

# 即興型ディベートにおける大局的な反論構造の可視化に基づく 議論の噛み合い度合いの基礎検討

福井雅弘<sup>1</sup> 中村聡史<sup>1</sup>

**概要:** 即興型ディベートで建設的な議論を行うためには、適切な反論を通じて議論を噛み合わせることが重要である。この競技では複数の論点が同時に提示され、それに対する反論も一括して行われるという特徴がある。そのため、反論の適切さには順序や間隔といった反論間の関係性が大きく関わり、議論が噛み合っているかどうかを判断するためには、反論の集合が持つ大局的な要素を考慮する必要がある。そこで本研究では、即興型ディベートのコーパスを構築し、試合の反論構造を可視化することで、その大局的な特徴に基づいて議論の噛み合いの度合いを評価できるか調査した。調査の結果、4つの大局的特徴量をもとに作成した指標と経験者による評価の相関係数が0.732となり、高い精度で噛み合い度合いを推定できることを明らかにした。

**キーワード:** 即興型ディベート, 反論, 議論構造, 可視化

## 1. はじめに

即興型ディベートとは、動議とよばれる賛否両論あるトピックに対して肯定側と否定側にランダムに分かれ、中立のジャッジを説得する知のスポーツである。この競技では各チームの立場と持ち時間、話す順序が明確に定められる(図1)という特徴がある。限られた時間の中でジャッジを論理と感情の両輪で納得させる必要があるため[1]、複数の論点に関する議論が同時並行で行われる。そのため、ディスカッションなどの対話形式と比べ、即興型ディベートでは複雑に入り組んだ議論がみられる。

この競技の目的は相手を言い負かすことではなく[2]、両者が互いの主張や意図を把握し、建設的な議論を深め合うことにある。そのためには、反論を通じて議論の対象を噛み合わせることが求められる。しかし、ある程度経験を積んだ者同士の試合であっても、説明や反論が不十分なまま終わってしまうことがある[3]。

このような議論構造を分析する手法として、議論マイニングがある。これは議論的な文章を Argumentative Discourse Unit(以下、ADU と略す) と呼ばれる主張や根拠などの議論単位に分割し、その支持や反論の関係を推定するものである。これまでの研究では、ディベートのコーパスから主張の説得力[4][5]や反論の質[6][7]を自動で評価する手法が提

案されている。また、これらの技術を応用し、ユーザが作成した論述文の議論構造を可視化することで、論理構造の把握を支援する研究[8][9]が行われている。

ここで、個々の反論は優れているものの、大局的に見ると反論が不十分で対話が深まらない試合がみられる。例えば、反論の対象が過度に偏って主要な議論を深めることができていない場合や、順序が整理されていない反論により双方が混乱し、議論の流れが断たれてしまう場合である。しかしこれまでの研究では、大局的な観点からみた議論構造の特徴が何を示唆しているのかは十分に明らかにされていない。

そこで本研究では、即興型ディベートの反論構造に着目し、その大局的な特徴をどのように議論の噛み合い度合いの評価に利用できるか調査することを目的とする。そのためにまず、試合記録の文字起こしからコーパスを作成する。次に大局的な反論構造の可視化を行い、仮説をもとに議論の特徴量をモデル化する。そして、ディベート経験者による議論の噛み合い度合いの評価と比較することで、モデルに基づく議論の評価精度を検証する。

## 2. 関連研究

### 2.1 議論マイニングによる議論の質の評価

機械学習などの自然言語処理技術を用いた議論マイニングの研究が行われている。Stab ら[10]は論述文における議論構造を自動推定する手法を提案した。具体的には、議論単位に分割されたコーパスから、単位間の支持や攻撃の関係を推定した。結果として、SVM を用いて 86.3%の精度で推定できることを明らかにした。

同様の手法で、様々な形式の討論における議論の質の分析が行われている。Fa-Hsuan ら[5]は、オンラインの掲示板において、複数ターンに渡るディベートにおける議論の質を評価する機械学習モデルを提案した。これにより、ター

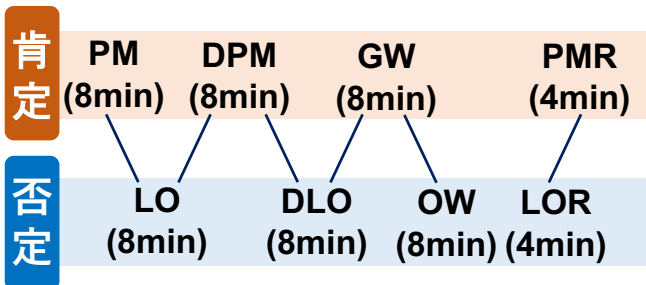


図1 World Schools Style 形式における  
各スピーチの名称、順序及び持ち時間

<sup>1</sup> 明治大学  
Meiji University

ンを跨いだ反論関係を学習することで勝者予測の精度が向上すること、相手の重要な主張への反論には、多くの理由づけが行われている主張を用いた方が効果的である可能性を明らかにした。また Ramon ら[4]は、長期間トピックに関する調査を行う形式のディベート大会の試合を記録し、コーパスを作成した[11]。そして、その議論の質を Argumentation Framework[12]と呼ばれる指標を用いて評価した[4]。結果として、TF-IDF や BERT などの自然言語処理技術のみを用いた手法より、Argumentation Framework と機械学習の両方を組み合わせた手法が高い精度で勝者予測を行える可能性を示した。

しかし、いずれの研究においても議論の質は主張や根拠の説得力の強さとして考えられており、対話として噛み合っているかという観点では検証されていない。また、反論間の関係性には言及されているものの、大局的な性質に関する分析は限定的である。本研究は、試合全体の反論間の関係性に着目することで、より頑健な議論の噛み合い度合いの評価を目指すものである。

## 2.2 可視化による議論構造の理解支援

Wambsganss ら[8]は、議論マイニング技術により推定した議論単位間の支持関係をグラフにより可視化し、学生が論述文の論理構造を把握する支援を行うシステムを提案した。結果として、システムを利用することで学生の文章の説得力が向上する可能性を明らかにした。また Xia ら[9]は、高く評価された投稿の議論構造のグラフや、ロゴス・エトス・パトスなどの論証の種類の割合を提示することで、その論証構造を理解する支援手法を実現した。その結果、システムの利用によりユーザがより説得力のある意見を考える能力が向上する可能性を明らかにした。

いずれの研究も議論構造の可視化にとどまっており、その特徴がどう解釈されるかは十分に分析されていない。本研究では、可視化の特徴に関して仮説を立てたうえで分析することで、結果の解釈可能性を高めるものである。

## 2.3 討論における反論の質の評価

Wachsmuth ら[6]は、オンラインディベート掲示板のコーパスに関して、特定の主張に対する複数の反論候補から、事前にアノテーターの評価により決定した最良の反論を検索する手法を提案した。その結果、平均 7.6 個の候補からトピックに依らず 60%の精度で最良の反論を推定できることを明らかにした。また、Shi ら[7]は、同様のタスクを行うための BERT ベースの Bipolar-encoder モデルを提案した。これにより、相手側の発言から 74.5%の精度で反論関係を推定できることを明らかにした。

いずれの研究においても、反論の質は反論先と反論元の内容に基づいて評価されており、他の反論との関係性及び反論の大局的な性質は考慮されていない。本研究では、試合中のすべての反論を対象に分析を行うことで、より広範囲の文脈を考慮した分析を行うものである。また、本研究で

取り組んでいる大局的な分析および可視化手法は、こうしたこれまでの研究と融合することでより良いものになると考えられる。

## 3. 即興型ディベートのモデル化手法

### 3.1 用語の定義

本研究では、ADU の定義を主張、理由づけ、及び Point of Information(以下、POI と略す)の 3 種類で構成される議論単位であるとする。POI とは、相手側のスピーチを中断して行われる質疑のことである。反論は、即興型ディベートの原則に倣い相手側の特定の発言に対抗した主張を述べたものとする。なお、相手側の発言を想定して予め言及しておくことや、単に相手の立場と対立する論点に関して発言することは反論とはみなさないこととする。また、POI の中でも相手側の特定の主張に対抗することを目的としたものは反論とみなす。

ここで、試合内で時系列順にみて  $i$  番目である ADU を  $ADU_i$  と表現する。ここで、ADU はいずれかのスピーチに分類されるため、 $S(ADU_i)$  を  $ADU_i$  が属するスピーチの番号を表す関数とした。また、 $N$  を試合における総スピーチ数と定義する。

次に、 $ADU_i$  から  $ADU_j$  への反論を  $Att(ADU_i, ADU_j)$  と表し、 $ADU_i$  を反論元、 $ADU_j$  を反論先と呼ぶ。さらに、試合内の ADU の数を  $num_{ADU}$ 、反論の数を  $num_{attack}$  と定義した。また、ある  $Att(ADU_i, ADU_j)$  において、 $ADU_i$  からみて相手側の何人前のスピーチに対する反論かという数値を  $D(ADU_i, ADU_j)$  と定義した。なお、集合  $A$  の要素数を  $n(A)$  と表す。

### 3.2 大局的特徴に基づく仮説

本節では、反論構造の大局的特徴に関し 4 つの仮説を立てた。

まず、反論において、直前の人に対する反論ではなく、相手側の 2 人以上前のスピーチに対する反論が頻繁に発生している場合、相手側の直前の発言に対応できていない可能性が高い。このような遠い反論の割合が大きければ、議論も噛み合っていない可能性が高いと考えられる。そこで、「遠い反論の頻度が高いほど議論の噛み合い度合いが小さい (仮説 1)」という仮説を立てた。

また反論への再反論、再反論への再々反論といったような「反論のラリー」が長く続いている場合、両チームが互いの新たな発言を反映して議論を構築している可能性が高いと考えられる。そこで、「反論のラリーが多く、長く続いているほど議論の噛み合い度合いが大きい (仮説 2)」という仮説を立てた。

次に、相手側の特定の ADU に対して複数の反論を行う際に、なるべく連続している方が望ましいと考えられる。ここで、別の話題に移ったあとに前の話題に戻るといったような一貫性のない反論が行われている場合、何に対して反論しているかが不明確になり、双方が混乱すると考えられる。そこで「特定の ADU に対する反論群の総間隔が大きいほど議論の噛み合い度合いが小さい（仮説 3）」という仮説を立てた。

最後に、反論が理路整然としていれば、相手側の発言の順序に沿ったものになる可能性が高いと考えられる。例えば、相手側が話題  $A \rightarrow B \rightarrow C$  の順で主張した場合には、同様に話題  $A \rightarrow B \rightarrow C$  の順で反論を行うと整理され、建設的な議論が展開しやすいと考えられる。逆に相手の発言の順序を無視した場合、議論が噛み合わなくなると考えられる。以上をもとに、「反論の順序が強く対応しているほど議論の噛み合い度合いが大きい（仮説 4）」という仮説を立てた。

### 3.3 大局的特徴量のモデル

3.2 節の 4 つの仮説に基づき、議論の大局的特徴量を試合ごとに定義する。

まず、仮説 1 の遠い反論の頻度に関するモデル化を行う。「遠い反論」は 3 番目以前のスピーチからは行えないため、頻度を求める際に 4 番目以降のスピーチからの反論の数で正規化が必要がある。以上から、仮説 1 に関する大局的特徴量を次の式で定義した。

$$Distance = \frac{n(\{Att(ADU_i, ADU_j) | D(ADU_i, ADU_j) \geq 2\})}{n(\{Att(ADU_i, ADU_j) | S(ADU_i) \geq 4\})}$$

次に、仮説 2 の反論のラリーについてのモデル化を行う。 $ADU_i$  を  $i = 1$  から順に探索し、反論しておらず、かつ反論されているものを root ノードとして順に反論元を子として追加していくことによりツリーを構築する。 $ADU_i$  を root ノードとするツリーを  $R_i$  とし、 $R_i$  の各 leaf ノードまでの深さの総和から全ての leaf ノードの個数を減算したものを  $R_i$  の「ラリー度合い」とする。これを  $depth_{total}(R_i)$  と表現する。例えば、図 2 のような構造の反論集合における「ラリー度合い」は 2 である。なお、試合全体の「ラリー度合い」を総合的に分析するにあたり、スピーチ数で正規化を行う。これは、スピーチ数が多いほどラリーが続きやすくなると考えられるためである。

$$Rally = \frac{\sum_i depth_{total}(R_i)}{N \cdot num_{attack}}$$

次に、「反論の間隔」を、同じスピーチ内における、同じ ADU に反論している反論集合の最初の反論元と最後の反論元の間にある ADU の個数とした。例えば、図 3 のような構造の反論集合において、肯定側 2 人目における「反論の間隔」は 2 である。またこの値を、該当のスピーチ内の総 ADU から 2 減算した値で正規化した。また正規化された値の和を求めることで、 $Interval$  を求めた。

最後に、反論の順序の対応に関するモデル化を行う。こ

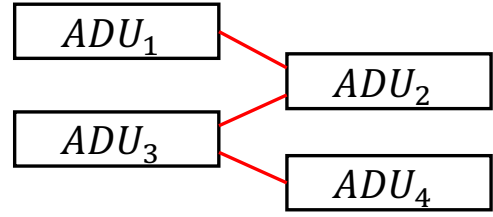


図 2 反論先と反論元が一致する反論集合の例

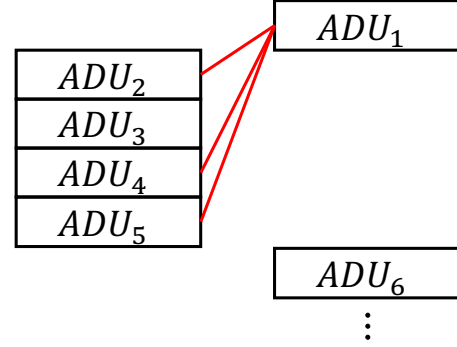


図 3 同じスピーチ内で反論先を共有する反論群の例

れには、反論構造のグラフにおけるエッジの交差数の総和  $num_{cross}$  を求め、仮説 4 に関する大局的特徴量を、次の式で定義した。

$$Order = \frac{num_{attack}}{num_{cross}}$$

## 4. 即興型ディベートの試合のコーパス構築

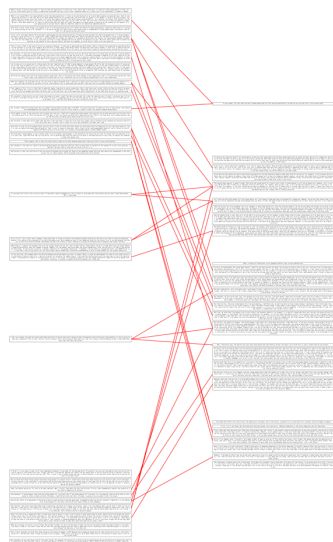
### 4.1 データ構造

本研究では、YouTube 上で公開されている 2 チーム生の即興型ディベートの試合動画 20 試合を対象とした。ここでは、高校生の練習試合や大学生の大会の予選など、十分な経験をもつ者が参加していると思われるものを中心に選定した。

次に、動画の音声から文字起こしを行い、ADU への分割及び反論関係の推定を行った。2.1 節で述べたように、ADU の分割及び反論関係を推定する手法は複数あるが、即興型ディベートの独特な言い回しなどの特徴はあまり考慮されていない。そのため、議論構造の推定は即興型ディベートの経験がある著者が人手により行った。ここで、各 ADU には試合ごとに時系列順に id を付与し、この id を用いて反論元 id と反論先 id のタプルとして反論関係を表現した。また、POI が行われている ADU の id のリストを作成した。

### 4.2 議論構造の可視化とその事例

構築したデータをもとに、各 ADU をノード、各反論をエッジとするグラフとして議論構造を可視化した。エッジ



禁欲的生活

図4 遠い反論が多くみられる事例



性産業国有化

厳格な子育ての禁止

図5 反論がまとまっている事例(左)と  
複雑に入り組んでいる事例(右)

の始点及び終点は、それぞれ反論元、反論先に対応する。また、ADUは上から時系列順に配置しており、肯定側の発言は左側、否定側の発言は右側に配置している。

図4はあるディベートの議論構造の可視化事例である。この図から、特に否定側の2人目と3人目が肯定側2人目に一切反論しておらず、肯定側の3人目も否定側3人目のスピーチに1度しか反論していないことが読み取れる。こうした試合では、直前の主張や理由づけを反映していない

議論が展開されていると考えられ、仮説1にあるように遠い反論が多く、議論が噛み合っていないと推察される。

図5は反論群がある程度まとまっている試合と、複雑に入り組んでいる試合における議論構造の可視化事例である。図5左の試合では、反論元と反論先が一致する反論が複数見られるため、反論のラリーが長続きした建設的な議論になっていたと考えられる。また、反論の分布もある程度集中しており、話題ごとに連続した組織的な反論が提示されていたと思われる。一方、図5右の試合では特に否定側3人目と4人目の反論が複雑に絡み合っており、どこでどの主張に対し反論していたかを把握することが困難であったと考えられる。このような試合では、対話を深めることが難しかったと推測される。

### 4.3 コーパスの概要

作成したコーパスの特徴を、Ramonら[10]が作成したコーパス「Vives Debate」と比較して述べる。まず、総単語数に関しては、我々のコーパスでは125,644単語であるのに対し、「Vives Debate」では139,756単語であり、全体の規模は同等であると言える。一方でADU数に関しては、我々のコーパスでは1,573個であるのに対し、「Vives Debate」では7,810個であった。このことから、我々のコーパスでは比較的ADUが粗いと言える。

## 5. 分析と考察

### 5.1 即興型ディベート経験者による評価

大局的特徴量のモデルを評価するにあたり、即興型ディベート経験者5名(全員が選手として2年以上の経験あり)が各試合の議論全体の噛み合い度合いや各チームの議論の質を評価した。評価の際には、著者が作成した各試合の文字起こしを読み、動画を視聴したうえでアンケートに回答するという方式をとった。最終的に、3試合について3名、4試合について2名、残りの13試合について1名からの評価があった。

即興型ディベート経験者による評価を分析するため、各試合で行われたディベートに対し、議論が噛み合っていたかなどについて評価値をつけてもらった。ここでは、「そう思う」を4、「どちらかというと思う」を3、「どちらかというと思うわない」を2、「そう思わない」を1とした。なお、複数の回答が得られた質問に関してはそれらの平均を評価値とした。

経験者によるアンケートでの「議論は噛み合っていましたか?」という質問の評価値(以下、議論の噛み合い度合いの評価値と呼ぶ)が、2.5以下の試合を噛み合っていない群、2.5より大きい試合を噛み合っている群と分類した。結果として、噛み合っていない群は12試合、噛み合っている群は8試合となった。



## 5.2 仮説に基づく議論の噛み合い度合いの評価

大局的特徴量と経験者による評価を比較し、3章で立てた仮説を検証する。まず、3章で4つの仮説に基づき定義した大局的特徴量を統合するため、正規化を行った。具体的には、次式のようにして正規化された指標を得た。なお、特徴量  $F$  について、 $F_{all}$  は全試合の結果の集合におけるその特徴量の値を指す。また  $\max(F_{all})$  は、特徴量  $F$  に関する全試合中の最大値を意味する。

$$\begin{aligned} score_{Rally} &= \frac{Rally}{\max(Rally_{all})} \\ score_{Order} &= \frac{Order}{\max(Order_{all})} \\ score_{Distance} &= \frac{Distance}{\max(Distance_{all})} \\ score_{Interval} &= \frac{Interval}{\max(Interval_{all})} \end{aligned}$$

次に、大局的特徴による議論の噛み合い度合いの評価指標  $score$  を試合ごとに次式で定義する。

$$score = score_{Rally} + score_{Order} + (1 - score_{Distance}) + (1 - score_{Interval})$$

図6は、各試合について、横軸に  $score$ 、縦軸に噛み合い度合いの評価値をとった散布図である。このグラフから、 $score$  が2.0未満の試合では噛み合っていない群が多く見られること、2.0以上2.5未満の試合では両群が混在すること、2.5以上の試合では全てが噛み合っている群に属することがわかる。 $score$  と議論の噛み合い度合いの評価値の相関係数を計算したところ0.732であり、正の相関が確認された。

また、どの特徴量が効いているかを確認するため、4つの特徴量から3つの特徴量を取り出して統合した指標に関して相関を調査した。その結果、 $score_{Distance}$  を除いたものが0.606、 $score_{Rally}$  を除いたものが0.619、 $score_{Interval}$  を除いたものが0.612、 $score_{Order}$  を除いたものが0.716であった。この結果より、今回採用した全ての大局的特徴量が精度向上に貢献しているが、 $score_{Order}$  を除いても十分な精度が得られることがわかった。以上のことより、反論の順序の対応度については、議論の噛み合いにおいて必ずしも必要ではないか、現時点のモデルに課題があることが明らかになった。

## 5.3 可視化事例に基づく仮説の検証

大局的なモデルがある程度適用可能であることが明らかになったが、本節では仮説とそのグラフに着目しつつ、その特性について深く分析を行う。

まず、仮説1について検証する。図7は、「反論の遠さ」に関して特徴的だった2つの試合の議論構造である。図7左の試合は遠くのADUに対して反論している傾向が強いのにに対し、図7右の試合は近くのADUに対して反論する傾向が強いものである。ここで特に図7左の試合では、肯定側の2人目の発言量が極端に少なく、結果として否定側

が肯定側の1人目に反論を集中してしまったと考えられる。また、肯定側の3人目も1人目に反論しているものが多かった。これは肯定側の2人目が対応できなかった内容をカバーするためであったと考えられる。このように、「反論の遠さ」が大きい場合には、全体として対話が崩れている可能性が考えられる。一方、図7右の試合では、否定側1人目は一切反論を行っていない。ここで経験者による評価では、この試合は「噛み合っている」と評価されており、相手の直前の発言への反論が続いていれば十分に対話が深められる可能性が示唆された。

次に、仮説2を検証する。図8は  $score_{Rally}$  について特徴的だった2つの試合の議論構造である。なお、ラリーが1回以上続いている反論のエッジは青、それ以外のエッジは赤で表されている。図8左の試合では反論のラリーが続いておらず、双方が相手の反論に対応できていないと考えられる。この試合では特定のポイントに極端に集中した反論が多く見られた。このように、ラリーが続かない試合では、

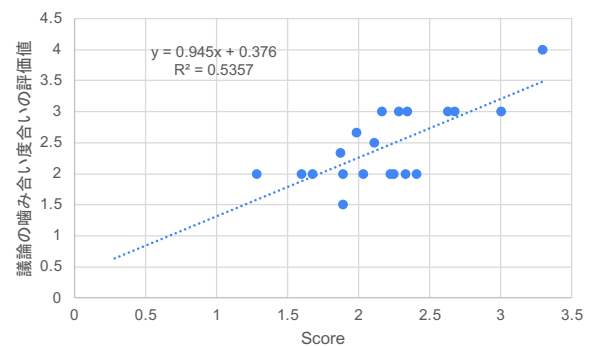
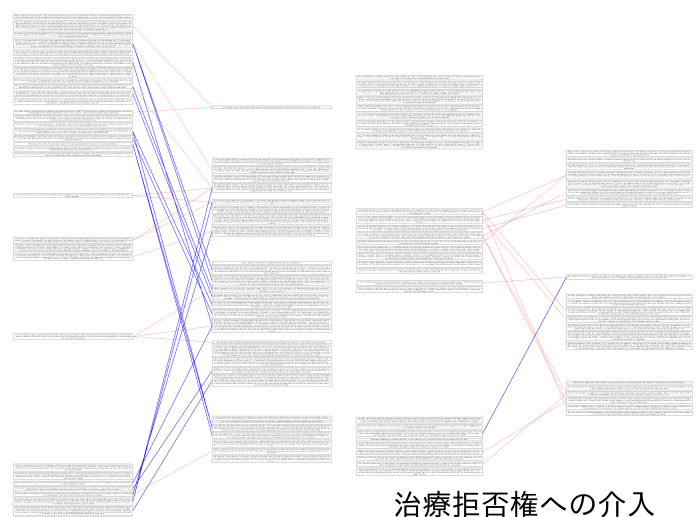


図6 統合指標のグラフ



禁欲的生活

治療拒否権への介入

図7 Distanceが最も大きい事例(左)と最も小さい事例(右)

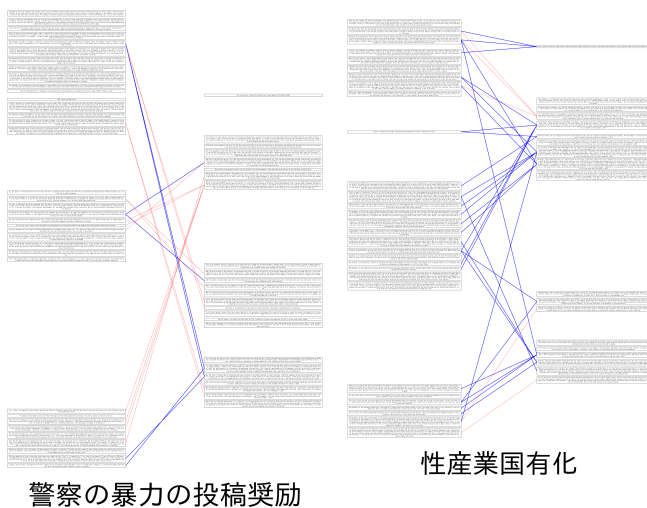


図8  $Score_{Rally}$ が最も小さい事例(左)と  
最も大きい事例(右)

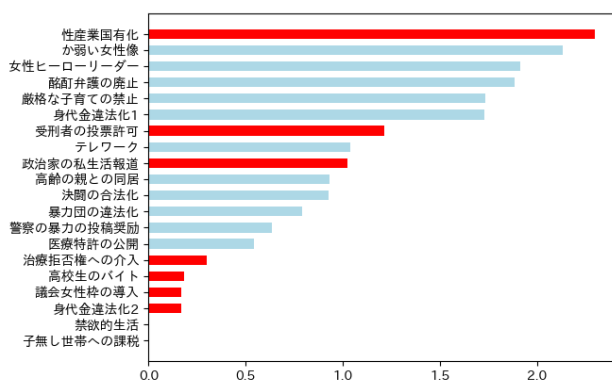


図9  $Interval$ が最も大きい事例(左)

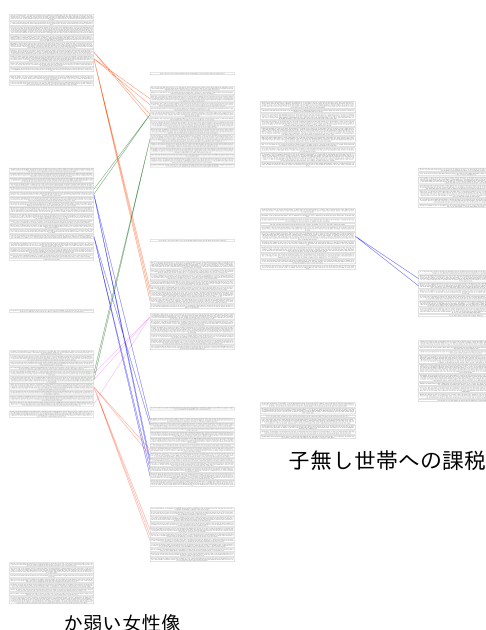


図10  $Interval$ が2番目に大きい事例(左)と  
最も小さい事例(右)

反論への対応や順序の遵守が軽視されている可能性が示唆された。一方、図8右の試合は $score_{Rally}$ が最も大きいものであった。この試合は、 $score_{Interval}$ が最大であり、 $score_{Order}$ も小さかったにも関わらず、噛み合っている群に属している。つまり、ラリーが長続きしていれば反論の間隔や反論の順序が乱れていても対話が深められる可能性が示唆された。

仮説3に関して、図9は試合ごとの $Interval$ のグラフである。このグラフから、反論の総間隔の下位6つの試合については噛み合っている群に属しており、上位6つの試合については1つを除き噛み合っていない群に属していることがわかる。この結果より、 $Interval$ は特徴量として有用であると考えられる。

一方、同じ反論先に複数の反論を行うこと自体が問題である可能性も示唆された。図10は $Interval$ が2番目に大きい試合と最も小さい試合の事例である。図10左は噛み合っていない群、図10右は噛み合っている群に属している。なお、エッジはひとつのスピーチ内で反論先を共有する反論のみを表示しており、反論先ごとに色分けされている。図10左では、特定のADUに対して反論が集中している箇所が多数見受けられる。一方、図10右では、この試合では一つのスピーチから同じADUに対する反論がほとんど行われていないことがわかる。これらのことから、 $Interval$ が小さく議論が噛み合っている試合では、連続した反論が行われているというより、特定のADUに複数の反論を行うこと自体が少ないと考えられる。なお、この傾向は図10に示された2試合に限られた話ではなく、同じスピーチ内で反論先を共有している反論は、全体では平均14.05個ある一方で、 $Interval$ が0.5未満であった6試合では平均4.16個、 $Interval$ が0.6以上であった6試合では17.83個と差が顕著であった。つまり、「反論先を同じスピーチ内で共有する反論」の個数が議論の噛み合い度合いに影響する可能性を示唆していると考えられる。そのため、今後はこうした特徴量についても検討を行っていく予定である。

次に、仮説4について、5.2節の分析から、 $Order$ による噛み合い度合いの評価精度は最も弱かったと考えられる。その主な理由は、整理された組織的な反論を適切に評価できなかったことである。図11は $Order$ が3番目に小さかった試合における可視化事例である。肯定側2人目や否定側の1人目と4人目に着目すると、エッジが交わることなく並んだ反論群が多く見られたことから、議論の全体像の把握が容易であったと考えられる。また、この試合は噛み合っている群に属しており、ディベート経験者から見ても対話が深まっていたと考えられる。しかし、 $Order$ の値は3番目に小さく、この特徴量では妥当な評価ができていなかった。これは、肯定側3人目が複雑に交差した反論を多数行ったことに影響されたためだと考えられる。そのため、今後は理路整然とした反論群の特徴が顕著に表れている構造



政治家の私生活報道

図 11 Order が 3 番目に小さい事例

を重点的に評価することが必要であると考えられる。

#### 5.4 今後の改善

今後の大局的特徴量の分析手法の改善案は主に 2 つある。1 つ目は、反論の順序の対応度の分析に際し、反論が交差する要因の違いを反映することである。今回は論点が提示された順序や、同じ論点の中の主張が提示された順序などを区別せずに扱っていたが、内容によっては必ずしも対応する必要がない可能性が考えられる。

2 つ目は、反論のラリーの分析に際し、ADU 間の支持関係を考慮することである。本研究では反論に直接反論していない場合ラリーが続いていないと判断したが、関係する ADU に反論することで間接的に対応することでラリーが続いている可能性が考えられる。このような改善により、今後更なる評価の分析を行う。

## 6. おわりに

本研究では、即興型ディベートにおける噛み合った議論の特性の基礎的な調査を目的に、反論構造の可視化により示唆される対話の大局的な特徴を分析した。具体的には、「相手側の 2 つ以上前のスピーチへの反論の割合は小さい方がよい」、「反論のラリーは、長く続くものが多く存在する方がよい」、「反論先を共有する反論の間隔は小さい方がよい」、「反論の順序は相手側の発言と対応している方がよい」という 4 つの仮説を立て、それぞれに関して特徴量を定義した。また、即興型ディベートに関するコーパスを構

築し、即興型ディベート経験者による評価に基づき試合ごとの議論の噛み合い度合いの分類を行い、仮説に基づく評価と比較することで分析した。分析の結果、4 つの仮説に基づく *score* と議論の噛み合い度合いの評価値の相関係数は 0.732 であり、正の相関が確認された。また、ラリーが長続きしている場合は「反論の間隔」や反論の順序の対応が弱くとも、対話が深まっている可能性が示唆された。

本研究の限界として、人手により ADU の分割や反論関係の推定を行ったため、見落としが生じていた可能性がある。また今回扱ったデータ構造では発言のタイミングに関する情報を反映できておらず、研究の再現性という観点から見てデータが不完全であったと言える。そこで今後は議論マイニングの技術を用いて議論のグラフ化を自動的に行うシステムを開発し、より客観的に解析できるようにする。また、ADU ごとのタイムスタンプ情報をデータ構造に加える予定である。

また、今後は、スピーチごとの役割の違いや、議論比較と呼ばれる即興型ディベートの独特な説得戦術などの今回扱われなかった様々な議論の要素を考慮して分析を行う。そのためにより多くの動画を収集し、サークルの練習や大会での試合の記録を行うことで、コーパスを拡充する。さらに、即興型ディベートの参加者に有益なフィードバックを提供することを目的に、議論構造の特徴をユーザに提示する手法について調査する予定である。

## 参考文献

- [1] 加藤彰(2020). 即興型ディベートの教科書 ~東大で培った瞬時に考えて伝えるテクニック あさ出版
- [2] 河野周(2021). 中学・高校 英語ディベート入門 三省堂
- [3] 小林良裕(2021). 英語ディベート練習ハンドブック [即興型] S.A.D. Works
- [4] Ruiz-Dolz, Ramon, Stella Heras, and Ana García-Fornes. "Automatic debate evaluation with argumentation semantics and natural language argument graph networks." *arXiv preprint arXiv:2203.14647* (2022).
- [5] Fa-Hsuan Hsiao, An-Zi Yen, Hen-Hsen Huang, and Hsin-Hsi Chen. 2022. Modeling Inter Round Attack of Online Debaters for Winner Prediction. In Proceedings of the ACM Web Conference 2022 (WWW '22). Association for Computing Machinery, New York, NY, USA, 2860–2869. <https://doi.org/10.1145/3485447.3512006>
- [6] Henning Wachsmuth, Shahbaz Syed, and Benno Stein. 2018. Retrieval of the Best Counterargument without Prior Topic Knowledge. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 241–251, Melbourne, Australia. Association for Computational Linguistics.
- [7] Shi, Hongguang & Cao, Shuirong & Nguyen, Cam-Tu. (2023). Revisiting the Role of Similarity and Dissimilarity in Best Counter Argument Retrieval.
- [8] Wambsganss, Thiemo and Niklaus, Christina and Cetto, Matthias and Söllner, Matthias and Handschuh, Siegfried and Leimeister, Jan Marco, An Adaptive Learning Support System for Argumentation Skills (April 25–30, 2020). Conference on Human Factors in Computing Systems (CHI), Available at SSRN: <https://ssrn.com/abstract=3911896>

- [9] Meng Xia, Qian Zhu, Xingbo Wang, Fei Nie, Huamin Qu, and Xiaojuan Ma. 2022. Persua: A Visual Interactive System to Enhance the Persuasiveness of Arguments in Online Discussion. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 319 (November 2022), 30 pages. <https://doi.org/10.1145/3555210>
- [10] Christian Stab and Iryna Gurevych. 2014. Identifying argumentative discourse structures in persuasive essays. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP '14*, pages 46–56, Doha, Qatar.
- [11] Ruiz-Dolz, R.; Nofre, M.; Taulé, M.; Heras, S.; García-Fornes, A. VivesDebate: A New Annotated Multilingual Corpus of Argumentation in a Debate Tournament. *Appl. Sci.* 2021, 11, 7160. <https://doi.org/10.3390/app11157160>
- [12] Dung, Phan Minh. “On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games.” *Artif. Intell.* 77 (1995): 321-358.