

## 基于可信度的辩论模型及争议评价算法<sup>\*</sup>

熊才权, 欧阳勇, 梅清

(湖北工业大学 计算机学院, 湖北 武汉 430068)

通讯作者: 熊才权, E-mail: x\_cquan@163.com

**摘要:** 辩论是智能主体间为了消除分歧的一种基于言语的交互行为. 由于知识的局限性, 争议以及争议内部的陈述通常存在不确定性, 因此在对辩论进行建模时需要考虑不确定信息处理问题. 提出一种基于可信度的辩论模型(CFA), 该模型将争议表示为由若干前提和一个结论组成的可废止规则, 并用对话树描述辩论推演过程. 为了表示不确定性推理, 引入可信度模型, 将争议前提的不确定性和争议之间的攻击强度统一用可信度因子表示. 在此基础上, 提出计算陈述可信度的争议评价算法, 并通过设定可信度阈值确定陈述的可接受性, 得出最终辩论结果. 最后, 用一个实例说明该方法的有效性. 该模型可以有效处理不确定信息条件下辩论推理过程, 其辩论算法建立在数值计算基础之上, 所得出的可接受陈述集在给定可信度阈值条件下是唯一的, 可以克服 Dung 的抽象辩论框架中扩充语义的不足.

**关键词:** 辩论模型; 不确定性; 可信度; 争议评价算法

**中图法分类号:** TP181

中文引用格式: 熊才权, 欧阳勇, 梅清. 基于可信度的辩论模型及争议评价算法. 软件学报, 2014, 25(6): 1225-1238. <http://www.jos.org.cn/1000-9825/4446.htm>

英文引用格式: Xiong CQ, Ouyang Y, Mei Q. Argumentation model based on certainty-factor and algorithms of argument evaluation. Ruan Jian Xue Bao/Journal of Software, 2014, 25(6): 1225-1238 (in Chinese). <http://www.jos.org.cn/1000-9825/4446.htm>

### Argumentation Model Based on Certainty-Factor and Algorithms of Argument Evaluation

XIONG Cai-Quan, OUYANG Yong, MEI Qing

(School of Computer Science, Hubei University of Technology, Wuhan 430068, China)

Corresponding author: XIONG Cai-Quan, E-mail: x\_cquan@163.com

**Abstract:** Argumentation is a kind of interaction between agents for resolving differences in speech. Due to the limitation of knowledge, arguments and its inner statements are usually uncertain, and therefore it is necessary to consider uncertain information processing when modeling the process of argumentation. This paper proposes an argumentation model based on the theory of certainty-factor, in which an argument is expressed as a defeasible rule with several premises but only one conclusion, and the process of argumentation is described as a dialogue tree. In order to do uncertain reasoning, the uncertainty of argument's premise and the strength of the attack that one argument makes against another are assigned with certainty-factors. Next, the algorithms of argument evaluation are provided to compute the credibility of statements. Using the algorithms the set of acceptable arguments can be determined by the credibility threshold and the result of argumentation can be obtained. Finally, an example is given to illustrate the validity of the method. The algorithms are based on numerical computation where the set of acceptable arguments is unique under a given credibility threshold, and can overcome the shortage of extension semantics of Dung's abstract argumentation framework.

**Key words:** argumentation model; uncertainty; certainty-factor; algorithms of argument evaluation

<sup>\*</sup> 基金项目: 国家自然科学基金(61075059, 61300127); 湖北省教育厅科研计划重点项目(D20101402); 湖北省自然科学基金(2013CFB021)

收稿时间: 2012-11-21; 修改时间: 2013-03-07; 定稿时间: 2013-06-21; jos 在线出版时间: 2013-11-01

CNKI 网络优先出版: 2013-11-01 14:59, <http://www.cnki.net/kcms/detail/11.2560.TP.20131101.1459.006.html>

辩论(argumentation)是智能主体(agent)间为了消除信念不一致的一种基于言语的交互行为.传统上,它是哲学、心理学、语言学、逻辑学和法律等领域的研究课题.但是随着计算机辅助辩论支持技术的发展<sup>[1,2]</sup>与多 Agent 系统的出现<sup>[3]</sup>,辩论开始成为人工智能领域的一个重要研究课题<sup>[4-7]</sup>.从逻辑推理过程来看,辩论是一种从可废止的前提推理出结论的非单调推理的直觉范型.辩论是通过主体之间的对话(dialogue)进行的,一条有明确主张和相应根据的发言(utterance)称为争议(argument).争议是由陈述(statement)组成的.陈述是对事物的肯定性表述,所有陈述统称为语言(language).由于主体间信念不一致或信息不对称,争议之间存在攻击和支持等关系,从而使争议的可防卫性和陈述的可接受性随着辩论推演的进行而产生变化.辩论模型是对辩论推演过程的形式化描述,其研究内容包括辩论空间构造(描述争议内部结构及辩论推演过程)以及辩论结果生成算法(确定争议的可防卫性和陈述的可接受性).

目前,对辩论模型的研究已有很多.在争议内部结构建模方面,影响较大的有 Toulmin 模型<sup>[8]</sup>,它把一条争议分为主张(claim)、根据(premise)、论证(warrant)、支援(backing)、模态限定(modality)和反驳(rebuttal)这 6 个部件.但 Toulmin 模型不能反映辩论推演过程,且 Toulmin 的模型元素过多,不利于辩论算法设计.不少研究对 Toulmin 模型做了简化和改进<sup>[9]</sup>.在辩论过程建模方面,影响较大的是 IBIS(issue-based information system)模型<sup>[10]</sup>.IBIS 模型主要用于描述问题求解过程,它将辩论过程中产生的发言信息构成一个以问题为中心的树型结构,将发言分为提出问题(issue)、表明立场(position)、给出理由(argument)这 3 种类型,但它没有提供对理由真伪的辩论,即没有表示辩论推演的层次关系,也没有提供相应的争议评价算法.对于争议之间的关系建模,影响较大的是 Dung 抽象辩论框架<sup>[6]</sup>,该框架将辩论空间表示为一个二元组: $AF=(AR,attacks)$ ,其中,AR 是争议集, $attacks \subseteq AR \times AR$  是 AR 上的二元关系.为了确定可防卫争议集,Dung 提出了优先扩充(preferred extension)、稳定扩充(stable extension)、完全扩充(complete extension)和基础扩充(grounded extension)等语义扩充.Dung 的抽象辩论框架把争议看做一个整体而忽视争议的内部结构,且没有表示争议之间的攻击强度,这与实际辩论存在较大差距.另外,Dung 虽然提出 4 种扩充语义及它们之间的关系,但他并没有指出哪种扩充是合适的解(即可接受争议集),而且判断一个争议是否在优先扩充或稳定扩充内是 NP 完全问题<sup>[7]</sup>.不少研究对 Dung 抽象辩论框架进行了扩展.文献[11]将争议扩展为  $h_1, \dots, h_n \Rightarrow h$ ,其中,  $h_i (i=1, 2, \dots, n)$  称为前提,  $h$  称为结论.文献[12]在 Dung 的争议攻击关系基础上增加了支持关系,提出双极辩论框架.文献[13]提出了一种辩论模型,该模型用简化的 Toulmin 模型表示争议内部结构,并增加了支持关系.为了进行争议评价计算,先对辩论框架进行约简得到对话树,使该模型与 Dung 抽象辩论框架得到统一,再设计算法求出对话树的可防卫争议集,最后回溯到辩论框架求出辩论框架的可防卫争议集和可授受陈述集.文献[14]提出了一种扩展双极辩论框架,增加了对争议关系的响应,其对辩论的刻画更为全面.

以上这些模型的一个共同特点是争议的提出者对争议的前提和争议之间的攻击或支持关系都是明确的,即争议的前提是完全可信的,争议之间关系要么是确定地支持,要么是确定地反对.

但是在实际辩论过程中,由于知识的不确定性、不完备性,争议前提和争议本身存在不确定性,因此,一个完备的辩论系统应该能处理前提的不确定性和争议之间的攻击强度.在这方面代表性的研究有 Amgoud 等人<sup>[15]</sup>提出的基于优先序的辩论框架(preference-based argumentation framework,简称 PAF)、Bench-Capon<sup>[16]</sup>提出的基于价值的辩论框架(value based argumentation frameworks,简称 VAF)、Dunne 等人<sup>[17]</sup>提出的赋权辩论框架(weighted argumentation systems,简称 WAS)、Haenni 等人<sup>[18]</sup>提出的概率辩论系统(probabilistic argumentation systems,简称 PAS)、Tang 等人<sup>[19]</sup>和 Das<sup>[20]</sup>提出的基于 Dempster-Shafer 理论的辩论框架(Dempster-Shafer based argumentation,简称 DSA)等.这些模型用不同方式描述了辩论模型中的不确定性,其中 PAF, VAF 和 WAS 在 Dung 的抽象辩论框架基础上增加了对争议强度或争议之间的攻击强度的表示,但仍用扩充语义来表示争议的可接受性,其计算依然很复杂.PAS 和 DSA 用主观概率表示争议前提的不确定性,并通过数值计算确定争议的可接受性,但它们对辩论推演过程的刻画不充分.

本文提出了一种基于可信度的辩论模型(certainty-factor based argumentation,简称 CFA),其基本思想是对辩论系统从定性和定量两个方面进行刻画:定性部分主要是描述辩论的推演过程,我们延用文献[13]中提出的辩

论框架;定量部分描述争议前提的可信度和争议的攻击强度,我们统一用可信度因子表示.在此基础上提出争议评价算法,该模型可以有效处理不确定信息条件下的辩论推理过程,其辩论算法建立在数值计算基础之上,所得出的可接受陈述集在给定的可信度阈值条件下是唯一的,可以克服 Dung 的抽象辩论框架中的扩充语义的不足.

本文第 1 节简要介绍基本辩论框架,基于可信度的辩论模型是对该辩论框架的扩展.第 2 节介绍基于可信度的辩论模型,包括争议前提的不确定性和争议之间的攻击强度的表示,以及争议评价算法.第 3 节用一个实例说明该模型和算法的有效性.第 4 节是相关工作比较,主要将本文提出模型和算法分别与 PAF, VAF, WAS 和 DSA 进行比较.最后一节是本文的结论与展望.

## 1 基本辩论框架

在引入可信度方法之前,先介绍基本辩论框架,该框架建立在 Dung 的抽象辩论框架基础之上,但对其进行扩展<sup>[13]</sup>.基本辩论框架主要描述争议结构及争议之间的关系.争议可分解为前提和结论两个部分,前提和结论统称为陈述.争议之间的关系统一用对话表示.

**定义 1(statement).** 陈述是对事物的肯定性表述,它是争议的基本组成单位.所有陈述的集合称为语言(language),记为 $\mathcal{L}$ .

**定义 2(consistency of two statements).** 如果两个陈述  $h_i, h_j (i \neq j)$  所表述的内容在逻辑上是一致的,则记为  $h_i \equiv h_j$ ; 如果它们所表述的内容在逻辑上是相反的,则记为  $h_i \equiv \neg h_j$ .

**定义 3(conflict-free subset of language).** 设有一个陈述子集  $S \subseteq \mathcal{L}$ , 如果  $\nexists h_i, h_j \in S, h_i \equiv \neg h_j$ , 则称  $S$  是相容的; 否则, 称  $S$  是不相容的.

**定义 4(argument).** 争议是一个二元组  $A=(H, h)$ , 其中,  $h \in \mathcal{L}$  是一个陈述,  $H=\{h_1, \dots, h_n\} \subseteq \mathcal{L}$  是一个陈述子集, 且满足: i)  $H$  是相容的; ii) 逻辑上  $H$  推理出  $h$ , 记为  $H \Rightarrow h$ ; iii)  $H$  是最小的, 即没有  $\{h_1, \dots, h_n\}$  的真子集满足条件 i) 和条件 ii).  $h_1, \dots, h_n$  等叫做争议的前提(premise), 记为  $Pre(A)=\{h_1, \dots, h_n\}$ , 即  $h_i \in Pre(A) (1 \leq i \leq n)$ ;  $h$  叫做争议的结论(conclusion), 记为  $Con(A)=h$ . 所有争议的集合记为  $\mathcal{A}$ .

**定义 5(abstract argument).** 如果忽略争议的内部结构, 则该争议称为抽象争议.

争议  $A=(\{h_1, \dots, h_n\}, h)$  可以用图 1(a) 表示, 图 1(b) 是它的抽象表达形式, 它省略了争议的内部结构.

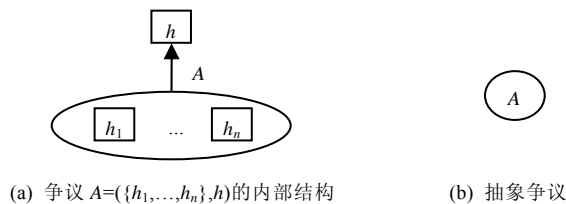


Fig.1 Argument node

图 1 争议节点

**定义 6(dialogue).** 如果一个争议  $B$  的结论是另一个争议  $A$  的一个前提, 即  $Con(B) \in Pre(A)$ , 则称为争议  $B$  对争议  $A$  的对话, 记为  $\langle B, A \rangle$ . 所有对话的集合记为  $\mathcal{D}$ .

**定义 7(argumentation framework).** 辩论框架是一个三元组  $AF=(\mathcal{L}, \mathcal{A}, \mathcal{D})$ , 其中,  $\mathcal{L}$  是陈述集,  $\mathcal{A}$  是争议集,  $\mathcal{D}$  是对话集.

**定义 8(head and rear).** 对于辩论框架  $AF=(\mathcal{L}, \mathcal{A}, \mathcal{D})$  的一个争议  $A$ , 如果  $\nexists B \in \mathcal{A}, \langle A, B \rangle \in \mathcal{D}$ , 即  $A$  不对任何其他争议进行响应, 则称  $A$  为首争议(head); 如果  $\nexists B \in \mathcal{A}, \langle B, A \rangle \in \mathcal{D}$ , 即  $A$  不被任何其他争议响应, 则称  $A$  为尾争议(rear).

**定义 9(abstract argumentation framework).** 忽略了争议内部结构的辩论框架称为抽象辩论框架. 抽象辩论框架是一个二元组, 记为  $AAF=(\mathcal{A}, \mathcal{D})$ .

抽象辩论框架可以用一个有向图表示, 称为对话树.

**定义 10(dialogue tree).** 对话树是  $n(n>0)$  一个抽象争议的有限集合.在任意一棵非空对话树中,有且仅有一个抽象争议节点称为根.当  $n>1$  时,除根之外的其余节点可以分为  $m(m>0)$  个互不相交的有限集  $T_1, T_2, \dots, T_m$ , 其中,  $T_i (1 \leq i \leq m)$  本身又是一棵树,称之为根的子对话树.每棵子树的根用一条由子树指向根的有向边相连.

如果考察辩论框架, 争议的对话子树都是针对该争议的某个前提.

例:设有一辩论框架,有如下争议:

$$A = (\{b, c\}, a), B = (\{d, e\}, b), C = (\{f\}, b), D = (\{g\}, c), E = (\{h, i\}, d), F = (\{j, k\}, f),$$

则该辩论框架中:陈述集  $\mathcal{L} = \{a, b, c, d, e, f, g, h, i, j, k\}$ , 争议集  $\mathcal{A} = \{A, B, C, D, E, F\}$ , 对话集  $\mathcal{D} = \{\langle B, A \rangle, \langle C, A \rangle, \langle D, A \rangle, \langle E, B \rangle, \langle F, C \rangle\}$ . 图 2(a)描述了本例的辩论过程,图 2(b)是本例的对话树.

本例中争议 A 有 3 个对话子树,其中  $\{B, E\}$  和  $\{C, F\}$  针对 A 的前提 b,  $\{D\}$  针对 A 的前提 c.

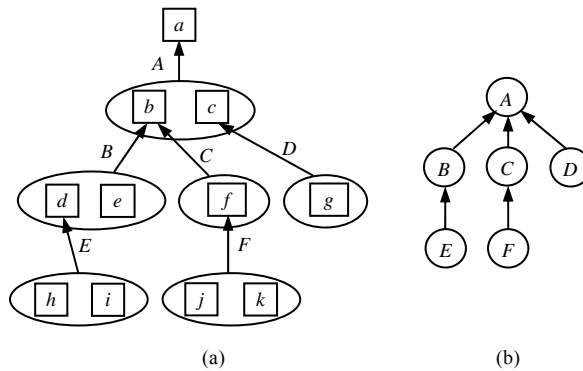


Fig.2 An example of argumentation

图 2 一个辩论例子

## 2 基于可信度的辩论模型(CFA)

### 2.1 可信度方法

可信度方法是由 Shortliffe<sup>[21]</sup>提出的一种表示和处理不确定性推理的有效方法,简称 C-F 模型,早期在基于规则的专家系统(如医疗系统 MYCIN 系统)中得到成功应用.

**定义 11(certainty factor,简称 CF).** 可信度是指专家根据经验对一个事物和现象判断为真的程度,它包括证据可信度和规则可信度.可信度用可信度因子表示.

可信度是一种主观经验判断,但对于某一领域的专家,他所给出的可信度可能会更接近于实际.由专家直接给出对证据、规则的可信度,再通过可信度的传递和合成来进行不确定性推理,可以避免对先验概率或条件概率的要求.不同系统对可信度的量化方法不同,主要有语言法(如绝对确定、可能、不确定、不太可能、绝对不可能等)和数字尺度(如 0~10,0~-1,或-1~1 等).

**定义 12(certainty factor of evidence).** 设 E 为证据(evidence),则 CF(E)为证据 E 的可信度因子,其取值范围在  $[-1, 1]$  内,即  $-1 \leq CF(E) \leq 1$ .

证据分初始证据和间接证据.初始证据的可信度由专家给出.如对它的所有观察都能肯定为真,则  $CF(E)=1$ ; 如能肯定为假,则  $CF(E)=-1$ ;若它以某种程度为真,则  $0 < CF(E) < 1$ ;若它以某种程度为假,则  $-1 < CF(E) < 0$ ;若不能确定其真假,则置  $CF(E)=0$ .间接证据是指当前推理的证据是上一次推理的结论,则其可信度由上次证据的可信度通过不确定性传递而计算出来.

当证据 E 是多个单一证据的合取时,即  $E = E_1 \wedge \dots \wedge E_n$ , 则  $CF(E) = \min\{CF(E_1), \dots, CF(E_n)\}$ , 即取单一证据可信度的最小值.当证据 E 是多个单一证据的析取时,即  $E = E_1 \vee \dots \vee E_n$ , 则  $CF(E) = \max\{CF(E_1), \dots, CF(E_n)\}$ , 即取单一证据可信度的最大值.

**定义 13(certainty factor of rule).** 设有规则 IF  $E$  THEN  $H$ , 其中,  $E$  是规则前提, 也称为证据,  $H$  是结论, 则规则的可信度表示为  $CF(H, E)$ , 其取值范围为  $[-1, 1]$ , 表示  $E$  为真时, 对  $H$  为真的支持程度,  $CF(H, E)$  值越大,  $E$  就越支持  $H$  为真.

**定义 14(certainty factor of conclusion).** 设有规则 IF  $E$  THEN  $H$ , 其中,  $E$  是证据,  $H$  是结论, 则:

$$CF(H) = CF(H, E) \times \max\{0, CF(E)\} \quad (1)$$

当证据确定为真时, 即  $CF(E) = 1$  时, 有  $CF(H) = CF(H, E)$ , 说明当证据存在且为真时, 则结论的可信度与  $CF(H, E)$  的可信度相同; 当证据以某种程度为假时, 即  $CF(E) < 0$  时, 有  $CF(H) = 0$ , 说明该模型没有考虑证据为假时对结论  $H$  的可信度所产生的影响.

**定义 15(combining of certainty factor).** 设有规则:

- IF  $E_1$  THEN  $H$  ( $CF(H, E_1)$ );
- IF  $E_2$  THEN  $H$  ( $CF(H, E_2)$ ).

则结论  $H$  的综合可信度因子可由如下的组合公式计算:

$$CF_{1,2}(H) = \begin{cases} CF_1(H) + CF_2(H) - CF_1(H)CF_2(H), & \text{当 } CF_1(H) \geq 0 \text{ \& } CF_2(H) \geq 0 \text{ 时} \\ CF_1(H) + CF_2(H) + CF_1(H)CF_2(H), & \text{当 } CF_1(H) \leq 0 \text{ \& } CF_2(H) \leq 0 \text{ 时} \\ \frac{CF_1(H) + CF_2(H)}{1 - \min\{|CF_1(H)|, |CF_2(H)|\}}, & \text{当 } CF_1(H) \text{ 和 } CF_2(H) \text{ 异号时} \end{cases} \quad (2)$$

这种组合方法能保证  $CF_{1,2}(H)$  的值在区间  $[-1, 1]$  内.

## 2.2 争议可信度表示

对辩论建模的目的是确定争议的可防卫性和陈述的可接受性. Dung 的抽象辩论框架及其相关研究通过定义一些扩充(如优先扩充、基础扩充、稳定扩充、完全扩充等)来求可防卫争议集, 但判断一个争议是否在优先扩充集或稳定扩充集是 NP 完全问题<sup>[7]</sup>, 因而在实际应用中存在较大的局限性. 因此, 对争议前提及争议之间的攻击强度进行量化, 是对 Dung 的抽象辩论框架进行改进的一个有效办法.

可信度因子是对专家信念不确定性的自然合理的建模. 考虑到专家对不同的初始陈述和争议攻击强度有不同的把握度, 我们引入可信度模型, 并通过设置可信度阈值确定可接受陈述集.

**定义 16(certainty factor of statement).** 陈述  $h$  的可信度记为  $CF(h)$ , 如  $h$  肯定为真, 则  $CF(h) = 1$ ; 如  $h$  肯定为假, 则  $CF(h) = 0$ ; 若它以某种程度为真, 则  $0 \leq CF(h) \leq 1$ .

陈述可信度对应可信度理论中的证据可信度, 但与证据可信度略有不同. 通常专家不会使用可信度为假的证据作为争议的前提, 因此我们假定所有陈述的可信度均为正值, 这将为以后的争议评价算法设计带来方便.

陈述初始可信度由专家直接设定, 它取决于专家的主观经验判断. 在辩论过程中, 陈述可能会遭到其他争议的响应(支持或反对), 从而使其可信度发生改变.

**定义 17(certainty factor of premise).** 争议  $A = (H, h)$  前提可信度记为  $CF(H)$ , 若  $H = \{h_1, \dots, h_n\}$ , 则争议前提可信度取  $H$  中陈述可信度的最小值, 即  $CF(H) = \min\{CF(h_1), \dots, CF(h_n)\}$ .

**定义 18(certainty factor of argument).** 争议  $A = (H, h)$  的可信度是指争议前提  $H$  支持争议结论  $h$  的强度, 记为  $CF(A)$  或  $CF(h, H)$ .

$CF(A)$  的取值范围为  $[-1, 1]$ . 当  $-1 \leq CF(A) < 0$  时, 前提的出现有利于降低结论的可信度, 特别地, 当  $CF(A) = -1$ ,  $CF(H) = 1$  时,  $CF(h)$  降低至 0. 当  $0 < CF(A) \leq 1$  时, 前提的出现有利于提高结论的可信度, 特别地, 当  $CF(A) = 1$ ,  $CF(H) = 1$  时,  $CF(h)$  升高到 1. 当  $CF(A) = 0$  时, 前提的出现对结论的可信度没有影响.

考虑了争议可信度后, 争议  $A = (H, h)$  可以表示为  $H \xrightarrow{CF(h, H)} h$ .

争议可信度由专家在给出争议时直接设定, 它取决于专家的主观经验判断. 争议可信度一旦给定, 在辩论的过程中不再改变.

**定义 19(certainty-factor based argumentation framework).** 基于可信度的辩论框架是一个五元组  $(\mathcal{L}, \mathcal{A}, \mathcal{D},$

$CP, CA$ ), 其中,  $\mathcal{L}$  是陈述集,  $\mathcal{A}$  是争议集,  $\mathcal{D}$  是对话集,  $CP: \mathcal{L} \rightarrow R_L$  是一个给争议前提赋予可信度的函数,  $CA: \mathcal{A} \rightarrow R_A$  是一个给争议赋予可信度的函数,  $R_L$  的值域为  $[0, 1]$ ,  $R_A$  的值域为  $[-1, 1]$ .

### 2.3 争议评价算法

在实际辩论过程中, 专家给出争议时不仅要设定前提和结论, 而且还要设定前提中每个陈述的可信度  $\{CF(h_1), \dots, CF(h_n)\}$  和争议可信度  $CF(h, H)$ . 争议结论的可信度依赖于其前提可信度和争议可信度. 随着辩论的进行, 争议前提可信度会受到其他争议的影响而发生变化, 从而使争议结论可信度发生变化. 争议评价算法就是计算辩论框架中作为争议前提和结论的各陈述的可信度.

#### 2.3.1 争议结论可信度计算

##### (1) 首争议的结论的可信度计算

**定理 1.** 设有争议  $A=(H, h)$  为首争议, 即  $h$  不属于任何争议的前提, 则其结论  $h$  的可信度为

$$CF(h)=CF(h, H)CF(H) \quad (3)$$

证明: 根据定义 16, 争议  $A=(H, h)$  的前提  $H=\{h_1, \dots, h_n\}$  的各陈述可信度满足:  $0 \leq CF(h_i) \leq 1, 1 \leq i \leq n$ . 根据定义 17 有  $0 \leq CF(H) \leq 1$ . 再根据定义 14 有  $CF(h)=CF(h, H) \times \max\{0, CF(H)\}=CF(h, H)CF(H)$ .  $\square$

当争议所有前提确定为真时, 即  $CF(H)=1$  时, 有  $CF(h)=CF(h, H)$ , 即争议结论的可信度与争议可信度相同; 当  $CF(H)=0$  时, 有  $CF(h)=0$ , 即结论为假; 当  $CF(h, H)=1$  时,  $CF(h)=CF(H)$ , 即结论  $h$  的可信度与前提  $H$  的可信度相同.

##### (2) 非首争议的结论的可信度计算

如果争议  $A=(H, h)$  不是首争议, 即  $h$  是某个争议  $B$  的一个前提, 即  $h \in Pre(B)$ , 由于  $h$  作为争议  $B$  的前提已经在之前设定了可信度 (设为  $CF(h)^0$ ), 则争议  $A$  的目的是修改  $h$  的可信度, 因此,  $h$  的可信度应是  $CF(h)^0$  与  $CF(h, H)CF(H)$  的合成.

**定理 2.** 设有争议  $A=(H, h)$ ,  $h$  为争议  $B$  的一个前提, 其初始可信度为  $CF(h)^0$ , 则  $h$  受到  $A$  响应后其可信度为

$$CF(h)=\begin{cases} CF(h)^0 + (1-CF(h)^0)CF(h, H)CF(H), & \text{当 } 0 < CF(h, H) \leq 1 \text{ 时} \\ CF(h)^0, & \text{当 } CF(h, H) = 0 \text{ 时} \\ CF(h)^0(1+CF(h, H)CF(H)), & \text{当 } -1 \leq CF(h, H) < 0 \text{ 时} \end{cases} \quad (4)$$

且  $CF(h)$  的取值在  $[0, 1]$  内, 满足陈述可信度规定的一致性.

证明: 先证明分段计算争议结论可信度的正确性.

##### 1) 如果 $0 < CF(h, H) \leq 1$ , 则

i) 当  $CF(H)$  不变时,  $CF(h)$  随  $CF(h, H)$  线性增大, 如图 3(a) 所示. 图中直线斜率  $k$  随  $CF(H)$  线性增大, 当  $CF(H)=1$  时,  $k$  达到最大值  $k=(1-CF(h)^0)$ . 此时, 如果  $CF(h, H)=1$ , 则  $CF(h)$  达到最大值 1. 当  $CF(H)=0$  时, 不管  $CF(h, H)$  为多少, 都不能改变  $h$  的可信度, 即  $CF(h)=CF(h)^0$ , 此时,  $k=0$ . 可见,  $k=(1-CF(h)^0)CF(H)$ . 于是有:  $CF(h)=CF(h)^0+kCF(h, H)=CF(h)^0+(1-CF(h)^0)CF(h, H)CF(H)$ .

ii) 当  $CF(h, H)$  不变时,  $CF(h)$  随  $CF(H)$  线性增大, 如图 3(b) 所示. 图中直线斜率  $k$  随  $CF(h, H)$  线性增大,  $k=(1-CF(h)^0)CF(h, H)$ . 于是有:  $CF(h)=CF(h)^0+kCF(h, H)=CF(h)^0+(1-CF(h)^0)CF(h, H)CF(H)$ .

可见, 公式(4)的第 1 项成立.

##### 2) 如果 $-1 \leq CF(h, H) < 0$ , 则:

i) 当  $CF(H)$  不变时,  $CF(h)$  随  $CF(h, H)$  的绝对值线性减小, 如图 4(a) 所示. 图中直线斜率随  $CF(H)$  线性增大, 当  $CF(H)=1$  时, 直线斜率达到最大值  $k=CF(h)^0$ . 此时, 如果  $CF(h, H)=-1$ , 则  $CF(h)$  达到最小值 0. 当  $CF(H)=0$  时, 不管  $CF(h, H)$  为多少, 都不能改变  $h$  的可信度, 即  $CF(h)=CF(h)^0$ , 此时, 直线的斜率为  $k=0$ . 可见,  $k=CF(h)^0CF(H)$ . 于是有:

$$CF(h)=CF(h)^0+k(CF(h, H)=CF(h)^0+CF(h)^0CF(h, H)CF(H)=CF(h)^0(1+CF(h, H)CF(H)).$$

ii) 当  $CF(h, H)$  不变时,  $CF(h)$  随  $CF(H)$  线性减小, 如图 4(b) 所示. 图中直线斜率  $k$  随  $CF(h, H)$  线性增大,  $k=CF(h)^0CF(h, H)$ . 于是有:

$$CF(h) = CF(h)^0 + kCF(H) = CF(h)^0 + CF(h)^0 CF(h, H) CF(H) = CF(h)^0 (1 + CF(h, H) CF(H)).$$

可见,公式(4)的第3项成立.

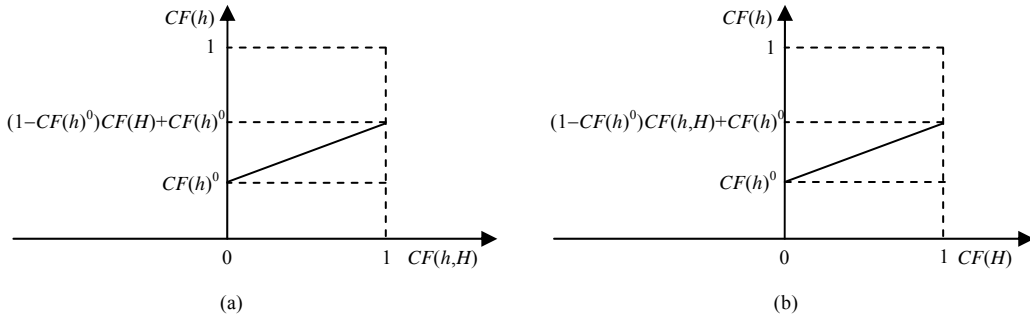


Fig.3 Change of  $CF(h)$  (when  $0 < CF(h, H) \leq 1$ )

图3  $CF(h)$ 变化曲线(当  $0 < CF(h, H) \leq 1$  时)

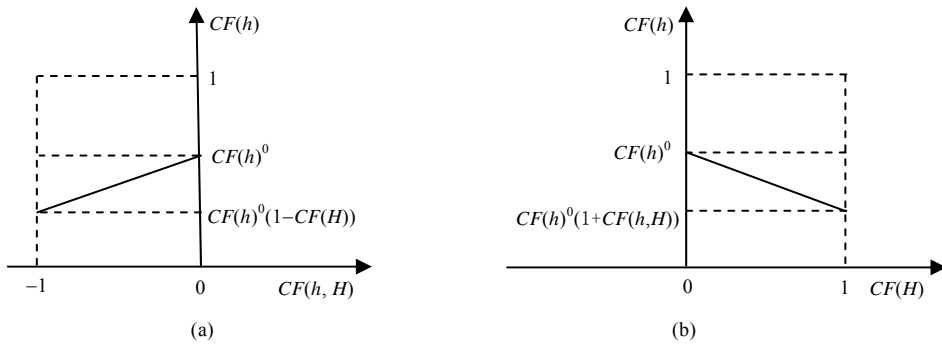


Fig.4 Change of  $CF(h)$  (when  $-1 \leq CF(h, H) < 0$ )

图4  $CF(h)$ 变化曲线(当  $-1 \leq CF(h, H) < 0$  时)

3) 如果  $CF(h, H) = 0$ , 则争议 A 虽然对争议 B 有响应, 但不会改变  $h$  的可信度, 即:  $CF(h) = CF(h)^0$ . 可见, 公式(4)的第2项成立.

再证明  $CF(h)$  的取值在  $[0, 1]$  内.

- 针对公式(4)的第1项:

$$\because 0 < CF(h, H) \leq 1, 0 \leq CF(H) \leq 1,$$

$$\therefore 0 \leq CF(h, H)CF(H) \leq 1,$$

$$\therefore 0 \leq (1 - CF(h)^0)CF(h, H)CF(H) \leq 1 - CF(h)^0,$$

$$\therefore CF(h)^0 \leq CF(h)^0 + (1 - CF(h)^0)CF(h, H)CF(H) \leq 1, \text{ 即 } CF(h)^0 \leq CF(h) \leq 1.$$

- 针对公式(4)的第2项, 有  $CF(h) = CF(h)^0$ .

- 针对公式(4)的第3项:

$$\because -1 \leq CF(h, H) < 0, 0 \leq CF(H) \leq 1,$$

$$\therefore -1 \leq CF(h, H)CF(H) \leq 0,$$

$$\therefore 0 \leq 1 + CF(h, H)CF(H) \leq 1,$$

$$\therefore 0 \leq CF(h)^0(1 + CF(h, H)CF(H)) \leq CF(h)^0, \text{ 即 } 0 \leq CF(h) \leq CF(h)^0.$$

$$\therefore 0 \leq CF(h) \leq 1,$$

$\therefore$  以上3种情况都满足  $0 \leq CF(h) \leq 1$ , 这与定义16的规定是一致的. □

### 2.3.2 可信度合成

如果一个争议的某个前提受到多个争议的响应,则需要进行可信度的合成.

先考虑两个争议的情况.设有争议  $A^1=(H^1,h)$  和  $A^2=(H^2,h), h \in Pre(B)$ , 即争议  $A^1$  和  $A^2$  同时对争议  $B$  的前提  $h$  作出响应, 设  $h$  的初始可信度为  $CF(h)^0$ . 可信度合成可采用以下步骤:

第 1 步:按公式(4)分别求出争议  $A^1$  对  $h$  的可信度的更新值  $CF(h)^1$  和争议  $A^2$  对  $h$  的可信度的更新值  $CF(h)^2$ ;

第 2 步:对  $CF(h)^1$  和  $CF(h)^2$  进行合成.可按公式(2)中的第 1 项进行合成,因为  $CF(h)^1$  和  $CF(h)^2$  的值均大于等于 0, 即有:

$$CF(h)=CF(h)^1+CF(h)^2-CF(h)^1CF(h)^2 \quad (5)$$

**定理 3.** 由公式(5)求得的可信度取值范围为在  $[0,1]$  内,能保证陈述可信度计算的一致性.

证明:由定理 2 知,  $0 \leq CF(h)^1 \leq 1, 0 \leq CF(h)^2 \leq 1$ ,

$$\therefore 0 \leq 1 - CF(h)^1 \leq 1, 0 \leq 1 - CF(h)^2 \leq 1,$$

$$\therefore 0 \leq (1 - CF(h)^1)(1 - CF(h)^2) \leq 1,$$

$$\therefore 0 \leq 1 - (1 - CF(h)^1)(1 - CF(h)^2) \leq 1,$$

$$\therefore 0 \leq CF(h)^1 + CF(h)^2 - CF(h)^1CF(h)^2 \leq 1. \quad \square$$

按公式(5)可以计算一个争议的某个前提受到多个争议响应的情况下的可信度合成.

### 2.3.3 可信度传递

可信度计算具有传递性.比如,如果存在对话  $(A^2, A^1)$ , 则争议  $A^2$  会改变争议  $A^1$  的某个前提的可信度,从而改变争议  $A^1$  的结论的可信度.随着辩论的进行,如果又有对话  $(A^3, A^2)$ , 争议  $A^3$  又会改变争议  $A^2$  的某个前提的可信度,进而改变争议  $A^2$  的结论的可信度(即争议  $A^1$  的某个前提的可信度),最后改变争议  $A^1$  的结论的可信度.可见,每增加一个新的争议时,从该争议到根节点的分支上对应的陈述的可信度都要发生改变.可信度传递从新产生的争议开始一直到首争议.可信度传递算法如下:

**算法.** 可信度传递算法.

*Argument\_Create*( $\cdot$ ) //产生争议节点

```

{
1  While (辩论没有结束)
   {
2    产生一个新的争议节点  $P$ ;
3    置  $P$  的前提集  $H^P = \{h_1^P, \dots, h_n^P\}$ ;
4    置  $P$  的各前提的初始可信度:  $CF(h_1^P)^0, \dots, CF(h_n^P)^0$ ;
5    定义数组  $CF(h_i^P) [1 \dots MAXSIZE]$ ;
       //  $CF(h_i^P) [k]$  保存前提  $h_i^P$  的第  $k$  个子树针对  $h_i^P$  的可信度,  $1 \leq i \leq n, 1 \leq k \leq MAXSIZE$ 
6    置  $P$  的各前提的子树标签初始值:  $f_1^P = 0, \dots, f_n^P = 0$ ;
       //  $f_i^P$  的值即为数组  $CF(h_i^P) [-]$  的下标的最大值,  $1 \leq i \leq n$ 
7    置  $CF(h_1^P) = CF(h_1^P)^0, \dots, CF(h_n^P) = CF(h_n^P)^0$ ; //  $CF(h_i^P)$  保存前提  $h_i^P$  的当前可信度
8    设置  $P$  的可信度  $CF(P)$ ;
9    IF ( $P$  是首争议)
10     置  $P$  的结论为  $C^P$ ;
11   ELSE
12     从已有争议中选择一个争议  $Q$  的某个前提  $h_j^Q$  作为  $P$  的结论;
       // 设  $Q = (\{h_1^Q, \dots, h_m^Q\}, C^Q), j=1, \dots, m$ 
13     置  $P$  的标签为 ++  $f_j^Q$ ; //  $f_j^Q$  是数组  $CF(h_i^Q) [-]$  的下标

```



```

14      $CF(h_j^0).MAXSIZE = f_j^0$ ; //  $CF(h_j^0).MAXSIZE$  保存数组  $CF(h_i^0)[\cdot]$  的下标最大值
15   ENDIF
16   Certainty-factor_revision( $P$ ); //调用修改争议节点可信度的函数
17 } ENDWHILE
}
Certainty-factor_revision(节点  $P$ ) //可信度更新
{
1  计算  $P$  的前提可信度:  $CF(H^P) = \min\{CF(h_1^P), \dots, CF(h_n^P)\}$ ;
2  IF ( $P$  是首争议)
3    计算  $P$  的结论  $C^P$  的可信度:  $CF(C^P) = CF(P) \times CF(H^P)$ ;
4  ELSE
5    读取  $P$  的结论  $h_j^Q$  的初始可信度  $CF(h_j^Q)^0$ ; //  $P$  的结论是  $P$  的父节点  $Q$  的某个前提  $h_i^Q$ 
6    读取  $P$  的标签  $f_j^Q$ ;
7    读取  $P$  的可信度  $CF(P)$ ;
8    IF ( $0 < CF(P) \leq 1$ )
9       $CF(h_j^Q)[f_j^Q] = CF(h_i^Q)^0 + (1 - CF(h_i^Q)^0)CF(P)CF(H^P)$ ;
10   ELSE
11      $CF(h_j^Q)[f_j^Q] = CF(h_j^Q)^0(1 + CF(P)CF(H^P))$ ;
12   ENDIF
13   for ( $k=2$ ;  $k \leq CF(h_j^Q).MAXSIZE$ ;  $k++$ )
14      $CF(h_j^Q)[k] = CF(h_j^Q)[1] + CF(h_j^Q)[k] - CF(h_j^Q)[1]CF(h_j^Q)[k]$ ; //多分枝可信度合成
15    $CF(h_j^Q) = CF(h_j^Q)[1]$ ; //重置  $h_j^Q$  的当前可信度
16   Certainty-factor_revision( $Q$ ); //递归调用可信度更新函数
17   ENDIF
}

```

算法由两个函数组成:产生争议节点的函数 *Argument\_Create*( $\cdot$ )和可信度更新函数 *Certainty-factor\_revision*( $\cdot$ ).*Argument\_Create*( $\cdot$ )的主要任务是产生一个新的争议节点,并对该争议的有关参数进行设置,包括争议前提、结论以及各前提的可信度和争议本身的可信度.如果该争议是首争议,则需要给出争议的结论;如果该争议不是首争议,则需要选择一个现有争议的某个前提作为该争议的结论.然后调用函数 *Certainty-factor\_revision*( $\cdot$ )完成从该争议到首争议结论路径上的陈述的可信度的更新.*Certainty-factor\_revision*( $\cdot$ )是一个递归函数,递归出口是首争议.在辩论过程中,每增加一个争议节点都要递归调用这个函数.可见,只要辩论没有结束,对话树中的所有陈述的可信度都有可能发生改变,这体现了辩论推理的可废止性.另外,如果一个陈述受到多个争议的响应,则其可信度值应该为多个可信度更新值的合成.

算法的时间复杂度分析:*Certainty-factor\_revision*( $\cdot$ )是一个递归函数,其递归次数为从首争议节点到该节点的分支数.*Certainty-factor\_revision*( $\cdot$ )每次运行的时间频次取决于响应陈述的争议分支数.因此,整个算法的时间复杂度为  $O(mn)$ ,其中,  $m$  为陈述的分支数,  $n$  为对话树的深度.

### 2.3.4 陈述的可接受性

辩论结束时,辩论系统中的每个陈述都有一个确定的可信度值,其值域为  $[0,1]$ .陈述的可信度越大,其可接受性就越大.

**定义 20(acceptability of statements).** 如果一个陈述的可信度超过了所设定的阈值,则该陈述就是可接受

的.所有可接受陈述组成的集合称为辩论结果(result of argumentation).

2.4 一致性与可行性分析

基于可信度的辩论模型(CFA)将可信度理论引入辩论建模中,能表示不确定信息条件下的辩论推理过程,是对实际辩论过程的合理建模.

CFA 建立在可信度理论基础之上,但将争议前提可信度的域值区间由原来的[-1,1]修改为[0,1],既可运用第 2.1 节公式(1)计算争议结论的可信度,又可保证可信度传递的合理性和一致性.首先,将争议前提中的证据可信度限定为[0,1]是合理的,因为专家一般不会用可信度为假的证据作为自己争议的前提,且[0,1]仍可表示证据的不确定性;其次,如果将争议前提中的证据可信度限定为[-1,1],按公式(1)计算的争议结论的可信度将全部为 0,不能反映争议前提可信度为假时对结论可信度的影响程度.而将争议证据可信度限定为[0,1]能保证可信度合成和可信度传递中可信度计算的一致性,如定理 2 和定理 3 中的证明.

CFA 的争议评价算法建立在可信度合成的基础之上,而不是简单的可信度替换,这体现了意见综合和集成,适合于对协商对话的建模.可信度合成体现在两个方面:一是非首争议结论的可信度的计算,其可信度融合了该陈述作为另一个争议前提时的初始可信度,如公式(4);二是当一个陈述受到多个争议响应时,其最终的可信度是多个争议的结论可信度的合成,如公式(5).

因此,CFA 中的争议评价算法是可行的,且满足可信度计算的一致性.

3 实例分析

下面举例说明该模型的应用.假设群体就某一主题展开辩论,所产生的争议及其可信度设置见表 1.

Table 1 Certainty-Factor of arguments and their premise

表 1 争议及其前提可信度设置

争议	争议前提集	争议结论	争议可信度	争议前提初始可信度
$A^1$	$\{h_1^1, h_2^1\}$	$h$	0.9	$CF(h_1^1)^0 = 0.8, CF(h_2^1)^0 = 0.9$
$A^2$	$\{h_1^2, h_2^2, h_3^2\}$	$h_1^1$	-0.9	$CF(h_1^2)^0 = 0.8, CF(h_2^2)^0 = 0.9, CF(h_3^2)^0 = 0.7$
$A^3$	$\{h_1^3, h_2^3\}$	$h_2^2$	-0.7	$CF(h_1^3)^0 = 0.6, CF(h_2^3)^0 = 0.8$
$A^4$	$\{h_1^4\}$	$h_1^1$	0.8	$CF(h_1^4)^0 = 0.6$
$A^5$	$\{h_1^5, h_2^5\}$	$h_2^1$	-0.8	$CF(h_1^5)^0 = 0.7, CF(h_2^5)^0 = 0.6$

首先(时间节点 1)产生第 1 个争议  $A^1 = (\{h_1^1, h_2^1\}, h)$ , 设置  $CF(h_1^1)^0 = 0.8, CF(h_2^1)^0 = 0.9, CF(A^1) = 0.9$ , 此时,

$$CF(h) = CF(A^1) \times (\min\{CF(h_1^1), CF(h_2^1)\}) = 0.9 \times 0.8 = 0.72.$$

此后(时间流水号 2)产生第 2 个争议  $A^2 = (\{h_1^2, h_2^2, h_3^2\}, h_1^1)$ , 设置

$$CF(h_1^2)^0 = 0.8, CF(h_2^2)^0 = 0.9, CF(h_3^2)^0 = 0.7, CF(A_2) = -0.9.$$

由于  $h_1^1 \in Pre(A^1)$ , 且  $CF(A^2)$  是负值, 因此,  $A^2$  是对  $A^1$  的反对. 调用可信度传递算法:

$$CF(h_1^1)[1] = CF(h_1^1)^0 (1 + (\min\{CF(h_1^2), CF(h_2^2), CF(h_3^2)\}) \times CF(A_2)) = 0.8 \times (1 + 0.7 \times (-0.9)) = 0.296,$$

然后, 将  $CF(h_1^1)[1]$  保存到  $CF(h_1^1)$  中. 由于  $A^1$  的前提的可信度被更新, 因此, 对  $A^1$  的结论的可信度要重新计算, 其值为 0.266 4. 可见,  $A^2$  产生后,  $h$  的可信度得以下降. 但争议  $A^3$  产生后,  $A^3$  的可信度也为负值, 降低了争议  $A^2$  的前提可信度, 进而降低了争议  $A^2$  对  $A^1$  反对强度, 最终使争议  $A^1$  的结论  $h$  的可信度上升为 0.381 7. 争议  $A^4$  产生后,  $A^2, A^4$  两个争议同时对  $h_1^1$  响应, 因此,  $h_1^1$  的当前可信度应为  $CF(h_1^1)[1]$  与  $CF(h_1^1)[2]$  的合成, 其值为 0.890 4, 使  $A^1$  的前提可信度得以提高, 进而又使  $h$  的可信度上升为 0.801 3. 争议  $A^4$  直接作用于  $A^1$  的前提  $h_2^1$ , 使  $h_2^1$  的可信度下降, 进而又使  $h$  的可信度下降为 0.421 2. 各陈述可信度的计算过程见表 2, 不同时间节点各陈述的可信度的值见表 3.

**Table 2** Computing process of Certainty-Factor**表 2** 可信度计算过程

时间节点	调用算法	陈述当前可信度值
1.1	<i>Argument_Create</i> ( $A^1$ )	$CF(h_1^1) = 0.8, CF(h_2^1) = 0.9; Con(A^1) = h$
1.2	<i>Certainty-factor_revision</i> ( $A^1$ )	$CF(H^1) = 0.8; CF(h) = 0.72$ (得到结论 $h$ 的可信度)
2.1	<i>Argument_Create</i> ( $A^2$ )	$CF(h_1^2) = 0.8, CF(h_2^2) = 0.9, CF(h_3^2) = 0.7; Con(A^2) = h_1^1; f_1^1 = 1$
2.2	<i>Certainty-factor_revision</i> ( $A^2$ )	$CF(H^2) = 0.7; CF(h_1^1)[1] = 0.296$ $CF(h_1^1) = 0.296$ ( $h_1^1$ 可信度被修改)
2.3	<i>Certainty-factor_revision</i> ( $A^1$ )	$CF(H^1) = 0.296$ ( $A^1$ 前提可信度被修改) $CF(h) = 0.2664$ ( $A^1$ 结论可信度第 1 次被修改)
3.1	<i>Argument_Create</i> ( $A^3$ )	$CF(h_1^3) = 0.6, CF(h_2^3) = 0.8; Con(A^3) = h_2^2; f_2^2 = 1$
3.2	<i>Certainty-factor_revision</i> ( $A^3$ )	$CF(H^3) = 0.6; CF(h_2^2)[1] = 0.552$ $CF(h_2^2) = 0.552$ ( $h_2^2$ 可信度被修改)
3.3	<i>Certainty-factor_revision</i> ( $A^2$ )	$CF(H^2) = 0.552; CF(h_1^1)[1] = 0.4242$ $CF(h_1^1) = 0.4242$ ( $h_1^1$ 可信度再次被修改)
3.4	<i>Certainty-factor_revision</i> ( $A^1$ )	$CF(H^1) = 0.4242$ ( $A^1$ 前提可信度再次被修改) $CF(h) = 0.3817$ ( $A^1$ 结论可信度第 2 次被修改)
4.1	<i>Argument_Create</i> ( $A^4$ )	$CF(h_1^4) = 0.6; Con(A^4) = h_1^1; f_1^1 = 2$
4.2	<i>Certainty-factor_revision</i> ( $A^4$ )	$CF(H^4) = 0.6; CF(h_1^1)[2] = 0.8096$ $CF(h_1^1) = 0.8096$ ( $CF(h_1^1)[1]$ 与 $CF(h_1^1)[2]$ 合成)
4.3	<i>Certainty-factor_revision</i> ( $A^1$ )	$CF(H^1) = 0.8096$ ( $A^1$ 前提可信度再次被修改) $CF(h) = 0.8013$ ( $A^1$ 结论可信度第 3 次被修改)
5.1	<i>Argument_Create</i> ( $A^5$ )	$CF(h_1^5) = 0.7, CF(h_2^5) = 0.6; Con(A^5) = h_2^3; f_2^3 = 1$
5.2	<i>Certainty-factor_revision</i> ( $A^5$ )	$CF(H^5) = 0.6; CF(h_2^3)[1] = 0.468$ $CF(h_2^3) = 0.468$ ( $h_2^3$ 可信度被修改)
5.3	<i>Certainty-factor_revision</i> ( $A^1$ )	$CF(H^1) = 0.468$ ( $A^1$ 前提可信度被修改) $CF(h) = 0.4212$ ( $A^1$ 结论可信度第 4 次被修改)

**Table 3** Certainty-Factor of statemnets in timeline**表 3** 不同时间节点各陈述的可信度的值

时间节点	$h$	$h_1^1$	$h_2^1$	$h_1^2$	$h_2^2$	$h_3^2$	$h_1^3$	$h_2^3$	$h_1^4$	$h_1^5$	$h_2^5$
1	<b>0.72</b>	<b>0.8</b>	<b>0.9</b>	—	—	—	—	—	—	—	—
2	<b>0.266 4</b>	<b>0.296</b>	0.9	<b>0.8</b>	<b>0.9</b>	<b>0.7</b>	—	—	—	—	—
3	<b>0.381 7</b>	<b>0.424 2</b>	0.9	0.8	<b>0.552</b>	0.7	<b>0.6</b>	<b>0.8</b>	—	—	—
4	<b>0.801 3</b>	<b>0.890 4</b>	0.9	0.8	0.552	0.7	0.6	0.8	<b>0.6</b>	—	—
5	<b>0.421 2</b>	0.890 4	<b>0.468</b>	0.8	0.552	0.7	0.6	0.8	0.6	<b>0.7</b>	<b>0.6</b>

表 3 中,黑体数字是在当前时间节点中被更新的可信度.从表 3 中可以看出:当产生一个新的争议并调用可信度传递算法时,只有从新产生的节点到首争议节点路径上的陈述的可信度才会被更新,而结论  $h$  的可信度则在每个时间节点都会被更新,其变化过程为  $0.72 \rightarrow 0.2664 \rightarrow 0.3817 \rightarrow 0.8013 \rightarrow 0.4212$ ,这与第 2.3.3 节中的分析是一致的.如果设定陈述可信度阈值为 0.6,则在第 5 个时间节点后,即发出第 5 个争议后,可接受陈述集为

$$\{h_1^1, h_1^2, h_2^3, h_1^3, h_2^3, h_1^4, h_1^5, h_2^5\}.$$

在这一时间节点上,结论  $h$  是不可接受的.可见,运用本方法可以正确计算陈述的可信度,并根据陈述可信度阈值确定最后的可接受陈述集.

#### 4 相关工作比较

Dung 的抽象辩论框架假定所有的争议都有相同的强度,争议之间只存在单一的攻击关系,且攻击强度相

同.这与实际辩论不甚相符,因为在实际辩论中,有的争议建立在确定信息基础之上,而有的争议则不然.针对这种情况,不少研究者对 Dung 的抽象辩论框架进行了扩展,本节将本文工作与它们进行比较.

Amgoud 等人<sup>[15]</sup>提出基于优先序的辩论框架(preference-based argumentation framework,简称 PAF),该模型将 Dung 抽象辩论框架<sup>[6]</sup>扩展为  $AF = \langle AR, attacks, Pref \rangle$ ,其中,  $Pref \subseteq AR \times AR$  表示争议之间的优先序关系,即给出争议之间的攻击关系的同时,还要给出它们之间的强度比较.其基本假定是:如果被攻击争议优先于攻击者,则该攻击失败.在此基础上提出争议可接受性算法,但这种模型只是表示了争议之间的强度比较,没有对争议强度进行量化,也没有表示争议之间的攻击强度,其争议评价方法仍然基于 Dung 的扩充语义. Bench-Capon 等人<sup>[16]</sup>提出一种基于价值的辩论框架(value based argumentation frameworks,简称 VAF):  $VAF = \langle AR, attacks, V, val, valprefs \rangle$ ,其中:  $V$  是争议价值集;  $val$  是从  $AR$  到  $V$  的映射函数,即不同的争议有不同的价值取向;  $valprefs \subseteq V \times V$  是价值取向的优先序关系.该方法用价值取向的优先序来修正扩充语义,且不同的观察者有不同的价值取向,从而使争议的可接受性与观察者的价值取向的优先序直接相关.但价值取向不是一个可量化的参数,因此其争议评价方法也不是基于数值计算.本文模型与 PAF 和 VAF 相比增加了对争议之间的攻击强度的度量,这种攻击强度也统一用可信度因子表示.

Dunne 等人<sup>[17]</sup>提出了一种赋权辩论系统(weighted argument system,简称 WAS),其基本思想是给争议之间的攻击关系赋予权重,表示攻击的强度. WAS 定义为一个三元组:  $WAS = \langle X, A, w \rangle$ ,其中,  $\langle X, A \rangle$  对应 Dung 抽象辩论框架  $\langle AR, attacks \rangle$ ;  $w: A \rightarrow \mathbb{R}_+$  是一个函数,它将攻击强度映射为一个大于 0 的实数.这个模型中一个关键概念是不一致预算(inconsistency budget)  $\beta \in \mathbb{R}_+$ ,它表示对不一致的忍受程度,当攻击强度没有达到  $\beta$  时,该攻击可以忽略不计.赋权辩论系统的争议可接受性定义如下:

- 假设  $\alpha_1, \alpha_2$  是  $X$  中的两个争议,  $R$  是  $\alpha_1$  对  $\alpha_2$  的攻击关系集合,  $\alpha_1$  对  $\alpha_2$  的攻击总强度定义为

$$wt(R, w) = \sum_{(\alpha_1, \alpha_2) \in R} w(\langle \alpha_1, \alpha_2 \rangle);$$

- 争议强度小于  $\beta$  的攻击关系子集定义为:  $sub(A, w, \beta) = \{R: R \subseteq A \ \& \ wt(R, w) \leq \beta\}$ ;
- 争议可接受集定义为:  $\varepsilon_\sigma^{WT}((X, A, w), \beta) = \{S \subseteq X: \exists R \in sub(A, w, \beta) \ \& \ S \in \varepsilon_\sigma(X, A \setminus R)\}$ . 其中,  $\varepsilon_\sigma(X, A)$  是 Dung 的抽象辩论框架的扩充语义,  $\sigma$  是可接受性标准.

可见,该方法基本思想是:先对攻击强度没有达到  $\beta$  的攻击关系进行删除,然后再按 Dung 的方法求解扩充语义.这种方法对 Dung 的辩论框架的改进主要体现在增大  $\beta$  时,可能消除 Dung 的辩论框架中的攻击环,从而可以求得非空基础扩充.但对于优先扩充的求解仍属于 NP 完全问题,且该方法没有表示争议本身的强度.本文模型与其不同之处是:将争议前提及争议攻击都用可信度因子表示,其争议评价是基于数值计算,而不是扩充语义.

Tang 等人<sup>[19]</sup>提出了一种基于 Dempster-Shafer 理论的辩论模型(dempster-shafer based argumentation,简称 DSA),该模型将争议定义为一个二元组:  $\langle h, E \rangle$ ,其中,  $h$  是争议的结论,  $E = \{e_1, \dots, e_n\}$  是争议的前提,  $h$  以及  $E$  中的元素  $e_i (1 \leq i \leq n)$  统称为公式,  $L$  为公式集.为了与 Dempster-Shafer 理论相一致,他们定义了一个文字集  $P$ ,  $h$  以及  $e_i (1 \leq i \leq n)$  都可以用解释函数  $I: L \rightarrow 2^P$  映射为  $P$  中的一个子集,  $\Omega = 2^P$  称为识别框架.如果有公式  $\theta \in L$ , 则有  $I(\theta) \subseteq \Omega$ . 然后定义概率分配函数  $m(E): E \rightarrow [0; 1]$ , 并约定

$$\sum_{i=1}^n m(E: e_i) = 1.$$

对于所有不属于  $E$  的证据  $\Phi$ ,  $m(E: \Phi) = 0$ . 在此基础上,定义信任函数  $b(h)$ 、不信任函数  $d(h)$  和不确定函数  $u(h)$  对争议结论  $h$  的信任值进行计算,其中,

$$b(h) = \sum_{I(e_i) \subseteq I(h)} m(E, e_i), d(h) = \sum_{I(e_i) \cap I(h) = \emptyset} m(E, e_i), u(h) = \sum_{I(e_i) \cap I(h) \neq \emptyset} m(E, e_i).$$

在此基础上,提出了证据合成方法.辩论框架定义为  $\langle h, \langle V^F, E^F \rangle \rangle$ ,其中  $h$  是最终结论,  $\langle V^F, E^F \rangle$  是规则网络.规则网络中至少有一个规则的结论是  $h$ ,所有规则的前提要么是知识库中一个公式,要么是另一个规则的结论,因此规则网络可以表示辩论推演过程.这种方法给出了将 Dung 的抽象辩论框架与 Dempster-Shafer 理论相结合的思路,通过数值计算对争议进行评价. Das<sup>[20]</sup>也提出将 Dempster-Shafer 理论应用于辩论模型,但它是基于 Toulmin 模型<sup>[8]</sup>,该方法没有表示辩论推演过程.本文引入可信度方法,直接对文献[13]模型的争议前提和争议之间的攻

击强度进行量化.与 DAS 相比,本文模型对辩论推演过程的表示更为直观。

## 5 结 论

基于不确定性推理的辩论模型应该考虑 3 个方面的因素:

- i) 辩论空间构造,包括争议及争议之间的关系表示;
- ii) 争议前提和争议之间的攻击强度的不确定性表示;
- iii) 争议评价算法.

本文针对这 3 个方面提出了一种基于可信度的辩论模型(CFA)及相应的争议评价算法,并通过设定陈述可信度阈值求解可接受陈述集.在辩论空间构造方面,将争议表示为一种可废止规则,即  $h_1, \dots, h_n \Rightarrow h$ , 其中,  $h_1, \dots, h_n$  等为争议的前提,  $h$  为争议的结论.除首争议外,其他争议的结论均为另一争议的一个前提,这样就构成了一颗对话树,它是对辩论推演过程的刻画.为了反映不确定性辩论推理,用可信度因子表示争议前提的不确定性和争议前提对争议结论的支持强度(即争议本身的可信度).随着辩论的进行,新增争议节点将递归改变其父争议节点某个前提的可信度,最终改变首争议的结论的可信度.争议评价算法包括争议结论可信度计算、可信度合成和可信度传递等 3 个方面.本文提出的算法建立在 Shortliffe 的可信度理论基础之上,能保证可信度计算的可靠性和完备性.

本文提出的模型是对辩论推演过程的比较全面的描述,其语义与建立在“击败”概念上的 Dung 的抽象辩论辩论框架的扩充语义有本质区别:首先,将争议进行了结构化分解,考虑了多个不确定性前提对结论的论证;其次,既能表示争议前提的不确定性,又能表示争议前提对结论的支持强度(可信度为正值)和反对强度(可信度为负值),反映了智能主体的理性思维;第三,争议评价算法建立在可信度合成基础之上,能够融合不同专家的意见,使得出的辩论结果更合理,比较适合于对协商对话的建模.本模型已在群体研讨支持系统中得到成功应用.下一步工作是将本模型应用于多 Agent 系统中的不确定性知识表示和信念修正中.

## References:

- [1] Verheij B. Artificial argument assistants for defeasible argumentation. *Artificial Intelligence*, 2003,150(1-2):291–324. [doi: 10.1016/S0004-3702(03)00107-3]
- [2] Reed C, Walton D, Macagno F. Argument diagramming in logic, law and artificial intelligence. *The Knowledge Engineering Review*, 2007,22(1):87–109. [doi: 10.1017/S0269888907001051]
- [3] Kakas A, Maudet N, Moraitis P. Modular representation of agent interaction rules through argumentation. *Autonomous Agents and Multi-Agent Systems*, 2005,11:189–206. [doi: 10.1007/s10458-005-2176-4]
- [4] Bench-Capon TJM, Dunne PE. Argumentation in artificial intelligence. *Artificial Intelligence*, 2007,171(10–15):619–641. [doi: 10.1016/j.artint.2007.05.001]
- [5] Reed C, Grasso F. Recent advances in computational models of natural argument. *Int'l Journal of Intelligent Systems*, 2007,22(1):1–15. [doi: 10.1002/int.20187]
- [6] Dung PM. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and  $n$ -person games. *Artificial Intelligence*, 1995, 321–357. [doi: 10.1016/0004-3702(94)00041-X]
- [7] Dunne PE, Bench-Capon TJM. Coherence in finite argument systems. *Artificial Intelligence*, 2002,141(1-2):187–203. [doi: 10.1016/S0004-3702(02)00261-8]
- [8] Toulmin SE. *The Uses of Argument*. Cambridge: Cambridge University Press, 1958.
- [9] Janssen T, Sage AP. Group decision support using Toulmin argument structures. *Systems, Man, and Cybernetics*, 1996,4: 2704–2709. [doi: 10.1109/ICSMC.1996.561367]
- [10] Kunz W, Rittel HWJ. *Issues as Elements of Information Systems*. Berkeley: University of California, 1970.
- [11] Caminada M, Amgoud L. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 2007,171:286–310. [doi: 10.1016/j.artint.2007.02.003]

- [12] Cayrol C, Lagasquie-Schieux MC. Gradual valuation for bipolar argumentation frameworks. In: Proc. of the 8th European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty. Springer-Verlag, 2005. 366–377. [doi: 10.1007/11518655\_32]
- [13] Xiong CQ, Li DH. Model of argumentation. Ruan Jian Xue Bao/Journal of Software, 2009,20(8):2181–2190 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3465.htm> [doi: 10.3724/SP.J.1001.2009.03465]
- [14] Chen JL, Wang CC, Chen C. Extended bipolar argumentation model. Ruan Jian Xue Bao/Journal of Software, 2012,23(6): 1444–1457 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4067.htm> [doi: 10.3724/SP.J.1001.2012.04067]
- [15] Amgoud L, Vesic S. A new approach for preference-based argumentation frameworks. Annals of Mathematics and Artificial Intelligence, 2011,63(2):149–183. [doi: 10.1007/s10472-011-9271-9]
- [16] Bench-Capon TJM. Persuasion in practical argument using value based argumentation frameworks. Journal of Logic and Computation, 2003,13(3):429–448. [doi: 10.1093/logcom/13.3.429]
- [17] Dunne PE, Hunter A, McBurney P, Parsons S, Wooldridge M. Weighted argument systems: Basic definitions, algorithms, and complexity results. Artificial Intelligence, 2011,175(2):457–486. [doi: 10.1016/j.artint.2010.09.005]
- [18] Haenni R. Probabilistic argumentation. Journal of Applied Logic, 2009,7(2):155–176. [doi: 10.1007/978-94-017-1737-3\_6]
- [19] Tang Y, Hang CW, Parsons S, Singh MP. Towards argumentation with symbolic Dempster-Shafer evidence. In: Verheij B, Szeider S, Woltran S, eds. Proc. of the Computational Models of Argument (COMMA 2012). Vienna: IOS Press, 2012. 462–469. [doi: 10.3233/978-1-61499-111-3-462]
- [20] Das S. Symbolic argumentation for decision making under uncertainty. In: Proc. of the 7th Int'l Conf. on Information Fusion (FUSION). 2005. 1001–1008. [doi: 10.1109/ICIF.2005.1591967]
- [21] Shortliffe EH, Buchanan BG. A model of inexact reasoning in medicine. Mathematical Biosciences, 1975,23(3-4):351–379. [doi: 10.1016/0025-5564(75)90047-4]

#### 附中文参考文献:

- [13] 熊才权,李德华.一种研讨模型.软件学报,2009,20(8):2181–2190. <http://www.jos.org.cn/1000-9825/3465.htm> [doi: 10.3724/SP.J.1001.2009.03465]
- [14] 陈俊良,王长春,陈超.一种扩展双极辩论模型.软件学报,2012,23(6):1444–1457. <http://www.jos.org.cn/1000-9825/4067.htm> [doi: 10.3724/SP.J.1001.2012.04067]



熊才权(1966—),男,湖北鄂州人,博士,教授,主要研究领域为人工智能,非单调逻辑,辩论模型.

E-mail: x\_cquan@163.com



梅清(1979—),男,讲师,主要研究领域为模式识别,智能系统.

E-mail: mking\_cn@126.com



欧阳勇(1977—),男,副教授,主要研究领域为模式识别,智能系统.

E-mail: oyywuhan@163.com