

## 雑談型対話システムにおける話題内容を考慮した話題転換応答の提案 Proposal of Topic Transition Responses Considering Topic Content in Chat-Oriented Dialogue System

小柳津賢人<sup>†</sup>      土屋誠司<sup>‡</sup>      渡部広一<sup>‡</sup>  
Masato Oyaizu      Seiji Tsuchiya      Hirokazu Watabe

### 1. はじめに

近年、AI との対話がコミュニケーションの手段の一つとして挙げられ、福祉の場などの様々な場面での活躍が期待されている。その際に、AI の応答が人間らしいものであれば、人間との意思疎通をより自然に行うことができると考えられる。

人間同士の会話において話題が変わるとき、直前に会話していた内容に関連するものに話題が遷移することがある。しかし、既存システム<sup>[1]</sup>では話題転換を行う際の応答文がそれまでの会話の内容を全く考慮していない、また話題転換の際の応答文の数が少ないという問題点がある。そこで、本研究では話題転換の際に会話の内容を考慮する応答を行うことで人間らしい会話を目指す。

### 2. 関連技術

#### 2.1 概念ベース

概念ベース<sup>[2]</sup>は単語を概念とし、概念を属性と重みの対の集合として定義している。属性は概念を特徴付ける概念ベースに存在する概念であり、重みは概念に対する属性の重要度を表す。重みは正の値をとる。概念ベースの例を表 1 に示す。

表 1 概念ベースの例

概念	属性, 重み
医師	(患者, 0.63) (医師, 0.58) (医療, 0.50) ...
治す	(医术, 0.56) (為政者, 0.31) (治める, 0.17) ...

#### 2.2 NTT シソーラス

NTT シソーラス<sup>[3]</sup>とは一般名詞の意味属性についてその上位下位関係、全体部分関係を木構造で示したものである。約 30 万語が登録されている。

#### 2.3 意味理解システム

意味理解システム<sup>[4]</sup>は、入力文を単語が持つ情報ごとに、7W1H と述語のフレームに分割するシステムである。7W1H とは英語の疑問詞に用いられる 6W1H に「誰と」を表す Who+フレームを追加したものである。例として、「私は昨日、水族館でイルカを観た。」という文を入力した場合を以下の表 2 に示す。

表 2 意味理解システム格納例

Who	What	When	Where	述語
私	イルカ	昨日	水族館	観た

#### 2.4 会話履歴フレーム

会話履歴フレームは、意味理解システムと直前のシステムの発話を用いて、会話履歴を 7W1H と述語に分割して格納したものである。表 2 の情報とシステムの「誰と観ま

表 3 会話履歴フレームの格納例

Who	Who+	What	When	Where	述語
私	友達	イルカ	昨日	水族館	観た

したか?」という発話、話者の「友達とだよ」という発話を基にした会話履歴フレームの例を表 3 に示す。

### 2.5 IDF

IDF (Inverse Document Frequency) とは、ある単語が出現する文書数と全文書数から、その単語に着目した特徴を表す指標である。単語が含まれる文書数が多いほど IDF の値は小さくなる。

### 3. 既存システム

#### 3.1 システム概要

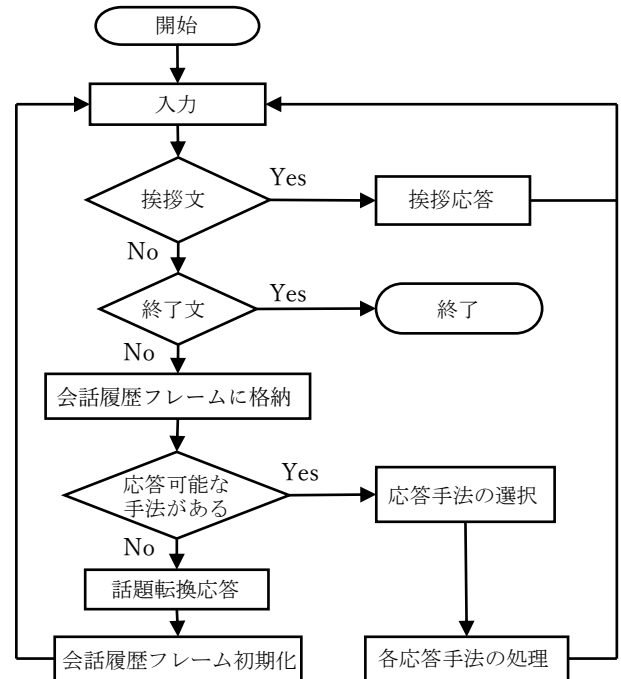


図 1 雑談型対話システムの流れ

既存の対話システムは、話者の発言に対して適切な応答手法を選択し、会話を展開することで、人間との自然なテキスト対話を目指したシステムである。雑談型対話システムによって展開される対話の流れを図 1 に示す。

システムを起動すると入力待機状態となり、そこで話者が文を入力する。まず入力が挨拶かどうか確認する。挨拶であった場合は挨拶応答を行い、最初の入力待機状態に戻る。挨拶でなければ入力が終了文（「今日もう寝る」）か

<sup>†</sup> 同志社大学大学院理工学研究科

<sup>‡</sup> 同志社大学理工学部インテリジェント情報工学科

どうか確認する。終了文であった場合はシステムを終了する。入力終了文でない場合は、入力文を、意味理解システムを用い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で応答可能かどうか判断する。応答可能な応答手法が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1つの場合はその応答手法を行う。複数あれば応答手法を重み付きランダムで選択し、その応答手法での処理が行われる。その後、入力待機状態に戻るという流れを繰り返し、対話を展開する。以下の表4に応答手法の一覧を示す。

表4 雑談型対話システムの応答手法一覧

応答手法	発話の種類	応答文の例
挨拶応答	挨拶	おはようございます。
未登録語応答	質問	インスタグラムとは何ですか？
感情判断応答	共感	よかったですね。
7WIH 応答	質問	どこで食べましたか？
掘り下げ応答	質問	友達の誰さんですか？
場所判断連想応答	質問	海に釣りに行ったのですか？
自己開示応答	自己開示	お菓子は甘いところが好きです。
話題連想応答	自己開示	SMAP といったら木村拓哉ですよ。
話題転換応答	質問	ところで明日は何をしますか？

### 3.2 話題転換応答

話題転換応答は、話題を変える応答である。7個の応答文候補からランダムに1つを選択し応答文を生成する。

## 4. 既存システムの問題点

既存システムの問題点として、話題転換応答の応答文の種類が少ないという点が挙げられる。応答文の種類が7種類しかないために、同じ応答文が連続で生成されるということが頻繁に起こってしまう。さらに、応答文の全てが普遍的な文であるため、user側に話を聞いていないのではないかと感じさせ、会話として不自然な印象を与えてしまう。

## 5. 提案手法

直前の会話履歴フレームのWhat格に格納されている語のシソーラス上での上位語と述語を用いて、「他にどんな{What格の上位語}+{助詞}+{述語}か?」という話題転換の応答文を生成する。具体的な応答例を図2に示す。

U: 昨日, 卓球をしました。  
S: 誰としましたか?  
U: 友達としました。  
:  
S: そうだったのですね。  
最近では, 他にどんなスポーツをしましたか?

図2 提案システムの応答

図2の対話例に関して、会話履歴フレームのwhat格には「卓球」、述語には「しました」が入っている。ここで、シソーラスを用いて「卓球」の上位語である「スポーツ」を取得する。上位語の「スポーツ」と述語の「しました」、userの発話内のWhat格に続く助詞である「を」を使って「他にどんなスポーツをしましたか?」という応答文を生成する。

また、シソーラス上での上位語が複数ある場合、使用する単語の選択方法としてIDFを用いる。IDFは概念ベースの3次属性までを使用し算出している。上位語が複数ある場合、それぞれの単語のIDFの値を比較し、その中で最も値が小さい単語を応答文に使用する。

## 6. 評価

提案した対話システムの快適さと自然さを測るために既存システムとの比較評価を行った。実験協力者10名にそれぞれのシステムと対話を行ってもらった。その後、6つの評価項目について、5段階による評価を行った。また、それぞれのシステムに同じ入力をしたときの対話例についても、自然さを5段階で評価してもらった。その結果を表5, 6に示す。

表5 対話システムの評価の平均値

	評価項目	提案	既存
快適さ	システムをまた使用したい	3.2	2.8
	システムを使用して楽しかった	3.5	3
	システムを使用して満足している	3.2	2.4
自然さ	システムが提供する話題を適切だった	3.2	2.4
	システムの応答は分かりやすかった	3.1	2.8
	システムの発話量は適切だった	3.5	3.7

6項目中、5項目で提案システムの評価の平均値が既存システムの評価の平均値を上回った。

表6 対話例の評価の平均値

	提案	既存
対話例の自然さ	3.96	2.56

提案システムの評価の平均値が既存システムの評価の平均値を上回った。

## 7. 考察

会話の話題を考慮して話題転換を行うことが雑談型対話システムの評価の向上につながった。しかし、提案システムの話題転換の応答文が生成されず、評価として差が出ない場合もあった。これは応答文の生成に会話履歴フレームのWhat格のみを活用していることなどが原因と考えられる。そのため、ほかの格に関する情報なども同時に活用することでuserの発話の内容を考慮した話題転換の応答文生成の頻度向上につながると考えられる。

## 8. おわりに

本研究では、話題転換を行う際にユーザとの会話履歴を活用した応答を行うことで、より自然に雑談型の対話ができるシステムの構築を目指した。提案した会話履歴とシソーラスを活用する話題転換応答によって、対話の自然さが向上する評価結果が得られた。しかし、話題転換の応答文生成の頻度の問題などもあり、これらを改善することで、システムのさらなる性能向上につながると考えられる。

### 参考文献

- [1] 小山翔平, 土屋誠司, 渡部広一, “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”, 情報科学技術フォーラム FIT2020, pp.127-128.
- [2] 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.44-45, 2007.
- [3] 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.
- [4] 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.