

基于颜色 Petri 网的有偿交互 QA 系统投诉模型

李彦成 曾庆田 鲁法明 薛洁

(山东科技大学信息科学与工程学院 青岛 266590)

摘要 有偿交互问答(QA)系统受到越来越多的关注,但是补偿的不公平性等问题可能会引起用户的投诉。通过分析有偿交互 QA 系统的投诉流程,首先利用颜色 Petri 网对投诉流程进行形式化建模;然后使用 CPN Tools 对所建立的形式化模型进行验证;最后对所建模型的状态空间报告进行分析,证明了系统各响应流程设计的正确性等特性。所提出的方法对有偿交互 QA 系统的综合形式化分析具有重要意义。

关键词 有偿交互问答系统,颜色 Petri 网,投诉模型,CPN Tools

中图分类号 TP311 文献标识码 A DOI 10.11896/j.issn.1002-137X.2014.07.022

Colored-Petri-Net-based Complaint Model for User-interactive Question Answering System

LI Yan-cheng ZENG Qing-tian LU Fa-ming XUE Jie

(College of Information Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract Due to users' complaints arising from such problems as unfair compensation, user-interactive question answering systems are catching more and more attention. By analyzing the complaint process of a user-interactive question answering system, the formal complaint model was presented using Colored Petri Net. Then the established formal model was verified by employing CPN tools. Finally, the state space report of the established model was analyzed, so the correctness and other features of the system design were demonstrated. This proposed method has special significance to the comprehensive formal analysis of the user-interactive question answering system.

Keywords User-interactive question answering system, Colored Petri net, Complaint model, CPN tools

1 引言

问答系统(Question Answering System, QA)被视为信息检索的一种高级形式,其能用准确、简洁的自然语言回答用户用自然语言提出的问题,是人工智能和自然语言处理领域中目前倍受关注并具有广泛发展前景的研究方向之一^[1]。近年来,多用户交互 QA 系统得到了研究者的关注^[1-5]。在多用户交互 QA 系统中,通过与用户的交互,可以提高搜索结果的实时性和效率^[1]。文献[2]定义了一种语义模板用于多用户交互 QA 系统中问题和答案的语义表达。文献[3]将多用户交互 QA 系统引入到个性化 E-Learning 系统中,用于获取用户的兴趣和能力。文献[4]研究了利用多用户交互 QA 获取 E-Learning 系统中用户知识需求的方法。文献[5]对有偿交互 QA 系统中用户的信誉模型进行了研究。

在多用户有偿交互 QA 系统中,投诉机制的引入可以有效弥补兑现的不公平性等问题。研究有偿交互 QA 系统的投诉模型,进而对有偿交互 QA 系统进行综合形式化分析,具有重要意义^[6]。

图 1 给出了有偿交互 QA 系统的框架图^[2,3]。

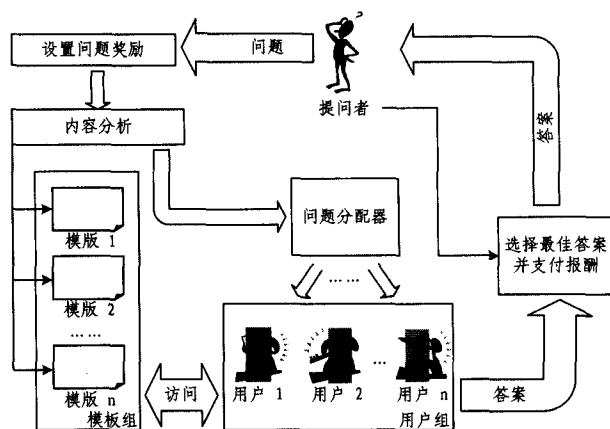


图 1 有偿交互 QA 系统的框架图

系统与用户交互的基本过程为:用户在系统中提出问题,并对问题设置一定的报酬,其他用户可以来回答问题。得到足够多的答案后,由提出问题的用户对其他用户的回答做出评价,评出最佳答案,并兑现相应的报酬。用户从所有答案中选择最佳答案时,可能会出现没有兑现给回答者相应的报酬,

到稿日期:2013-04-07 返修日期:2013-05-22 本文受国家自然科学基金(61170079,61202152),公益性行业(农业)科研专项(201303107),教育部高等学校博士学科点专项科研基金(20103718110007),山东科技大学杰出青年基金(2010KYJQ101)资助。

李彦成(1988—),男,硕士生,主要研究方向为人工智能与信息处理,E-mail:yancheng_lee@163.com;曾庆田(1976—),男,博士,教授,博士生导师,主要研究方向为个性化信息服务、流程挖掘、Petri 网等,E-mail:qtzeng@163.com(通信作者);鲁法明(1981—),男,讲师,主要研究方向为 Petri 网、业务流程建模等;薛洁(1989—),女,硕士生,主要研究方向为 Petri 网、Web 服务组合。

或者其他回答者认为提问者对最佳答案的选择存在不合理的情况。这些情况一旦出现,就会导致问题回答者对提问者的投诉。

颜色 Petri 网^[7-9]在描述系统的并行、分叉、同步等行为以及处理数据方面相比其他建模方法具有强大优势,可以对模型的多个实例进行并发的模拟,在描述系统流程中得到了广泛的应用。同时,由于颜色 Petri 网拥有自身的建模和分析工具——CPN Tools^[10-12],使得颜色 Petri 网在实际应用中具有独特的优势。

本文在已有工作的基础上^[2-6],基于颜色 Petri 网研究有偿交互 QA 系统中的投诉模型,使用 CPN Tools 对所建立的模型进行形式化验证和分析,得到了所建模型的状态空间报告,证明了系统各响应流程设计的正确性等特性。

2 有偿交互 QA 系统的投诉流程框架

用户可通过系统提供的投诉机制进行问题投诉。图 2 给出了有偿交互 QA 系统中的投诉流程框架,共包括 4 类角色:投诉者、提问者、投诉管理员和仲裁者。

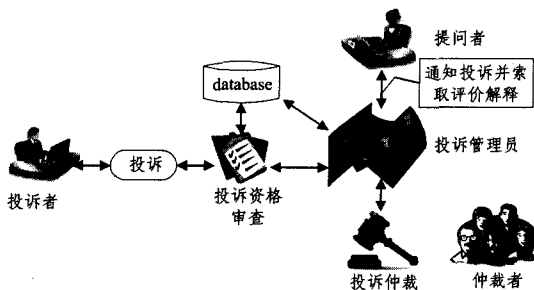


图 2 有偿交互 QA 系统中的投诉流程框架

一次完整的投诉流程具体包括如下步骤:

- (1) 投诉者向系统的投诉管理员提出投诉申请。
- (2) 投诉管理员对投诉者的投诉请求进行审核。

这个步骤是为了避免某些恶意用户随意提交投诉拖慢整个系统。对投诉请求进行审核的粗略的方法是根据投诉者和提问者的信誉值来判断是否接收投诉请求。判定方法如下:

(a) 检测系统数据库。如果系统数据库中不存在投诉者的多条投诉记录,并且投诉者的信誉值 $< \beta$,判定为投诉失败,系统将拒绝投诉者的投诉请求。

(b) 否则,系统接收投诉者的投诉请求,并由系统进行记录处理。

其中我们设定的 β 称为投诉系数。一般情况下,可以依据不同系统的需要来调整投诉系数。

(3) 投诉管理员向相应提问者索取与投诉相关的数据信息,包括提问的问题、当前最佳答案、投诉答案和提问者的评价解释。

(4) 投诉管理员对投诉进行处理,容易判断的投诉直接由系统仲裁人员进行判定。有争议的投诉,将相关的数据信息发到投诉讨论区,并设定投诉仲裁的期限,由仲裁者进行投票仲裁。到达投诉仲裁期限时,由系统对投票结果进行整理、建

档、存储,汇总评价,得出结论。

(5) 如果投诉成功,规定提问者付给投诉者相应的报酬,扣掉提问者的信誉值;如果投诉失败,规定扣除投诉者的信誉值,同时把投诉结果回馈给提问者和投诉者。

3 形式化模型

颜色 Petri 网^[7-9]在经典 Petri 网的基础上对 Token 进行了颜色扩展,将库所中的 Token 染色。染色的方法既能体现个性,又能体现共性,其实质就是对 Token 进行分类,以实现网系统的折叠。这也是目前 CPN 称为应用主流模型的原因之一。Token 由于其值不同而具有不同的颜色,一个 Token 的所有颜色称为一个颜色集,这样一个库所中就可以包含几个不同颜色的 Token。要使变迁能够发生,必须指定消耗库所中的哪一类 Token,与变迁的有向弧相连的输入库所含有要求的 Token,若满足变迁表示的条件,则变迁被激发,产生新的 Token 送到变迁的输出库所中。

3.1 颜色 Petri 网定义

颜色 Petri 网是一个 9 元组^[12], $CPN = (S, P, T; F, N, C, G, E, I)$, 其中 $(P, T; F)$ 是一个 Petri 网结构;

(1) S 是一个非空的类型有限集,称为颜色集 (Color Sets);

(2) P 是一个库所的有限集;

(3) T 是一个变迁的有限集, $P \cap T = \emptyset$;

(4) F 是一个弧的有限集,满足 $P \cap F = T \cap F = \emptyset$;

(5) N 是一个节点函数 (Node function), 是定义从 F 到 $P \times T \cup T \times P$ 的函数;

(6) C 是一个颜色函数 (Color function), 定义为 $\forall p \in P$ 都有 $C(p) \in S$;

(7) G 是一个守卫函数 (Guard function), 是定义在 T 上的表达式, 并满足 $\forall t \in T: [Type(G(t)) = Boolean \wedge Type(Var(G(t))) \subseteq S]$;

(8) E 是一个弧表达式函数 (Arc expression function), 是定义在 F 上的表达式, 并满足 $\forall f \in F: [Type(E(f)) = C(p)_{MS} \wedge Type(Var(E(f))) \subseteq S]$;

(9) I 是一个初始化函数 (Initialization function), 是定义在 P 上的封闭表达式 (Closed expression), 并满足 $\forall p \in P: [Type(I(p)) = C(p)_{MS}]$ 。

3.2 有偿交互 QA 系统的投诉模型建模

在对有偿交互 QA 系统的投诉模型进行建模时,使用 CPN Tools^[10-12]作为颜色 Petri 网的建模和分析工具。CPN Tools 可以用于颜色 Petri 网的编辑和实现对颜色 Petri 网模型的仿真,还可以对所建立的颜色 Petri 网模型的性能(如活性、有界等)进行分析,是一个应用广泛的颜色 Petri 网自动化分析工具。

3.2.1 颜色集及变量

根据颜色 Petri 网的理论,针对有偿交互 QA 系统的投诉模型,首先定义了颜色集与变量,如表 1 所列。

表 1 颜色集与变量定义

颜色集	变量	备注
color USERID=int with 1...4;	userid,k,j;	有界整数类型,表示用户 ID
color QID=int with 1...5;	qnum	有界整数类型,表示问题 ID
color CREDIT=int with 1...10;	credit	有界整数类型,表示用户信誉值
color RESULT=with Acc Ref Suc Fail Undo;	result	枚举类型,表示投诉处理结果
color QDATA=string;	qdata	字符串类型,表示问题相关数据
color QTYPE=with Easy Complex;	qty	枚举类型,表示问题复杂程度
color CPACK=record uid;USERID * qid;QID * crd;CREDIT * res;RESULT;	cp	记录类型,表示投诉信息集合
color QPACK=record qid;QID * qdata;QDATA * qty;QTYPE;	qp	记录类型,表示问题信息集合
color UNITPACK=product CPACK * QPACK;	up	积类型,表示全部投诉信息集合
color RANDOM=int with 1...2;	random	有界整数类型,表示随机数

3.2.2 基于颜色 Petri 网的有偿交互 QA 系统的投诉模型
利用 CPN Tools 对有偿交互 QA 系统的投诉模型进行

建模,模型如图 3 所示。
在图 3 中,库所和变迁的说明分别如表 2 和表 3 所列。

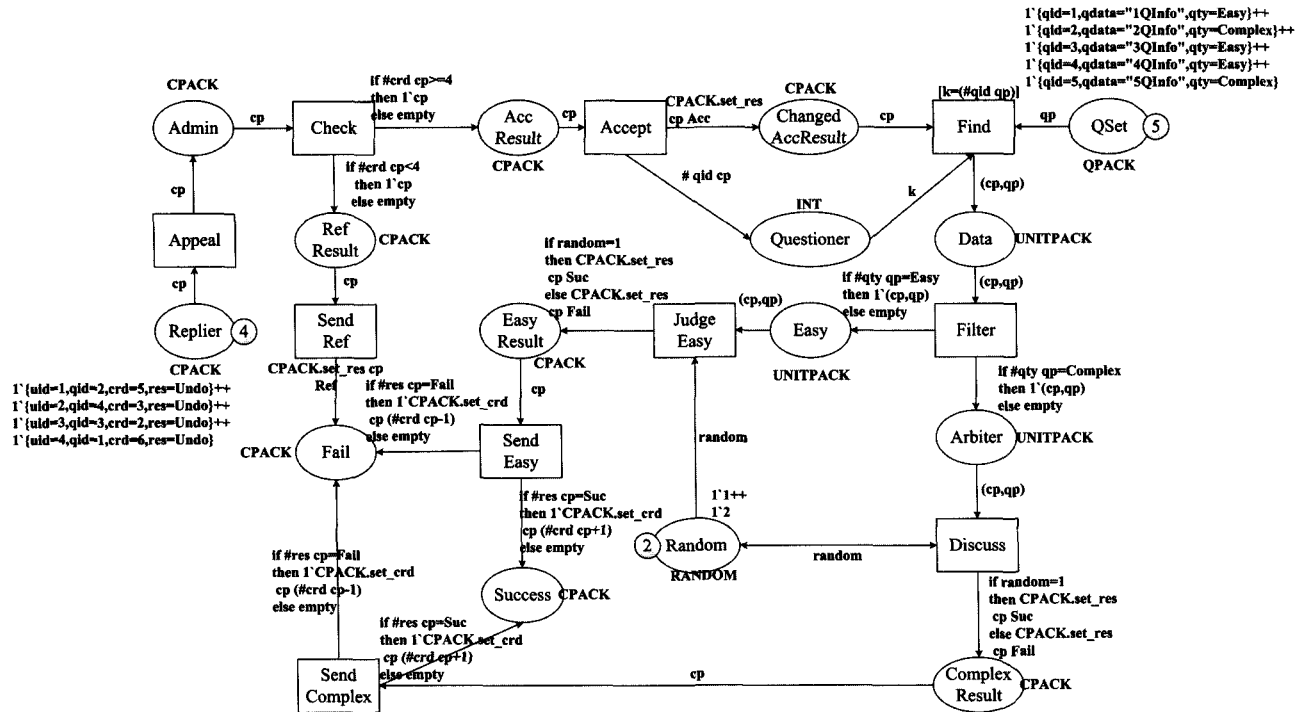


图 3 基于颜色 Petri 网的有偿交互 QA 系统的投诉模型

表 2 库所说明

库所	颜色集	说明
Replier	CPACK	投诉者
Admin	CPACK	投诉管理员
RefResult	CPACK	拒绝的投诉请求
AccResult	CPACK	接受的投诉请求
ChangedAccResult	CPACK	改变处理结果后的投诉请求
QSet	QPACK	问题集合
Questioner	INT	提问者
Data	UNITPACK	投诉请求的完整数据
Easy	UNITPACK	筛选结果为容易的投诉请求
EasyResult	CPACK	修改后的投诉请求
Arbitr	UNITPACK	仲裁讨论区
ComplexResult	CPACK	筛选结果为复杂的投诉请求
Random	RANDOM	随机处理结果
Fail	CPACK	投诉失败
Success	CPACK	投诉成功

表 3 变迁说明

变迁	说明
Appeal	投诉
Check	系统审查
SendRef	发送拒绝处理结果
Accept	接受投诉请求
Find	查找投诉相关问题信息
Filter	筛选投诉问题
JudgeEasy	投诉管理员判断投诉问题
Discuss	仲裁讨论区讨论投诉问题
SendEasy	发送简单投诉处理结果
SendComplex	发送复杂投诉处理结果

其中,库所 Replier 中有 4 个托肯,代表 4 个投诉请求,模拟 4 次投诉。Admin 代表管理员接收到投诉请求,通过变迁 Check,判断投诉者的信誉值是否满足系统要求,如果不满足,进入 RefResult 库所,发送失败信号,通知投诉者投诉请求被拒绝,进入 Fail 库所,完成投诉流程;如果投诉者信誉值满足系统要求,则接受投诉请求,进入 AccResult 库所,系统接受投诉,进入 ChangedAccResult 库所,同时通知提问者,向提问

者索取相关问题数据,提问者从问题集合中查找投诉的问题,得到投诉所需的完整数据进入 Data 库所。管理员对投诉数据进行筛选,经过 Filter 变迁,如果投诉问题比较容易,进入 Easy 库所,投诉管理员对投诉请求进行判定,得到投诉结果,进入 EasyResult 库所,通过 SendEasy 变迁增加或者扣除投诉者的信誉值,相应地进入 Fail 库所或者 Success 库所,完成投诉流程;如果投诉问题比较复杂,难以判断,则进入 Arbiter 库所,仲裁讨论区对投诉问题进行讨论,得到投诉结果,进入 ComplexResult 库所,经由 SendComplex 变迁增加或者扣除投诉者的信誉值,相应地进入 Fail 库所或者 Success 库所,完成整个投诉流程。

4 模型性能分析

利用 CPN Tools 对有偿交互 QA 系统的投诉模型建模进行仿真模拟,确定该模型是否可以表示所描述的投诉流程。同时对所建立的模型进行结构分析,得到标准的状态空间报告,从而确定是否存在死锁。对 CPN Tools 得出的标准状态空间进行统计数字分析、活性分析和有界性分析。

(1)状态空间分析

表 4 列举了有偿交互 QA 系统的投诉模型仿真得到的统计信息。State Space 给出了状态空间的节点与连接弧的个数,在该模型中分别为 2304 和 7680,构造该状态空间过程花费了 2 秒的时间;而 Scc Graph 为强连接图,其节点和连接弧的个数与状态空间图相等,符合建模要求。

表 4 状态空间报告——统计数字分析

统计	节点	弧	所用时间(秒)	状态
State Space	2304	7680	2	Full
Scc Graph	2304	7680	0	

(2)活性分析

表 5 列出活性分析结果。Dead Markings 指的是死标识状态,在该标识下任何赋值变迁都是非使能的。Dead Transition 指的是死的变迁,即该变迁在该模型中的任何发生序列中都无法发生,一般情况下死变迁的出现预示着模型设计的不合理,可以去掉死变迁而不影响模型的动态行为。由结果看出,本实验中的模型不存在死变迁。Live Transition 指的是活的变迁,即在任何可达标识下,该变迁都在一个可发生序列中。表 5 显示本实验中的模型不存在活的变迁,这很正常,因为该模型存在死标识状态,该标识下任何变迁都是不可以发生的,实际上我们知道这时的状态就是模拟 4 次投诉请求处理结束的状态。这说明模型除了结束标识外,其它的标识都不是死标识,说明模型不存在死锁的状态。

表 5 状态空间报告——活性分析

Dead Markings	[2301,2302,2303,2304]
Dead Transition Instances	None
Live Transition Instances	None

(3)有界性分析

表 6 给出了每个库所拥有的托肯的数量区间,如库所 Replier 最多有 4 个托肯,代表初始情况下有 4 个投诉请求,最少可以没有托肯,代表投诉请求已经提交系统,其它库所也是有穷的。从结果可以看出,该模型中的库所是符合实际应用要求的。

表 6 状态空间报告——有界性分析

最佳整型边界	上界	下界
New_Page'Acc_Result 1	2	0
New_Page'Admin 1	4	0
New_Page'Arbiter 1	1	0
New_Page'Changed_AccResult 1	2	0
New_Page'Complex_Result 1	1	0
New_Page'Data 1	2	0
New_Page'Easy 1	1	0
New_Page'Easy_Result 1	1	0
New_Page'Fail 1	4	0
New_Page'QSet 1	5	3
New_Page'Questioner 1	2	0
New_Page'Random 1	2	2
New_Page'Ref_Result 1	2	0
New_Page'Replier 1	4	0
New_Page'Success 1	2	0

结束语 本文利用颜色 Petri 网对有偿交互 QA 系统中的投诉模型进行了形式化建模,并利用针对颜色 Petri 网的自动验证工具 CPN Tools 对所建立的模型进行了验证。通过颜色 Petri 网,描述了系统的整体性和个体性,描述了有偿交互 QA 系统中用户之间的交互行为,并能够对系统进行正确性分析。在本文工作的基础上,尚需对投诉模型中信誉度评价和投诉讨论区仲裁流程进行细化;也可尝试用其他形式化建模方法对所建立的模型进行分析验证,同时比较它们之间的优缺点。

参 考 文 献

[1] 刘里,曾庆田. 自动问答系统研究综述[J]. 山东科技大学学报, 2007,26(4):73-76

[2] Hao Tian-yong, Hu Da-wei, Liu Wen-yin, et al. Semantic Patterns for User-Interactive Question Answering, Concurrency and Computation[J]. Practice and Experience,2008,20(7):783-799

[3] Hu Da-wei, Chen Wei, Zeng Qing-tian, et al. Using a User-interactive QA System for Personalized e-Learning[J]. The International Journal of Distance Education Technologies,2008,6(3): 1-22

[4] Zeng Qing-tian,Zhao Zhong-ying,Liang Yong-quan. Course Ontology-based User's Knowledge Requirement Acquisition from Behaviors within E-Learning Systems[J]. Computers & Education,2009,53(3):809-818

[5] Chen Wei,Zeng Qing-tian,Liu Wen-yin,et al. A user reputation model for a user-interactive question answering system[J]. Concurrency Computation Practice and Experience,2007,19(15): 2091-2103

[6] Li Yan-cheng, Duan Hua, Zeng Qing-tian, et al. A π -calculus-based Complaint Model for a User-interactive Question Answering System[C]//Proceedings of the 2012 2nd International Conference on Computer and Information Application (ICCIA 2012). 2012:97-100

[7] 吴哲辉. Petri 网导论[M]. 北京:机械工业出版社,2006

[8] Jensen K. Colored Petri Nets: A High Level Language for System Design and Analysis [M]. Heideberg, Berlin:Springer,2006

[9] Jensen K. An introduction to the theoretical of colored Petri nets [J]. Computer Science,1994,83:230-272

[10] Lu Xi. Scop design with CPN Tools [D]. Eindhoven University of Technology,2005

[11] 管华,应时,贾向阳,等. 一种基于 CPN 的 BPEL 异常处理逻辑的开发方法[J]. 计算机科学,2013,40(1):150-156

[12] Jensen K L,Kristensen M,Wells L. Colored Petri Nets and CPN Tools for Modeling and Validation of Concurrent Systems[J]. International Journal on Software Tools for Technology Transfer,2007,9(3/4):213-254