

# 間接的攻撃をもつ議論システム

## Argumentation Systems with Indirect Attacks

高橋 和子<sup>1\*</sup>

Kazuko TAKAHASHI<sup>1</sup>

<sup>1</sup> 関西学院大学理工学部

<sup>1</sup> School of Science and Technology, Kwansei Gakuin University

**Abstract:** We propose two argumentation frameworks: the one with why-questions and the other with support relations. They are the extensions of Dung's argumentation framework so that actual argumentation can be handled. A why-question is regarded as a kind of attack relation, and a support is an answer to an un-presented why-question. We construct these argumentation frameworks by defining arguments and binary relations over the arguments from a pair of knowledge bases. These argumentation frameworks can be converted from each other, and we define the semantics using this property. We also discuss properties of these frameworks.

### 1 はじめに

マルチエージェント間の意思決定において、各エージェントの意見がわかれると、エージェント同士が互いに主張を繰り返しながら時には交渉や説得をして合意に至る。エージェントは相手の意見に納得しなければ反論し、自分の意見に対して十分な理由とともに反論されると相手の意見を受け入れる。このような意思決定の過程における論理構造を明確にするために議論モデルが提案されている。

もともと議論 argumentation は論理に基づくものなので、論理をベースとしたモデルをつくるのは自然である。近年、哲学者 Toulmin は議論の論理構造を示すモデルを提唱したが、これは以後の多くの研究の基礎となった [16]。彼は、論証をデータ Data, 主張 Claim, それらをつなぐ論拠 Warrant, 論拠の裏付け Backing, それに加えて反駁 Rebuttal, 限定詞 Quantifier の 6 要素から成り立つものとし、論証内あるいは論証間でのこれらの要素間の関係をダイアグラムとして示すことで議論の進展する構造を明確化した。

Toulmin のモデルは哲学の分野だけでなく、人工知能や計算機科学の分野にも影響を与え、これらの分野における議論の研究を活性化させた。Dung は論理プログラミングの観点から抽象議論フレームワークを提案し、その意味論を外延という形で定義した [7]。この研究を機に、特に論理プログラミングや非単調推論を中心とする分野で議論の研究がさかんに行われることになった [2, 6, 15]。

Dung の議論フレームワークは論証の集合、その上の 2 項関係である攻撃関係の 2 項組として定義される。攻撃関係は通常相手の論証を否定する反論として具体化される。しかしながら、現実の議論においては、このような直接的な攻撃だけでなく、自分の主張を補強したり相手の主張の理由を尋ねることで間接的な攻撃をするような例も多く見受けられる。このような発言は議論を段階的に深いものにしていくと考えられる。また、Dung の議論フレームワークでは論証間の攻撃は反論のみであることから、具体化する際に Toulmin のモデルの論拠や裏付けを表現するには工夫が必要であり、その意味論はあまり明確ではない。

たとえば、新しく購入するプリンタの機種を決める場合を考える。美しさ、使いよさ、速さ、コスト等様々な点で議論がなされ、ほぼ同じような機種について最後にコスト面だけから判断をすることになったとする。このとき、ある人は「新型の方がコストが安い」、別の人は「新型の方がコストが高い」と主張したとする。同じ「新型だ」という事実から「コストが安い」と「コストが高い」という反対の結論が出されたわけである。この場合、「高い」「安い」の議論ではなく、エージェントはなぜそのような結論が出たのか推論の理由を問はずである。あるいはなぜ自分がそのような結論を出したのか理由を説明する。たとえば、一方が「新型のトナーの代金が高いので利用コストがかかる」、もう一方が「新型だと頻繁にトナー交換になるので利用コストがかかる」という理由をあげ、それについてまた議論が進む。このように理由を尋ねたり主張を補強する発言を取り入れた議論フレームワークの重要性はすでに指摘されており [17, 13]、二極議論フレームワーク

\*連絡先: 関西学院大学理工学部情報科学科  
〒 669-1337 兵庫県三田市学園 2-1  
E-mail: ktaka@kwansei.ac.jp

(BAF) などが抽象的なフレームワークとして提案されている [1]。しかし抽象レベルでは論証の内容が具体化されないので論証間の 2 項関係の意味するところが不明であり、最終的にどのような合意に至るかも明確ではない。

本発表では、理由を尋ねるための質問発言をもつ議論フレームワークと主張を補強する支持発言をもつ議論フレームワークの二つを、異なる知識ベースをもつ 2 体のエージェントの議論フレームワークとして生成する方法を述べ、それらの満たす性質を示す。議論を対話してみると、非論理的な側面も多く含まれるが、ここでは論理的な側面のみに着目し、特に論理的な枠組みの中で現実的な議論をどのように扱うのが妥当かという点に注目する。

抽象議論フレームワークが議論の枠組みをグラフ上の問題のみに帰着しているのに対して、我々は具体的な論理式の集合である知識ベースから論証や攻撃関係を構築することで論証の中身を論理式で具体化したモデルを生成する。このようにして構成したフレームワークは抽象フレームワークの部分集合になっており、外延の無矛盾性などよい性質を満たす。

本発表の構成は以下の通りである。2 節では Dung の抽象議論フレームワークおよび基本概念について述べる。3 節では与えられた知識ベースから質問をもつ抽象議論フレームワークを生成する方法とその性質について述べる。4 節では 3 節で生成した議論フレームワークを支持をもつ議論フレームワークへ変換する方法と得られたフレームワークの性質について述べる。5 節では関連研究との比較を示し、最後に 6 節でまとめと今後の課題を述べる。

## 2 議論フレームワーク

### 2.1 Dung の議論フレームワーク

定義 1 (Dung の議論フレームワーク) 議論フレームワーク *Argumentation Framework (AF)* は 2 項組  $\langle \mathcal{A}, \mathcal{R} \rangle$  として定義される。ここで、 $\mathcal{A}$  は論証の集合、 $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$  は攻撃関係である。

議論フレームワークは論証をノード、攻撃関係をエッジとするグラフで表現できる。

例 1 図 1 に  $AF = \langle \{A, B, C, D\}, \{(A, B), (A, C), (B, D), (C, A), (D, A)\} \rangle$  の例を示す。

外延 *extension* は与えられた議論フレームワークの受理される論証の集合である。Dung をはじめとして複数の研究者がいくつかの種類の外延を定義しているが、ここでは選好外延をとりあげ、特に断らなければ

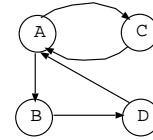


図 1: 議論フレームワークのグラフ表現例

本発表においては外延とは選好外延をさすものとする。他の外延についても同様の議論が可能である。

定義 2 (無衝突, 許容可能, 外延)  $\langle \mathcal{A}, \mathcal{R} \rangle$  を議論フレームワークとし、 $A, B \in \mathcal{A}$  とする。 $\mathcal{E} \subseteq \mathcal{A}$  に対して (1)  $(A, B)$  が攻撃関係であるような  $A, B$  が  $\mathcal{E}$  に含まれないならば、 $\mathcal{E}$  は  $\langle \mathcal{A}, \mathcal{R} \rangle$  において無衝突 *conflict-free* という。

(2)  $\mathcal{E}$  の要素  $A$  に対して、それを攻撃する論証を攻撃する要素が  $\mathcal{E}$  に含まれるならば  $\mathcal{E}$  は  $\langle \mathcal{A}, \mathcal{R} \rangle$  において  $A$  を防御する *defend* という。

(3)  $\mathcal{E}$  が無衝突かつその任意の要素を防御するとき、 $\mathcal{E}$  は  $\langle \mathcal{A}, \mathcal{R} \rangle$  において許容可能 *admissible* という。

(4)  $\mathcal{E}$  は  $\subseteq$  に関して極大許容可能集合のとき、 $\langle \mathcal{A}, \mathcal{R} \rangle$  の選好外延 *preferred extension* という。

例 2 図 1 のような  $AF$  において選好外延は  $\{B, C\}$  である。

### 2.2 基本概念

AF に理論を付与する。

定義 3 (無矛盾)  $\mathbf{L}$  を命題論理式の集合とする。 $\psi \in \mathbf{L}$  と  $\neg\psi \in \mathbf{L}$  を同時に満たす  $\psi$  が存在しなければ、 $\mathbf{L}$  は無矛盾 *consistent* であるという。

定義 4 (c 無矛盾 [12])  $\mathbf{L}$  を命題論理式の集合とする。 $\phi \Rightarrow \psi \in \mathbf{L}$  および  $\phi \Rightarrow \neg\psi \in \mathbf{L}$  を同時に満たす  $\phi$  および  $\psi$  のペアが存在しなければ  $\mathbf{L}$  は c 無矛盾 *c-consistent* という。ここで、 $\Rightarrow$  は *logical entailment* を表す。

$\mathbf{L}$  を命題論理式の集合とする。知識ベース  $\mathbf{K} \subseteq \mathbf{L}$  は無矛盾な命題論理式の有限集合である。各エージェントは独自の知識ベースをもち、そこに格納された知識を使って論証を組み立てる。 $\mathbf{K}$  は演繹的に閉じていなくてもよい。すなわち、 $\phi, \phi \Rightarrow \psi \in \mathbf{K}$  および  $\psi \notin \mathbf{K}$  が同時に成り立っている場合もある。また、 $\neg\neg\psi$  は  $\psi$  と同等と見なす。また、 $\phi \Rightarrow \psi$  が  $\neg(\phi \Rightarrow \psi)$  だけでなく  $\phi \Rightarrow \neg\psi$  も攻撃するものとするために  $\sim$  を導入する。 $\alpha$  を  $\phi \Rightarrow \psi$  とする。ただし  $\phi$  は  $\top$  でもかまわない。このとき、 $\sim\alpha$  は  $\neg(\phi \Rightarrow \psi)$  または  $\phi \Rightarrow \neg\psi$  を表す。

エージェント  $p$  の知識ベースを  $K_p$  と記述する .

定義 5 (共通知識)  $K_a, K_b$  をそれぞれエージェント  $a, b$  の知識ベースとする . 式  $\psi$  が  $K_a$  および  $K_b$  の双方に属するとき ,  $\psi$  を共通知識 common knowledge と呼ぶ .

### 3 質問をもつ議論フレームワーク

与えられた知識ベースのペアから質問をもつ議論フレームワークを生成する .

まず論証と直接的な攻撃関係をもつ議論フレームワークを構成し , それに対して質問による間接的な攻撃を定義する . エージェントは相手の言ったことに納得しない場合 , すなわち相手の論証に含まれる論理式が自分の知識ベースにないが , それに反論できる十分な根拠を持たない場合には質問をするものとする .

#### 3.1 知識ベースに関する議論フレームワーク

定義 6 (論証)  $\phi_1, \dots, \phi_n$  および  $\psi$  を  $K_p$  に属する式とする .  $K_p$  の論証 *argument* は 3 項組  $(Data, Warrant, Claim)$  である . ただし ,  $Data = \phi_1, \dots, \phi_n, Warrant = \phi_1 \wedge \dots \wedge \phi_n \Rightarrow \psi, Claim = \psi$  とする .

$K_p$  の論証  $P = (Data, Warrant, Claim)$  に対して ,  $Dat(P), Wrr(P)$  および  $Clm(P)$  はそれぞれ  $Data, Warrant$  および  $Claim$  を表す .  $Fml(P)$  は式の集合  $\{Dat(P)\} \cup \{Wrr(P)\} \cup \{Clm(P)\}$  を表す . 簡単化のために , 論証は単一データのみから成るものとする . すなわち論証は  $(\phi, \phi \Rightarrow \psi, \psi)$  と表されるものとする . ただし ,  $\phi, \phi \Rightarrow \psi, \psi \in K_p, \phi = \top$  の場合は , 論証は  $(\psi)$  と表す .

定義 7 (直接攻撃)  $A, B$  をそれぞれ  $K_a, K_b$  の論証とする .

1.  $Clm(A) \Leftrightarrow \sim Clm(B)$  かつ  $Dat(A) \neq \top$  ならば ,  $(A, B)$  は  $A$  から  $B$  への *rebut* という .
2.  $Clm(A) \Leftrightarrow \sim Dat(B)$  または  $Clm(A) \Leftrightarrow \sim Wrr(B)$  でありかつ  $Dat(A) \neq \top$  ならば ,  $(A, B)$  は  $A$  から  $B$  への *undercut* という .
3.  $(A, B)$  が  $A$  から  $B$  への *rebut* または *undercut* ならば ,  $(A, B)$  は  $A$  から  $B$  への直接攻撃 *direct attack* という .

定義 8 ( $K_a$  と  $K_b$  に関する議論システム)  $K_a, K_b$  をそれぞれエージェント  $a, b$  の知識ベース ,  $\mathcal{A}_a, \mathcal{A}_b$  をそれぞれ  $K_a, K_b$  の論証の集合とする . また ,  $\mathcal{R}_a$  を  $A$  から  $B$  への直接攻撃の集合 ,  $\mathcal{R}_b$  を  $B$  から  $A$  への直接攻撃の集合とする . このとき ,  $K_a$  と  $K_b$  に関する議論システム *argumentation system* は 4 項組  $(\mathcal{A}_a, \mathcal{A}_b, \mathcal{R}_a, \mathcal{R}_b)$  で表される .

$K_a$  と  $K_b$  に関する議論システムに対する質問および回答を以下のように定義する .

$p$  はエージェント  $a$  または  $b$  を表すものとし ,  $q$  はその相手を表すとする .

(1)  $Q \in \mathcal{A}_q$  に対して ,  $Wrr(Q) \notin K_p$  ならば ,  $p$  の新しい論証  $A_{whyp} = (\neg Wrr(Q))$  および新しい攻撃  $(A_{whyp}, Q)$  を生成する . また , もし  $Clm(Q') \Leftrightarrow \neg Wrr(A_{whyp})$  を満たす論証  $Q' \in \mathcal{A}_q$  が存在すれば ,  $q$  の新しい攻撃  $(Q', A_{whyp})$  を生成する . これは質問に対する回答になる .

(2)  $Q \in \mathcal{A}_q$  に対して ,  $Dat(Q) \notin K_p$  ならば ,  $p$  の新しい論証  $A_{whyp} = (\neg Dat(Q))$  および新しい攻撃  $(A_{whyp}, Q)$  を生成する . また , もし  $Clm(Q') \Leftrightarrow \neg Dat(A_{whyp})$  を満たす論証  $Q' \in \mathcal{A}_q$  が存在すれば ,  $q$  の新しい攻撃  $(Q', A_{whyp})$  を生成する . これは質問に対する回答になる .

$A_{whyp}$  は  $p$  の質問論証 *why-argument* と呼ばれ ,  $(A_{whyp}, Q), (Q', A_{whyp})$  はそれぞれ  $p$  から  $q$  への質問 *why-question* ,  $q$  から  $p$  への回答 *why-answer* と呼ばれ ,  $p$  から  $q$  への質問および  $p$  から  $q$  への回答は  $p$  から  $q$  への質問攻撃 *why-attack* と呼ばれる .

例 3 図 2 に質問論証と質問攻撃の例を示す .  $F, G, H, F \Rightarrow G, H \Rightarrow (F \Rightarrow G) \in K_q$  , および  $F \Rightarrow G \notin K_p$  と仮定する . 図において各枠が 1 つの論証であり , その上段が *Claim* , 下段左が *Data* , 下段右が *Warrant* である . まず , 論証  $Q, Q'$  が知識ベースから作成される (図 2(a)) . 次に , 質問論証  $A_{whyp}$  ,  $p$  から  $q$  への質問  $(A_{whyp}, Q)$  と  $q$  から  $p$  への回答  $(Q', A_{whyp})$  が生成される (図 2(b)) .

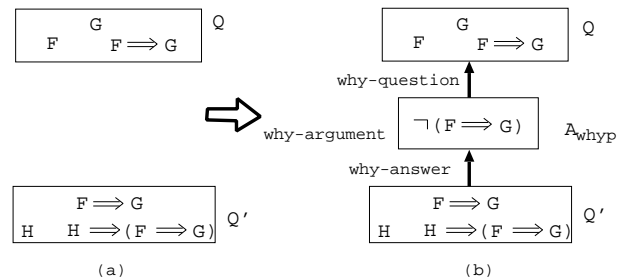


図 2: 質問論証と質問攻撃の例



$\mathcal{A}_{why_a}, \mathcal{R}_{why_a}$  をそれぞれ  $a$  の質問論証の集合および  $a$  から  $b$  への質問攻撃の集合とする。また,  $\mathcal{A}_{why_b}, \mathcal{R}_{why_b}$  をそれぞれ  $b$  の質問論証の集合および  $b$  から  $a$  への質問攻撃の集合とする。このとき, 質問をもつ議論システム  $AS_2$  は  $AS_2 = (\mathcal{A}_a \cup \mathcal{A}_{why_a}, \mathcal{A}_b \cup \mathcal{A}_{why_b}, \mathcal{R}_a \cup \mathcal{R}_{why_a}, \mathcal{R}_b \cup \mathcal{R}_{why_b})$  となる。

定義 9 (質問をもつ議論フレームワーク)  $\mathcal{A}$  を  $\mathcal{A}_a \cup \mathcal{A}_{why_a} \cup \mathcal{A}_b \cup \mathcal{A}_{why_b}$ ,  $\mathcal{R}$  を  $\mathcal{R}_a \cup \mathcal{R}_{why_a} \cup \mathcal{R}_b \cup \mathcal{R}_{why_b}$  とする。このとき,  $\langle \mathcal{A}, \mathcal{R} \rangle$  は  $K_a$  と  $K_b$  に関する質問をもつ議論フレームワークといい,  $AF_{AS}$  と記述する。

$AF_{AS}$  における攻撃は直接攻撃か質問攻撃のいずれかである。 $AF_{AS}$  は Dung の議論フレームワークの具体例となる。

### 3.2 $AF_{AS}$ の性質

命題 1  $AF_{AS}$  はそれを構成するノード数が奇数個のループを持たない。すなわち,  $(A_{i-1}, A_i)$  ( $\forall i. 1 \leq i \leq n; A_n = A_0$ ) を攻撃とすると,  $n$  は偶数である。

証明)  $A_{j-1}, A_j$  がともに同一エージェントの論証であるような  $j$  ( $1 \leq j \leq n$ ) が存在したとする。 $Fml(A_{j-1}) \cup Fml(A_j)$  の任意の要素はそのエージェントの知識ベースに含まれる。しかし,  $A_{j-1}$  が  $A_j$  を攻撃することから,  $\psi, \sim\psi$  がともに  $Fml(A_{j-1}) \cup Fml(A_j)$  に属するような  $\psi$  が存在する。これらは知識ベースの無矛盾性および  $c$  無矛盾性に反する。

Caminada らは議論システムの外延が満たすべき合理的真理を提唱し [4], 外延がこの真理が満たすようにシステムを構築すべきだと主張した。 $AF_{AS}$  はこの中の直接無矛盾性は満たすが, 閉包性と間接無矛盾性は満たさない。それは, 我々が知識ベースの演繹閉包を仮定していないためである。以下で直接無矛盾性の証明を示す。

命題 2  $\mathcal{E}$  を  $AF_{AS}$  の外延とする。 $(i) \cup_{A \in \mathcal{E}} \{CIm(A)\}$  は無矛盾かつ  $c$  無矛盾である。 $(ii) \cup_{A \in \mathcal{E}} Fml(A)$  は無矛盾かつ  $c$  無矛盾である。

証明) (i) 外延の定義から明らか。  
(ii)  $\psi$  と  $\neg\psi$  がともに  $\mathcal{E}$  に含まれるとする。(i) より  $\psi$  と  $\neg\psi$  の少なくとも一方は主張として出現しない。 $\psi$  が  $a$  の論証のデータとして出現したとする。このとき,  $\psi \notin K_b$  が成り立つ。したがって, 主張が  $\neg\psi$  であるような質問論証が存在する。これは矛盾である。 $c$  無矛盾性も同様に証明できる。

## 4 二極議論フレームワークの具体例

### 4.1 $AF_{AS}$ から $BAF_{AS}$ への変換

支持 support は他の論証を補強する論証である。これは, 質問なしで回答が自発的に述べられた質問攻撃だと考えることができる。この考えに基づき, 我々は  $AF_{AS}$  から支持をもつ議論フレームワークである  $BAF_{AS}$  を生成する方法を示す。

$AF_{AS} = \langle \mathcal{A}_a \cup \mathcal{A}_{why_a} \cup \mathcal{A}_b \cup \mathcal{A}_{why_b}, \mathcal{R}_a \cup \mathcal{R}_{why_a} \cup \mathcal{R}_b \cup \mathcal{R}_{why_b} \rangle$  を  $K_a$  と  $K_b$  に関する質問をもつ議論フレームワークとする。

$S_a, S_b$  をそれぞれ エージェント  $a, b$  の支持の集合とする。

[ $AF_{AS}$  から  $BAF_{AS}$  への変換  $T$ ]

$S_a, S_b$  を  $\emptyset$  に初期化する。

$p$  を  $a, b$  いずれかのエージェント,  $q$  をその相手とする。

各質問論証  $Q \in \mathcal{A}_{why_q}$  と  $(P', Q) \in \mathcal{R}_{why_p}, (Q, P) \in \mathcal{R}_{why_q}$  を満たす論証  $P'$  に対して以下を行う。

(i)  $(P', P)$  を  $S_p$  に追加する。(ii)  $(P', Q)$  を  $\mathcal{R}_{why_p}$  から除去する。(iii)  $(Q, P)$  を  $\mathcal{R}_{why_q}$  から除去する。

この結果, 知識ベースに関する支持をもつ議論フレームワークが得られる。

定義 10 (支持をもつ議論フレームワーク) 上記変換  $T$  の結果  $AF_{AS}$  から最終的に得られた  $\mathcal{A}$  を  $\mathcal{A}_a \cup \mathcal{A}_{why_a} \cup \mathcal{A}_b \cup \mathcal{A}_{why_b}$ ,  $\mathcal{R}$  を  $\mathcal{R}_a \cup \mathcal{R}_{why_a} \cup \mathcal{R}_b \cup \mathcal{R}_{why_b}$ ,  $S$  を  $S_a \cup S_b$  とする。このとき,  $\langle \mathcal{A}, \mathcal{R}, S \rangle$  を  $K_a$  と  $K_b$  に関する支持をもつ AF といい,  $BAF_{AS}$  と記述する。

$T$  において, 質問およびその回答のペアは支持の関係に変換されるが, 回答のない質問は質問攻撃としてそのまま残る。 $AF_{AS}$  と  $BAF_{AS}$  は 1 対 1 対応しているので, その逆変換として  $T^{-1}$  が定義できる。

例 4 図 3(a) は  $AF_{AS}$  を示す。この図において,  $X, Y, Z$  を質問論証, これらのノードの入出力エッジを質問攻撃とする。図 3(b) は  $T$  によって (a) から得られた  $BAF_{AS}$  を示す。この図において,  $\rightarrow$  は攻撃関係を,  $\Rightarrow$  は支持関係をそれぞれ示す。

### 4.2 二極議論フレームワーク

二極議論フレームワーク Bipolar Argumentation Framework (BAF) は支持関係をもつ議論フレームワークとして定義されている。ここでは, 二極議論フレームワークを  $T^{-1}$  によって対応する議論フレームワークに変換し, それを利用してもとの二極議論フレームワークの外延を定義する方法を示す。

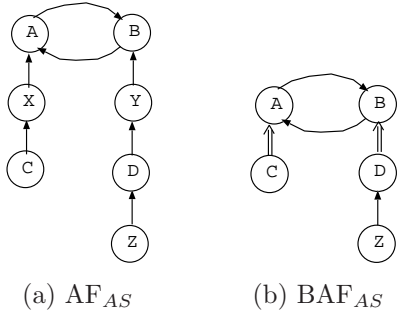


図 3:  $AF_{AS}$  から  $BAF_{AS}$  への変換例

定義 11 (BAF [1]) 二極議論フレームワーク BAF は,  $A$  を論証の集合,  $\mathcal{R} \subseteq A \times A$  を攻撃関係,  $\mathcal{S} \subseteq A \times A$  を支持関係とする 3 項組  $\langle A, \mathcal{R}, \mathcal{S} \rangle$  で表される.

例 5 図 3(b) は  $BAF = \langle \{A, B, C, D, Z\}, \{(A, B), (B, A), (Z, D)\}, \{(C, A), (D, B)\} \rangle$  を示す.

命題 3 二極議論フレームワークは  $T^{-1}$  によって *Dung* の議論フレームワークに変換できる.

[二極議論フレームワークから議論フレームワークへの変換]

$\langle A, \mathcal{R}, \mathcal{S} \rangle$  を二極議論フレームワークとする.  $(A, B) \in \mathcal{S}$  を満たす任意のノードペアに対して, 新しい論証  $C$  および新しい攻撃  $(A, C), (C, B)$  を導入する. 結果として  $A' = A \cup \{C \mid C \text{ は新しい論証} \}$ ,  $\mathcal{R}' = \mathcal{R} \cup \{(A, C), (C, B) \mid C \text{ は新しい論証}, (A, B) \in \mathcal{S}\}$  であるような  $AF = \langle A', \mathcal{R}' \rangle$  を得る.

この  $AF$  の外延  $\mathcal{E}_{AF}$  は定義 2 で定義されるものである. 二極議論フレームワークの外延を  $\mathcal{E}_{AF}$  を用いて以下のように定義する.

定義 12 (二極議論フレームワークの外延)  $\langle A, \mathcal{R}, \mathcal{S} \rangle$  を二極議論フレームワークとする.  $AF$  を対応する *Dung* の議論フレームワークとする.  $\mathcal{E}_{AF}$  を  $AF$  の外延とするとき,  $BAF$  の外延を  $\mathcal{E}_{AF} \cap A$  と定義する.

この定義から, 二極議論フレームワークでは以下の性質が成り立つ.

命題 4  $\mathcal{E}_{AF}$  を  $K_a$  と  $K_b$  に関する質問をもつ議論フレームワーク  $AF_{AS}$  の外延とする. このとき,  $\mathcal{E}_{BAF} \subseteq \mathcal{E}_{AF}$  を満たす  $BAF_{AS}$  の外延  $\mathcal{E}_{BAF}$  は  $T$  によって  $AF_{AS}$  から得ることができる.

例 6 図 3 において, (b) の  $BAF_{AS}$  は  $T^{-1}$  によって (a) の  $AF_{AS}$  に変換される. この例では  $AF_{AS}, BAF_{AS}$  とも単一の外延しか持たない.  $AF_{AS}$  の外延は  $\{A, C, Y, Z\}$ ,  $BAF_{AS}$  の外延は  $\{A, C, Z\}$  である.

### 4.3 $BAF_{AS}$ の性質

二極議論フレームワーク自身は抽象フレームワークとして定義されている. したがって攻撃が循環するような論証列を含むこともあれば一つの論証ペアが攻撃関係かつ支持関係をもつことも許される. さらに, どのエージェントがどの論証を与えたかや, 論証を述べる順序やタイミングについても考慮していない. そのため実際の議論において外延のもつ内部的な意味は不明確である. それに対して,  $AF_{AS}$  から変換  $T$  を施して得られた  $BAF_{AS}$  は以下のようにいくつかのよい性質を満たす二極議論フレームワークの部分集合になっている.

まず,  $BAF_{AS}$  は以下のように  $AF_{AS}$  のもつ性質を保存する.

命題 5  $BAF_{AS}$  は奇数個のノードがなすループを含まない. すなわち,  $(A_{i-1}, A_i) (\forall i. 1 \leq i \leq n; A_n = A_0)$  が攻撃ならば  $n$  は偶数である.

命題 6  $\mathcal{E}$  を  $BAF_{AS}$  の外延とする. (i)  $\{Clm(A) \mid A \in \mathcal{E}\}$  は無矛盾かつ  $c$  無矛盾である. (ii)  $A \in \mathcal{E}$  である  $Fml(A)$  は無矛盾かつ  $c$  無矛盾である.

さらに,  $BAF_{AS}$  に関して以下の命題が成り立つ.

命題 7  $\langle A, \mathcal{R}, \mathcal{S} \rangle$  を  $BAF_{AS}$  とする.  $\mathcal{R} \cap \mathcal{S} = \emptyset$ .

この命題が成り立つことは攻撃関係と支持関係の定義から明らかである.

命題 8  $\langle A, \mathcal{R}, \mathcal{S} \rangle$  を  $BAF_{AS}$  とする.  $(A', A)$  が  $\mathcal{S}$  に属するならば, その時に限り,  $Clm(A') \Leftrightarrow Wrr(A)$  または  $Clm(A') \Leftrightarrow Dat(A)$  が成り立ち, さらに  $Clm(A')$  は共通知識ではない.

この命題が成り立つことは質問論証の定義から明らかである. また, この命題は  $BAF_{AS}$  が  $AF_{AS}$  を介さずに  $K_a$  と  $K_b$  から直接生成できることを示している.

命題 9  $\langle A, \mathcal{R}, \mathcal{S} \rangle$  を  $BAF_{AS}$  とする.  $A, A', B$  を  $A$  の論証とする.  $(A', A)$  が  $\mathcal{S}$  の要素であり, かつ,  $(B, A)$  が  $\mathcal{R}$  の要素であれば, (i)  $(B, A')$  は  $\mathcal{S}$  の要素ではなく, (ii)  $(A', B)$  は  $\mathcal{S}$  の要素ではない.

証明) (i)  $(B, A') \in \mathcal{S}$  と仮定する. すると, 命題 8 より  $Clm(B)$  は  $Dat(A')$  または  $Wrr(A')$  のいずれかと等しい. これは  $Clm(B) \in Fml(A')$  であることを示す. 一方  $(A', A) \in \mathcal{S}$  から  $Clm(A')$  は  $Dat(A)$  または  $Wrr(A)$  のいずれかと等しい. したがって,  $Fml(A') \subseteq K_a$  なので  $Clm(B) \in K_a$ . 一方,  $(B, A) \in \mathcal{R}$  から  $\neg Clm(B) \in Fml(A)$ . したがって,  $\neg Clm(B)$  と

$Clm(B)$  の双方とも  $K_a$  に含まれることになるが、これは  $K_a$  の無矛盾性に反する。

(ii)  $(A', B) \in S$  と仮定する。すると、 $Clm(A')$  は  $Dat(B)$  または  $Wrr(B)$  のいずれかと等しい。これは  $Clm(A') \in Fml(B)$  であることを示す。一方  $(A', A) \in S$  から  $Clm(A') \in Fml(A)$ 。攻撃関係の定義から  $A \in K_a$ ,  $B \in K_b$ 。したがって、 $Clm(A') \in K_a \cap K_b$ 。しかし、命題 8 より  $Clm(A')$  は共通知識ではない。これは矛盾である。

この命題は  $B$  が  $A$  を攻撃し、 $A'$  が  $A$  を支持するときに  $B, A'$  が互いに支持関係にある、という直観に反する場合を考慮しなくてもよいことを示している。

## 5 関連研究

### 5.1 二極議論フレームワーク

Amgoud らは抽象的な二極議論フレームワーク BAF を提案したが、知識ベースから論証の中身を構成する方法については述べていない [1]。そのため、彼らは BAF の支持関係にある論証の中身が互いにどのような関係をもつのかについては議論していない。

Cayrol らは支持関係にある論証をまとめて coalition という概念でとらえ [5]。この概念を使ってメタ的な議論フレームワークをつくることで BAF の意味論を定義した。しかしながら、その定義は複雑で Dung の意味論との整合性がとれていない。

Boella らはこの欠点を補うために支持関係を介して間接的に行われる攻撃の方向を Cayrol らのものとは逆にした調整された攻撃 mediated attack を定義し Dung の意味論がそのまま使えることを示した [3]。彼らは論証をまとめることなく、個々の論証間の関係をいかすことで Dung の意味論との整合性を示している。彼らはまた、攻撃や支持関係が論証間だけでなく攻撃や支持に対しても適用可能なメタ的な議論フレームワークを構築した。彼らのアプローチと我々のものとは論証の粒度が異なる。我々が Data, Claim, Warrant をまとめて一つの論証を構成しているのに対して彼らは Data, Claim それぞれを単独の論証とし、Warrant を支持関係としている。したがって、彼らがメタ攻撃と呼んでいるものは我々のフレームワークでは論証間の攻撃に相当する。

さらに、Boella らのフレームワークも Amgoud, Cayrol らと同様、抽象レベルの話であるため、メタ攻撃の意味や調停された攻撃の方向の意味について曖昧さが残る。たとえば、本論文の図 3(b) の  $BAF_{AS}$  において、彼らの定義を適用すると、 $A$  を含み  $C$  を含まない外延が存在することになり、これは直観に必ずしも一致しない。それに対して、我々は知識ベースから具

体例を構築するアプローチをとっているため生成される二極議論フレームワークは直観に一致するものしかない。

### 5.2 拡張議論フレームワーク

Modgil は Dung's のフレームワークを拡張し、攻撃に対する攻撃であるメタ攻撃を導入した拡張議論フレームワーク Extended Argumentation Framework (EAF) を提案した [11, 10]。多くの議論システムでは、論証同士が互いに攻撃しあっている場合、あらかじめ論証間に与えられた選好関係によっていずれかの攻撃のみが成功するように設定するが、EAF ではメタ攻撃をもつ攻撃は有効でない、という考えのもとに攻撃の強さを決定する。EAF も一種の間接的攻撃と考えられる。

### 5.3 Defeasible Logic Program における議論

Defeasible Logic Program (DeLP) は非単調性を陽に記述できるような論理プログラミング言語である。García らの一連の仕事では、議論フレームワークを使い弁証において提案が正当化される理由を説明する仕組みをつくっている [9, 8]。彼らのシステムでは、議論フレームワークは DeLP で具体化されるため、推論規則は strict rules か defeasible rules のいずれかで記述される。エージェントは共通の知識ベースをもち、攻撃は別の論証の主張を否定するものと見なされる。抽象レベルではなく具体例として論証が定義されているのは我々のシステムと同じである。DeLP における strict rule や公理は我々のフレームワークでは共通知識として扱われる。

### 5.4 構造をもつ議論フレームワーク

構造をもつ議論フレームワーク Structured Argumentation Framework (SAF) は上記のように多数提案されてきたフレームワークを統合し一般化した形で扱ったものである [4, 14, 12]。SAF において知識ベースは DeLP を使って記述され各事実や推論規則は defeasibility に応じて分別される。SAF では各論証が構造をもっていることが特徴である。その構造は、主張をトップレベルとし、主張を直接的/間接的に導出するのに使ったすべての推論規則および事実を含むようになっている。したがって、主張の理由がすべて一つの論証に陽に組みこまれていることになる。また、攻撃は主張に対してだけでなく、それを導出するのに使った推論規則の中間ゴールに対しても可能である。このような攻撃は一種の間接的な攻撃ととらえられるが、論

証間の因果関係は曖昧であり、いつ、なぜ、攻撃が起こったのかが不明確である。たとえば、ある論証が提示されてはじめてその理由に対する攻撃が生起するのが自然だが、SAF ではこのような仕組みまでは表現していない。つまり、SAF は議論の静的な構造や側面をとらえているが動的な側面は表現していない。それに対して、我々の  $AF_{AS}$  は「質問」という形で明確な因果関係を表現した。

SAF と  $AF_{AS}$  には他にも違いがある。 $AF_{AS}$  では各論証は一つの推論規則 (Warrent) しかもたないが SAF では複数の推論規則をもつので、SAF においては  $AF_{AS}$  に比べて同一論証間の攻撃の数が多い。また、SAF ではゴール (Data) に対する攻撃のみが間接攻撃として扱われているが、 $AF_{AS}$  では推論規則そのものへの攻撃も扱っている。

## 6 結論

本発表では、Dung の議論フレームワークの具体例として、与えられた知識ベースから質問をもつ議論フレームワークを生成する方法について述べ、その外延が無矛盾であることを示した。さらに、これを支持をもつ議論フレームワークに変換し、その満たす性質について述べた。

本発表の主な貢献は以下の通りである。

1. 多くの議論システムが共通の知識ベースを使っているのに対して、ここでは各エージェントが異なる知識ベースをもつものとした。
2. 従来具体化されていなかった質問を含む議論フレームワークを知識ベースから生成する方法を示した。
3. 二極議論フレームワークの外延として新しい単純な定義を示すとともに、知識ベースから生成した二極議論フレームワークがよい性質をもつことを証明した。

最初の 2 点は現実の議論を扱う上で有効である。

今後は議論の進行とともに知識ベースが変化していくようなシステムについても考慮する予定である。

## 参考文献

- [1] Amgoud, L., Cayrol, C., and Lagasquie-Shiex, M.-C. (2008). On bipolarity in argumentation frameworks. *International Journal of Intelligent Systems*, 23(10):1062–1093.
- [2] Bench-Capon, T. and Dunne, P. (2007). Argumentation in artificial intelligence. *Artificial Intelligence*, 171:619–641.
- [3] Boella, G., Gabbay, D., Torre, L. d., and Villata, S. (2010). Support in abstract argumentation. In *COMMA2010*, pages 111–122.
- [4] Caminada, M. and Amgoud, L. (2007). On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5-6):286–310.
- [5] Cayrol, C. and Lagasquie-Shiex, M.-C. (2010). Coalitions of arguments: A tool for handling bipolar argumentation frameworks. *International Journal of Intelligent Systems*, 25(1):83–109.
- [6] Chesñevar, C., Maguitman, A., and Loui, R. (2005). Logical models of argument. *ACM Computing Surveys*, 32(4):337–383.
- [7] Dung, P. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77:321–357.
- [8] García, A., Chesñevar, C., Rotstein, N., and Simari, G. (2007). An abstract presentation of dialectical explanations in defeasible argumentation. In *ArgNMR2007*, pages 17–32.
- [9] García, A. and Simari, G. (2004). Defeasible logic programming: an argumentative approach. *Theory and Practice of Logic Programming*, 4(1):95–138.
- [10] Modgil, S. (2009). Reasoning about preferences in argumentation frameworks. *Artificial Intelligence*, 173(9-10):901–1040.
- [11] Modgil, S. and Bench-Capon, T. (2008). Integrating object meta-level value based argumentation. In *Proc. of 2nd Int. Conf. on Computational Models of Argument*, pages 240–251.
- [12] Modgil, S. and Prakken, H. (2011). Revisiting preference and argumentation. In *IJCAI2011*, pages 1021–1026.
- [13] Prakken, H. (2011a). Argumentation without arguments. *Argumentation*, 25(2):171–184.
- [14] Prakken, H. (2011b). An overview of formal models of argumentation and their application in philosophy. *Studies in logic*, 4(1):65–86.
- [15] Rahwan, I. and Simari, G., editors (2009). *Argumentation in artificial intelligence*. Springer.
- [16] , Toulmin, S. (1958). *The Use of Argument*, Cambridge University Press.
- [17] Walton, D. and Krabbe, E. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. SUNY Press.