

表層的・内容的判定機能を有する一対一通話型討論システムの実現方式

石井 雄大 神野 満里奈 黒木 友瑛 石橋 直樹 佐々木 史織

武蔵野大学データサイエンス学部 〒135-8181 東京都江東区有明 3-3-3

E-mail: {s2022052, s2022043, s2022010}@stu.musashino-u.ac.jp, {n-ishi, ssasaki}@musashino-u.ac.jp

あらまし 本稿では、ユーザ同士の通話によるインタラクティブな討論の発言内容から、討論の勝敗または優劣を自動判定する「一対一形式通話型討論システム」の実現方式を示す。本システムのユーザは、要論のテーマに対して賛成または反対を選択し、賛成者と反対者が揃った時点でビデオ通話を通じた討論を開始する。本システムは、両者の発言内容を1分区分切りの時系列データとして記録し、その内容に対して、「表層的観点」と「内容的観点」からアプローチすることで討論の結果を判定する。本システムの特徴は、音声通話での発言内容データから討論終了直後に自動的な勝敗/優劣判定を行う点にある。このことにより、ユーザはユーザ自身の背景に左右されることなく、より平等な討論を実現することができる。

キーワード ソーシャルメディア、自然言語処理、web 情報システム、言論、討論、スピーチ、インタラクティブ、ストリーム

1. はじめに

討論とは特定の論争的な議題を巡り対立する二つの立場に分かれて交互に主張や反論を展開し、議決を問う議論形式のことである[1]。

日本には古来より豊かな議論の伝統があり、明治大正期になると盛んに討論会などが行われてきた[1]。明治6年、慶應義塾大学創立者である福沢諭吉が「士族の家録なるものは、一体プロパーチーであるかサラリーであるか」という論題で討論を展開したことを灯火に、日本で「討論会」が盛んに行われるようになった[2]。自由民権の時代は、日本における第一の討論時代であったといえる。

1946年頃には、「朝日討論会」と呼ばれる学生を対象とした朝日新聞社主催の競技討論が行われ[3]、この時期から「討論」が教育の場に導入された。これは日本における第二の討論時代といえる。

そして現在、スマートフォンやPCの目覚ましい普及によりこれまでのような対面形式での討論だけでなく、オンライン上での討論やSNSを用いたオンラインチャット形式による討論も行われるようになってきている。また、テレビにおいては、討論形式によるニュース番組やバラエティ番組での企画によるエンターテインメントとしての討論も登場し、討論というものがより身近な存在になりつつある。これは日本が第三の討論時代を迎えているともいえる。

一方で、オンライン討論を行っている者は、SNSや動画配信サービスで多くの人から支持を受けているインフルエンサーやメディアなどで注目されている著名人とどまっている。さらにそのような場合において必ずしも純粋な主張の内容で結果が決まるというわけではなく、彼らの肩書きや支持率に結果が左右されるケースも存在する。このような現状において、多くの

人は議論の参加者としてではなく、聴講者という立場として議論を聞くことがほとんどである。

そこで、本研究では、多くの人がフォロワー数や社会的地位などの個人の「背景」を気にすることなく討論を行うことができるプラットフォームを構築する。そして、ユーザ同士の通話によるインタラクティブな討論の発言内容から、討論の勝敗または優劣を自動判定する「一対一形式通話型討論システム」を実現する。ここでいう「背景」とは、本来の討論での判定の際にバイアスになり得る討論内容以外の参加者自身が持つSNSにおけるフォロワー数や社会的地位などの要素のことを指す。

討論をサポートするサービスとしては、2020年にリリースされた討論アプリUooV[4]がある。これは投稿されたテーマに対してSNSに近い形での討論を可能とする。しかし、SNS形式での討論では、結論が出ないことや、討論自体の盛り上がり欠けることが懸念される。また討論内容とは関係のない要素である「タイピングの速さ」で不利になってしまう可能性が存在する。

本方式では、通話形式を採用することで、個人のタイピング能力に依存せず、かつ対面での討論と同じような感覚で参加することのできるシステムを実現する。また、判定の際は発言の内容のみを対象とするため、ユーザの背景に左右されず、より平等な討論を実現することが可能となる。

本システムのユーザは、討論のテーマに対して賛成または反対を選択し、賛成者と反対者が揃った時点でビデオ通話を通じた討論を開始する。本システムは、両者の発言内容を1分区分切りの時系列データとして記録し、その内容に対して、「表層的観点」と「内容的観点」からアプローチすることで討論の結果を判定する。

インターネットや SNS の普及によりオンライン上での人々の繋がりが強くなった現在において、あえて純粋な主張の内容のみに着目したシステムを実現することにより、新たな討論の形を形成することが可能となると考える。

本稿は次の構成からなる。2 節では、本方式の関連研究について紹介する。3 節では、表層的・内容的判定機能を有する一対一通話型討論システムの実現方式の詳細を述べる。4 節では、本方式を実現する実験システムを構築した上で行なった実験について述べる。5 節では、結論と今後の展望について述べる。

2. 関連研究

本節では、提案方式に関連する研究について述べる。

2.1 では、ユーザの背景差によるバイアスに関する研究について示す。2.2 では、討論の形式に関する研究について示す。2.3 では、討論の自動判定に関する研究について示す。2.4 では、本研究の位置付けについて示す。

2.1. ユーザの背景差によるバイアスに関する研究

本節では、ユーザの背景差によるバイアスに関する研究について述べる。

D.Westerman ら[5]や J.Weismueller ら[6]は、ソーシャルメディアにおけるフォロワー数と情報の信頼性の関連について示唆している。

D.Westerman ら[5]は、ソーシャルメディアである Twitter におけるフォロワー数と情報の信頼性の関連について調査し、関連性があることを示した。

J.Weismueller ら[6]は、ソーシャルメディアである Instagram におけるフォロワー数と情報の信頼性の関連について調査し、関連性があることを示した。

J.Fraser ら[7]は、役職と情報の信頼性の関連性について示唆している。

J.Fraser ら[7]は、環境保全に関するメッセージを発信する広報担当者の役職が与える信頼性への影響を調査した。また、その結果から一部の役職が他の役職よりも信頼性が高いことを示した。

これらの研究は、社会的に支持されている人や一部の職業や役職の人の発信する情報の信頼性が高くなることを示している。これらは討論という場合においてもバイアスとなると考えることができる。

2.2. 討論の形式に関する研究

本節では、討論の形式に関する研究について述べる。

S.Jane ら[8]は、教育分野においてオンライン上でのチャット形式を用いた討論をデザイン、実施し、オンラインチャットによる討論の有用性について示した。

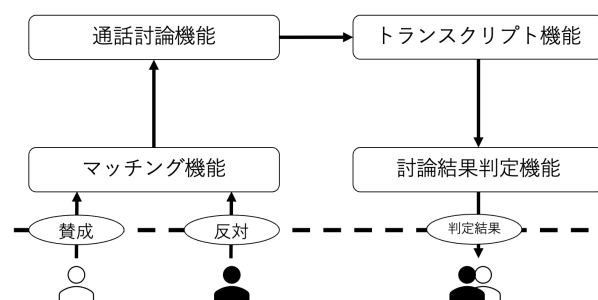


図1 本システムの概要

J.Son ら[9]は、教育分野におけるオンライン上でのビデオ通話形式を用いた討論をデザイン、実施し、オンラインビデオ通話による討論の有用性について示した。

これらは、オンライン上での討論の有用性を示す事例である。

2.3. 討論の自動判定に関する研究

本節では、発言内容の判定に関する研究について述べる。

喜多ら[10]は、討論における発言者の話者態度、根拠の信頼性、討論技術の3観点からの優劣度の判定方式を実現した。

2.4. 本研究の位置付け

本方式では、参加者自身の背景に影響されないオンラインビデオ通話による討論システムを実現し、その発言内容にのみ着目し討論の技術に左右されない判定方式を実現する。

2.1 では、ユーザの背景差によるバイアスに関する研究をソーシャルメディアにおけるフォロワー数と職業や役職の2つの観点から取り上げた。本研究ではこれらの討論内容以外の討論結果に影響するバイアスを排除した形での討論プラットフォームの実現を目指す。

2.2 では、討論の形式に関する研究として、オンラインチャット形式とオンラインビデオ通話形式の2つを取り上げた。本研究では、より議論の活発化を目指しオンラインビデオ通話形式を採用する。

2.3 では、発言内容の判定に関する研究として、喜多ら[10]の研究を取り上げた。本研究では討論技術を判定の対象としない、純粋な発言内容のみを対象とした方式を実現する。

3. 表層的・内容的判定機能を有する一対一通話型討論システムの実現方式

本節では、提案方式である表層的・内容的判定機能を有する一対一通話型討論システムの実現方式について

て提示する。

本システムは、各テーマに対しユーザが賛成または反対を選択し、両者が通話によってインタラクティブに討論を行ったその発言内容から、どちらが優勢であるかの結果を判定する。

本節では、本研究における提案手法の概要を述べる。提案システムの概要と機能群を図 1 に示す。

本システムの入力、テーマに対するユーザの賛成または反対という表明と討論開始後の発言内容であり、出力は、討論の勝敗/優劣に関する判定結果である。ユーザは、討論のテーマに対して賛成または反対を選択し、賛成者と反対者が揃った時点でビデオ通話を通じた討論を開始する。本システムは、両者の発言内容を 1 分区分切りの時系列データとして記録し、その内容に対して、「表層的観点」と「内容的観点」からスコアリングすることで討論の結果を判定する。

本システムは、1) マッチング機能、2) 通話討論機能、3) トランスクリプト機能、4) 討論結果判定機能で構成される。このうち、4) 討論結果判定機能は、さらにサブ機能として「表層的判定機能」と「内容的判定機能」を有している。

1) マッチング機能は、賛成意見を持つユーザと反対意見を持つユーザをマッチングさせる機能である。2) 通話討論機能は、通話形式で、賛成意見を持つユーザと反対意見を持つユーザがインタラクティブに討論を行う機能である。3) トランスクリプト機能は、討論における両者の発言内容を 1 分区分切りの時系列データとして記録する機能である。4) 討論結果判定機能は、トランスクリプト機能で記録された発言内容に対して、表層的観点と内容的観点からアプローチし、討論の結果を判定する機能である。以下に各機能の詳細について述べる。

3.1. マッチング機能

マッチング機能とは、テーマに対して賛成派のユーザと反対派のユーザをマッチングさせる機能である。

この機能は、賛成または反対を選択したユーザを待機させ、賛成派と反対派がそれぞれ 1 人ずつ以上揃った段階で賛成派と反対派のペアを決定する。そのタイミングで、もし片方の主張の人物が複数人存在した場合は先に待機していたユーザが優先され、残りのユーザは引き続き待機させる。この機能によって、討論の参加者を確定させ、一対一の討論環境を実現することができる。

3.2. 通話討論機能

通話討論機能とは、賛成派のユーザと反対派のユーザが通話形式でインタラクティブに討論を行う機能で

表 1 討論の経過時間と討論フェーズ

討論開始からの経過時間	討論フェーズ
0-1 分	賛成意見の主張
1-2 分	反対意見の主張
2-8 分	討論タイム
8-9 分	賛成意見の主張
9-10 分	反対意見の主張

ある。

この機能は、各ユーザのカメラ映像と音声を互いの画面上に表示させることによってインタラクティブな討論を実現する。討論には、10 分の時間制限が設定されており、議論の活発化促進のために時間ごとにフェーズを設定し、経過時間と共に画面に表示する。各時間の表示内容を表 1 に示す。

この機能によって、背景差に考慮した討論とそのスムーズな進行の促進を実現することができる。

3.3. トランスクリプト機能

トランスクリプト機能とは、入力された音声を文字起こしする機能である。この機能は、各ユーザの発言内容を文字化し、1 分区分切りの時系列データとして出力する。

この発言内容はリアルタイムでユーザの画面にも表示される。これによりユーザは自分の発言内容を確認しながら討論に参加することができる。また、この発言内容のデータを用いて討論結果判定機能において討論の結果を判定する。

3.4. 討論結果判定機能

討論結果判定機能とは、トランスクリプト機能によって出力された発言内容データを対象として、討論結果を判定する機能である。この機能は、「表層的判定機能」と「内容的判定機能」から構成される。

表層的判定機能は、発言内容を対象に表層的観点から賛成派・反対派の両者のスコアを算出する。本研究では、一定時間においてより多く発言している人が優勢であると定義する。

STEP1: 賛成派と反対派の各ユーザの発言について、一定時間のフレーム（たとえば 1 分間）の発言内容に区切った文を生成する。

STEP 2: STEP 1 で生成した文に対して形態素解析を行い、名詞、動詞、形容詞、副詞のみを抽出する。

STEP 3: 各分ごとに単語数をカウントし、数が多かった方のユーザのスコアに 1 ポイント追加する。

STEP 4: STEP 2 の計測・計算を全フレームについて（たとえば制限時間 10 分間に計 10 回）行い、各ユーザの合計スコアを算出する。同数の場合は、両者ポイントなしとする。（制限時間 10 分間の場合は、両者合計のスコアは最大で 10 となる。）

内容的判定機能は、発言内容を対象に内容的観点から両者のスコアを算出する。本研究では、テーマに含まれるワードに近いワードを多く用いた人が優勢であると定義する。

STEP1: 賛成派と反対派の各ユーザの発言について、一定時間のフレーム（たとえば 1 分間）の発言内容に区切った文を生成する。

STEP 2: STEP 1 で生成した文に対して形態素解析を行い、名詞のみを抽出し、重複した語を除外する。

STEP 3: 発言内容の各名詞とテーマごとに設定される 1 つの主要ワードに対して類似度計量を行う。算出された類似度が閾値（たとえば 0.5）以上のものを類似単語であるとし、この類似単語の数をカウントする。

STEP 4: カウントした類似単語の数を名詞の数で割り、割合を算出する。

STEP 5: 両ユーザの類似単語の割合が算出されたところで、フレーム（1 分）ごと両者の類似単語数の割合を比較し、フレームごとに割合が大きい方に 1 ポイントとしてスコアを算出する。（制限時間 10 分間の場合は、両者合計のスコアは最大で 10 となる。）

表層的判定機能と内容的判定機能の実行後、各機能によって算出されたスコアの合計を最終的なスコアとする。（制限時間 10 分間の場合は、両者合計のスコアは最大で 20 となる。）スコアが大きかった方が優勢として出力する。両者のスコアが同数だった場合は、「引き分け」として表示する。

以上の一連の機能によって、参加者の背景差に考慮した一対一形式の通話型討論システムが実現する。

なお、本研究の内容的判定機能では STEP 3 の類似度計量に Word2vec[11]を用いる。Word2vec のモデルに



図 2 プロトタイプシステムの討論参加画面

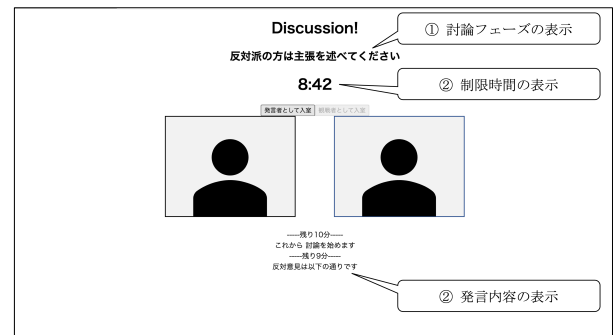


図 3 プロトタイプシステムの討論画面



図 4 プロトタイプシステムの結果画面

は「日本語 Wikipedia エンティティベクトル」[12]を用いる。

4. プロトタイプ実装

3 節で述べた実現方式により、プロトタイプシステムを実装した。

本プロトタイプシステムの討論参加画面を図 2 に示す。討論参加画面は、討論テーマと賛成・反対ボタンから構成される。ユーザは表示された討論テーマに対し、「賛成」または「反対」を選択する。その後、マッチング機能によって討論画面へと遷移する。

本プロトタイプシステムの討論画面を図 3 に示す。討論画面は、討論フェーズと制限時間、発言内容から構成される。これらの内容は、3.2.通話討論機能と 3.3.トランスクリプト機能において述べた通りである。

本プロトタイプシステムの結果画面を図 4 に示す。

表 2 各討論の参加者とそのテーマ

No	賛成	反対	テーマ
1	A	B	iphone の充電ケーブルはライトニングであるべきか
2	C	A	貯金はすべきか
3	B	C	朝型であるべきか
4	F	E	iphone の充電ケーブルはライトニングであるべきか
5	G	F	貯金はすべきか
6	E	G	朝型であるべきか

表 3 各討論の結果

No	賛成スコア	反対スコア
1	9	10
2	10	7
3	4	15
4	11	8
5	14	5
6	12	7

結果画面では、討論結果判定機能によって結果を自動判定し、勝敗と両者のスコアを表示する。

5. 実験

本節では、4 節で実装したプロトタイプシステムを用いて、その実現可能性と有効性を検証するための実験を行なった。

実験では、本システムを用いた被験者実験を行うことで、有効性を検証する。

5.1. 実験環境

本実験は、20 代の男女 6 名の協力のもと行なった。被験者 3 人を 1 グループとし、グループ内の総当たり形式で 3 回の討論を実施した。これを 2 グループ分実施した。各討論の参加者とそのテーマを表 2 に示す。なお、参加者名に関しては、A-G に匿名化して記載する。討論後、被験者は 3 つのアンケートに回答する。設問 1 は「今回の判定に納得しているか」、設問 2 は「スムーズに討論できたか」、設問 3 は「本システム全体に満足しているか」である。被験者は、毎討論の終了後、設問 1 と設問 2 に回答し、全ての討論の終了後、設問 3 に回答する。

5.2. 実験結果

本システムのプロトタイプを使用して実施した各討論の結果を表 3 に示す。全体として、賛成側が 4 勝、

反対側が 2 勝していることが分かる。

設問 1 の回答結果を表 4 に示す。表 4 を見ると優勢側の 6 回中 5 回が「納得している」と回答していることが分かる。劣勢側は 6 回中 3 回が「納得している」と回答していることが分かる。「納得していない」と回答した被験者からは「結果のスコアが直観的でない」などの意見が得られた。

設問 2 の回答結果を表 5 に示す。表 5 を見ると、全 12 回答中 10 回答が「スムーズだった」と回答していることが分かる。「スムーズだった」と回答した被験者からスムーズであると感じた要因に関して「フェーズの表示」が大きかったという意見が得られた。反対に、「スムーズでない」と回答した被験者からは「初めての場、フェーズの全体像を把握できない」や「フェーズにより時間配分などの進行面は機能していたが、代わりに時間が余ってしまった」などの意見が得られた。

設問 3 の回答結果を図 2 に示す。図 2 を見ると 8 割以上が本システムを使用したいと答えたことが分かる。「はい」と回答した被験者からは「偏見による考えで多数の人が一方の意見に賛成することで偏って判定となることを避けることができるという点で優れている」という意見が得られた。「いいえ」と回答した被験者からは「将来性は感じるが、現時点では不十分であると感じた」という意見が得られた。

5.3. 考察

表 4 問 1 の解答結果

No	優勢側	劣勢側
1	納得している	納得している
2	納得している	納得していない
3	納得していない	納得していない
4	納得している	納得している
5	納得している	納得している
6	納得している	納得していない

本節では、実験に対する考察を行う。

設問 1 では、討論結果判定機能の有効性の検証を目的に質問した。問 1 の結果より、6 割以上が「納得している」と回答していることが分かる。このことから発言内容を対象とした判定機能は妥当性があるといえる。

「納得していない」という回答の理由として「スコアが直感的でない」ことが挙げられる。スコアの詳細を表示することで改善できると考える。具体的には、2 つの観点別のスコアを付随させて表示させることで、結果の表示に説得力を持たせることを目指す。

設問 2 では、討論通話機能の有効性の検証を目的に質問した。問 2 の結果より、ほとんどの回答が「スムーズだった」であり、討論フェーズがその大きな要因となっていることが分かる。このことから討論フェーズの表示には有用性があると考えられる。しかし、「スムーズでない」と回答した要因もこの討論フェーズにあると考えられる。討論フェーズの全体像の把握に関しては、画面上に全体像を表示させることで解決できると考えられる。時間が余ってしまうという点に関しては、沈黙が多く発生したフェーズが 6 分間と一番時間の長い「討論タイム」であったことから、討論フェーズに加えて、発言内容に応じて参加者に話を振るようなファシリテーションシステムを実現することで解決できると考える。

表 5 問 2 の回答結果

No	優勢側	劣勢側
1	スムーズだった	スムーズでない
2	スムーズだった	スムーズだった
3	スムーズでない	スムーズだった
4	スムーズだった	スムーズだった
5	スムーズだった	スムーズだった
6	スムーズだった	スムーズだった

このシステムを使いたい

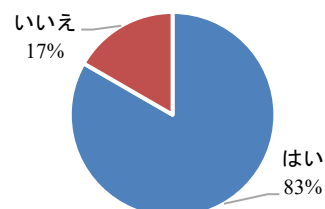


図 4 問 3 の回答結果

また、討論全体の時間を短くするなどの検討を必要であるとする。

設問 3 では、本システム全体の需要の確認を目的に質問した。設問 3 の結果より、実際に本システムを使用した被験者の中に、本方式を使用したいという意見が多いことが分かった。このことから、本システムには需要が認められるといえる。

6. おわりに

本稿では、表層的・内容的判定機能を有する一対一通話型討論システムの実現方式について示した。

本方式は各テーマに対しユーザが賛成または反対を選択し、両者が通話によってインタラクティブに討論を行い、その発言内容からどちらが優勢であるかの

結果を判定する。また、提案方式を実装し、有効性を検証するための実験を行なった。

今後の課題としては、討論結果や討論フェーズの表示方法の改善とファシリテーションシステムの実装の2つが挙げられる。これにより、討論の更なる活発化とユーザの感覚と結果の一致の促進が可能となる。また、カメラ映像のアバター化や音声処理による匿名化によって更なる討論結果へのバイアス最小化が可能であると考えられる。

本システムによって、ユーザの立場や肩書きに結果が左右されることのない、純粋な内容のみの討論とその判定を実現することができる。

参 考 文 献

- [1] 師岡淳也, "近代日本における討論の史的研究に関する予備的考察", 2011.
- [2] 平井一弘, "福沢諭吉と演説".
- [3] 和井田清司, "戦後日本における「朝日式討論」の盛衰", 全国社会科教育学会 社会科学研究第 54 号, 2001.
- [4] UooV (<https://uooov06.webnode.jp/>).
- [5] D. Westerman, P. Spence and B. Heide "A social network as information: The effect of system generated reports of connectedness on credibility on Twitter", Computers in Human Behavior, pp.199-206, 2012.
- [6] J. Weisueller, P. Harrigan, S. Wang and G.N. Soutar "Influencer Endorsements: How Advertising Disclosure and Source Credibility Affect Consumer Purchase Intention on Social Media", Australasian Marketing Journal, pp.160-170, 2020.
- [7] J. Fraser, A. Taylor, E. Johnson and J. Sickler "The Relative Credibility of Zoo-Affiliated Spokespeople for Delivering Conservation Messages", Curator: The Museum Journal, pp.407-418, 2008.
- [8] S.J. Lin and S.Y. Crawford "An online debate series for first-year pharmacy students", American journal of pharmaceutical education, 2007.
- [9] S.Jane, D. Vitaly, K.Sara and W.M. Timothy "Group processes in computer-mediated communication", Organizational Behavior and Human Decision Processes, pp.157-187, 1986.
- [10] 喜多唯, 末政貴弘, 原田実, "意味解析に基づく討論支援システム DESSY の開発", 情報処理学会研究報告, 2012.
- [11] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado and J. Dean "Distributed Representations of Words and Phrases and their Compositionality", Advances in Neural Information Processing Systems 2013, pp. 3111-3119, 2013.
- [12] 鈴木正敏, 松田耕史, 関根聡, 岡崎直観, 乾健太郎, "Wikipedia 記事に対する拡張固有表現ラベルの多重付与", 言語処理学会第 22 回年次大会 (NLP2016), 2016.