

## 一种研讨模型<sup>\*</sup>

熊才权<sup>1,2+</sup>, 李德华<sup>2</sup>

<sup>1</sup>(湖北工业大学 计算机学院,湖北 武汉 430068)

<sup>2</sup>(华中科技大学 图像识别与人工智能研究所,湖北 武汉 430074)

### Model of Argumentation

XIONG Cai-Quan<sup>1,2+</sup>, LI De-Hua<sup>2</sup>

<sup>1</sup>(School of Computer, Hubei University of Technology, Wuhan 430068, China)

<sup>2</sup>(Institute for Pattern Recognition and Artificial Intelligence, Huazhong University of Science and Technology, Wuhan 430074, China)

+ Corresponding author: E-mail: x\_cquan@163.com

**Xiong CQ, Li DH. Model of argumentation. Journal of Software, 2009,20(8):2181-2190.** <http://www.jos.org.cn/1000-9825/3465.htm>

**Abstract:** This paper proposes a model of argumentation, in which the internal structures of arguments are expressed with the simplified Toulmin model, and the relations between arguments are defined with the method of Dung's abstract argumentation framework. It also gives the algorithms of the defensibility of an argument and the acceptability of a statement. This model is used to reconstruct the examples in the published literature, and results show that this model can accurately calculate the acceptability of statements and draw a conclusion from argumentation. This research is to construct models for real-world practices of argumentation, but the conclusion will contribute to the modeling of formal nonclassical logics systems.

**Key words:** argumentation; argument structure; argumentation graph; argument evaluation

**摘要:** 提出了一种研讨模型.该模型用简化的 Toulmin 模型表示争议内部结构,用 Dung 的抽象辩论框架的方法定义争议之间的关系,给出了争议可防卫性和陈述可接受性算法.用该模型对已有文献中的实例重新建模,结果表明,该模型能够准确计算陈述可接受性并得出研讨结果.该模型研究出发点是对实际群体研讨建模,但也可以用于非经典逻辑形式系统建模.

**关键词:** 研讨;争议结构;争议图;争议评价

**中图法分类号:** TP18 **文献标识码:** A

研讨是群体激活思维、解决争端、谋求共识的重要手段.研讨可分为协商研讨和劝说研讨两种<sup>[1]</sup>.协商研讨的目标是探寻复杂问题的求解方案,而劝说研讨的目标则是试图说服对方接受自己的主张.劝说研讨存在于协商研讨之中,当一个主张遭到反驳或攻击时,就转入劝说研讨.研讨是通过智能主体之间的对话进行的,一条有明确主张的发言称为争议(argument).由于知识的不完备性和不确定性,争议之间可能存在冲突.研讨的基本过

\* Supported by the National Natural Science Foundation of China under Grant No.69775022 (国家自然科学基金); the Hubei Provincial Natural Science Foundation of China under Grant No.2007ABA025 (湖北省自然科学基金)

Received 2008-05-06; Accepted 2008-08-11

程是,首先智能主体通过对话构造争议集合以及争议之间的支援或攻击关系,然后采用某种准则确定争议的可防卫性,最后根据争议的可防卫性确定所有陈述的可接受性。

研讨模型是计算机群体研讨支持系统的理论基础.它包括争议和争议之间关系描述,以及研讨结果生成算法等内容.研讨模型决定了研讨支持系统的智能性和有效性.研讨包括协商和辩论等多种活动,是非单调逻辑推理的直觉范型<sup>[2,3]</sup>,因此研讨模型在人工智能领域也有广泛应用,如逻辑程序设计<sup>[4]</sup>、多 AGENT 对话建模<sup>[5,6]</sup>等.

目前研讨模型的研究目标大致可分为两类.一类是面向群体研讨支持系统<sup>[7]</sup>,其目标是对争议信息进行组织和管理,如争议结构化分解<sup>[8,9]</sup>、对话过程可视化<sup>[10,11]</sup>等,这类模型的主要问题是争议评价没有提出很好的办法.另一类是面向人工智能领域,其着眼点是用辩论模式对形式系统建模<sup>[12]</sup>,这类模型强调争议之间的攻击关系,并能计算争议的可接受性,但将争议抽象为一个节点,而忽略了争议的内部信息.Gordon 等人<sup>[1]</sup>提出了一种称为 Carneades 的辩论模型,它不仅描述了争议的组成部件,而且还表示了争议之间的关系,并提出了争议评价算法,但它主要是面向劝说研讨.

本文受文献[1]的启发提出了一种新的研讨模型,该模型将协商研讨与劝说研讨统一起来,不仅描述了争议的内部结构,而且还描述了争议之间的相互关系,并通过计算争议可防卫性和陈述可接受性得出最终的研讨结果.本文第 1 节简要介绍相关研究工作.第 2 节详细描述我们提出的研讨模型.第 3 节分析基于该模型的争议评价技术及争议可防卫性和陈述可接受性算法.第 4 节是实例分析.第 5 节给出本文的结论与展望.

## 1 相关研究工作

在争议结构的研究中,影响较大的是 Stephen Toulmin 提出的论证的法律学模型,史称 Toulmin 模型<sup>[13]</sup>.Toulmin 模型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)6 个部件,其中,“支援”是对“根据”的补充性论证,“反驳”是对“模态限定”的削弱,阻止从“论证”得出“结论”.Toulmin 模型对争议进行了结构化处理,有利于计算机对争议信息进行管理和处理.不少研究者对 Toulmin 模型作了简化和改进,并开发了基于该模型的决策支持系统<sup>[9]</sup>.Toulmin 模型存在的问题是没有表示争议之间的关系,即无法表示对话状态的迁移.另一个比较重要的模型是 IBIS 模型<sup>[11]</sup>,其基本思路是,首先引出问题(issue),然后人们针对该问题表达立场(position),如支持(pro)或反对(con)等.IBIS 模型的主要问题是提供对辩论的支持,它主要用于实际决策中的协商研讨.

Dung 将辩论思想用于形式系统建模,提出了抽象辩论框架<sup>[12]</sup>.该框架表示为  $AF = \langle \mathcal{AR}, attacks \rangle$ ,其中  $\mathcal{AR}$  是争议集,attacks 是  $\mathcal{AR}$  上的二元关系,即  $attacks \subseteq \mathcal{AR} \times \mathcal{AR}$ .对于两个争议  $A, B$ ,如果  $attacks(A, B)$  成立,则表示  $A$  攻击  $B$ .Dung 认为,一个理性的 Agent 是否相信某种断言,取决于支持该断言的争议是否能够成功地防卫那些对这个争议进行攻击的争议.基于这一思想,Dung 定义了无冲突争议集、可接受争议、可接纳争议集等几个重要概念,并提出了争议可接受性计算方法.Dung 抽象辩论框架的本质是用辩论模式对形式系统建模,但是 Dung 认为他的理论也可以用于描述实际辩论过程.Nieves 等人<sup>[14]</sup>对 Dung 抽象辩论框架进行了改进,并构建了基于该辩论框架的决策支持系统.Dung 方法的最大特点是把争议看作一个整体而忽视争议的内部结构.但在实际辩论中,对争议的攻击往往是针对争议中的前提或论证,不考虑争议的内部结构是不行的.所以目前基于 Dung 抽象辩论框架的研讨支持系统很少.

Gordon 等人提出了一种称为 Carneades 的争议框架<sup>[1]</sup>.他们把争议表示为一个三元组  $\langle c, d, p \rangle$ ,其中  $c$  是争议的结论, $d$  是方向,表示该争议对  $c$  是支持还是反对, $p$  是前提集,并将结论和前提统一称为陈述(statement).Carneades 用一个争议图表示辩论推理过程,主要符号有矩形框、圆形框和有向边,其中矩形框表示陈述,圆形框表示争议,有向边表示前提与争议和争议与结论之间的关系.在此基础上,Carneades 提出了陈述可接受性、证明标准满足度、争议防卫度和前提支持度的计算方法.Carneades 的争议框架既表示了争议的内部结构,又反映了争议之间的关系,是对实际辩论过程的全面描述.Carneades 的争议可防卫性计算完全不同于 Dung 抽象辩论框架,它是通过计算前提可支持度来确定争议可防卫性的.Carneades 的主要问题是争议抽象为一个节点,而争议的部件(如前提和结论)则处于该节点之外,使争议表示不够直观.另外,Carneades 对前提分类过细,由于争议

可防卫性决定于前提支持度,使争议可防卫性计算变得十分复杂.

本文在以上研究基础上,提出了一种新的研讨模型.该模型以 Toulmin 模型和 Dung 的抽象辩论框架为基础,用简化的 Toulmin 模型表示争议的内部结构,用改进的 Dung 方法表示争议之间的攻击和支援关系,并能通过计算争议可防卫性和陈述可接受性得出最终的研讨结果.

### 2 研讨模型

研讨模型的基本组成单位是争议.我们将争议分解为根据、论证和结论 3 个部分,并将根据、论证和结论统称为陈述.争议之间的关系有支援、攻击、反驳 3 种.支援是对某个争议的根据或论证进行补充性解释,增强该争议所支持的结论的可接受性,当某个争议的根据和论证遭到质问时,就需要支援;攻击是削弱某个争议的根据和论证,降低该争议所支持的结论的可接受性.反驳是对某个争议的结论持反对意见.支援、攻击、反驳也有相应的根据和论证,它们也是争议.

**定义 1(statement).** 陈述是争议的基本组成单位.如果陈述是对事物的肯定性表述则称为原子陈述(atom);如果陈述是形如  $h_1, h_2, \dots, h_n \Rightarrow h_m$  表达式,其中  $h_i (1 \leq i \leq n, \text{或 } i=m)$  也是陈述,  $\Rightarrow$  表示推理出,则称为规则(rule).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  $h_i \equiv \neg h_j$ .所有陈述的集合称为语言(language),用  $\mathcal{L}$  表示.

我们用  $h_1, h_2, \dots, h_n \Rightarrow h_m$  表示规则,是为了与传统的一阶谓词逻辑中的严格公式  $h_1, h_2, \dots, h_n \rightarrow h_m$  相区别,因为在研讨系统中,规则也是可被攻击的.

**定义 2(conflict-free subset of language).** 设有一个陈述子集  $\mathcal{H} \subseteq \mathcal{L}$ , 如果不存在  $h_i, h_j \in \mathcal{H}, h_i \equiv \neg h_j$ , 则称  $\mathcal{H}$  是相容的, 否则称  $\mathcal{H}$  是不相容的.

由于在研讨过程中往往存在意见的不一致性,所以  $\mathcal{L}$  一般是不相容性的.

**定义 3(argument).** 争议是一个二元组  $A=(\mathcal{H}, h)$ , 其中  $h \in \mathcal{L}$  是一个陈述,  $\mathcal{H}=\{h_1, h_2, \dots, h_n\}$  是  $\mathcal{L}$  的一个子集, 且满足:(1)  $\mathcal{H}$  是相容的;(2) 逻辑上  $\mathcal{H}$  推理出  $h$ , 记为  $\mathcal{H} \Rightarrow h$ ;(3)  $\mathcal{H}$  是最小的, 即没有  $\mathcal{H}$  的真子集满足(1)和(2).  $\mathcal{H}$  叫做争议的前提,  $h$  称为争议的结论.如果  $h$  不属于任何其他争议的前提, 则称  $A$  为首争议(head); 如果不存在争议  $(\mathcal{H}', h')$ , 使  $h' \in \mathcal{H}$  或  $\neg h' \in \mathcal{H}$ , 则称  $A$  为尾争议(rail).所有争议的集合记为  $\mathcal{A}$ .

争议中的结论  $h$  对应 Toulmin 模型中的主张,  $\mathcal{H}$  中的原子陈述对应 Toulmin 模型中的根据,  $\mathcal{H}$  中的规则对应 Toulmin 模型中的论证.

争议  $A=(\mathcal{H}, h)$  可以用图 1(a) 来表示. 如果  $A=(\mathcal{H}, \neg h)$ , 则用图 1(b) 来表示. 图 1(c) 是争议的抽象表达形式, 它省略了争议的内部结构, 称为抽象争议节点(以下简称争议节点).

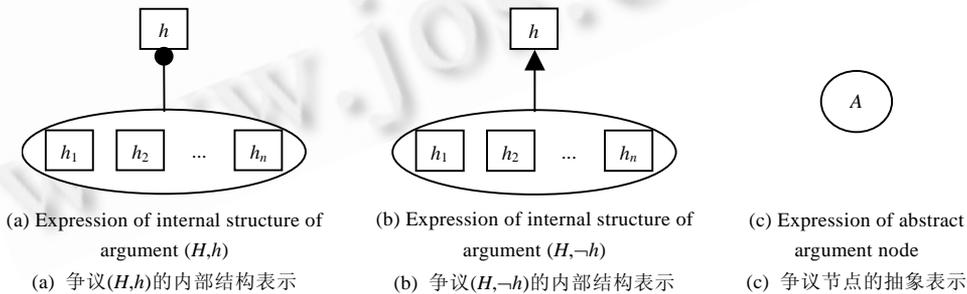


Fig.1 Expression of argument node

图 1 争议节点表示

**定义 4(relation of arguments).** 设有两个争议  $A=(\mathcal{H}_1, h_1) \in \mathcal{A}, B=(\mathcal{H}_2, h_2) \in \mathcal{A}$ ,

(1) 如果  $h_1 \equiv \neg h_2$ , 则  $A$  反驳  $B$  或  $B$  反驳  $A$ , 记为  $A \leftrightarrow B$ ;

(2) 如果  $\exists h \in \mathcal{H}_2, h_1 \equiv \neg h$ , 则  $A$  攻击  $B$ , 记为  $A \triangleright B$ ;

(3) 如果  $\exists h \in \mathcal{H}_2, h_1 \equiv h$ , 则  $A$  支援  $B$ , 记为  $A \triangleleft B$ .

当一个争议的前提的某个陈述受到质问时,就需要支援.争议之间的关系如图 2 所示,每个图右边的小图是该关系的抽象表示.

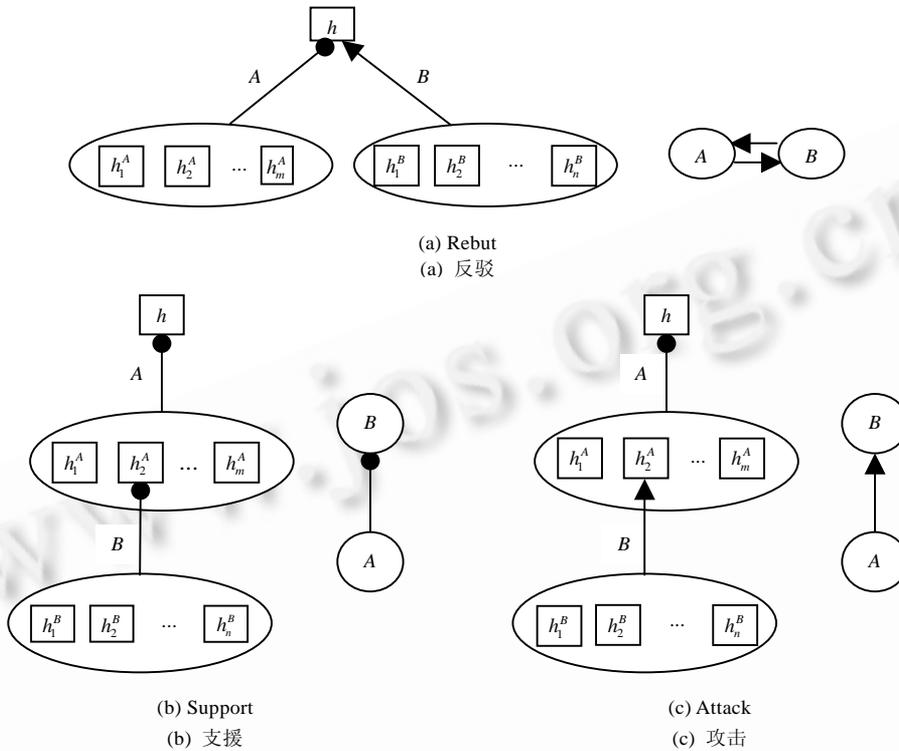


Fig.2 Relation of arguments

图 2 争议之间的关系

例 1: 设有 3 个争议  $A = (\{a, b\}, c), B = (\{d, d \rightarrow \neg c\}, \neg c), C = (\{e\}, a), D = (\{f\}, \neg(d \rightarrow \neg c))$ , 则有  $A \leftrightarrow B, C \triangleleft A, D \triangleright B$ .

**定义 5(argumentation framework).** 一个争议框架(记为  $AF$ )是一个三元组  $AF = (\mathcal{L}, \mathcal{A}, \mathcal{R})$ , 其中,  $\mathcal{L}$  是陈述集;  $\mathcal{A}$  是争议集;  $\mathcal{R}$  是争议之间的二元关系, 它包含反驳、攻击、支援 3 种关系,  $\mathcal{R} = \{\leftrightarrow, \triangleright, \triangleleft\}$ .

**定义 6(abstrate argument graph).** 抽象争议图(简称争议图)定义为  $AG = (\mathcal{A}, \mathcal{R})$ ,  $\mathcal{A}$  是争议节点集,  $\mathcal{R} = \{\leftrightarrow, \rightarrow, \bullet\}$  是有向边集, 表示争议之间的关系, 它包括 3 类有向边,  $\leftrightarrow$  代表反驳,  $\rightarrow$  代表攻击,  $\bullet$  代表支援. 对于争议  $A \in \mathcal{A}$ , 如果  $A$  是首争议, 则在争议图中对应的节点称为根节点(root); 如果  $A$  是尾争议, 则在争议图中对应的节点称为叶子节点(leaf). 对于两个争议  $A, B \in \mathcal{A}$ , 如果  $A \leftrightarrow B$ , 则称  $A, B$  为平行节点; 如果  $A \triangleright B$  或  $A \triangleleft B$ , 则称  $A$  为  $B$  的直接下级节点,  $B$  为  $A$  的直接上级节点.

争议图省略了争议的内部结构, 只反映争议之间的关系.

例 2: 设争议集中有争议  $A = (\{a, b\}, c), B = (\{d, d \rightarrow \neg c\}, \neg c), C = (\{e\}, a), D = (\{f\}, \neg b), E = (\{g, h\}, d), F = (\{a \rightarrow \neg d, i\}, \neg(d \rightarrow \neg c)), G = (\{e, b\}, \neg f), H = (\{b \rightarrow \neg h\}, \neg g)$ , 则  $AG = (\{A, B, C, D, E, F, G, H\}, \{A \leftrightarrow B, C \triangleleft A, D \triangleright A, E \triangleleft B, F \triangleright B, G \triangleright D, H \triangleright E\})$ .

争议图可以用一个有向图表示. 例 2 的争议图如图 3 所示, 其中  $A, B$  为根节点,  $C, G, H, F$  为叶子节点;  $A, B$  互为反驳, 为平行节点;  $C$  和  $D$  是  $A$  的直接下级节点;  $E$  和  $F$  是  $B$  的直接下级节点;  $G$  是  $D$  的直接下级节点;  $H$  是  $E$  的直接下级节点.

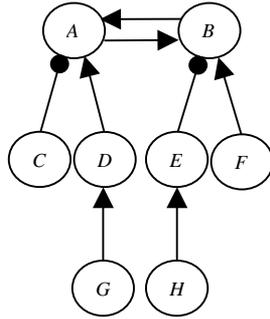


Fig.3 An example of argumentation graph

图3 争议图实例

### 3 争议评价及其计算

**定义 7(defensibility of arguments).** (1) 一个争议节点是可防卫的,如果它没有受到任何节点的攻击或支援,或者攻击它的所有节点都是不可防卫的且支援它的所有节点都是可防卫的。(2) 一个争议节点是不可防卫的,如果它的直接下级节点中存在可防卫的攻击节点或不可防卫的支援节点。

一个争议的可防卫性随着研讨的进行是动态变化的.人类辩论遵循这样一个简单原则:“最后一个说话的得到最后胜利”<sup>[12]</sup>,即如果一个争议没有受到其他争议的攻击,则它赢得胜利.因此,所有争议的可防卫性的确定必须要等到研讨结束.

**引理 1.** 叶子节点都是可防卫的.

**引理 2.** 任意节点可防卫性只与其直接下级节点的可防卫性有关,而与上级节点无关.

在研讨过程中,人们最为关注的是首争议节点的可防卫性及其所支持的结论,因此有必要对争议图中的其他节点进行简化处理.

设  $A, B$  是争议图中的两个争议节点,且  $B \triangleleft A$ ,删除支援节点  $B$  是指从争议图中直接去掉节点  $B$  以及从  $B$  指向  $A$  的有向边,并将  $B$  的所有直接下级节点的有向边改为指向  $A$ .

**定理 1.** 删除争议图的所有支援节点,争议图的其他节点的可防卫性不变.

证明:设  $B$  是一个支援节点,根据引理 2,删除  $B$  节点不会影响  $B$  的直接下级节点的可防卫性.依此类推, $B$  的所有下级节点的可防卫性都不会改变.

现在考察  $B$  的直接上级节点  $A$  的可防卫性取决于  $B$  的可防卫性和  $A$  的其他直接下级节点的可防卫性.如果  $A$  的其他直接下级节点中存在可防卫的攻击节点或不可防卫的支援节点,则  $A$  必是不可防卫的, $A$  的可防卫性与  $B$  无关,删除  $B$  显然不会影响  $A$  的可防卫性.如果  $A$  的其他直接下级节点中的所有攻击节点都是不可防卫的,所有支援节点都是可防卫的,则  $A$  的可防卫性仅取决于  $B$ .分以下两种情况:

(1) 如果  $B$  是可防卫的,则  $A$  应是可防卫的.根据定义 7, $B$  的所有直接下级节点中的攻击节点都是不可防卫的,支援节点都是可防卫的,删除  $B$  节点后,这些节点直接作用于  $A$ ,使  $A$  成为可防卫的.

(2) 如果  $B$  是不可防卫的,则  $A$  应是不可防卫的.根据定义 7, $B$  的直接下级节点中必然存在可防卫的攻击节点或不可防卫的支援节点,删除  $B$  节点后,这些节点直接作用于  $A$ ,将使  $A$  成为不可防卫的.

由于  $A$  的可防卫性不变, $A$  的直接上级节点的可防卫性也不会变,依此类推, $A$  的所有上级节点的可防卫性都不会变. □

**定义 8(dialogue graph).** 删除支援节点的争议图称为对话图,定义为  $DG=(\mathcal{D}, \mathcal{R})$ ,其中  $\mathcal{D} \subseteq \mathcal{A}, \mathcal{R}=\{\leftrightarrow, \rightarrow\}$ .

**定义 9(dialogue tree).** 删除所有反驳关系的对话图称为对话森林,它由对话树组成.对话树定义为  $DT=(\mathcal{T}, \mathcal{R})$ ,其中  $\mathcal{T} \subseteq \mathcal{D}, \mathcal{R}=\{\rightarrow\}$ .

由图 3 的争议图得到的对话图  $DG$  如图 4(a)所示.由于图中包含争议之间的反驳关系,可以对该图进行分解,

得出如图 4(b)所示的两棵对话树  $DT_1, DT_2$ .

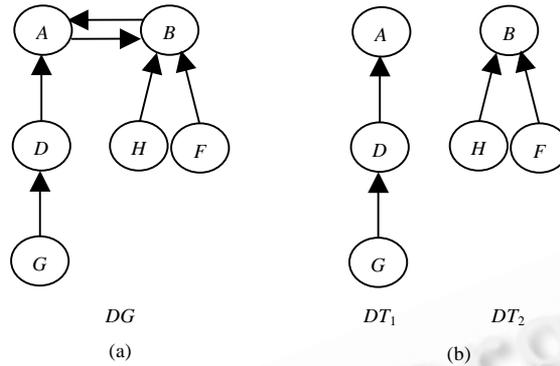


Fig.4 Dialogue graph and dialogue tree

图 4 对话图与对话树

争议图简化为对话树后就可以借用 Dung 的方法计算争议的可防卫性.对话树中的争议评价如下:

**定义 10(conflict-free subset of arguments).** 在一棵对话树  $DT$ 中,设有一个争议集  $S \subseteq \mathcal{T}$ ,如果不存在  $A, B \in S, A \triangleright B$  或  $B \triangleright A$ ,则称  $S$  是相容的.

**定义 11(defensive subset of arguments).** (1) 设有一个争议集  $S \subseteq \mathcal{T}$ ,和一个争议  $A \in \mathcal{T}$ ,如果  $\forall B \in \mathcal{T}, B \triangleright A$ ,则  $\exists C \in S, C \triangleright B$ ,称  $A$  被  $S$  防卫.

(2) 如果  $S \subseteq \mathcal{T}$  是相容的,且  $S$  中所有争议都能被  $S$  防卫,则称  $S$  是可防卫的.

**定义 12(characteristic function).** 设  $S \subseteq \mathcal{T}$ ,一个对话树  $DT=(\mathcal{T}, \mathcal{R})$  的特征函数定义为  $F_T: 2^{\mathcal{T}} \mapsto 2^{\mathcal{T}}$  (映射),且  $F_T(S) = \{A | A \in \mathcal{T}, \text{且 } A \text{ 被 } S \text{ 防卫}\}$ .

**定理 2.** 如果  $S$  是相容的,则  $F_T(S)$  也是相容的.

证明:用反证法.假设  $F_T(S)$  是不相容的,则  $\exists A, B \in F_T(S), A \triangleright B$ .由定义 12,当  $B \in F_T(S)$  时,  $B$  被  $S$  防卫,当  $A \triangleright B$  时,必  $\exists C \in S, C \triangleright A$ ;另一方面,  $A \in F_T(S)$ ,  $A$  被  $S$  防卫,当  $C \triangleright A$  时,必  $\exists D \in S, D \triangleright C$ .这样,  $S$  就不是相容的,与前提相矛盾. □

**定理 3.** 如果  $S$  是可防卫的,则  $S \subseteq F_T(S)$ .

证明:  $\forall A \in S$ , 因为  $S$  是可防卫的,所以  $A$  被  $S$  防卫,根据定义 12,  $A \in F_T(S)$ , 所以  $S \subseteq F_T(S)$ . □

**定义 13.** 对于  $DT=(\mathcal{T}, \mathcal{R})$ , 定义如下的争议子集序列:

$F_T^0 = \{A | A \in \mathcal{T}, \text{且 } A \text{ 是叶子节点}\};$

$F_T^i = \{A | A \in \mathcal{T}, \text{且 } A \text{ 被 } F_T^{i-1} \text{ 防卫}\}.$

对话树的最大可防卫争议集为  $S^T = \bigcup_{i=0}^{\infty} F_T^i$ . 在计算  $S^T$  时, 存在一个不动点  $0 \leq k < \infty$ , 使  $F_T^k = F_T^{k+1} = \dots = F_T^{\infty}$ .

**定理 4.** 如果  $A \in \mathcal{T}$  是可防卫的, 则必有  $A \in S^T$ .

证明: 由于  $A \in \mathcal{T}$  是可防卫的, 则  $A$  要么没有受到任何争议的攻击, 要么攻击它的所有节点都是不可防卫的.

用数学归纳法证明. 如果  $A$  没有受到任何争议的攻击, 则  $A$  必是叶子节点, 即  $A \in F_T^0$ , 显然  $A \in S^T$ . 如果  $A$  不是叶子节点, 则  $\exists B \triangleright A$ , 因为  $A$  是可防卫的, 所以  $B$  是不可防卫的,  $B$  必然受到叶子节点或可防卫的节点  $C$  的攻击. 如果  $C \in F_T^0$ , 则  $A \in F_T^1$ ; 如果  $C \in F_T^{i-1}$ , 则有  $A \in F_T^i$ , 即  $A \in S^T$ . □

**算法 1.** 求对话树的最大可防卫争议集  $S^T$ .

1. INPUT:  $DT=(\mathcal{T}, \mathcal{R})$ .
2. OUTPUT:  $S^T$ .
3. BEGIN
4.  $S^T = \emptyset$ ;

5.  $F_T^0 = \{A | A \in \mathcal{T}, \text{且 } A \text{ 是叶子节点}\};$
6.  $i=0;$
7. do {
8.  $S^i = S^{i-1} \cup F_T^i;$
9.  $i=i+1;$
10. 求  $F_T^i;$  }
11. Until ( $F_T^i = F_T^{i-1}$ )
12. END

算法先依次从对话树的叶子节点开始搜索,初始时  $F_T^0 = \{A \in \mathcal{T} | A \text{ 是叶子节点}\}$ . 算法复杂度为  $O(n)$ ,  $n$  为对话树的最大深度.

**定义 14.** 对话图的最大可防卫争议集为  $S^D = \bigcup_{i=1}^m S^{T_i} - \{U | \exists V \in \bigcup_{i=1}^m S^{T_i}, V \leftrightarrow U\}$ . 即  $S^D$  为各对话树的最大可防卫争议集之和,但不包含存在反驳关系的争议.

**定义 15.** 争议框架的最大可防卫集  $S^A = S^D \cup \{\text{可防卫支援节点集}\}$ .

**算法 2.** 求争议框架的最大可防卫集  $S^A$ .

1. INPUT:  $AF = (\mathcal{L}, \mathcal{A}, \mathcal{R})$ .
2. OUTPUT:  $S^A$ .
3. BEGIN
4. 将争议框架  $AF = (\mathcal{L}, \mathcal{A}, \mathcal{R})$  转化为对话图  $DG = (\mathcal{D}, \mathcal{R})$ ;
5.  $\mathcal{P} = \mathcal{A} - \mathcal{D}$ ; //  $\mathcal{P}$  为支援争议集
6. 求  $DG$  的最大可防卫集  $S^D$ ; // 调用算法 1
7.  $\mathcal{Q} = \emptyset$ ; //  $\mathcal{Q}$  为不可防卫的支援节点集
8. while ( ) { // 依次考察从叶子到根的每条路径上的支援节点  $X$
9. IF ( $\exists X$  的直接下级攻击节点  $\subseteq S^D$  或  $\exists X$  的直接下级支援节点  $\subseteq \mathcal{Q}$ )
10.  $\mathcal{Q} = \mathcal{Q} + X$ ;
11.  $S^A = S^D \cup (\mathcal{P} - \mathcal{Q})$ ;
12. END

算法 2 先求不可防卫的支援节点集  $\mathcal{Q}$ , 再求可防卫的支援节点集  $\mathcal{P} - \mathcal{Q}$ , 最后求  $\mathcal{P} - \mathcal{Q}$  与  $S^D$  的并集. 根据引理 2, 每个支援节点的可防卫性只与其直接下级节点的可防卫性有关, 所以, 算法 2 也是从叶子节点开始沿从叶子到根的路径进行搜索的.

例 3: 求图 3 争议框架中的最大可防卫集.

(1) 将争议图转化为对话图, 再将对话图转化为对话树, 如图 4 所示.  $\mathcal{D} = \{A, B, D, F, G, H\}$ ,  $\mathcal{T}_1 = \{A, D, G\}$ ,  $\mathcal{T}_2 = \{B, F, H\}$ .

(2) 求各对话树的最大可防卫争议集:  $S^{T_1} = \{G, A\}$ ,  $S^{T_2} = \{H, F\}$ .

(3) 求对话图的最大可防卫争议集:  $S^D = \{G, A, H, F\}$ .

(4) 考察图 3 中的支援争议. 对于  $C$ , 由于  $C$  没有受到任何争议的攻击或支援, 所以  $C$  是可防卫的; 对于  $E$ , 由于  $E$  受到  $H$  的攻击, 而  $H \subseteq S^D$ , 所以  $E$  是不可防卫的.

(5) 求争议框架的最大可防卫集:  $S^A = \{G, A, H, F, C\}$ .

从整个研讨来看,  $B$  争议所在的一方失败.

**定义 16 (acceptability of statements).** 一个陈述  $h$  是可接受的, 如果: (1) 不存在争议  $A = (\mathcal{H}_1, h_1)$ ,  $h_1 \equiv h$  或  $h_1 \equiv \neg h$ , 即陈述不受到任何争议的支援或攻击; (2)  $\forall A = (\mathcal{H}_1, h_1)$ ,  $h_1 \equiv \neg h$ ,  $A$  是不可防卫的, 即攻击  $h$  的争议是不可防卫的; 或  $\forall A = (\mathcal{H}_1, h_1)$ ,  $h_1 \equiv h$ ,  $A$  是可防卫的, 即支援  $h$  的争议是可防卫的.

一个争议是可防卫的,意味着可以由它的前提推理出结论,即它的前提中的所有陈述都是可接受的;一个争议是不可防卫的,意味着不能由它的前提推理出结论,它的前提中至少有 1 个陈述是不可接受的。

对于两个相互反驳的争议  $A=(\mathcal{H}_1,h),B=(\mathcal{H}_2,-h)$ ,如果  $A$  是可防卫的,而  $B$  是不可防卫的,则  $h$  是可接受的;如果  $A$  是不可防卫的,而  $B$  是可防卫的,则  $h$  是不可接受的;如果  $A,B$  都是可防卫的,或都是不可防卫的,则  $h$  的可接受性无法确定。

**定义 17(result of discussion).** 对话森林中的所有对话树的可防卫根节点的结论所组成的陈述集称为研讨结果。

我们提出的研讨模型是对研讨过程的全面描述,从单棵对话树来看是劝说研讨,从全局来看则是协商研讨。

#### 4 实例分析

下面用两个例子来说明该模型的应用。

实例 1:Toulmin 逻辑中 Harry 的例子<sup>[12]</sup>.原例用 Toulmin 模型分析如下:

Claim: 亨利是英国人.( $a$ )

Datum: 亨利出生于百幕大.( $b$ )

Warrant: 出生于百幕大的人一般是英国人.( $c$ )

Backing: 英国法律 123 条规定一个出生于英国本土的人可有英国国籍.( $d$ )

Exception: 亨利父母是美国人,他取得美国国籍.( $e$ )

在我们的模型中,Datum,Claim,Warrant,Backing,Exception 都是陈述,分别用  $a,b,c,d,e$  表示.基本争议是  $A=(\mathcal{H}_1,h_1)$ 其中  $\mathcal{H}_1=\{b,c,b\wedge c\Rightarrow a\},h_1=a,\mathcal{H}_1$ 中的  $b$  是大前提, $c$  是小前提,  $b\wedge c\Rightarrow a$  是一个规则.Backing 是对 Warrant 的支援,可以把它当作另一条争议  $B=(\mathcal{H}_2,h_2)$ ,其中  $\mathcal{H}_2=\{d\},h_2=c$ ;Exception(或称 rebuttal)是对争议  $A$  的攻击,可以把它当作又一条争议  $C=(\mathcal{H}_3,h_3)$ ,其中  $\mathcal{H}_3=\{e\},h_3=\neg(b\wedge c\Rightarrow a)$ .其争议图如图 5 所示。

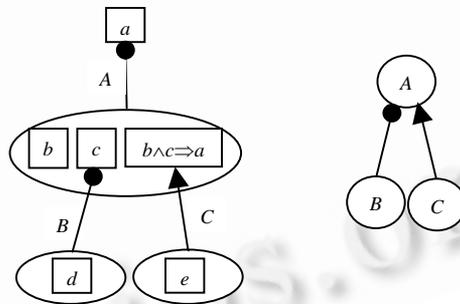


Fig.5 Reconstruct Toulmin's logic structure

图 5 重建 Toulmin 逻辑结构

由上图可以得出可防卫争议集为  $\{B,C\}$ ,可接受陈述集为  $\{d,e,b,c\}$ .可见  $a$  是不可接受的陈述.争议  $C$  是对 Toulmin 模型中的例外的表示,它的作用是削弱由 Datum 和 Warrant 得出 Claim 的力量.显然要使  $a$  为可接受陈述,需要另一个争议否定  $e$ 。

实例 2:本模型也可以用于法律辩论.下面的实例来自文献[15], $p$  为原告, $d$  为被告。

$p$ : (1) 证人约翰和比尔说已经有了要约和承约.( $b$ )

(2) 约翰和比尔是可靠证人.( $c$ )

(3) 如果可靠证人说某事发生了,那么这事就发生了.( $d$ )

因此,(4) 已经有了要约和承约.( $e$ )

(5) 如果已经有了要约和承约,那么合同就生效.( $f$ )

因此,(6) 合同生效.( $a$ )

被告 d 对原告 p 进行攻击:

- d: (7) 约翰和比尔以前说过谎.(g)
- (8) 如果人们以前说过谎,那么他们就是不可靠证人.(h)
- 因此,(9) 约翰和比尔是不可靠证人.(-c)

该辩论过程的争议框架如图 6(a)所示,抽象争议图如图 6(b)所示,对话图如图 6(c)所示,它只有一棵对话树.现根据定义 14 计算对话图的最大可防卫集:在图 6(c)中,C 是可防卫的,因而 B 是不可防卫的,A 也是不可防卫的.C 所支持的  $\neg c$  是可接受的,即 c 是不可接受的,B 所支持的 e 是不可接受的,A 所支持的 a 是不可接受的.这样,可接受陈述集  $S=\{f,b,d,g,h\}$ ,辩论结果集为  $\{-a\}$ ,即合同不能生效.

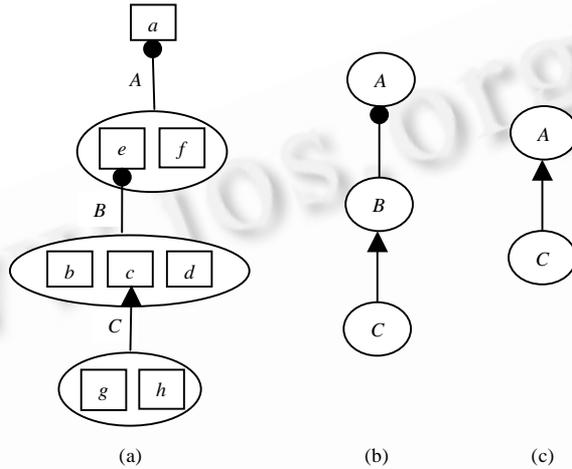


Fig.6 Modeling for legal reasoning  
 图 6 法律辩论建模

### 5 结论与展望

本文将辩论理论用于群体研讨系统建模,提出了一种能将协商研讨与劝说研讨统一起来的研讨模型.该模型用简化的 Toulmin 模型表示争议的内部结构,用 Dung 的抽象辩论框架的方法定义争议之间关系,并给出了争议的可防卫性和陈述的可接受性算法.本文提出的研讨模型有以下特点:(1) 对争议结构的分解仍然遵循 Toulmin 的逻辑表示,并用矩形框表示陈述(即争议的内部部件),用圆圈表示争议,圆圈框内嵌矩形框,使争议的逻辑结构不被破坏.(2) 在争议评价计算时,忽略争议的内部结构,将争议抽象为一个节点,并对支援节点进行约简,使争议图退化为对话图,这种表示与 Dung 抽象辩论框架一致,填补了 Toulmin 模型与 Dung 抽象辩论框架的断层,从而可以借助 Dung 的方法求解争议的可防卫性.(3) 提出了对话森林的概念,将协商研讨与劝说研讨统一起来,从单个对话树来看是劝说研讨,从对话森林来看则是对整个研讨过程的描述.(4) 提出了基于该模型的争议可防卫性和陈述可接受性计算方法,对话森林中所有对话树的可防卫根节点的结论组成的陈述集合即为最终的研讨结果,即问题求解的备选方案.

目前人工智能领域对辩论理论的研究主要沿着两个方向进行,一是用辩论模式描述形式系统,如逻辑程序设计、多 AGENT 对话等,二是构造模型描述现实世界中人与人之间的实际辩论过程.前者称为用辩论建模,后者称为对辩论建模<sup>[3]</sup>.本文研究的出发点是对实际研讨过程建模,但本文提出的方法同样可用于形式系统的建模,这是下一步研究的主要工作.

## References:

- [1] Gordon TF, Prakken H, Walton D. The Carneades model of argument and burden of proof. *Artificial Intelligence*, 2007,171(10-15): 875–896.
- [2] Chesñevar CI, Maguitman AG, Loui RP. Logical models of argument. *ACM Computing Surveys*, 2000,32(4):337–383.
- [3] Reed C, Grasso F. Recent advances in computational models of natural argument. *Int'l Journal of Intelligent Systems*, 2007,22(1): 1–15.
- [4] Prakken H, Sartor G. Argument-Based logic programming with defeasible priorities. *Journal of Applied Non-Classical Logics*, 1997,7:25–75.
- [5] Kakas A, Maudet N, Moraitis P. Modular representation of agent interaction rules through argumentation. *Autonomous Agents and Multi-Agent Systems*, 2005,11(2):189–206.
- [6] Amgoud L, Maudet N, Parsons S. Modelling dialogues using argumentation. In: Durfee E, ed. *Proc. of the 4th Int'l Conf. on MultiAgent Systems (ICMAS 2000)*. Los Alamitos: IEEE Computer Society Press, 2000. 31–38.
- [7] Sillince JAA, Saeedi MH. Computer-Mediated communication: Problems and potentials of argumentation support systems. *Decision Support Systems*, 1999,26(4):287–306.
- [8] Tweed C. Supporting argumentation practices in urban planning and design. *Computers, Environment and Urban Systems*, 1998, 22(4):351–363.
- [9] Janssen T, Sage AP. Group decision support using Toulmin argument structures. *Systems, Man, and Cybernetics*, 1996,4(4): 2704–2709.
- [10] Gelder TV. Argument mapping with Reason!Able. *The American Philosophical Association Newsletter on Philosophy and Computers*, 2002,2:85–90.
- [11] Kunz W, Rittel HWJ. Issues as elements of information systems. Working Paper, WP131, Berkeley: Institute of Urban and Regional Development, University of California, 1970.
- [12] Dung PM. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and  $n$ -person games. *Artificial Intelligence*, 1995,77(2):321–357.
- [13] Toulmin SE. *The Uses of Argument*. Cambridge: Cambridge University Press, 1958.
- [14] Nieves JC, Osorio M, Cortés U, Olmos I, Gonzalez JA. Defining new argumentation-based semantics by minimal models. In: *Proc. of the 7th Mexican Int'l Conf. on Computer Science (ENC 2006)*. Washington: IEEE Computer Science Press, 2006. 210–220.
- [15] Prakken H. On formalising burden of proof in legal argument. In: Herik HJ, *et al.*, eds. *Legal Knowledge Based Systems (JURIX 1999)*, the 12th Conf. Nijmegen: Gerard Noodt Instituut, 1999. 85–97.



熊才权(1966—),男,湖北鄂州人,博士,副教授,主要研究领域为人工智能,模式识别.



李德华(1946—),男,教授,博士生导师,主要研究领域为思维科学,人工智能.