

対話行為タグ付きコーパスの作成支援

洪 陽 杓 白 井 清 昭

北陸先端科学技術大学院大学 情報科学研究科

{y-kou, kshirai}@jaist.ac.jp

1 はじめに

本研究では、談話タグの一つである対話行為をタグ付けしたコーパスの構築を目的とする。対話行為を自動的に決定する手法も数多く提案されてはいるが [2, 3], その精度は十分高いわけではなく、自動的に対話行為タグ付けコーパスを構築することは困難である。そのため、自動的に決定された対話行為を修正するなど、人手による作業が必要不可欠である。本研究は、人間による対話行為のタグ付け作業を支援する様々な手法を提案する。また、提案手法を実装した対話行為タグ付きコーパス作成支援ツールを構築する。

2 提案システム

2.1 概要

本研究では、対話行為のタグ付けを支援する方法として、以下の4つを考える。

1. 対話を発話単位に自動的に区切る。また、発話単位を簡単に変更することのできるインターフェースを提供する。詳細は2.2項で述べる。
2. 作業がある発話に対する対話行為を決める際、対話行為タグの候補を提示する。また、対話行為の信頼度を求め、その高い順に作業者に提示する。作業者は、表示された対話行為のリストの中から正しいものを選択する。詳細は2.3項で述べる。
3. 作業がある発話に対する対話行為を決める際、その発話に似ている発話と、それに対する過去の対話行為のタグ付け結果を表示する。作業者は、対話行為の選択に迷ったとき、過去のタグ付け結果を参照することができる。詳細は2.4項で述べる。
4. 実際の発話の音声データを再生する。対話行為を決定する際に、韻律のような音声から得られる情報は重要な手がかりとなる。

本論文では、上記1~3の機能を備えたGUIによる支援ツールを作成した。ツールのプログラミング言語としてPerl/Tkを用いた。タグ付けを行う対象となる対話コーパスとして交通経路案内タスクのコーパスを使用した。また、対話行為の定義として、荒木らによって提案

された22個の対話行為 [4] を用いた。以下、上記1~3の処理の詳細について述べる。

2.2 発話単位への分割

対話行為タグを付与する最初のステップは、対話を発話単位に区切ることである。本研究では、堂坂ら [5] が提案した発話単位の定義に準じて対話を発話単位に分割し、個々の発話単位に対して対話行為を付与する。発話単位への自動分割は以下のように行う。

- 話者情報を事前に入力し、話者交替があったとき自動的に発話単位に分割する。
- 節の境界で発話単位を分割する。節の境界を検出するために、「けれども」「ので」や句点のような節の境界を示唆するキーワードを15種類用意した。
- フィラーやあいづちの前後で発話単位を分割する。フィラーやあいづちを検出するために、「えーっと」や「はい」のようなキーワードを40種類用意した。
- 二人の人が同時に発話したとき、発話者毎に発話単位を分割する。

予備実験の結果、本手法の発話単位区切りの精度は97.9%、再現率は90.9%であった。また、発話単位への分割は常に正しいわけではないので、マウス操作によって発話単位の分割やマージを簡単にできるインターフェースを実装した。

2.3 対話行為タグの候補の提示

本項では、作業者に対話行為タグの候補を提示する手法について述べる。まず、少量のタグ付きコーパス(訓練コーパス)が存在すると仮定し、発話の対話行為を自動的に決定する決定木を学習する。次に、対話行為を決めたい発話 u' を決定木に入力し、 u' が到達する決定木の葉 l を求める。訓練コーパスにおける発話のうち、 l に分類される発話の集合を U_l とする。そして、 U_l の各発話に付与されている対話行為の集合 D_l を作業者に提示する対話行為の候補とする。また、 D_l における対話行為の確率分布 $P(d|D_l)$ を求め、これを対話行為 d の信頼度とし、この順に対話行為を並べて作業者に提示する。また、支援ツールでは、決定木による分類が正しく

ない場合も考えられるので、 D_l 以外の対話行為も信頼度 0 として対話行為候補のリストの最下位に提示した。

2.3.1 決定木の学習

決定木の学習には C4.5 [1] を用いた。学習に用いた素性は以下の通りである。

- 前/後の発話との話者交替
発話の話者が前/後の発話の話者と同じか否か。
- 話者の役割
対話行為を決める発話の話者がシステム側かユーザ側か。本研究で対象とする対話はタスク指向型の対話であり、一人が情報を提供するシステムの役割を、もう一人は情報を受け取るユーザの役割を果たす。
- 前/後の発話の対話行為
直前または直後の発話の対話行為を素性の値とする。但し、直前または直後の対話行為が「フィラー」または「あいづち」である場合、「フィラー」または「あいづち」以外の対話行為が出現するまで前または後の発話を辿る。これは、「フィラー」や「あいづち」が比較的どの場所にも出現し、対話行為を決める手がかりとはならないためである。
- 文末の単語
発話を茶筌¹によって形態素解析し、句点や「？」以外で最も末尾に出現する単語を素性の値とする。また、名詞については、単語そのものを素性とする。過学習を起こす恐れがあると考え、「一般名詞」「数名詞」「固有名詞」のいずれかに抽象化する。
- 特定の対話行為の発話にだけ出現するキーワード
特定の対話にのみ出現するキーワードのリストをあらかじめ用意し、そのキーワードを含むか否かを素性の値とする。キーワードは、訓練データにおける出現頻度がある閾値 O_t 以上で、かつ式 (1) のエントロピーが閾値 O_h 未満である単語とする [2]。

$$H = - \sum P(d|w) \log_2 P(d|w) \quad (1)$$

$P(d|w)$ は単語 w を含む発話の対話行為が d である確率である。本研究では $O_t = 3$ 、 $O_h = 0.5$ とした。

- 前の発話と同じ名詞が含まれているか
相手の言ったことを確認する [確認] という対話行為を識別するための素性である。確認をする際には前の発話の単語を繰り返す可能性が高く、この素性が有効に働くと考えられる。
- 疑問語が含まれているか

相手に対して質問を行う [未知情報要求] という対話行為を識別するための素性である。

2.3.2 決定木の使い分け

本研究では、作業員に対話の先頭の発話から順番に対話行為タグを付与することを強制するのではなく、任意の順序で対話行為を付与することを許している。このとき、決定木の素性として前後の発話の対話行為を用いることは問題となる。なぜなら、ある発話の対話行為を決める際、前後の対話行為が常にタグ付けされているとは限らないからである。前後の対話行為が未決定のときはそれらを決定木の素性として用いることはできない。そこで、 $T_{0,0}$ 、 $T_{0,1}$ 、 $T_{1,0}$ 、 $T_{1,1}$ の 4 つの決定木を学習することにした。ここで、 $T_{i,j}$ の i は前の発話の対話行為を素性として用いる (1) か用いない (0) かを、 j は後の発話の対話行為を素性として用いる (1) か用いない (0) かを表わす。実装したツールでは、状況に合わせて 4 種類の決定木を使い分け、対話行為の候補を提示する。

2.4 類似発話の提示

作業員による対話行為の選択を助けるため、過去のタグ付け事例をユーザに提示する。ここでのタグ付け事例とは、対話行為を決めたい発話に類似した発話と、それに付与された対話行為タグを指す。例として、「じゃあ、相模大野まで来ていただいて」という発話に対して提示される類似発話を表 1 に示す。類似発話は付与されている対話行為タグ毎に分類する。表 1 の [] は発話に付与された対話行為を表わす。また、前後の文脈も対話行為を決めるための重要な手がかりになるので、前後の発話も同時に表示する。表 1 で * のついている発話が類似発話、それ以外は前後の発話である。さらに、発話は類似度の高い順に並べて提示する。表 1 の対話は交通経路案内コーパスからの抜粋だが、このコーパスでは、作業員が対話行為タグ付けを行う際、対話行為が [示唆] か [情報伝達] かで迷うことが多い。しかし、表 1 のように類似発話を示せば、元の発話と似た発話「ではそちらの駅の方へ、来ていただいて、」に [示唆] の対話行為がついていることから、元の発話の対話行為も [示唆] であると判断できる。

提示する類似発話を決定する手法は以下の通りである。

1. 類似発話を提示する対話行為の選択

全ての対話行為タグについて類似発話を提示すると、提示する発話の量が多くなり、作業員にとってわかりづらい。そこで、作業員が判断に迷いそうな対話行為タグをいくつか選択し、それらの対話行為に対する発話を提示する。本研究では、2.3 項の決定木

¹<http://chasen.aist-nara.ac.jp/hiki/ChaSen/>

表 1: 類似発話の提示

[示唆]
S: あ,
S*: ではそちらの駅の方へ、来ていただいて、
U: はい
U: えーと、
S*: 折れていただいて、
U: はい
[情報伝達]
S: で、武蔵野住宅前だとー、バス停降りてー、
S*: しばらく、歩いて行って、
U: はい
U: はい.
S*: の実験で来ましたというふうに言えば、通してくれますので、
U: はい

によって得られる対話行為タグの候補について類似発話を提示する。学習した決定木は、表層表現や前後の文脈などが似ている発話を自然に分類できると考えたからである。また、2.3.2 で述べたように、状況に応じて 4 種類の決定木を使い分ける。

2. 提示する発話の決定

1. で決めた対話行為について、類似度の大きい N 個の発話を提示する。二発話間の類似度は以下のように定義する。

(a) 決定木における共通上位ノードの深さ

決定木において近い位置に分類される発話に高い類似度を与える。

(b) 単語ベクトル間の類似度

発話に含まれる単語に対して、式 (2) のように単語ベクトルを作成する。

$$\vec{v} = (x_1, x_2, \dots, x_n) \quad (2)$$

x_i は単語 w_i の重みであり、 w_i が助詞、助動詞などの付属語のときは 1、それ以外は 2 とする。発話間の類似度は式 (2) の単語ベクトルのコサイン類似度とする。

発話の順位付けは、まず類似度 (a) にしたがって行う。(a) の類似度が同じときは類似度 (b) にしたがう。(a),(b) の類似度がともに同じときは、ランダムに順位付けする。そして、対話行為毎に上位 N 個の類似発話を選択する。

3 評価実験

本節では提案手法の評価実験について述べる。以下、3.1 項では対話行為タグを自動的に決定する決定木 (2.3 項) の評価を行う。3.2 項では、類似発話の提示 (2.4 項) がどれだけ作業者に有効であるかを評価する。

表 2: 決定木の評価

決定木	上位一位	全て
$T_{0,0}$	68.1%	83.5%
$T_{0,1}$	68.1%	83.9%
$T_{1,0}$	74.6%	86.0%
$T_{1,1}$	72.8%	85.3%
ベースライン	24.0%	

3.1 決定木の評価

まず、2.3.1 で述べた素性を用いて、対話行為を決定する決定木を学習した。訓練データとして、正しい対話行為タグが付与された 1,224 発話からなる対話コーパスを用いた。2.3.2 で述べたように、前後の発話の対話行為を素性として用いるか否かによって 4 種類の決定木を学習した。次に、訓練データとは異なる 279 発話からなるコーパスをテストデータとし、決定木が出力する対話行為タグの評価を行った。結果を表 2 に示す。表 2 において、「上位一位」は決定木が出力する上位一位の対話行為が正解である割合を、「全て」は決定木が出力する全ての対話行為の候補の中に正解が含まれる割合を表わす。「ベースライン」は訓練データにおける最頻出対話行為 ([肯定・受諾]) を常に選択したときの正解率である。なお、決定木が出力する対話行為の候補数の平均は 3.05 であった。

表 2 から、4 種類の決定木の中では $T_{1,0}$ の正解率が一番高いことがわかる。 $T_{1,0}$ の上位一位の対話行為の正解率は 74.6% である。これは、例えば Reithinger らの手法 [2] による 74.7% に匹敵する値である。対話行為タグの種類、訓練データ量、言語などの実験条件が異なるため単純に比較することはできないが、学習した決定木は現在の機械学習手法によって得られる高い水準に達しているといえよう。また、前の発話の対話行為を素性として使う決定木と使わない決定木では大きな差があるが、後の発話の対話行為については正解率の差はほとんどない。したがって、今回の実験では、後の発話の対話行為は前の発話の対話行為ほど有効な素性ではなかったといえる。

3.2 類似発話の評価

3.1 項で決定木の評価に用いた 279 発話について、2.4 項の手法にしたがって類似発話を提示し、それが正しい対話行為の判定の参考になるかどうかを調べた。提示すべき過去のタグ付け事例として、決定木の学習に用いた 1,224 発話を用いた。決定木は $T_{1,0}$ を用いた。結果を

表 3: 類似発話の評価

	発話数
(A) 類似発話なし	5 (1.8%)
(B) 正解の対話行為なし	36 (12.9%)
(C) 参考になる発話がない	10 (3.6%)
(D) 参考になる発話がある	228 (81.7%)

表 3 に示す。

表 3 の (A) は類似する発話が 1 つも見つからない場合を表わす。また、(B) は提示した類似発話の中に正解の対話行為が存在しない場合を表わす。言い換えれば、ある発話について決定木を辿ったとき、(A) は到達した葉の中に分類される過去のタグ付け事例が 1 つもないことを、(B) は過去のタグ付け事例の中に正解の対話行為を持つ発話がないことを示している。いずれの場合も、正解となる対話行為に関して類似発話を表示できていないので、作業者にとって有効ではない。一方、表 3 の (C) は、正解となる対話行為について類似発話を提示しているが、その対話行為が正しいかどうかの参考にならない場合を表わす。これは、提示した類似発話が元の発話とあまり似ていないため、そのタグ付け結果を参考にしてよいかかわからない場合である。この割合を低くするためには、まずタグ付け事例の量を増やすことが必要である。参照する事例の数が増えれば増えるほど、類似した発話が見つかる可能性は高くなる。また、発話間の類似度の定義を改良することも必要であろう。例えば、式 (2) の単語ベクトルを作成する際、文末の単語の重みを重くしたり、前後の対話行為をベクトルに加えたりすることによって、類似する発話の順位が上位になる可能性がある。

表 3 の (D) は対話行為タグ付けの参考となる発話が提示できた場合を表わす。約 8 割の発話が (D) に該当することから、本研究で提案する類似発話の提示が対話行為タグ付け作業の有効な支援になりうるといえる。

4 関連研究

対話行為タグ付きコーパスを作成する手法のひとつは、発話の対話行為を自動的に決定するシステムを構築することである。近年では、教師あり機械学習の手法を用いて対話行為を自動的に決定する研究がいくつか行われている [2, 3]。Reithinger らは、発話に含まれる単語の n-gram や前の発話の対話行為などを用いて、対話行為を決定する確率モデルを学習した [2]。彼らのモデルによる対話行為タグ付けの正解率は 74.7%であった。一方、Samuel らは、対話行為の決定に Transformation-

Based Learning (TBL) という機械学習手法を用いた [3]。Reithinger らと同じ対話コーパスを用いて実験を行い、彼らよりも少ない訓練データで 71.2%の正解率が得られたと報告している。しかし、これらの手法の精度は自動的に対話行為タグ付けコーパスを構築できるほど高くはない。したがって、本研究のように人間の対話行為のタグ付け作業を支援する技術もまた重要である。

対話行為タグ付けを支援することを目的とした研究として駒谷らの研究がある [6]。彼らもまた、本研究と同じように、対話行為タグ付けを支援する GUI ツールを作成している。駒谷らのツールでは、作業者に提示する対話行為の候補はルールベースの手法によって決定している。これに対し、本研究では機械学習の手法を用いている。機械学習を用いる利点としては、対話行為推定の精度が高いこと、作業が進むにつれて利用可能な訓練データの数が増えることなどが挙げられる。訓練コーパスとして正しい対話行為がタグ付けされたコーパスが必要だが、コーパス作成作業の最初の段階で少量のタグ付きコーパスを用意することはそれほど大変ではない。また、類似発話を提示する機能を実装している点も本研究の特徴のひとつである。

5 おわりに

本研究では、対話行為タグ付けコーパスの作成を支援するツールについて述べた。また、支援ツールの機能のうち、対話行為を自動的に決定する決定木と類似発話を提示する手法の評価を行い、いずれも対話行為タグ付けの支援に有効であることがわかった。今後の課題として、決定木の正解率を向上させることや、音声情報の再生機能を実装することなどが挙げられる。

参考文献

- [1] J. Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, 1993.
- [2] Norbert Reithinger and Martin Klesen. Dialogue act classification using language models. In *Proceedings of the EuroSpeech*, pp. 2235–2238, 1997.
- [3] Ken Samuel, Sandra Carberry, and K. Vijay-shanker. Dialogue act tagging with transformation-based learning. In *COLING-ACL 98*, pp. 1150–1156, 1998.
- [4] 荒木雅弘, 伊藤敏彦, 熊谷智子, 石崎雅人. 発話単位タグ標準化案の作成. 人工知能学会誌, Vol. 14, No. 2, pp. 251–260, 1999.
- [5] 堂坂浩二, 島津明. タスク指向型対話における漸次的発話生成モデル. 情報処理学会論文誌, Vol. 37, No. 12, pp. 2190–2200, 1996.
- [6] 駒谷和範, 荒木雅弘, 堂下修司. 対話コーパスにおける発話単位タグの一推定手法. 人工知能学会誌, Vol. 14, No. 2, pp. 273–281, 1999.