

# 対話システムにおける用語間の関係性を用いた 話題誘導応答文生成の拡張と評価\*

山内 祐輝, Graham Neubig, Sakriani Sakti, 戸田 智基, 中村 哲 (奈良先端大)

## 1 はじめに

対話システムにおける応答文生成は、自然で円滑な対話を実現する上で、重要な要素技術の一つである。システム構築の汎用性を高める手法として、テンプレートを用いた手法 [1] やフレーズ単位に基づく統計的手法 [2] が提案されており、主にユーザ側の目的に沿った対話での応答を対象として、その有効性が確認されている。これに対し、我々はシステム側が特定の目的を持ち対話を進める際の応答に着目し、現在の話題から意図した別の話題へと誘導するための応答文（話題誘導応答文）の生成法を提案している [3]。生成される応答文単体に対する評価において、その有効性を確認している。

本稿では、実際の対話において、提案法による話題誘導が可能であるかを調査する。単純な雑談対話を対象とし、ユーザが入力した話題からシステムが目標とする話題へと誘導する話題誘導対話システムを構築する。実験的評価により、その有効性を示す。

## 2 テンプレートによる話題誘導応答文生成

現在の話題に関する用語と誘導する話題に関する用語を用いて、テンプレートにより応答文を生成する。自然な誘導を可能とする用語ペアを得るために、概念辞書に基づく手法と Web 検索を用いる手法を提案している [3]。

### 2.1 概念辞書に基づく手法

概念辞書とは、用語間の関係を記述した辞書である。概念辞書に記述のある用語間の関係を基に応答文テンプレートを作成し、現在の話題に関する用語と誘導する話題に関する用語を入れることで、応答文を生成する。例えば、上位語から下位語へ誘導する応答文テンプレートとして「〈Current〉の研究には〈Target〉もありますよね。」を作成する。実験的評価により、概念辞書内の距離が近い用語ペアほど自然性の高い応答文が生成できることが確認されている。

### 2.2 Web 検索を用いる手法

Web 検索により取得した用語間の関係性を利用して、応答文テンプレートに対して埋め込み対象となる用語ペアを抽出する。手法の手順を以下に示す。

1. 話題誘導応答文に利用できる可能性が高い関係を想定した応答文テンプレートを作成

例：応答文テンプレート：

「〈Current〉は〈Target〉に使われていますよね。」

2. 応答文テンプレートを基に検索テンプレートを作成

例：検索テンプレート：

「〈Current〉を用いた〈Target〉」

3. 検索結果を基に応答文テンプレートに当てはまる関係のある用語ペアを決定

（ヒット件数が 0 の場合は関係がないと判断）

誘導に利用する用語ペアを決定する際の尺度として、用語間の相互情報量を用いる。実験的評価により、相互情報量が高い用語ペアの応答文ほど自然性が高くなる傾向が確認され、上位の用語ペアは概念辞書に基づく手法の子孫関係（例：上位語の上位語）にある用語ペアの応答文と同等の自然性が得られている。

## 3 話題誘導対話システム

### 3.1 概要

話題誘導対話システムとは、システム側が目標とする話題を持って対話を進め、最終的に目標とする話題へ話を誘導するシステムである。本研究で扱うタスクとして、情報科学研究についての雑談対話を対象とする。対話の流れは、ユーザが話したい話題を入力することで対話を開始し、それ以降はシステムがその話題に関係のある話題へ誘導しながら対話を進め、システムが目標とする話題へ誘導できたら対話が終了する。

システムの応答は以下の 3 つである。

- 現在の話題と関係のある話題への誘導
- 誘導先の利用語についての説明
- 現在の用語と誘導先の利用語の関係の説明

1 つ目の応答は、2 節で述べた話題誘導応答文生成法を利用する。2 つ目の応答は、辞書に登録されている各用語を説明する応答文を手で用意し、3 つ目の応答は、関係性を取得する際の情報源となる Web ページを表示することで応答する。対話開始時はユーザが入力した話題を現在の話題とし、それ以降は、生成された話題誘導応答文に対してユーザが同意した場合に、誘導先の話題〈Target〉を現在の話題とする。また、一度誘導を試みた用語間の誘導は行わない。

システムが認識できるユーザの発話を以下に示す。

- 話題誘導に対するの同意（「はい」、「そうですね」等を想定）
- 話題誘導に対するの否定（「いいえ」、「関係ないと思います」等を想定）
- 誘導先の利用語について質問
- 現在の用語と誘導先の利用語の関係について質問
- 一つ前の話題へ戻る
- 対話の終了（話題の打ち切り）

### 3.2 強化学習による誘導経路の最適化

対話開始時に入力された話題からシステムが目標とする話題へと誘導する上で、複数の誘導経路が存在する。自然な誘導を実現するためには、誘導経路を適切に決定する必要がある。この問題に対して、マルコフ決定過程 [1] によるモデル化と、強化学習による最適化を適用する。システムが認識可能な話題を状態  $s \in S$  とし、他の話題への誘導を行動  $a \in A$  とする。状態  $s$  において行動  $a$  をとり、状態  $s'$  に遷移する確率は、ユーザが話題誘導に対して同意する確率とし、ユーザモデルとして与える。同意が高確率で得られかつ効率的に誘導が行われるように報酬を設定し、強化学習を行うことで、最適な方策（すなわち誘導経路）を決定する。

代表的な強化学習法として、Q 学習 [4] を用いる。全ての状態  $S$  と行動  $a$  の組み合わせに対して行動価値関数（Q 値）を設け、以下の式に従って更新する。

$$Q(s_t, a_t) \leftarrow (1 - \alpha)Q(s_t, a_t) + \alpha(r_t + \gamma \max_{a \in A} Q(s_{t+1}, a)) \quad (1)$$

ここで、 $Q(s_t, a_t)$  は時刻  $t$ 、状態  $s_t$  において行動  $a_t$  を選ぶ場合の Q 値を表し、 $r_t$  は状態  $s_t$  で行動  $a_t$  をとった際に得られる報酬、 $\alpha$  と  $\gamma$  は学習率と割引率と呼ばれるパラメータである ( $0 < \alpha < 1, 0 < \gamma < 1$ )。以下のアルゴリズムにより Q 値を学習する。

\* An Extension and Evaluation of Answer Sentence Generation using Relationships between Terms for Guiding Users to New Topics in Spoken Dialog Systems by YAMAUCHI, Yuki, NEUBIG, Graham, SAKTI, Sakriani, TODA, Tomoki, NAKAMURA, Satoshi (NAIST)

1. 全ての  $Q(s, a)$  の値を 0 で初期化
2. 全状態の中からランダムで状態  $s_0$  を選択 ( $t = 0$ )
3. 状態  $s_t$  において取り得る行動  $a_t$  の中から  $\epsilon$ -greedy 法により行動  $a_t$  を選択
4. ユーザモデルを基に次の状態  $s_{t+1}$  への遷移と  $Q(s_t, a_t)$  の更新を実施  
(状態  $s_{t+1}$  が目標状態の場合, 状態  $s_0$  へ遷移)
5. 3, 4 を一定回数実行した後, 2 へ戻る
6. 2~5 を一定回数実行した後, 終了

$\epsilon$ -greedy 法とは, 確率  $1 - \epsilon$  で状態  $s$  において行動価値関数  $Q(s, a)$  が最大となる行動  $a$  を選択し, 確率  $\epsilon$  で  $Q(s, a)$  によらずランダムで行動  $a$  を選択する手法である. 学習後, 各状態  $s$  に対して  $Q(s, a)$  が最大となる行動  $a$  を選ぶことによって誘導経路が決定する.

ユーザモデルは, システムが状態  $s$  において行動  $a$  を選択して応答する場合に想定されるユーザの同意確率  $P(R = \text{yes}|s, a)$  を定義する. 本稿では, [3] で得られた話題誘導応答文の自然性に基づき初期化する. 本システムでは, 運用して得られた実際のユーザデータに基づいて, 誘導経路を最適化することも可能である. その際には, ユーザモデルを以下の式に基づいて更新する.

$$P(R = \text{yes}|s, a, D) = \frac{N_{s,a,\text{yes}} + h\mu}{N_{s,a} + h} \quad (2)$$

$P(R = \text{yes}|s, a, D)$  は実データ  $D$  が入力された際の状態  $s$  で行動  $a$  が選択されたときのユーザが同意する発話  $R = \text{yes}$  が行われる確率で,  $N_{s,a}, N_{s,a,\text{yes}}$  は実データ内で状態  $s$  で行動  $a$  が選択された回数とその際にユーザが同意した回数,  $h$  は更新する際のハイパーパラメータ,  $\mu$  は初期値である. 更新されたユーザモデルを用いて Q 学習を行うことで, 実データを考慮した誘導経路の最適化を行う.

## 4 実験的評価

### 4.1 実験条件

システムをユーザに利用してもらい, その対話データを分析する. システムの目的は, 奈良先端大の情報科学研究科の研究に関する雑談対話を通して特定の話題(研究分野)へ誘導することである. システムが目標とする話題については, ユーザは未知という条件で実験を行う. システムが認識可能な話題は, 人手で作成した概念辞書に登録されている用語 150 語で, ユーザはこの中から好きな研究分野を 1 つ選択することで対話を開始する. システムの目標とする話題は, 事前に設定された 3 用語の内ランダムで 1 用語を選択する. 音声認識誤り, 言語理解誤りの影響をなくすために, ユーザはあらかじめ指定した発話内容の番号に従って入力する. また, システムの応答は生成された文章を表示することで行う.

Q 学習におけるパラメータは, 学習率  $\alpha = 0.1$ , 割引率  $\gamma = 0.8$ ,  $\epsilon = 0.2$  と設定する. 報酬に関しては, 目標とする話題へ誘導できた場合は 100, 話題誘導に対してユーザが同意した場合は -5, 否定した場合は -10 を与える. 誘導経路は目標とする話題に対して 200 万回学習した結果を用いる. ユーザモデルに関しては, 先行研究 [3] の評価結果において, 自然性の平均が 3 の手法ではユーザは必ず同意し, 1 の手法ではユーザは必ず否定すると仮定して, 同意確率の初期値を決定する.

システムの利用者数は 6 名, 総対話数は 47 対話のデータを評価する. 初期値のユーザモデルにより最適化したシステムを用いる. 評価基準を以下に示す.

- 定量的評価
  - 誘導経路学習の過程における利得(割引率を考慮した報酬の総和)の期待値
- 定性的評価
  - 1 対話中の平均話題誘導回数

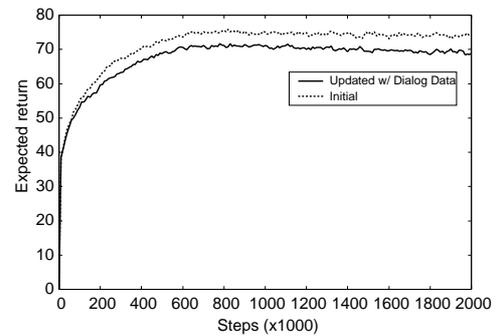


Fig. 1 Change of expected return in Q-learning

Table 1 Result of subjective evaluation

Guide	Yes	No	Success
6.68	2.32	4.36	95.7%

- 1 対話中の話題誘導に対する平均同意数
- 1 対話中の話題誘導に対する平均否定数
- システムが目標とする話題への誘導成功率

### 4.2 定量的評価

誘導経路を学習する際の学習過程における利得の期待値の推移を図 1 に示す. Initial は初期値のユーザモデルを利用した結果, Updated w/ DialogData は本実験の対話データで更新したユーザモデル(式 (2) において  $h = 5$  に設定)を利用して学習した結果である. 学習回数が 60 万回程度で利得の期待値が概ね収束していることが確認できる.

### 4.3 定性的評価

システムとユーザとの対話の分析結果を表 1 に示す. Guide は 1 対話中にシステムが話題誘導を行う回数の平均, Yes, No はそれぞれ 1 対話中にシステムが話題誘導に対してユーザが同意する回数と否定する回数の平均, Success はシステムが目標とする話題への誘導の成功率である.

表 1 より, 95.7% の確率でシステムが目標とする話題へ誘導できている. このことから, 提案する話題誘導応答文生成法を用いることで, 辞書に登録されている話題間に関しては, ユーザの許容範囲内での話題誘導が概ね可能であることが分かる. 一方で, 話題誘導回数とユーザの否定回数を見ると, システムが話題誘導を行う際に約 65% の確率でユーザにとって不自然な話題誘導が生成されている. この原因として, ユーザモデルの初期値が実際のユーザデータと一致していないことなどが挙げられる. そのため, 図 1 で示したような, 実データを用いて更新したユーザモデルにより最適化したシステムの検討が必要である.

## 5 おわりに

本稿では, テンプレートによる話題誘導応答文生成法の有効性を実際の対話を通して調査するため, 話題誘導対話システムを構築し, その評価を行った. その結果, ユーザの許容範囲内でシステムが目標とする話題へと誘導できることが確認された. 一方で, ユーザにとって不自然な応答が約 65% の確率で生成されることも分かった. 今後は実データを用いたシステムの最適化に取り組む.

## 参考文献

- [1] 河原 他, “音声対話システム.” オーム社, 2006.
- [2] F. Mairesse et al., Proc. ACL 2010, pp. 1552–1561, 2010.
- [3] 山内 他, 情報研報, Vol. 2012-SLP-94, No. 3, pp. 1–7, 2012.
- [4] C. J. C. H. Watkins et al., Machine Learning, Vol. 8, pp. 279–292, 1992.