

# 発話に対する拡張談話タグ付与

関野嵩浩 井上雅史 (山形大学)

本研究では、機械学習により自動構築された分類器を用いて2者間の対話における各発話に対して自動で談話タグを付与する手法を提案する。先行研究で用いられているタグセットでは対話の流れをおおまかにしか表すことができない。それに対して、本研究では細かな対話の流れが表現可能となるタグセットを用いた。タグ付与に用いる素性として発話文字数、内容語数、発話順番を追加した。実験の結果、タグ付与の一致度が最大33.3%と意味的に許容できるものを含めた一致度が最大57.3%であった。これにより、素性の追加によるタグ一致度の増加を確認できた。

## 1. はじめに

現在、談話タグが付与された対話コーパスを用いた研究が広く行われている。タグ付与された対話コーパスは、様々な目的に利用できる。例えば、発話に付与したタグの系列は、音声認識の認識候補を絞り込んだり、音声対話システムの誤り修復に利用したりできる。また、働きかけや応答のような局所的な構造を調べることで、対面非対面のコミュニケーションの特徴を明らかにしたり、局所構造や局所構造間の関係が変化したときのコミュニケーションの効率や課題達成率の変化を明らかにしたりできる。磯村ら[1]は、発話毎にタグを付与し、タグの系列を対話の流れとして人間同士の対話とどれだけ近いかで対話システムを評価する手法を提案している。磯村らは課題としてタグ付与方法の改善を挙げており、大規模コーパスにおけるタグ付与を機械的に行うことが望まれている。また、複数のタスク指向型対話コーパスを分析し、タスクに依存しない発話を抽出することでタスクに共通な発話の機能を考察する研究が行われている。しかし、これらの研究で用いられているコーパスは人手でタグ付けされたものであり、問題としてタグ付与の際の労力が挙げられている。

こういった問題の解決策としてタグを自動的に付与する研究が行われている。駒谷らは、発話者間の「働きかけ」→「応答」という「やりとり単位」を推定し、「働きかけ」部に対してキーワードマッチングでタグを決定し、「働きかけ」部のタグに対応して「応答」部のタグを決定するという手法を用いている。しかし、この手法では、表現の変化に対して柔軟性に乏しく、正しい日本語で書かれたテキスト対話以外には有効に作用しない。磯村らの手法は 1-gram, 2-gram, 前の発話のタグを素性として CRF による機械学習で発話に対するタグを決定する [2]。しかし、この研究で用いられたタグセットは SWBD-DAMSL タグをカテゴリ毎にまとめた簡易

DAMSL タグである。そのため、表1のように対話の大まかな流れは表現できるが、細かな範囲で対話の流れを表現することができない。

本研究では、機械学習により自動構築された分類器を用いて 2 者間の対話における 各発話に対して自動で談話タグを付与する手法を提案する。2 者間の対話を対象とした理由は、3 者以上の対話の場合、タグ付与を行う発話の前後に関係性が少ない場合が多いためである。また、簡易 DAMSL タグではなくオリジナルの SWBD-DAMSL タグ(表 2)を用いることで細かな対話の流れを表現することを目指した [3]。本研究では後述の先行研究で利用されている CRF(Conditional Random Fields)の素性を、1-gram, 2-gram, 前の発話のタグに加え、新たに内容語の数、発話文字数を加えることでタグセット 拡張によるタグ付与の正解率が増加した。以降では第 2 章で研究で用いた、対話コーパス、タグセット、タグ付与に用いた手法について述べる。第 3 章では、本手法と先行研究で行われた手法の実験方法と結果について述べ、第 4 章でその考察について示す。そして第 5 章で、本研究の総括を述べる。

表 1: タグ付与例

話者	発話	簡易 DAMSL タグ	SWBD-DAMSL タグ
L	あそこのたこ焼きはおいしいですよ	Statement	Statement-opinion
R	本当ですか?	Question	Backchannel in question form
L	自家製のソースを使っています	Statement	Statement-opinion
R	今度行ってみます	Statement	Statement-opinion

表 2:SWBD-DAMSL タグ

SWBD-DAMSLタグ	内容	短縮タグ
3rd-party-talk	第三者への発話	t3
Self-talk	独り言	t1
Uninterpretable	解釈不能	%
Task-Management	タスク管理	^t
Communication-Management	コミュニケーション管理	^c
Statement-Non-opinion	意見のない発話	sd
Statement-opinion	意見のある発話	sv
Open-option	開けた付属	oo
Action-directive	行動命令	ad
Yes-No-Question	yn疑問文	qy
Declarative Yes-NoQuestion	yn平叙疑問文	qy^d
Wh-Question	Wh疑問文	qw
Declarative Wh-Question	Wh平叙疑問文	qw^d
Open-Questions	最初の質問	qo
"or"-questions	Or疑問文	qr
Or-Clause	複数句	qrr
declarative question	平叙疑問文	^d
Tag-Question	付加疑問	^g
Rhetorical-Questions	修辞疑問文	qh
Offers	提供	co
Commit	委託	cc
Openings	開始	fp
Closing	終了	fc
Thanking	感謝	ft
Welcome	歓迎	fw
Exclamation	叫び	fe
eXplicit performative	遂行的	fx
Other-forward-function	その他の前進機能	fo
Apology	謝罪	fa
Agree/Accept	同意	aa
Maybe/Accept-part	許可	aap/am
Reject	拒否	ar
Reject-part	拒否	arp
Signal-Non-understanding	理解していないシグナル	br
Acknowledge	事実を認める	b
Backchannel in question form	質問によるあいづち	bh
Response Acknowledgement	謝辞応答	bk
Repeat-phrase	繰り返し	b^m
Summarize/reformulate	要約	bf
Collaborative Completion	共同補完	^2
Appreciation	賞賛	ba
Downplayer	軽視した発言	bd
sympathy	同情	by
Correct-misspeaking-by-other-speaker	他話者による訂正	bc
affirmative answers	肯定的回答	ny
Affirmative Non-yes answers	肯定的回答	na,ny^e
negative answers	否定	nn
Negative Non-no answers	ネガティブな肯定	ng,nn^e
Hold before Answer/agreement	回答、契約の前の待機	^h
Other answers	その他の回答	no
Quotation	引用	^q
Hedge	言い逃れ	h

## 2. データと手法

### 2.1. 対話コーパス

現在、対話コーパスは様々な研究で利用されている。森川らの研究では、予約案内やテレフォンショッピングなど複数の対話コーパスに対して分析を行い、それぞれ

のタスクに依存しない発話を抽出した。また、対話コーパスの構築も広く行われており、心的状態を表す対話行為タグを付与したテキスト対話コーパスの構築や人間同士のマルチモーダルな対話コーパスの構築等がある。

本研究では様々な2者間での対話の流れを表現するため、対話の流れが対話毎に変化する非タスク指向型対話を対象とするため日本語話し言葉コーパス(Corpus of Spontaneous Japanese:CSJ)を使用した。CSJには学会講演や、インタビュー等様々なカテゴリのデータが存在する。その中から自由対話の転記テキストを使用した。

CSJでは対話をスラッシュ単位で転記している。しかし、本研究では、発話毎にタグを付与するためターン構成単位(turn constructional unit: TCU)に置換した。これにより、2者間の対話の流れを発話単位で表せるようになった。置換はTCU認定に従って行った[4]。図1はCSJの書き下し文である。CSJでは発話番号、発話時間、話者の情報の後に発話のスラッシュ単位で記述されている。図1の書き下し文をTCUに置換したものが図2である。

```
0021 00027.160-00030.962 R:
多分                & タブン
さっきのでも        & サッキノデモ
そういう            & (W ソ;ソー)(W ユ;ユウ)
ことが              & コトガ
あったんじゃないかというような          & アツタンジャンイカト(笑 ユー(? ヨー)ナ
感じです            & カンジデス)<笑>
```

図 1:CSJ の一部

```
R:多分さっきのでもそういうことがあったんじゃないかというような感じです
```

図 2:CSJ の置換

### 2.2. タグセット

談話タグは研究目的が異なれば必然的に異なる。そのため談話タグを構築する研究は広く行われている。京都案内対話コーパスで用いられているタグは、発話がどのような内容について言及しているかという意味内容を表す意味内容タグ、どのような意図で発話されているかという情報伝達機能を表す発話行為タグである。また、荒木らはデータ流通の観点から標準化を目的としたタグセットを設計した。これはやりとりを大きな分類とし、その下位分類として発語内行為を位置づけたタグセットである。

本研究では談話のつながりを対話の自然さと定義する。談話のつながりを表現するため、対話の浅い構造、すなわち発話間の隣接を近似できるという特徴があるSWBD-DAMSL タグを用いた。この SWBD-DAMSL タグは、話者の目的、発話の焦点や主題といった対話の

深い構造は扱わない。本研究とほぼ同様の手法を用いる磯村らの研究では最低限の対話の自然さを比較評価できるように SWBD-DAMSL タグをカテゴリ毎にまとめた簡易 DAMSL タグを用いている。しかし、簡易 DAMSL タグを用いると表1のように発話に対して大まかなタグしか付与できず、対話の流れが曖昧な形しか表現できない場合が存在する。そこで、本研究ではより自然な対話の流れを表現するため簡略化されていないオリジナルの SWBD-DAMSL タグを用いる。今回実験に用いる対話コーパス中に出現する SWBD-DAMSL タグ毎の頻度を図3に示す。図3では横軸が SWBD-DAMSL タグ、縦軸がタグの出現頻度を表している。また、出現頻度 10 未満のタグについてはその他にまとめた。

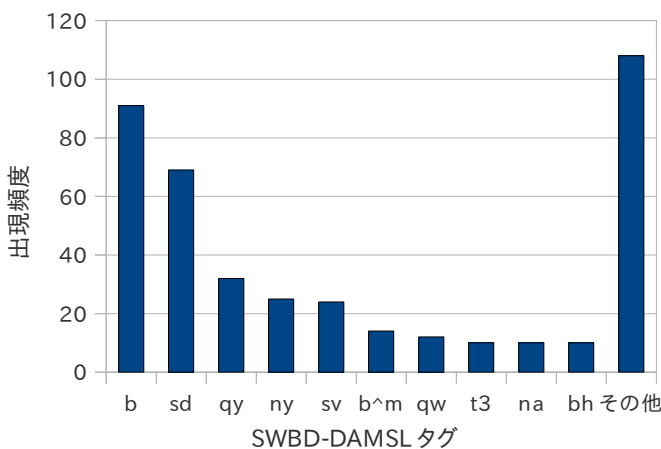


図3: コーパス中の SWBD-DAMSL タグ出現頻度

### 2.3. タグ自動付与手法

本研究では自動談話タグ付与に CRF(Conditional Random Fields)を用いた。CRFは特徴の独立性の過程が必要なく、単語よりも細かいレベルでの特徴設計ができ、このタスクで精度が高いことが知られている。系列ラベリング問題を解くのに適した識別モデルであり、入力発話系列  $x$ (表1発話列)、タグ系列  $y$ (表1タグ列)の対応関係を条件付き確率  $P(x|y)$ で表現する。自動タグ付与では、 $x$ と $y$ は同じ長さであり、例えば長さ  $n$ の入力発話系列では、

$$X_n = X_1 X_2 X_3 \cdots X_n, Y_n = Y_1 Y_2 Y_3 \cdots Y_n \text{ である。}$$

$X_i$  は  $i$  番目の発話を表し、 $Y_i$  はその発話に付与する談話タグを示す。CRFでは素性を素性関数  $f_x$ 、それに対応する重みパラメータを  $\lambda$  で表し、 $k$  番目の素性に対する素性関数を  $f_k$ 、重みパラメータを  $\lambda_k$  とする。このとき  $P(y|x)$  は次式で表される。

$$P(y|x) = \frac{1}{Z_x} \exp\left(\sum_{i=1}^n \sum_k \lambda_k f_k(x, y, i)\right)$$

ただし、 $Z_x$  は全系列の総和を1にするための正規化

項であり、

$$Z_x = \sum_y \exp\left(\sum_{i=1}^n \sum_k \lambda_k f_k(x, y, i)\right)$$

となる。パラメータ  $\lambda$  は、最尤推定法で求めることができる。素性関数は、0,1の2値を返す関数を用いられる。

以上を用いて、入力発話系列  $x$  に対して、最適な SWBD-DAMSL タグ系列  $\hat{y}$  は、

$$\hat{y} = \operatorname{argmax}_y P(y|x)$$

により求めることができる。

### 3. 評価実験

実験は手動タグ付与された2対話 405 発話の中から、ランダムで連続する5発話を抜きだしてテストデータとし、残りを学習データとして用いた。データによって偏りが生じないようにこれを素性の組み合わせ毎に15回繰り返す。その結果から人手で付与したタグと学習により付与されたタグが一致している一致度と意味的に許容できるものも含めた一致度を求めた。意味的に許容できるものとは、例えば、質問によるあいづち bh と yn 疑問文 qy など人手で付与されたタグと異なっても会話の流れが自然なものを表す。また、手動で付けられているタグに対して学習によりどのタグが付与されたかを求めた。

#### 3.1. 予備調査

予備調査として先行研究と同様に学習データに2回以上出現した形態素の 1-gram, 2-gram, 前の発話のタグを素性として CRF によるタグ付与を簡易 DAMSL タグと SWBD-DAMSL タグについて行った。表3は 1-gram, 2-gram, 前の発話のタグを素性として CRF によるタグ付与を行ったタグの一致度と意味的に許容できるものも含めた一致度である。

表3: 1-gram, 2-gram, 前の発話のタグを用いた一致度

一致度	意味的に許容できるものも含めた一致度	一致したタグ種類数
25.3%	45.3%	2

予備調査の結果、正しく付与されたタグはあいづち b と意味のない発話 sd のみであった。あいづち b と意味のない発話 sd は図1を見るとコーパス中で最も出現頻度の高いタグであることがわかる。このことから、タグ1つあたりの 1-gram, 2-gram のデータが少ないために他のタグが一致しなかったと考えられる。

そこで本研究では発話の内容に依存しない発話文字数、内容語数、発話順番を素性として加えることで一致度の向上を図った。

### 3.2. 素性の有効度調査

素性の組み合わせは以下の通りである.1-gram,2-gram,前の発話のタグをまとめて旧素性とする.

- ① 旧素性+発話文字数
- ② 旧素性+内容語数
- ③ 旧素性+発話順番
- ④ 旧素性+発話文字数+内容語数
- ⑤ 旧素性+発話文字数+発話順番
- ⑥ 旧素性+内容語数+発話順番
- ⑦ 旧素性+発話文字数+内容語数+発話順番

実験結果を以下に示す.表6は①から⑦を素性としてCRFによるタグ付与を行ったタグの一致度と意味的に許容できるものも含めた一致度である.

表 4:追加素性とタグ一致度

素性	一致度	意味的に許容できるものも含めた一致度	一致したタグ種類数
①	32.0%	50.7%	5
②	30.7%	57.3%	4
③	32.0%	50.7%	5
④	28.0%	53.3%	3
⑤	26.7%	50.7%	2
⑥	33.3%	56.0%	5
⑦	28.0%	53.3%	3

### 4. 考察

表4を見ると,素性を追加したことで一致度が全体的に向上していることが分かる.また,最も一致度が高い素性が⑥旧素性+内容語数+発話順番で33.3%,意味的に許容できるものも含めた一致度が最も高いのが②旧素性+内容語数で57.3%となった.

⑥に対して④,⑤の一致度が低いことから発話文字数は内容語数と発話順番に関係が薄いことが分かる.逆に⑥は②,③に対して一致度が上昇していることから内容語数と発話順番に何らかの関係があることがわかる.また,意味的に許容できるものも含めた一致度は⑥より②の方が高くなっているが,正しく付与できているタグの種類は⑥の方が多いため,学習データを増やしていくと,⑥の方が様々なタグに対応できるものと考えられる.

全体として,タグの一致度が3割弱,意味的に許容できるものも含めた一致度でも5割強程度で今回用いられたタグ付与手法は実用的でない.これは,それぞれのタグに対する学習データが少ないためだと考えられる.今後の課題として,学習データを増加しタグ1つ当たりの出現頻度を増やし,タグ付与一致度を向上させることが挙げられる.

### 5. おわりに

これまで,対話の流れを表すために簡易 DAMSL タグを用いた自動タグ付与が行われていた.しかし,簡易 DAMSL タグでは対話の詳細な流れを表すことが困難であった.

そこで,本研究では簡易 DAMSL タグを拡張したオリジナルの SWBD-DAMSL タグを用いて自動タグ付与を行った.簡易 DAMSL タグによる先行研究で用いられていた素性は学習データ中に2回以上出現した1-gram,2-gram,前の発話のタグであった.この素性を用いてオリジナルの SWBD-DAMSL タグでタグ付与を行ったところ,一致度は25%,意味的に許容できるものも含めた一致度は45%であった.

これに対して,発話文字数,内容語数,発話順番,これらを組み合わせたものを素性として追加したところ,一致度は内容語と発話順番を素性として追加したもので33.3%,意味的に許容できるものも含めた一致度は内容語を素性として追加したもので57.3%となった.本研究の結果から素性に内容語または内容後と発話順番を追加したことによりタグ自動付与一致度が向上することが分かった.

タグ付与が正しく行われたもののほとんどが学習データ中の出現頻度が高いものに偏った.このことから今後の課題として,学習データを増加し,学習データ中のタグの出現頻度を増加させて今回正しく付与されなかったタグの一致度の向上を図る.また,増加したデータ中からそれぞれのタグと関連性の深い素性を調査し,タグ付与の一致度の向上を図る.

### 参考文献

- [1] 磯村直樹, 鳥海不二夫, 石井健一郎: HMMによる非タスク指向型対話システムの性能の比較評価(言語理解のためのコーパスからの知識獲得), 電子情報通信学会技術研究報告. NLC, 言語理解とコミュニケーション, vol.107, no.246, pp. 1-6 (2007).
- [2] 磯村直樹, 鳥海不二夫, 石井健一郎: 対話エージェント評価におけるタグ付与の自動化(エージェントデザイン, <特集>人とエージェントのインタラクション論文), 電子情報通信学会論文誌. A, 基礎・境界, vol. 92, no. 11, pp. 795-805, (2009).
- [3] D. Jurafsky, E. Shriberg, and D. Biasca: Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13, *University of Colorado, Institute of Cognitive Science, Tech. Rep*, pp. 97-101, (1997).
- [4] 榎本美香: 会話・対話・談話研究のための分析単位: ターン構成単位(TCU)(<連載チュートリアル>多人数インタラクションの分析手法[第4回]), 人工知能学会誌, vol. 23, no. 2, pp. 265-270, (2008).