, 0.47)\}$

方法2 由定理4求解

(1)由 $\beta=0.3, k=1$ 可以得到 $k/(1-\beta)=10/7, k/\beta=10/3$, 计算结果如下:

$posR_{0.3 \wedge V1}^{\geq} \tilde{X} = \emptyset$

$negR_{0.3 \wedge V1}^{\geq} \tilde{X} = \{(x_1, 0.18), (x_2, 1), (x_3, 1), (x_4, 1), (x_5, 1), (x_6, 1), (x_7, 0.24), (x_8, 1), (x_9, 0.15), (x_{10}, 1), (x_{11}, 0.37), (x_{12}, 1), (x_{13}, 0.15), (x_{14}, 0.15), (x_{15}, 1), (x_{16}, 1), (x_{17}, 1), (x_{18}, 1), (x_{19}, 0.15), (x_{20}, 0.15)\}$

$UbnR_{0.3 \wedge V1}^{\geq} \tilde{X} = \{(x_{13}, 0.85), (x_{14}, 0.85), (x_{19}, 0.85)\}$

$LbnR_{0.3 \wedge V1}^{\geq} \tilde{X} = \{(x_1, 0.82), (x_7, 0.76), (x_9, 0.85), (x_{11}, 0.63), (x_{20}, 0.47)\}$

(2)利用(1)所求结果及定理2可以计算得到:

$R_{\beta \wedge V k}^{\geq} \tilde{X} = \{(x_{13}, 0.85), (x_{14}, 0.85), (x_{19}, 0.85)\}$

$R_{0.3 \wedge V1}^{\geq} \tilde{X} = \{(x_1, 0.82), (x_7, 0.76), (x_9, 0.85), (x_{11}, 0.63), (x_{20}, 0.47)\}$

由上述可知,两种结果是一致的,但是其求解过程却不一样,所以可以通过定义的方式进行求解,也可以利用定理4进行求解。在实际的工程应用中可以根据掌握的数据类型,适当地选择求解的方式,也可以根据需求选择求解方式,例如,若只需求解正域、负域、左边界域、右边界域以及边界域中的一种或者几种,则可以选择定理求解,而无需求解上、下近似;如果只需求解上、下近似,则可以从定义出发,以减少计算的步骤。

结束语 本文在序信息系统中,在分明的近似空间中,以模糊概念为背景,将变精度与程度“逻辑且”和“逻辑或”融合起来,建立了一种新的粗糙模糊集模型,然后深入研究了该模型的基本性质;最后通过对实际案例的计算,比较了通过定义进行计算和通过定理进行计算的优缺点,在实际的工程应用中,可以按需求选择适合的计算方法,从而减少计算步骤,达到高效计算的目的。

参考文献

- [1] Zadeh L A. Fuzzy Sets and Information Granularity[J]. Advances in Fuzzy Set Theory & Application, 1996; 3-18
- [2] Pawlak Z. Rough Sets[J]. International Journal of Computer and Information Sciences, 1982, 11(5): 341-356
- [3] Zhang Xian-yong. Study on several rough set models and their algorithms based on logical combinations of precision and grade [D]. Chengdu: Sichuan Normal University, 2011 (in Chinese)

张贤勇. 基于精度与程度逻辑组合的几类粗糙集模型及其算法研究[D]. 成都: 四川师范大学, 2011

- [4] Zhang Xian-yong, Xiong Fang, Mo Zhi-wen. Properties of Approximation Operators of Logical Difference Operation of Grade and Precision[J]. Mathematics in Practice and Theory, 2011, 18(41): 209-213
- [5] Zhang Xian-yong, Mo Zhi-wen, Xiong Fang, et al. Comparative study of variable precision rough set model and graded rough set model[J]. International Journal of Approximate Reasoning, 2012, 53(1): 104-116
- [6] Zhang Xian-yong, Miao Duo-qian. Two basic double-quantitative rough set models of precision and grade and their investigation using granular computing [J]. International Journal of Approximate Reasoning, 2013, 54(8): 1130-1148
- [7] Yu Jiang-huang, Xu Wei-hua. Rough Set Based on Logical Disjunctoperation of Variable Precision and Grade in Ordered Information System[J]. Journal of Frontiers of Computer Science and Technology, 2015, 9(1): 112-118 (in Chinese)
- 余建航, 徐伟华. 序信息系统中变精度与程度的“逻辑或”粗糙集[J]. 计算机科学与探索, 2015, 9(1): 112-118
- [8] Yao Y Y, Lin T Y. Generalization of rough sets using modal logics[J]. Intelligent Automatic and Soft Computing, 1996, 2: 103-120
- [9] Zhang W X, Wu W Z, Liang J Y, et al. Rough set theory and method [M]. Beijing: Science press, 2001 (in Chinese)
- 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001
- [10] An A, Shan N, Chan C, et al. Discovering rules for water demand prediction: an enhanced rough-set approach[J]. Engineering Application of Artificial Intelligence, 1996, 9: 645-653
- [11] Xu Wei-hua, Liu Shi-hu, Wang Qiao-rong. The First Type of Grade Rough Set Based on Rough Membership Function[C]// 2010 Seventh International Conference System and Knowledge Discovery(FSKD2010). 2010: 1922-1926
- [12] Zhang Xian-yong, Xiong Fang, Mo Zhi-wen. Rought Set Model Based on Logical Or Operation of Precision and Grade[J]. Pattern Recognition and Artificial Intelligence, 2009, 17(9): 151-155 (in Chinese)
- 张贤勇, 熊方, 莫志文. 精度与程度的逻辑或粗糙集模型[J]. 模式识别与人工智能, 2009, 17(9): 151-155
- [13] Ziarko W. Variable precision rough set model [J]. Journal of Computer and System Sciences, 1993, 46(1): 39-59
- [14] Dubois D, Prade H. Rough fuzzy sets and fuzzy rough sets[J]. International Journal of General Systems, 1990, 17: 191-209
- [15] Xu Wei-hua. Ordered information system and rough set[M]. Beijing: Science press, 2013 (in Chinese)
- 徐伟华. 序信息系统与粗糙集[M]. 北京: 科学出版社, 2013

(上接第97页)

- [12] Kong W Z, Liu Y Q, Zhang M, et al. Answer quality analysis on community question answering[J]. Journal of Chinese Information Processing, 2011(1): 3-8 (in Chinese)
- 孔维泽, 刘奕群, 张敏, 等. 问答社区中回答质量的评价方法研究[J]. 中文信息学报, 2011(1): 3-8
- [13] Yang H T, Wang J, Ling H F. Question sentence similarity computing based on multi-features fusion in community question an-

swering[J]. Journal of Jiangxi Normal University, 2013, 37(2): 125-129 (in Chinese)

- 杨海天, 王健, 林鸿飞. 基于特征融合的社区问答问句相似度计算[J]. 江西师范大学学报, 2013, 37(2): 125-129
- [14] Zhou Z M. The study of models and features for non-factoid question answering[D]. Shanghai: East China Normal University, 2012 (in Chinese)
- 周志敏. 非事实类问题问答模型和特征的研究[D]. 上海: 华东师范大学, 2012