

雑談型対話システムにおける自己開示発話の割合増加手法の提案

Proposal of a method for increasing the proportion of self-disclosure responses in chat type dialogue system

小山翔平<sup>†</sup> 土屋 誠司<sup>‡</sup> 渡部 広一<sup>‡</sup>  
 Shohei Koyama Seiji Tsuchiya Hirokazu Watabe

1. はじめに

ロボットと人間が円滑な意思疎通を行うための手段として自然言語による対話がある。その中で雑談に着目した際、自分の考えや嗜好を相手に伝えることは大切である。そこで、システムが対話内に出てきた事柄について主観を含んだ表現を行う自己開示応答を取り入れた雑談型対話システム<sup>[1]</sup>が提案されている。しかし、自己開示応答が十分に行えておらず、人間の対話における実際の自己開示発話の割合とは乖離しており、不自然さを感じる原因となっている。そこで本研究では、ある事柄からそれと関連のあるものを思い浮かべる連想を用いて自己開示応答を行うことで、人間の発話の割合に近づけた対話システムを提案した。

2. 関連技術

2.1 Mecab

Mecab<sup>[2]</sup>とは、入力された文に対して形態素解析を行うシステムである。形態素解析は文を言語として意味を持つ最小単位である形態素に分割し、またそれぞれの形態素の品詞を判別するものである。対話システムにおいてシステムからの質問に対するユーザの応答を形態素解析する際に用いる。表 1 に Mecab の実行例として「学校に行く。」を入力した場合を示す。

表 1 Mecab の実行例

形態素	解析結果
学校	名詞, 一般, *, *, *, *, 学校, ガッコウ, ガッコー
に	助詞, 格助詞, 一般, *, *, *, に, ニ, ニ
行く	動詞, 自立, *, *, 五段・カ行促音便, 基本形, 行く, イク, イク
.	記号, 句点, *, *, *, *, ., ., ., .

2.2 意味理解システム

意味理解システム<sup>[3]</sup>は、入力文の情報を 7W1H (誰が, 誰と, 何を, いつ, どこで, なぜ, 誰に, どのように) と述語のフレームに分割して格納するシステムである。意味理解システムに「昨日, 私は水族館でイルカを観た。」という文を入力した例の結果を表 2 に示す。

表 2 意味理解システム格納例

Who	Who+	What	When	Where	Whom	How	Why	述語
私		イルカ	昨日	水族館				観た

2.3 概念ベース

概念ベース<sup>[4]</sup>とは、様々な語 (概念) が、それを特徴付ける語 (属性) とその重要度を表す数値 (重み) の対の集合によって定義されている知識ベースである。概念には属性と重みが付与されており、約 9 万語の概念が収録されている。概念ベースの例を表 3 に示す。

表 3 概念ベースの例

概念	属性, 重み
医者	(医師, 0.34)(患者, 0.11)(病院, 0.08) . . .
治す	(治療, 0.43)(医療, 0.21)(病气, 0.13) . . .
⋮	⋮

2.4 NTT シソーラス

NTT シソーラス<sup>[5]</sup>とは日本語語彙体系から作成された同義語, 類義語を整理した約 2700 個の意味属性の上位下位関係, 全体部分関係を木構造で示すものである。シソーラスには約 13 万語が登録されている。

3. 既存システム

雑談型対話システムは、話者の発言に対して適切な応答手法を選択し、話題について会話を展開することで、人間との自然な雑談型対話を目指したシステムである。既存システムの全体フローを図 1 に示す。

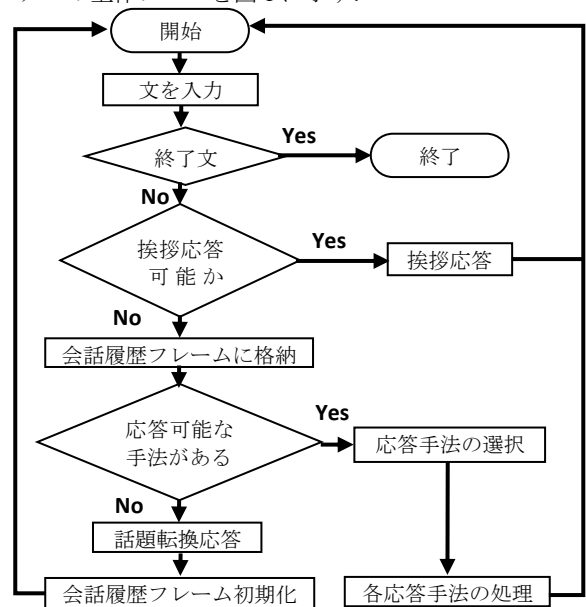


図 1 既存システムの全体フロー

話者の入力終了文であるか、また、システムが挨拶応答が可能かどうか判断した後、入力文を意味理解システムによって会話履歴フレームに格納する。会話履歴フレームの状態から、話題転換応答以外の各応答手法が応答可能かどうか確認する。このとき、応答可能な手法が 1 つもなければ話題転換応答が行われ、会話履歴フレームを初期化した後に、開始時の状態に戻る。応答可能な応答手法があれば、応答手法をあらかじめ設定した確率により選択し、応答処理を行う。発話の種類毎の設定確率を表 4 に示す。

表 4 発話の種類毎の設定確率

発話の種類	設定確率
自己開示	57.5%
質問	29.7%

表 4 の設定確率は、人間の会話をテキストに起こした名大会話コーパス<sup>[6]</sup>を用いて、人間の発話を分類し、その割

<sup>†</sup> 同志社大学大学院理工学研究科  
 Graduate School of Science and Engineering, Doshisha University  
<sup>‡</sup> 同志社大学理工学部  
 Faculty of Science and Engineering, Doshisha University

話を用いている。表 5 に応答手法の一覧を示す。

表 5 雑談型対話システムの応答手法一覧

応答手法	発話種類	応答文の例
挨拶	挨拶	おはようございます。
未登録語	質問	ポケモンとは何ですか？
7WIH	質問	どこで食べましたか？
掘り下げ	質問	友達の誰さんですか？
場所判断連想	質問	海に泳ぎに行ったのですか？
自己開示応答	自己開示	犬は可愛くて好きです
話題転換応答	質問	明日は何をしますか？

### 3.1 自己開示応答

自己開示応答の生成条件は、話者の発話に含まれる名詞から話者主観知識ベースを用いて感覚知覚語を取得することである。感覚知覚語とは五感によって感じとった内容を表す形容詞、形容動詞と五感以外で感じる印象を表す形容詞、形容動詞をあわせたものと定義する。また、話者主観知識ベースとは修飾語と名詞と感覚知覚語の関係を表したものである。話者主観知識ベースには約 1300 語の名詞と 186 種の感覚知覚語が格納されている。自己開示応答の例としては、話者の入力が「今日、犬を観ました」である場合、「犬」に対する感覚知覚語「可愛い」が取得できるため「犬は可愛くて好きです」という応答文が生成される。

### 4. 既存システムの問題点

既存システムの問題点を検討するために、システムと人間の発話を種類ごとに分類し、比較したものを表 6 に示す。

表 6 人間と既存システムの発話比較

発話者	自己開示発話の割合	質問発話の割合
人間	57.5%	29.7%
システム	8.0%	86.7%

表 6 からシステムは、人間に比べると自己開示発話の割合が非常に少なく、質問発話が多いことが分かる。これは、感覚知覚語を取得できる語が少ないため、自己開示発話の割合が低いと考えられる。

### 5. 提案システム

提案システムでは話者主観知識ベースに加え、感覚判断システム<sup>[7]</sup>を用いて対話中の名詞から感覚知覚語を取得し、自己開示発話を行った。感覚判断システムとは入力した名詞に対して、該当する感覚知覚語を出力するシステムである。感覚判断システムは入力した名詞のシソーラスにおける親ノードや概念ベースにおける属性の感覚知覚語を取得することができる。例えば、「温泉」という入力語から、関連が強いと考えられる「入浴」という語を連想し、これに対する感覚知覚語「快い」を出力する。既存システムでは話者主観知識ベースに格納されていない名詞に対しては感覚知覚語を取得できず、自己開示応答は行えない。しかし、提案システムでは話者主観知識ベースに格納されていない名詞に対して関連が強い語を連想し、自己開示応答を行い、自己開示発話の割合増加に繋がると考えた。

### 6. 評価

被験者 5 名はシステムとテキスト対話を行い、一つ一つの応答に対して、また対話全体の流れについて自然さの評価を行った。尚、対話 10 往復を 1 セットとし、被験者一人につき 3 セット行った。以下の表 7 にその結果を示す。

またシステムの発話の種類毎の割合についての結果を表 8 に示す。表 7,8 における各数値の単位は%である。

表 7 対話の自然さについての評価結果

システム	応答毎			対話全体		
	○	△	×	○	△	×
既存システム	46.0	29.3	24.7	13.3	80.0	6.7
提案システム	37.3	34.7	28.0	13.3	73.3	13.3

表 8 発話の割合についての評価結果

システム	発話の種類		
	質問	自己開示	その他
既存システム	86.7	8.0	5.3
提案システム	69.3	24.0	6.7

### 7. 考察

既存システムに対して提案システムは自己開示発話の割合が 16.0% 増え、質問発話の割合が 17.4% 減った結果となった。これは感覚判断システムを用いることで既存システムに比べて、多くの感覚知覚語を取得することに成功したためであると考えられる。しかし、感覚想起システムを用いる際に、話題となっている語に対して不自然さがある感覚知覚語を取得している例が目立った。例えば、「火曜日」に対して「危ない」といった不自然な感覚知覚語を取得してしまった。また、「どこ」や「いつ」などの指示代名詞に対して感覚知覚語を取得してしまうことがあった。そのため、感覚判断システムの改善が必要であると考えられる。

### 8. おわりに

本研究では、対話システム中の自己開示発話と質問発話を行う割合を人間に近づけることで、自然な雑談型の対話ができるシステムの構築を目指した。その結果、自己開示発話の割合が増加した。しかし、不自然な応答も多く、対話全体を見て自然だと評価される割合は 13.3% に留まっており、さらなる改善が必要である。今後の展望として、感覚判断システムの改善によって自然な自己開示発話の割合を増加させることや、感情判断などを実装することができれば発話のバリエーションが増え、更なる精度向上が期待できるのではないかと考えられる。

### 謝辞

本研究の一部は、JSPS 科研費 16K00311 の助成を受けて行ったものです。

### 参考文献

- [1] 北川智裕, 土屋誠司, 渡部広一, “自己開示発話を取り入れた雑談型対話システムの提案”, 第 113 回知識ベースシステム研究会, 13-18, 2018-03-01.
- [2] MeCab, <<http://taku910.github.io/mecab/>>, (参照 2020-02-04)
- [3] 小谷涼, 吉村枝里子, 土屋誠司, 渡部広一, “入れ子構造による文章の意味理解手法の提案”, 情報科学技術フォーラム FIT2016, pp.259-260, 2016.
- [4] 奥村紀之, 土屋誠司, 渡部広一, 河岡司, “概念間の関連度計算のための大規模概念ベースの構築”, 自然言語処理, Vol.14, No.5, pp.41-64, 2007.
- [5] NTT コミュニケーション科学研究所, “日本語語彙体系”, 岩波書店, 1997.
- [6] 藤村逸子・大曾美恵子・大島ディヴィッド義和, 2011, 「会話コーパスの構築によるコミュニケーション研究」, 藤村逸子, 滝沢直宏編『言語研究の技法: データの収集と分析』, pp. 43-72, ひつじ書房.
- [7] 米谷彩, 堀口敦史, 河岡司, “語の共起情報を考慮した感覚連想メカニズムに関する研究”, 情報処理学会研究報告, 2005-NL-166-(9), Vol.2005, No.22, pp.63-70, 2005 年 3 月.