

話題の継続に着目した対話者間の発話に対する逐次的親疎判別

三浦拓人^{†1} 金井秀明^{†1}

概要: 近年、対話システムの有用性を向上させる等の理由を踏まえて、対話から得られる情報から話者間の親密度を推定する研究が盛んに行われている。これらの研究の多くは、対話が一通り終了した後の情報を対象として、事後的に話者間の親密度を推定している。しかし、親密度推定の結果を対話システムに活かすためには、対話を通じて変化する話者間の親密度を逐次的に推定する必要がある。本研究では、逐次的な親密度推定手法の実現における第一歩として、話者間が親しい関係であるか・疎い関係であるかを逐次的に判別する親疎判別手法を提案する。既存研究では、親疎判別に寄与する特徴量として敬語の有無が同定されていた。しかし、敬語の有無は親密度を表層的にしか捉えることができず、敬語の有無のみに頼った親疎判別では頑健性が低くなってしまうことが考えられる。そこで、本研究では、提案手法として、トピック類似度に基づく話題の継続に着目した特徴量を作成することで、既存の逐次的親疎判別手法が抱える問題の解消を試みた。評価実験の結果、提案手法はベースラインの手法よりも優れた結果を得ることができた。

キーワード: 親疎判別, 話題の継続, トピック類似度, 非タスク指向型対話

Sequential intimacy classification for utterances between talkers focusing on topic continuity

TAKUTO MIURA^{†1} HIDEAKI KANAI^{†1}

Abstract: In recent years, many studies have been conducted to estimate the degree of intimacy between talkers based on the information obtained from dialogues, to improve the usability of dialogue systems. These studies have tried to estimate the degree of intimacy between talkers after the information has been completed. However, to utilize the results of intimacy estimation in a dialogue system, it is necessary to estimate the intimacy between talkers that changes through dialogue sequentially. In this research, as the first step toward realizing a sequential intimacy estimation method, we propose a method for determining whether the relationship between talkers is close or loose. Previous studies have identified the presence or absence of honorifics as a feature that contributes to the determination of intimacy. However, the presence or absence of honorifics can only superficially perceive intimacy. The robustness of intimacy classification based on only the presence or absence of honorifics might be low. Therefore, in this research, we tried to solve the problem of the existing sequential determination of intimacy by creating a feature value that focuses on topic continuity based on topic similarity. As a result of the evaluation experiment, the proposed method obtained better results than the baseline method.

Keywords: Classification of Intimacy, Topic Continuity, Topic Similarity, Non-task-oriented Dialogue

1. はじめに

近年、ユーザからの情報取得やユーザへの行動支援等において、ユーザとシステムが自由な形式で対話を行う非タスク指向型対話システムが注目されている。このような対話システムの満足度を高めるためには、ユーザと親密な関係を築くことが有効であるとされている[1, 2]。親密化を図るためには、ユーザとの現在の仲の良さ（親密度）を定量化してそれらに基づいた発話生成を行うことが重要である[2]。対話システムがユーザとの現在の親密度を定量化するためには、発話等の対話中の情報から親密度を推定することが必要となる。

今日までに対話から得られる情報から話者間の親密度を推定する研究が盛んに行われている[3, 4, 5, 6]。しかし、これらの研究では対話後の情報を対象として事後的に話者間の親密度を推定している。そのため、これらの手法では、

対話システムが対話を通じて逐次的に変化する話者間の親密度を推定し、発話生成へと活かすことが困難である。

以上のことを踏まえて、我々は、話者の各発話を対象に話者間の親密度を逐次的に推定する手法の実現を目指す。そこで、本研究では、逐次的親密度推定の実現に向けて、話者間が親しい関係であるか・疎い関係であるかを逐次的に判別する親疎判別手法を提案する。提案する親疎判別手法には、話者間の親密度と同話題での全発話数の間に何らかの有意な関係性が示唆された研究[9, 10]を踏まえて、話題の継続に関する特徴を活用する。

以降では、第2章で関連研究について述べ、本研究の位置付けを整理する。第3章で親密度と話題の継続に関する予備分析について述べ、第4章では、提案する親疎判別手法について述べる。第5章で提案手法の評価実験について述べ、第6章で本論文をまとめる。

^{†1} 北陸先端科学技術大学院大学
Japan Advanced Institute of Science and Technology

2. 関連研究

既存の親密度推定手法，ならびに逐次的親密度推定手法に関する研究を取り上げ，その問題点を指摘する．また，話題の継続は話者によって差異が発生しうる対話構造に関する特徴の1つであると捉え，話題の継続と親密度との関連性を調査した研究を取り上げる．

2.1 親密度推定

西原らは，対話中に含まれる各発話文の役割を助詞・助動詞の関係に基づいて同定し，対話者間の仲の良さや上下関係を推定する手法の実現を試みている[3]．また，Jiaらは，短い対話に対しても対話者間の人間関係を高精度で推定するための手法として，複数の対話事例を基に推定を行うモデルの構築を目指した[4]．具体的には，各発話を埋め込みベクトルとして特徴量化し，それらを LSTM[16]や BERT[17]等の時系列モデルによって学習するというタスクを定義していた．さらに，1回の対話を対象としたタスクと複数の対話を対象としたタスクの2段階にタスクを分けることで精度の向上を目指すという手法を提案していた．

松本らは，演劇台本におけるシーン別の対話を対象として，発話役割の同定結果，辞書や条件付き確率場に基づいた感情状態の同定結果，発話・返答回数の特徴量として利用し，対話者間の親密度を推定する手法の実現を試みている[5]．

C. S. Rodriguezらは，発話時の音響特性を特徴量として，チームメンバー間の信頼関係や親密度を推定する手法の実現を試みている[6]．

これらの研究では，対話後の情報を取得し，特徴量を作成しているため，話者間の親密度推定が事後的に行われている．そのため，対話を通じて変化する対話者との親密度を逐次的に推定できず，それらを活かした発話生成を行う対話システムに対応することができない．

そこで本研究では，発話毎に特徴量を作成し，親密度推定を発話単位で行うことによって，逐次的な親密度の推定を試みる．

2.2 逐次的親密度推定

木下らは，対話中の感情表出・文書特徴量・話題の内容等を用いて発話対に対する逐次的な親疎判別手法の実現を試みている[7]．また，Chibaらは，対話中の言語的情報・韻律情報・視覚的情報から有効な特徴量を同定し，それらを用いた逐次的な親密度推定手法の実現を試みている[8]．

これらの研究では，推定精度において一定の成果が挙げられている．しかし，これらの研究で親密度推定への寄与が明示されたテキストベースの特徴量は，敬語の有無のみである．例1の対話例のように，親密な関係性でも敬語を使用する場合もある．よって，敬語の有無だけでは親密度

を表層的にしか捉えることができないと考えられる．そのため，敬語の有無のみを用いた親密度推定手法では頑健性が低くなってしまふことが示唆される．

そこで，本研究では，敬語の有無ではないテキストベースの特徴量を作成する．具体的には，話者によって差異が発生しうる対話構造の特徴である話題の継続に関する特徴量を作成する．

A：残業のため外食で済みます

了解です：B

A：ケーキ買って帰るけど？

私モンブランで：B

A：了解、C子は？

遠足だったのでもう寝ました：B

例1 家族での対話で敬語を使用する例

Example 1 An example of using honorifics in family conversations

2.3 親密度と話題の継続について

話者間の親密度と対話について調査を行った研究[9, 10]では，話者間の親密度と同話題での全発話数の間に何らかの有意な関係性があることを示唆している．これらの研究では，有意差検定の結果から，会話の目的・種類と話者間の親密性が1会話での全発話数に何らかの影響を及ぼすことを明示している．特に[9]では，会話の題材と話者間の親密性の関係を調査しており，目的がない非タスク指向型会話において，話者間の親密性が高いほど発話数が多くなり，話者間の親密性が低いほど発話数が少なくなるという結果を示している．

そこで，本研究では，対話中の話題の継続に着目して，現在の発話とそれ以前の発話における話題の類似度合いに関する特徴量を作成する．

3. 予備分析

親密度推定への寄与が期待される，話題の継続に着目した特徴量に関する分析について述べる．

3.1 話題の継続に関する指標について

本研究では雑談対話における話題変遷の検出手法[11]を参考として，発話中に含まれる語間の類似度に基づいた話題の継続に関する指標を話題継続性とし，以下のように作成した．同様に，話題の飛躍に関する指標を話題飛躍性とし，以下のように作成した．話題の飛躍は，対話構造において話題の継続と対照的な関係性があり，話題の継続と同じく親密度推定への寄与が期待できると考えた．

3.1.1 話題継続性について

はじめに、話題と関連のない語を話題の継続に関する指標の作成に利用しないようにするため、指標の作成に使用する品詞や使用しない語（ストップワード）をあらかじめ設定した。使用する品詞は名詞、特に普通名詞と固有名詞とし、ストップワードは Slothlib[12]に含まれる語とした。さらに、トピックモデル洗練化手法に関する研究[13]と同様に、全文書中における当該語を含む文書頻度（DF 値）の閾値を設定して抽出した語とそれらの類似語も前述と併せてストップワードとした。ストップワードを除いた普通名詞と固有名詞の語を有効語として以降の処理に用いた。

次に、対話中の隣接発話同士を対象として、一方の発話に含まれる有効語ともう一方の発話に含まれる有効語の全組の類似度を Word2Vec[15]によって算出し、それらの平均値をトピック類似度（TopicSim）として算出した（式 1）。Word2Vec は、テキストデータに対して、単語とその周辺語の出現確率を基に単語間の関連性を学習し、各単語に特徴ベクトルを付与する手法である。なお、Word2Vec の作成には日本語 Wikipedia 全文を用いた。

$$TopicSim = \frac{\sum_a^A \sum_b^B \cos(w_a, w_b)}{\max(A, B)} \quad (\text{式 1})$$

A : 1 つ目の発話に含まれる単語数

B : 2 つ目の発話に含まれる単語数

w_q : q 番目の単語の特徴ベクトル

$\cos(r, s)$: r と s のコサイン類似度

上記のトピック類似度の閾値を設定して、対話中の隣接発話同士のトピック類似度が閾値以上の時、話題の継続として計数した（話題継続性と定義）。なお、発話者の交代が行われるまでの発話を当該話者の 1 発話とした。

最後に、対話中に含まれる複数の話題ごとに継続した発話の回数を算出し、対話ごとの話題継続性の平均値と最大値を算出した。なお、1 つの対話に含まれる全発話数が対話毎に異なることを踏まえて、当該対話の全発話数を用いて話題継続性の平均値と最大値を正規化した。

3.1.2 話題飛躍性について

前述の通り、話者間の親密度と同話題での全発話数の間に何らかの有意な関係性が示唆されたことを踏まえて、本研究では、話題の継続に関する指標だけではなく、話題の継続と対照的な関係性にある、対話中の話題の飛躍に関する指標も対象とした。

3.1.1 と同様の処理を行い、対話中の隣接発話同士のトピック類似度が閾値以下の時、話題の飛躍として計数した（話題飛躍性と定義）。そして、対話中の話題飛躍性の回数を算出し、対話ごとの話題飛躍性の合計とトピック類似度が最

小となる時の値を算出した。なお、3.1.1 と同様に、当該対話の全発話数を用いて話題飛躍性の合計を正規化した。

3.2 親密度と話題継続性・話題飛躍性に関する検定

前節で述べた話題継続性の平均値と最大値、話題飛躍性の合計とトピック類似度最小値を対象として、親密度と話題継続性、親密度と話題飛躍性の間の有意な関係に関する検定を試みた。なお、検定の手法として対応なし t 検定を用いることとした。

検定に用いるコーパスは BTSJ 日本語自然会話コーパス（トランスクリプト・音声）2022 年 3 月 NCRB 連動版[14]とした。ただし、話者同士の関係が初対面である対話を非親密対話、友人や家族等の知り合いである対話を親密対話として、それらに該当する対話のみを用いた（表 1）。

表 1 検定用コーパスの対話数と全発話数

Table 1 Number of dialogues and the total number of utterances in the test corpus

	非親密	親密	合計
対話数	85	71	156
全発話数	30,707	33,552	64,259

3.3 検定結果と考察

3.3.1 話題継続性に対する分析

対話中の隣接発話を話題継続性として計数する際の閾値として、0.1~0.9 を用意し、0.1 刻みに 9 回の検定を行った。

表 2 話題継続性に関する検定結果

Table 2 Topic Continuity Test Results

閾値	平均値	最大値	傾向
0.1	有意差あり	有意差なし	非親密
0.2	有意差あり	有意差あり	非親密
0.3	有意差なし	有意差なし	傾向なし
0.4	有意差なし	有意差なし	傾向なし
0.5	有意差なし	有意差なし	傾向なし
0.6	有意差なし	有意差なし	傾向なし
0.7	有意差なし	有意差なし	傾向なし
0.8	有意差あり	有意差なし	親密
0.9	有意差あり	有意差なし	親密

表 2 より、トピック類似度の閾値を 0.2 とした時、話題

継続性の平均値・最大値共に、非親密対話と親密対話の間で有意な差が見られた。さらに、非親密対話の方が親密対話と比べて、より話題が継続する傾向があると示唆された。

また、トピック類似度の閾値を 0.8 以上とした時、話題継続性の平均値のみに、非親密対話と親密対話の間で有意な差が見られた。さらに、閾値 0.2 の時とは異なり、親密対話の方が非親密対話と比べて、より話題が継続する傾向があると示唆された。

これらの結果は、単語間の類似度に基づくトピック類似度が発話内容の具体性に比例する可能性が高いことが要因であると考察する。例えば、「太平洋」に対する「大西洋」・「遠洋」・「水圏」について、「太平洋」は具体的な単語であり「大西洋」との相性が良い一方で、他単語との類似度は「太平洋」と「大西洋」の類似度ほど高まらないことが考えられる。一方、「海洋」に対する「大西洋」・「遠洋」・「水圏」について、「海洋」は「太平洋」と比べると抽象的な単語であり、どの単語ともある程度類似することが考えられる一方で、それらの類似度は「太平洋」と「大西洋」の類似度ほど高まらないことが考えられる。これらのことから、非親密な話者間ではより抽象的な内容の話題、親密な話者間ではより具体的な内容の話題が継続しやすいと考察できる。

以上のことを踏まえて、話題継続性に関する特徴量は親疎判別に対する寄与が期待できると示唆された。また、トピック類似度の閾値を変えて同じ特徴量を作ることで、意義の異なる特徴量を作ることができると示唆された。以降では、トピック類似度の閾値が 0.2 の時、抽象的な内容における話題継続性（抽象的_話題継続性）、トピック類似度の閾値が 0.8 の時、具体的な内容における話題継続性（具体的_話題継続性）とする。

3.3.2 話題飛躍性に対する分析

3.3.1 の結果を踏まえて、対話中の隣接発話を話題飛躍性として計数する際の閾値として、 $0.2 \cdot 0.1 \cdot 0.05 \cdot 0.04 \cdot 0.01$ を用意し、5 回の検定を行った。

表 3 話題飛躍性に関する検定結果

Table 3 Topic Leap Test Results

閾値	合計	類似度 最小値	傾向
0.2	有意差なし	有意差あり	親密
0.1	有意差なし		親密
0.05	有意差なし		親密
0.04	有意差あり		親密
0.01	有意差なし		親密

表 3 より、トピック類似度の閾値を 0.04 とした時、話題飛躍性の合計・トピック類似度最小値共に、非親密対話と親密対話の間で有意な差が見られた。さらに、親密対話の方が非親密対話と比べて、より話題が飛躍する傾向があると示唆された。

話題継続性に関する検定と話題飛躍性に関する検定の双方を踏まえて、非親密対話の方が親密対話と比べて、抽象的な内容で話題が継続しやすく、話題が飛躍しにくいと考察できる。また、親密対話の方が非親密対話と比べて、具体的な内容で話題が継続しやすく、話題が飛躍しやすいと考察できる。

以上のことを踏まえて、話題飛躍性に関する特徴量は親疎判別に対する寄与が期待できると示唆された。

4. 提案手法

4.1 特徴量について

前述の結果を踏まえて、本研究では話題継続性と話題飛躍性に関する特徴量を次のように作成した。

(1) 話題継続性の特徴量

・平均値

話題継続性の平均値 (ContinuityMean) は、現在発話までの話題継続性の合計を、現在までの発話数で正規化したものとする。

$$ContinuityMean = \sum_{i=1}^t c_i / u_t \quad (\text{式 2})$$

c_t : 発話 t 時点までの話題継続性

u_t : 発話 t 時点までの全発話数

・最大値

話題継続性の最大値 (ContinuityMax) は、現在発話までの話題継続性の最大値を、現在までに計数した最大値の時の発話数で正規化したものとする。

$$ContinuityMax = \max(c_t) / u_{\text{argmax}(c_t)} \quad (\text{式 3})$$

$\max(x)$: 集合 x の最大値

$\text{argmax}(x)$: 集合 x の最大値のインデックス

・分散

話題継続性の分散 (ContinuityVar) は、現在発話までの話題継続性の分散を、現在までに計数した最大値の時の発話数で正規化したものとする。

$$ContinuityVar = \text{Var}(c_t) / u_{\text{argmax}(c_t)} \quad (\text{式 4})$$

$\text{Var}(x)$: 集合 x の分散

ただし、トピック類似度の閾値が 0.2 の時、抽象的_話題継続性の平均値・最大値・分散、トピック類似度の閾値が 0.8 の時、具体的_話題継続性の平均値・最大値・分散とする。

(2) 話題飛躍性の特徴量

・合計

話題飛躍性の合計 (LeapMean) は、現在発話までの話題飛躍性の合計を、現在までの発話数で正規化したものとする。

$$LeapMean = \sum_{i=1}^t l_i / u_t \quad (式 5)$$

l_t : 発話 t 時点までの話題飛躍性

・トピック類似度の最小値

話題飛躍性のトピック類似度の最小値 (LeapMin) は、現在発話までのトピック類似度の最小値とする。

$$LeapMin = \min(TS_t) \quad (式 6)$$

$\min(x)$: 集合 x の最小値

TS_t : 発話 t 時点までのTopicSimの集合

ただし、正規化に用いる発話数に関して、挨拶などの対話初期に発生する有効語を持たない発話は発話数としてカウントしないこととした。

4.2 逐次的親疎判別モデルについて

前述した特徴量を用いて、話者間の親疎を逐次的に判別するモデルを構築する。

逐次的な親疎判別の方法について、全ての対話における 1 つ 1 つの発話を対象とし、各発話において前述の特徴量を導出する。なお、挨拶などの対話初期に発生する有効語を持たない発話の各特徴量について、各々の特徴量が取りうる範囲に含まれない任意の定数を与えることとした。抽象的_話題継続性の平均値・最大値・分散、具体的_話題継続性の平均値・最大値・分散、話題飛躍性の合計に対しては 0、話題飛躍性のトピック類似度の最小値に対しては 10 を与えた。

親疎ラベルについて、当該発話が含まれる対話における話者間の関係性が知り合い以上である場合、当該発話の親疎ラベルを「親」とし、当該発話が含まれる対話における話者間の関係性が初対面である場合、当該発話の親疎ラベルを「疎」とした。

以上のような特徴量、親疎ラベルを用いて、発話ごとに親疎を判別する逐次的親疎判別モデルを構築する。モデルの学習・評価に用いる手法やコーパスについては、第 5 章で記述する。

5. 評価実験

本研究では、提案手法の有用性を評価するために、ベースラインと提案手法との判別精度の比較を行った。なお、評価のために使用するデータは予備分析で使用したデータ (表 1) と同一であり、親疎判別モデルの学習に 50%、親疎判別モデルの評価に 50%を使用することとした。

5.1 ベースラインについて

提案手法との比較に用いるベースラインの手法としては、以下の 3 つを用いる。

(1) 敬語の有無による親疎判別

本研究では、敬語の有無のみを親密度推定に用いることの問題について指摘していたことから、ベースラインとして採用した。

(2) Doc2Vec による親疎判別

[7]でのベースラインとして、Word2Vec を文書長に拡張した手法である Doc2Vec による発話文ベクトルを用いていた。本研究でも、同様に Doc2Vec による発話文ベクトルをベースラインとして採用した。なお、Doc2Vec の作成には日本語 Wikipedia 全文を用いた。

(3) 敬語の有無および Doc2Vec による親疎判別

上記の (1) と (2) を併用した手法であり、ベースラインとして採用した。

5.2 親疎判別モデルについて

本研究では、親疎判別モデルとして、決定木、サポートベクターマシン (SVM)、多層ニューラルネットワーク (NN) のそれぞれを学習に用い、比較することとした。

ただし、ベースラインの (1) に関して、1 入力 (敬語の有無) ・ 1 出力 (親疎) の関係にあることから、モデルによる学習ではなく、発話文中の敬語使用の有無とその発話の親疎との一致率を導出することとした。

5.3 評価指標について

表 4 混同行列

Table 4 Confusion Matrix

		実際の親疎	
		疎	親
予測した親疎	疎	TP	FP
	親	FN	TN

混同行列について、TP (True Positive : 真陽性) は、実際の親疎と予測した親疎が共に「疎」であることを表している。FN (False Negative : 偽陰性) は、実際の親疎が「疎」、

予測した親疎が「親」であることを表している。FP (False Positive : 偽陽性) は、実際の親疎が「親」、予測した親疎が「疎」であることを表している。TN (True Negative : 真陰性) は、実際の親疎と予測した親疎が共に「親」であることを表している。

本研究では、親疎判別モデルの評価指標として、混同行列 (表 4) に基づいて導出される次の4つの指標、正解率 (Accuracy), 適合率 (Precision), 再現率 (Recall) および F 値 (F-score) を用いた。

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (式 7)$$

$$Precision = \frac{TP}{TP+FP} \quad (式 8)$$

$$Recall = \frac{TP}{TP+FN} \quad (式 9)$$

$$F - score = \frac{2Precision*Recall}{Precision+Recall} \quad (式 10)$$

5.4 結果と考察

表 5 正解率
 Table 5 Accuracy

	決定木	SVM	NN
提案手法	0.973	0.718	0.827
敬語の有無	0.669		
Doc2Vec	0.560	0.636	0.643
敬語の有無+Doc2Vec	0.609	0.690	0.687

表 6 適合率
 Table 6 Precision

	決定木	SVM	NN
提案手法	0.983	0.726	0.843
敬語の有無	0.620		
Doc2Vec	0.580	0.614	0.626
敬語の有無+Doc2Vec	0.628	0.640	0.672

表 7 再現率

Table 7 Recall

	決定木	SVM	NN
提案手法	0.964	0.739	0.823
敬語の有無	0.950		
Doc2Vec	0.575	0.818	0.792
敬語の有無+Doc2Vec	0.620	0.931	0.787

表 8 F 値

Table 8 F-score

	決定木	SVM	NN
提案手法	0.974	0.733	0.833
敬語の有無	0.750		
Doc2Vec	0.578	0.702	0.699
敬語の有無+Doc2Vec	0.624	0.758	0.725

前述の結果より、提案手法を用いた決定木による親疎判別モデルが4つの指標全てで最も良い精度を導出した。すなわち、提案手法が親疎判別において、有用であると言える。

第3章で行った予備分析の結果からも考察できるように、非親密な話者間の対話では、出身や年齢等の基本的な情報が互いに未共有であるため、抽象的な話題が継続しやすく、さらに現在の話題を基として連鎖的に話題が展開される傾向がある。一方で、親密な話者間の対話では、前述したような基本的な情報が互いに共有済みであるため、共通の趣味や仕事等の具体的な話題が継続しやすく、さらに現在の話題と関連性がない話題へと飛躍するような話題展開が行われる傾向がある。このような話者間の関係性に応じた対話の構造的特徴は、口調や言葉遣い等の表層的な特徴とは異なり、話者間の関係性を本質的に捉えられる特徴であると言える。以上のことから、対話の構造的特徴である話題の継続・飛躍に着目した提案手法の判別精度が良くなったと考える。

一方、敬語の有無に関して、発話中に含まれる第三者の発言における敬語と話者本人の発言における敬語を区別することが難しい。また、前述したように、親密な話者同士で敬語を使用する状況は十分に考えられる。他方で非親密な話者同士でも「私は途中で (やめました)。」のように、述語の省略等の理由で敬語が使用されない発話が見られた。以上のことから、敬語の有無のみを用いた手法では判

別精度が高まらなかったと考察する。

また、Doc2Vec による発話ベクトルに関して、前述したように Doc2Vec は、テキストデータに対して、文書とその周辺文書の出現確率を基に文書間の関連性を学習し、各文書ならびに各単語に特徴ベクトルを付与する手法である。すなわち、Doc2Vec による発話ベクトルは、当該発話がどのような内容、話題の場合に用いられやすいかを表現している。このことを踏まえて、前述の結果より、対話中に出現する話題の内容自体では、話者間の親疎を明らかにすることはできないと考察する。

以上のことを踏まえて、非タスク指向型対話を行う話者間の親疎を逐次的に判別する手法として、対話の構造的特徴である話題の継続や飛躍を用いた提案手法が有効であると明示された。ただし、提案手法の課題として、対話初期に発生する挨拶などの有効語を持たない発話に対しては、親疎判別が困難となるという点が挙げられる。これらの課題を改善する方法として、対話初期の発話に対しては有効語に依存しない敬語の有無による手法を利用し、以降の発話に対しては提案手法を利用するという方法などが考えられる。

6. おわりに

本研究では、逐次的親密度推定の手始めとして、話者間が親しい関係であるか・疎い関係であるかを判別する親疎判別手法を提案した。評価実験の結果、提案手法を用いた親疎判別モデルが最も良い評価を得た。

今後は、本研究を踏まえて、話者間の親疎関係を細分化した仲の良さの度合い（親密度）を話者の各発話を対象として逐次的に推定する手法の実現を目指す。さらに、構築した親密度推定モデルを用いて、意図的にユーザと親密な関係を築くことができるような非タスク指向型対話システムの実現を目指す。また、構築した親密度推定モデルを用いて、対話を行いながら初対面の人同士で親密化を図ることができる対話支援システムの実現を目指す。

参考文献

- [1] 宮澤幸希, 常世徹, 榊井祐介, 松尾智信, 菊池英明. 音声対話システムにおける継続欲求の高いインタラクションの要因. 電子情報通信学会論文誌 A, Vol. 95, No. 1, pp. 27-36, 2012.
- [2] Cory D, Kidd and Cynthia Breazeal. Robots at home: Understanding long-term human-robot interaction. In Proceedings of International Conference on Intelligent Robots and Systems, pp. 3230-3235, 2008.
- [3] 西原陽子, 砂山渡, 谷内田正彦. 発話テキストからの人間の仲の良さと上下関係の推定. 電子情報通信学会論文誌. D, 情報・システム, J91-D(1), pp.78-88, 2008.
- [4] Jia, Q., Huang, H., & Q. Zhu, K. DDrRel: A New Dataset for Interpersonal Relation Classification in Dyadic Dialogues. *Proceedings of the AAAI Conference on Artificial*

- Intelligence*, 35(14), pp.13125-13133, 2021.
- [5] 松本和幸, 秋田恭佑, 任福続, 吉田稔, 北研二. 演劇台本における登場人物間の親密度推定手法. 知能と情報 日本知能情報ファジィ学会, 30(3), pp.591-604, 2018.
- [6] C. S. Rodriguez, A. R. Paganiban, M. N. Stolar, R. S. Bolia and M. Lech, "Prediction of Inter-Personal Trust and Team Familiarity From Speech: A Double Transfer Learning Approach," in *IEEE Access*, vol. 8, pp. 225437-225447, 2020.
- [7] 木下泰輝, 楠和馬, 蒲原智也, 波多野賢治. 逐次的親密度推定に向けた発話対の親疎判別方の提案. 言語処理学会年次大会発表論文集, 24, pp.7-12, 2018.
- [8] Yuya Chiba, Yoshihiro Yamazaki, and Akinori Ito. Speaker Intimacy in Chat-Talks: Analysis and Recognition based on Verbal and Non-Verbal Information. *Proceedings of the 25th Workshop on the Semantics and Pragmatics of Dialogue*, September 20-22, 2021, Potsdam / The Internet.
- [9] 浦光博, 桑原尚史, 西田公昭. 対人的相互作用における会話中の質的分析. 実験社会心理学研究, 26(1), pp.35-46, 1986.
- [10] 西田公昭, 浦光博, 桑原尚史, 梶野潤. 対人的相互作用に及ぼす会話の媒介的影響. 社会心理学研究 / 日本社会心理学会「社会心理学研究」編集委員会 編, 3(2), pp.46-55, 1988.
- [11] 関萌水, 林篤司, 岩下志乃. 雑談対話における話題遷移の検出. 日本知能情報ファジィ学会 第37回ファジィシステムシンポジウム 講演論文集, 37, pp.338-341, 2021.
- [12] Slothlib, <http://svn.sourceforge.jp/svnroot/slothlib/CSharp/Version1/SlothLib/NLP/Filter/StopWord/word/Japanese.txt>
- [13] 東和幸, 高橋仁, 中川博之, 土屋達弘. 単語の出現頻度と類似度に基づいたトピックモデルの洗練化. コンピュータソフトウェア, 36(4), pp.25-31, 2019.
- [14] 宇佐美まゆみ監修 (2022) 『BTSJ 日本語自然会話コーパス (トランスクリプト・音声) 2022年3月 NCRB 連動版』、国立国語研究所、日本語教育研究領域「日本語学習者のコミュニケーションの多角的解明」、サブ・プロジェクト「日本語学習者の日本語使用の解明」(リーダー: 宇佐美まゆみ)
- [15] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13). Curran Associates Inc., Red Hook, NY, USA, 3111-3119.
- [16] Zhou P, Shi W, Tian J, Qi Z, Li B, Hao H, and Xu B. 2016. Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 207-212.
- [17] Devlin J, Chang M, Lee K, and Toutanova K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1 Long and Short Papers), 4171-4186.