

## 知乎话题结构协同构建中的冲突与协作分析

An Analysis of Conflicts and Collaboration in Social Construction of Zhihu Topic Structure

张鹏翼 杨玉宇

(北京大学信息管理系, 北京, 100871)

[摘要] 随着社交问答网站的发展,越来越多的用户开始通过社会化问答平台获取和分享内容。以知乎为例,同一个问题会被贴上不同的标签,标签之间存在着上下位关系,形成一个类似于主题词表的话题结构,标签与标签之间的关系均由用户创造。该话题结构对于知乎站内检索起着重要作用。本文通过分析用户知识构建的特点,对知乎话题结构构建过程中用户之间的冲突进行研究。以话题编辑日志为研究对象,提出了一个基于知识构建的分析框架,抽取了文学、哲学、物理学和社会学四个领域的 1,037 个话题的 10,778 条日志进行研究。结果表明,知乎的话题结构存在着用户间知识构建水平低、用户语言交流未能达到探索性讨论的要求、用户交流缺乏提供理性依据的理论家等问题,话题知识构建水平仍需提高,知乎系统也需要进一步改进。本研究有利于了解用户自发构建主题词表过程中的冲突行为特点,进而根据这种行为特点探索出最佳的冲突解决模式。

[关键词] 知识协作 知识构建冲突 大众分类法 协作角色 语言功能分析 社会化标签 知乎

[中图分类号] G252.0 [文献标识码] A [文章编号] 1003-2797(2017)03-0108-10 DOI: 10.13366/j.dik.2017.03.108

[Abstract] More and more users start to use social question answering (SQA) sites as sources for information seeking and sharing. As an example, zhihu.com allows users to edit topic tags and the relationships among them, resulting a hierarchical taxonomy of topics which is essential to the organization and retrieval of its content. This research aims to analyze the conflicts and collaboration during the social construction of topic structures in Zhihu.com. Using an analysis framework based on knowledge construction, we analyzed 10,778 editing logs of 1,037 topics in four domains including literature, philosophy, physics and sociology. Results show that the knowledge construction process of topic structure in Zhihu.com is unintentional and shallow; users didn't involve in higher level of collaboration partially due to the lack of collaborative functions of the system. The results are useful in understanding the conflicts and collaboration among users in constructing the topic taxonomy, and future research will expand the data set to get more generalizable results and to explore the best ways to resolve the conflicts.

[Key words] Knowledge collaboration Knowledge construction conflicts Folkonomy Collaborative roles Communication function Social tagging Zhihu

## 1 引言

随着社交问答网站的发展,越来越多的用户开始

通过社会化问答平台获取和分享内容。以社会化方式组织内容,例如协同构建话题结构、等级体系等,也

[基金项目] 本文系项目 OCLC/ALISE 图书馆学情报学研究基金项目“Collaborative Sensemaking in Online Knowledge Groups”和国家自然科学基金青年项目“知识社区在线协作动态研究:构成、因素与演变”(71603012)的研究成果之一。

[作者简介] 张鹏翼,副教授,博士,研究方向:信息与知识组织,信息行为与认知,交互性信息检索,以用户为中心的设计与评价,Email: pengyi@pku.edu.cn; 杨玉宇,通讯作者,硕士生,研究方向:信息与知识组织,信息行为与认知,Email: yangyuyu@pku.edu.cn。

逐渐成为一种趋势。以知乎为例,问题通过话题(标签)组织,话题之间存在着上下位关系,这些上下位关系将知乎话题联结成了一个巨大的“话题结构”。知乎话题结构几乎都由用户建立,用户自发创建话题,自发给话题添加上下位关系,众多用户的话题编辑行为共同组成了协同式知识构建行为。用户在话题构建过程中会产生冲突,具体表现在父话题的添加和删除上。

目前为止,国内学者有少量关于用户协同过程中冲突的研究,例如网络百科编辑过程中的用户冲突研究<sup>[1]</sup>,但几乎没有对于知乎话题树这样涉及上下位概念关系的协同构建和冲突的研究。研究在网络协作环境中,用户对不同标签之间上下位关系的理解存在差异,了解他们协作的过程中如何处理这些差异和冲突,不仅有利于了解网络用户之间冲突的特点与模式,也有利于探究协同模式下用户间冲突对群体认知结果(话题树)的影响。本文的研究试图解释以下问题:  
①话题结构协同构建中的冲突是如何产生和解决的?  
②用户和系统在冲突过程中承担着什么样的角色?  
③知乎在话题结构协同构建的机制方面有哪些不足?

## 2 相关研究

知乎话题树实际上是一种社会化标签(Social Tagging),用户自发创建标签,建立标签后为这些标签之间建立联系。建立话题树需要用户之间的协作(Collaboration),而协作过程中可能会产生冲突。在用户协作构建话题树的过程中,知乎网站为用户提供平台、编辑工具,同时还会制定编辑准则,以保证话题树的质量,在这方面,可以借鉴许多本体构建工具的设计经验。因而,本文从“社会化标签”、“协同学习”和“本体构建工具”三方面切入进行分析。

### 2.1 社会化标签与大众分类法

社会化标签(Social Tagging)是“用户以关键词形式给分享的资源添加元数据”的过程<sup>[2]</sup>。自2003年英文标签网站delicious正式上线以来,社会标签不断发展。研究表明,社会化标签对于资源检索和分享有一定帮助,但与受控词表相比,也存在着缺乏词汇控制、缺少等级结构等不足<sup>[3,4]</sup>。Lemieux等提出,大众分类法(Folksonomy)和词表(Taxonomy)有4种结合方式:①共存:即独立使用大众分类法和分类系统,二

者之间不产生联系;②大众分类导向的分类系统:大众分类法向分类系统提供新概念,分类系统将这些新概念纳入自己原有的分类体系中;③分类系统导向的大众分类法:当用户给资源添加标签时,系统会向用户提供分类系统中的受控词汇;④大众分类本体:有两种类型,第一种是用户自发在标签之间建立结构关系,另一种是系统通过算法自动在标签之间建立等级联系<sup>[5]</sup>。Yang等探究了如何通过语义分析自动建立标签簇之间的等级关系<sup>[6]</sup>,Zhou等则提出了“覆盖率(Coverage Rate)”的概念以建立标签等级<sup>[7]</sup>。知乎话题结构是对标签进行结构化的结果,用户创建标签并且在标签之间建立联系,属于Lemieux提出“大众分类本体”,因此从标签的角度看待话题节点有利于加深对话题结构的理解<sup>[8]</sup>。

### 2.2 协同学习

知乎的话题结构是用户共建的结果,用户之间经过冲突争议之后达成共识,协作理论对于研究用户之间的协作水平高低很有帮助。计算机支持下的协同学习(Computer-supported Collaborative Learning,以下简称CSCL)使得用户可以随时随地参与讨论,不再需要面对面的交流<sup>[9]</sup>。Onrubia等通过用户实验对计算机环境下大学生完成报告过程中的协作策略和协作阶段进行研究,他们发现在知识构建过程中,存在着四种阶段:分别是开题阶段(Phase of Initiation)、探索阶段(Phase of Exploration)、协商阶段(Phase of Negotiation)和建设阶段(Phase of Construction)。在开题阶段中,小组成员只发表自己的观点,不对他人的观点进行审视质疑;在探索阶段,小组成员会审视其他小组成员的观点,但是不会提出不同意见;在协商阶段,小组成员会对他人的观点进行质疑批判,提出建设性意见,小组成员之间的知识共享程度得到提高;在建设阶段,通过对任务文档的认同、质疑、修改等行为,小组成员对知识概念达成一致<sup>[10]</sup>。Schellens等研究了在CSCL中,角色分配对知识构建的影响。研究表明,角色分配小组的知识构建水平、期末成绩明显高于非角色分配小组,同时过于简单和过于复杂的任务都会降低学生知识构建的水平<sup>[11]</sup>。Arvaja等探索了情景(Context)在网络学习讨论中所起到的作用,并从交流

功能 (Communication Function) 和背景资源 (Contextual Resource) 两个方面对讨论内容进行分析,发现不同的情景确实影响了两组学生知识使用和交流方式,表明了学习受环境和阐释方法影响的特点<sup>[12]</sup>。此外,甘永成设计了一套内容分析的编码系统,包括提问、解释、冲突、辩护、共识、综合、评估、反思、引用、社交、技术、协调服务等内容,并通过这套框架对首都师范大学某门课程上学生之间的协作进行分析<sup>[13]</sup>。

### 2.3 本体构建工具

知乎话题结构的建立过程和本体的建设过程非常的相似,用户需要在不同的话题节点之间建立联系,使其符合真实世界的逻辑。同时不同用户之间还需要对本体达成共识,以便更好地引导用户进行浏览查找。Gruder 将本体定义为“某一主题领域的明确概念化”<sup>[14]</sup>。在语义网中,本体表达需要有明确的语义。然而本体并不仅仅是某领域的形式表达,更是群体对形式表达的认同。因此在本体构建的过程中“如何使多人对本体达成一致”一直是研究者探讨的问题<sup>[15]</sup>。Gendarmi 等则设想了一种新型的、能够将用户参与引入本体设计中的写作系统。在该系统中,用户可以将自己感兴趣的引文数据置于自己的收藏夹 (Binder) 中,同时给引文数据设定关键词自行确定所给关键词

在知识本体中的位置。结果表明,社区和本体构建相结合,能够降低本体设计的难度和花费,同时还能够增加本体更新的速率<sup>[16]</sup>。Noy 等对现存的知识构建工具进行了研究,通过这些研究了解了用户对知识构建工具的期望和需求,结果表明:①在信任、可靠性和本体维护方面,可以通过对用户进行投票或者对用户的修改行为进行打分来判断一个用户的信誉,用户必须知道谁对内容进行了修改,同时能够阅读修改者对修改行为的解释,了解改动的作者和原因;②在本体理解方面,本体在构建成功之后需要用可视化方式表达出来,以便用户更好地理解;③在用户界面设计方面,要有一个良好的界面设计,将软件功能简洁的列出,便于用户理解<sup>[17]</sup>。Hamasaki 等则借鉴了社会标签机制,将以往由专家完成的知识构建转化为能够让普通大众参与的行为,提出了一种新型的知识构建机制——“社会属性标签 (Social Property Tagging)”,即模拟传统的社会标签,用户在给资源贴标签时给出“属性-值”的二元组<sup>[18]</sup>。

### 3 分析框架

本文依据 Onrubia<sup>[19]</sup>、Schellens<sup>[20]</sup>、Arvaja<sup>[21]</sup>、甘永成<sup>[22]</sup>等的模型,整合提出知乎话题协同构建冲突的分析框架,具体如图 1。

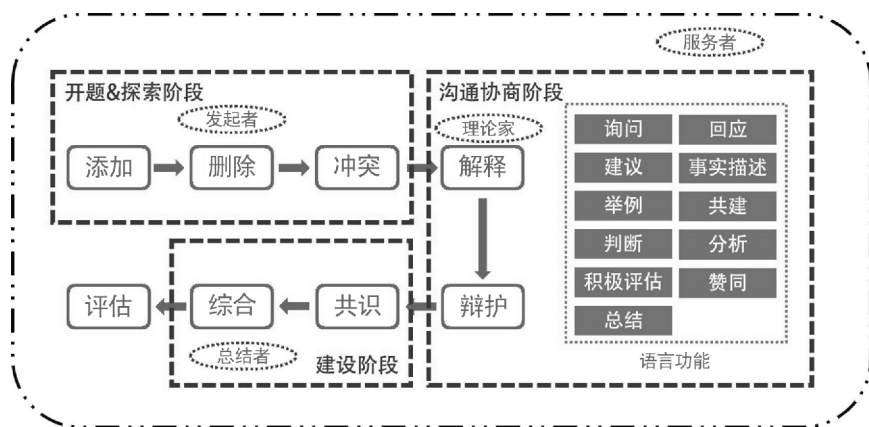


图1 分析框架

该分析框架从“协作步骤”、“协作阶段”、“语言功能”和“协作角色”四个角度对话题构建冲突进行分析。

(1) 协作步骤: 描述的是建立稳定话题结构的步

骤。由甘永成<sup>[23]</sup>的内容分析编码系统转化而来,并且在具体分析话题冲突流程的基础上进行了改进。

(2) 协作阶段: 描述的是知识构建达到的阶段,目

的是反应协作水平的高低。参考了 Onrubia<sup>[24]</sup>的知识构建的四个阶段,并根据知乎话题构建的特点修改为三个阶段。①开题 & 探索阶段:不同的用户会浏览审视话题结构,并通过添加父话题和删除父话题表达自己的修改意见,然后部分修改意见之间会产生冲突;②沟通协商阶段:用户之间会对自己的观点进行辩护、同时对他人的意见进行质疑批判。在这个阶段中,用户之间的共享知识会增加,双方对话题结构的理解程度会得到提升;③建设阶段:通过双方之间的质疑和辩护,双方对话题结构的理解达成一致,会给出一个最终的修改方案,并且这个方案会得到双方的承认和认同。

(3)语言功能:不同的发言代表了参与者参与谈话的不同方式<sup>[25]</sup>,有利于知识构建的探索性讨论(Exploratory Talk),仅在参与者对他人的观点进行有建设意义的批判时产生<sup>[26]</sup>。语言功能分析的目的在于分析话题冲突的过程中,用户的发言是否达到了探索性讨论的要求,同时在多大程度上促进了知识构建。内容包括询问、回应、建议、事实描述、共建、判断、分析、积极评估、赞同、总结等。

(4)协作角色:以往的角色分配研究都是先行赋予研究对象角色,然后再进行内容分析,主要研究角色分配是否能提高交流质量<sup>[27,28]</sup>。而本文套用角色分析目的则是研究知乎用户协作的特点和模式,所以本文是在进行内容分析的基础上依据参与者的发言特点分析其角色,以此观察交流内容和质量。在话题结构构建过程中,一共有四个关键的角色:发起者、理论家、总结者和服务者。

## 4 研究方法

### 4.1 数据抓取与清理

本文抓取了知乎的话题日志页,并利用 perl 语言对网页内容进行处理,将日志转化为可分析的结构化数据。本文共选取了 4 个话题:哲学、文学、社会学和物理学,抓取了知乎建立至 2015 年 10 月中旬之间所有的话题编辑数据共 10778 条(见表 1)。

每条日志包括日志 ID,时间,日期,话题 1ID,话题 1 名称,话题 2ID,话题 2 名称,操作,用户 ID,修改理由等项。例如“话题 1:哲学 话题 2:人文学科 操作:添

加了父话题”意为:对“哲学”话题进行操作,使“人文学科”成为其父话题。表 2 展示了操作类型的分布占比及占比。

表 1 研究数据集

话题领域	文学领域	哲学领域	物理学领域	社会学领域	总计
话题个数	536	148	240	113	1037
记录条数	5731	1669	1949	1429	10778
参与用户数	756	268	371	236	1631

表 2 用户操作类型分布

操作	频次	百分比(%)
添加了父话题	4679	43.41
移除了父话题	2147	19.92
创建了话题	1037	9.62
编辑了话题描述	895	8.30
修改了话题图片	725	6.73
添加了话题别名	689	6.39
编辑了话题名称	467	4.33
移除了话题别名	125	1.16
编辑了话题弱关系	13	0.12
总计	10778	100

### 4.2 数据分析

统计分析:利用 SPSS 对日志数据进行了包括频次分析、描述分析、幂律分布检验、非参数独立检验在内的统计分析。

内容分析:在协作阶段分析中,以每对话题冲突为分析单位,对话题冲突进行编码(见表 3)。在语言功能分析中,以有注释内容的单条日志为分析单位,对用户话题修改意见进行了分析编码(见表 4)。在协作角色分析中,以单条日志为分析单位,对用户角色进行编码(见表 5)。

表 3 协作阶段编码

协作阶段	含义
开题 & 探索阶段	用户浏览话题结构,提出意见,发现自己与他人意见不同
沟通协商阶段	用户之间进行对话协商、有意见交换
建设阶段	知识理解达成一致,得出双方认同的解决方案

表4 语言功能编码

语言功能类型	含义
询问	向对方发出疑问
回应	回应对方疑问
建议	用户表达自己的观点,但是却不提供具体理由
事实描述	用事实作为意见依据
共建	用户对另一个用户的观点进行补充和升华
判断	用户对其他用户的观点进行简单粗暴的定论,并且不提供理由
分析	在建议的基础上,提供详细的理由和依据
积极评估	对其他用户的观点进行肯定赞扬
赞同	接受其他用户的观点
总结	对冲突双方的意见进行梳理概括

表5 协作角色编码

角色	含义
发起者	对已有话题结构发起质疑的人
理论家	交流过程中提供理论依据的人
总结者	对双方的观点进行总结的人
服务者	为交流过程提供支持的人或者系统

## 5 研究结果

### 5.1 协作步骤分析

冲突记录描述分析。①冲突的界定:本文仅对话题结构冲突进行研究,即本文只研究“添加、移除父话题”操作。本文将冲突定义为“用户对已有话题结构的质疑和修改”,即当用户首次移除已有的父子话题结构时,冲突便产生了。②描述分析:从日志中共分析出话题冲突 1719 对,冲突记录 4161 条,占总记录数的 38.7% 左右。话题冲突次数分布见图 2,将数据取对数后进行线性拟合,拟合结果  $\text{Sig.} = 0.00, R^2 = 0.837$ 。即只有少部分冲突持续了多次,大部分冲突只持续了很少次。话题冲突持续天数分布见图 3,可以看出大部分的冲突持续时间在 200 天以内,占总数的 50% 以上。冲突参与人数分布见图 4,可以看出 60% 以上的冲突仅有 2 名用户参与,多人参与的用户冲突是少数(图 4 中有一部分的冲突仅有 1 人参与,表示用户添加父话题之后又自行删除,这并非协作冲突的一部分)。

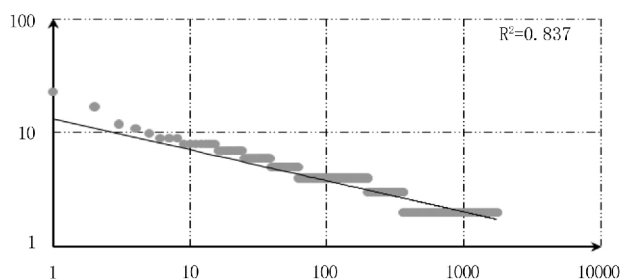


图2 冲突次数分布

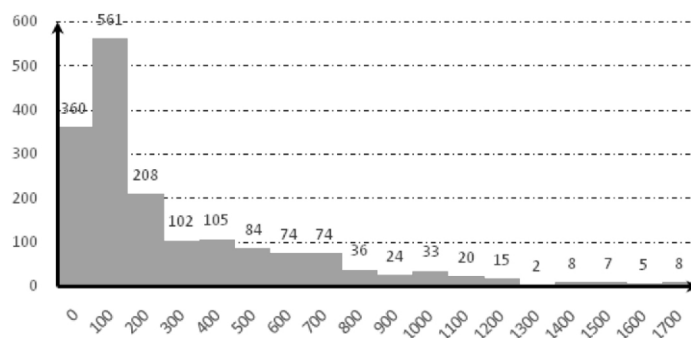


图3 冲突持续天数分布

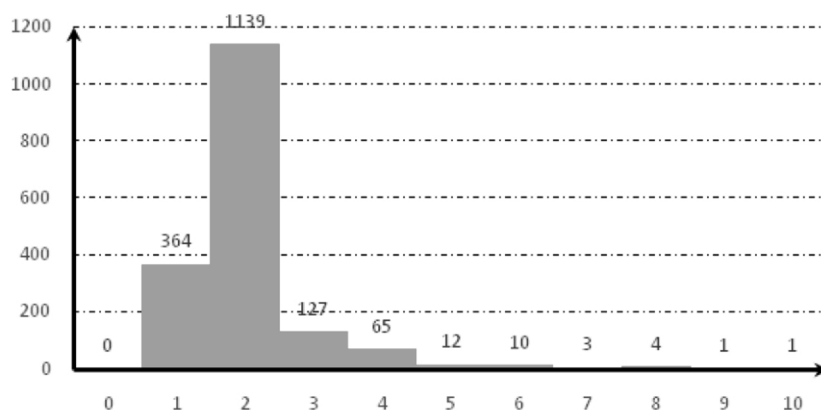


图 4 冲突参与人数分布

## 5.2 协作阶段分析

知乎的话题结构是知识构建的产物,用户根据目前的趋势创建新话题,然后根据自己的理解将话题结构定位在话题结构中的某一个位置。不同的用户对话题节点的位置会产生分歧,此时便需要用户之间进行交流协作。

模型中提出三个协作阶段:开题 & 探索阶段、沟通协商阶段、建设阶段。

在开题探索阶段,用户浏览话题结构,提出自己的修改意见,进而发现与他人意见不同。达到这个认知阶段的关键有两点,其一是用户浏览别人的观点,其二是用户发表自己的观点。

知乎的话题结构冲突基本上都达到了这个阶段,

因为用户只有在浏览了话题结构、浏览了其他用户的修改行为之后才能够发表自己的修改意见。例如,用户 AL 首先将哲学话题设置为伦理话题的父话题,表达了自己的观点。用户 TA 在浏览用户 AL 的操作和观点后,不赞同用户 AL 的观点,移除了哲学话题和伦理话题之间的父子话题关系,表达了自己的观点。

但在话题结构中,用户观点内容是十分单薄的,如表 6,用户 AL 和用户 TA 一般,大多数用户并不会对自己的操作行为进行详细的解释,有注释的话题结构操作仅占总操作数的 2.8%,更多人只是单纯地进行添加删除操作,直到他人对自己操作提出质疑,才会有少部分用户进行解释。

表 6 开题探索阶段示例

日志 ID	话题 1 名称	话题 2 名称	操作	用户 ID	修改原因
17643321	伦理	哲学	添加了父话题	用户 AL	
17644563	伦理	哲学	移除了父话题	用户 TA	
53955012	伦理	哲学	添加了父话题	用户 MA	
54100923	伦理	哲学	移除了父话题	用户 PB	

在沟通协商阶段,用户对自己的观点进行辩护,同时对他人的观点进行质疑,通过不断地沟通协商,用户之间的共享知识扩大,概念理解的一致程度也会上升。达到此认知阶段的关键在于,用户之间要进行对话协商、需要有意意见交换。鉴于知乎异步讨论 (Asynchronous Discussion) 的特点,用户之间无法进行及时的回复,用户之间的协商与意见交换就表现

在“修改理由”中。如表 7,用户 CH 对用户 AN 将哲学作为伦理父话题的观点提出了质疑——“相关问题太少,无关问题太多”,而用户 AN 则对用户 CH 进行了回应,对自己的观点进行了辩护——“可以试着移除“伦理”话题下的无关问题。伦理学是哲学的核心领域之一”。

本文不将没有注释的“添加”“删除”操作视作用

户之间的意见交换,在这个前提下话题结构构建中仅有 2 对话题冲突达到了沟通协商阶段,仅占话题冲突总数 (1719 对) 的 0.12%,总体知识构建的水平较低,同时大部分注释内容也没有很好的达到沟通效果。

表 7 沟通协商阶段示例

日志 ID	话题 1 名称	话题 2 名称	操作	用户 ID	修改原因
129217445	伦理	哲学	移除了父话题	用户 UN	
129217550	伦理	哲学	添加了父话题	用户 UN	
132694871	伦理	哲学	移除了父话题	用户 CH	相关问题太少,无关问题太多
148745874	伦理	哲学	添加了父话题	用户 AN	可以试着移除“伦理”话题下的无关问题。伦理学是哲学的核心领域之一

在建设阶段,用户对话题结构的理解达成一致,对话题冲突提出解决方案,且方案得到双方的认可。但是所有的话题冲突都没有到达这个阶段,反应用户之间缺乏共识,同时也是因为知乎网站没有提供一个确认冲突参与者都满意的机制。话题冲突记录共有 4161 条,其中含有用户注释的仅有 181 条,仅占总数的 4.4%,话题结构构建过程中的交流缺乏由此可见。

5.3 语言功能分析

本文对用户注释内容进行语言功能分析,对交流内容进行编码分类,以了解用户交流的特点和对知识构建的作用,各类型的语言功能示例请见表 8。依据此框架对含有用户注释的 181 条日志进行分析。编码一共由两人进行,初次标引得到的一致率结果为

93%,经过讨论,解决不一致后,编码结果见图 5。10 个语言功能分析类别中,判断型最多,分析型其次,无意义内容再次、然后是建议型、事实描述型、询问型、共建型。其余类型如回应、共建、积极评估、赞同、总结等均未在用户注释中出现。

知乎话题结构共建是一个异步交流的过程,客观上制约了用户之间的问答,导致缺少询问型(0 条)、回应型内容(1 条)和共建型内容(1 条),也反映出用户间缺乏互动意识;缺少积极评估型和赞同型内容反映出用户对他人观点的关注有限,或者用户缺乏达成一致观念的意识;缺乏总结型内容反映出用户之间的交流层次较浅,没有人对冲突观点进行梳理,也无法形成一个双方都接受的方案。

表 8 语言功能分析示例

语言功能类型	示例
询问	话题 1:哲学 话题 2:人文学科 操作:添加了父话题 哲学不是人文学科?
回应	无
建议	话题 1:唯物主义 话题 2:物质 操作:移除了父话题 话题父子关系不明显
事实描述	话题 1:时间 话题 2:物理量 操作:移除了父话题 时间下问题基本没有和物理相关的
共建	话题 1:哲学 话题 2:伦理 操作:添加了父话题 可以试着移除“伦理”话题下的无关问题。伦理学是哲学的核心领域之一。(回应:相关问题太少,无关问题太多)
判断	话题 1:热力学 话题 2:物理学 操作:添加了父话题 恶意编辑
分析	话题 1:科普 话题 2:文学 操作:移除了父话题 科普可以运用各种传播媒介、各种艺术手段,文学只是其中之一,不能作为父话题
积极评估	无
赞同	无



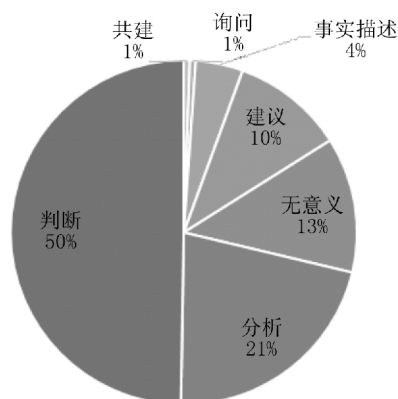


图5 语言功能分析结果

通过语言功能分析可以看出,在话题结构冲突的过程中,用户之间的交流没有很好地促进共识。用户注释中,简单粗暴的判断大约占一半(49.7%),理性分析仅占四分之一(分析 21.6%,事实描述 4.4%),而询问、回应、共建、总结等能提高用户知识共享水平的内容却几乎没有,且大部分对话题结构的修改操作是没有注释的。由此看出,知乎话题结构的知识构建是非常浅层次的、低质量的。

#### 5.4 协作角色分析

**发起者:**本文将首次对话题结构进行修改的人视为冲突的“发起者”,经过分析,发起者共 203 位,占参与话题结构编辑总人数(1315 人)的 15.4%。

已知居住地的发起者中,大部分人居住在北京、上海、成都、广州等一线城市或省会城市,这些城市大多经济发达、文化教育水平高。已知行业分布中,大部分人从事互联网、计算机软件、高新科技等科技文化相关的职业,反应出发起者较高的文化水平。已知工作机构分布中,有 6%的发起者填写了“知乎公共领域编辑计划”,3%的发起者填写了知乎公司。“知乎公共领域编辑计划”是一个自发的民间组织,而非知乎的官方机构。这个组织的人参与了大量的话题结构编辑。从工作机构分布中可以看出,知乎也安排了一些员工对话题结构进行编辑维护。从已知职位分布中可以看出,知乎公共领域编辑计划的志愿者占 7%,反应出志愿者在话题编辑方面的活跃。

通过曼-惠特尼 U 检验发现,发起者和知乎普通

用户在获得赞同数( $c^2 = 80702$ , Sig. = 0.00)、获得感谢数( $c^2 = 81196$ , Sig. = 0.00)、回答问题数( $c^2 = 95061$ , Sig. = 0.03)、专栏文章数( $c^2 = 93835$ , Sig. = 0.00)、收藏数( $c^2 = 92902$ , Sig. = 0.00)、公共编辑数( $c^2 = 61929$ , Sig. = 0.00)、关注人数( $c^2 = 95156$ , Sig. = 0.03)、被关注人数( $c^2 = 84205$ , Sig. = 0.00)和个人主页被浏览数方面( $c^2 = 76989$ , Sig. = 0.00),存在着显著的不同。

**理论家:**理论家是在解释辩护过程中能够用科学的理论支持自己观点的用户。本文将能够提出“分析”类发言的用户视作“理论家”,通过数据分析处理,共分析出理论家 9 人,占参与话题编辑总人数的 0.68%,有关 9 名用户的描述数据参见表 9。

通过非参数独立检验发现,理论家和知乎普通用户在获得赞同数( $c^2 = 1125.5$ , Sig. = 0.00)、获得感谢数( $c^2 = 1030.5$ , Sig. = 0.00)、回答数( $c^2 = 2927.5$ , Sig. = 0.01)、专栏文章数( $c^2 = 1955.5$ , Sig. = 0.00)、公共编辑数( $c^2 = 1607.0$ , Sig. = 0.00)、被关注数( $c^2 = 1267.5$ , Sig. = 0.00)和个人主页被浏览数( $c^2 = 1417.5$ , Sig. = 0.00)方面存在着显著不同。

表9 理论家用户描述统计

项目	描述统计
性别	女性 2 名,男性 7 名
居住地	北京 4 名,美国 1 名,蓝鲸 1 名,未知 3 名
行业	高等教育 1 名,科研 1 名,医疗服务 1 名,未知 6 名
工作机构	知乎公共领域编辑计划 1 名,其他 3 名,未知 5 名
职位	志愿者 1 名,其他 3 名,未知 5 名
教育背景	北京大学 1 名,内布拉斯加大学林肯分校 1 名,蓝鲸大学 1 名,南京大学 1 名,其他 1 名,未知 2 名

**总结者:**总结者是对冲突双方的意见进行总概括的用户。经过分析发现,参与话题编辑的用户当中,没有总结者类型的人。分析原因有二,其一是编辑话题结构的用户没有争议进行总结的意识;其二是知乎并未能提供一个争议总结机制,例如对争议进行投票评论等。除此之外,通过分析发现,绝大部分的话题冲突都是以“移除了父话题”结束的。详情见图 6。



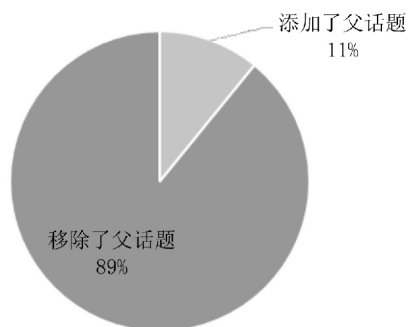


图6 话题冲突结束操作分布

**服务者:**服务者是指对整个冲突流程提供讨论平台或支持的用户、系统。在话题冲突过程中,并没有承担“服务者”角色的用户出现,担任服务者角色的主要是知乎系统。知乎系统为持有不同意见的用户提供了交流平台,让冲突者能够以异步交流的方式交换双方的意见,但是这无法进一步提高知乎话题结构的知识构建水平。例如,为了让用户能够更加详细的说明修改意见和观点,用户修改话题结构时系统可以强制用户添加注释;再比如,为了让用户对最终的修改意见达成共识,系统可以建立一个打分投票机制,让用户对修改意见和最终解决方案进行打分和投票。然而以上功能系统都没有提供,所以知乎系统作为“服务者”还有很多需要改进的地方。

## 6 结论与讨论

### 6.1 结论

从协作步骤来看,添加、删除父话题操作远远超出创建话题操作的数量,反应出用户对话题结构进行大量的编辑和优化。从日志整体来看,添加父话题操作多于移除父话题操作,而从冲突结束局部来看,移除父话题操作多于添加父话题操作,由此推测,一旦话题父子关系被质疑,那么这对父子关系更有可能以“移除父话题”操作结束。

从协作阶段来看,只有极少部分冲突发展到了沟通协商阶段,而所有的冲突都没能达到建设阶段,反应出用户之间缺乏共识,缺少用户对争议双方的意见进行综合,同时系统缺陷也导致了建设阶段的缺失。

从语言功能分析来看,有注释的日志仅占总日志总数的很少一部分,限制了用户之间的交流。语言功

能分析结果表明,绝大部分的语言内容没有与其他用户产生互动,而真正具有互动意义、促进交流的内容仅占总数的1%。

从协作角色来看,发起者与理论家在知乎的答题数量与质量、对公共编辑做出的贡献方面高于知乎普通用户,同时也比一般用户受到更多的关注。发起者用户普遍居住在经济文化水平较高的城市,从事的行业或与互联网相关,或要求较高的文化水平。此外,理论家角色占编辑用户总数的比例非常低,表明在冲突过程中进行理论分析的用户较少。没有用户扮演“总结者”的角色,反应话题冲突缺乏总结升华。此外,承担“服务者”角色的知乎系统也存在着一系列问题,需要进一步改进。

综上所述,知乎的话题结构存在着用户间知识构建水平低、用户语言交流未能达到探索性讨论的要求、用户交流缺乏提供理性依据的理论家等问题,话题知识构建水平仍需提高,知乎系统也需要进行相应的改进。

### 6.2 局限及未来研究方向

本文仅选择了四个学科话题作为研究样本,不同的话题及其编辑用户可能有着不同的特点,因此研究结论是否具有普适性有待考察。本文的模型存在一定程度上的简化,无法全面涵盖话题编辑过程中的各种细节,例如对同一个父子话题关系进行质疑的可能是不同的用户、存在用户自己添加父话题之后又删除父话题的情况等。最后,社会文化资本对用户编辑行为和用户角色可能会存在一定的影响<sup>[29]</sup>,例如具有较高文化资本的用户可能会更热衷于话题结构的编辑、有更高的几率成为理论家等。

后续研究可考虑增加话题样本个数和多样性,提高研究的普适性;也可以研究个体用户在话题冲突过程中的互动,研究冲突过程中个体用户知识结构的变化;同时还可以将社会文化资本理论引入语言功能分析和协作角色分析研究中,分析社会文化资本高低对用户语言和角色分析的影响;同时还可以尝试与国外的知识问答网站 Quora 进行对比研究,观察中外用户知识构建的差异;再比如对知乎最终的话题结构进行分析评价,分析知识构建水平的高低是否会影响话题结构的质量等。

## 参考文献

- 1 黄令贺, 张迎军, 程靖淇, 等. 网络百科内容生产过程中的用户冲突研究——以“Px 词条保卫战”为例[J]. 图书情报工作, 2016, 60(3): 114-120.
- 2 Golder S A, Huberman B A. Usage Patterns of Collaborative Tagging Systems[J]. Journal of Information Science, Sage Publications, 2006, 32(2): 198-208.
- 3 Goh D H-L, Chua A, Lee C S, et al. Resource Discovery Through Social Tagging: A Classification And Content Analytic Approach[J]. On-line Information Review, 2009, 33(3): 568-583.
- 4 Lu C, Park J, Hu X. User Tags Versus Expert-Assigned Subject Terms: A Comparison of Librarything Tags And Library of Congress Subject Headings[J]. Journal of Information Science, Sage Publications Sage Uk: London, England, 2010, 36(6): 763-779.
- 5,8 Lemieux S. Hybrid Approaches to Taxonomy And Folksonomy[J]. Semantic Technology, 2009.
- 6 Yang K, Shi Z. Exploring Web Resources in Social Annotation Services in Hierarchy[J]. 8th International Conference on Grid and Cooperative Computing, Gcc 2009, 2009: 182-187.
- 7 Zhou T C, King I. Automobile, Car And Bmw: Horizontal And Hierarchical Approach in Social Tagging Systems[J]. Proceeding of the 2nd Acm Workshop on Social Web Search And Mining, 2009: 25-32.
- 9 Kreijns K, Kirschner P A, Jochems W. The Sociability of Computer-Supported Collaborative Learning Environments[J]. Educational Technology & Society, Jstor, 2002, 5(1): 8-22.
- 10,19,24 Onrubia J, Engel A. Strategies for Collaborative Writing and Phases of Knowledge Construction in Cscw Environments[J]. Computers And Education, Elsevier Ltd, 2009, 53(4): 1256-1265.
- 11,20 Schellens T, Van Keer H, De Wever B, et al. Scripting by Assigning Roles: Does It Improve Knowledge Construction in Asynchronous Discussion Groups? [J]. International Journal of Computer-Supported Collaborative Learning, 2007, 2(2-3): 225-246.
- 12,21,26 Arvaja M. Contextual Perspective in Analysing Collaborative Knowledge Construction of Two Small Groups in Web-Based Discussion [J]. International Journal of Computer-Supported Collaborative Learning, 2007, 2(2-3): 133-158.
- 13,22,23 甘永成. 虚拟学习社区的知识建构分析框架[J]. 中国电化教育, 2006(2): 27-31.
- 14 Gruber T R. Toward Principles for the Design of Ontologies Used for Knowledge Sharing? [J]. International Journal of Human-Computer Studies, Elsevier, 1995, 43(5-6): 907-928.
- 15 Hepp M, Bachlechner D, Siorpaes K. Ontowiki: Community-Driven Ontology Engineering And Ontology Usage Based on Wikis[C]//Proceedings of the 2006 International Symposium on Wikis. Acm, 2006: 143-144.
- 16 Gendarmi D, Abbattista F, Lanubile F. Fostering Knowledge Evolution Through Community-Based Participation[J]. Ceur Workshop Proceedings, 2007, 273.
- 17 Noy N F. The Ckc Challenge: Exploring Tools for Collaborative Knowledge Construction[J]. 2011.
- 18 Hamasaki M, Goto M, Takeda H. Social Infobox: Collaborative Knowledge Construction By Social Property Tagging[J]. Proc. Cscw, 2011: 641.
- 25 Mercer N. The Quality of Talk in Children's Collaborative Activity in the Classroom[J]. Learning And Instruction, Elsevier, 1996, 6(4): 359-377.
- 27 Dillenbourg P, Jermann P. Designing Integrative Scripts[G]//Scripting Computer-Supported Collaborative Learning. Springer, 2007: 275-301.
- 28 Kollar I, Fischer F, Slotta J D. Internal And External Collaboration Scripts In Web-Based Science Learning At Schools[C]//Proceedings of Th 2005 Conference on Computer Support for Collaborative Learning: Learning 2005: The Next 10 Years! International Society of the Learning Sciences, 2005: 331-340.
- 29 张鹏翼, 张璐. 社会资本视角下的用户社交问答行为研究——以知乎为例[J]. 情报杂志, 2015, 34(12): 186-191.

(收稿日期: 2016-12-14)