

# 対話における応答文の候補文検索型生成法

富浦 洋一 , 柴田雅博, 西口友美

九州大学

E-mail : {tom,shibata,nisiguti}@agent.kyushu-u.ac.jp

## 1 はじめに

バス路線案内や天気予報などに関する対話システムでは、対話の流れがある程度定型であるため、パターンマッチによってユーザ発話を理解(解析)し、対象領域の知識ベースを検索して応答文を生成している。しかし、現段階では、ユーザ発話を完全に理解して応答文を生成する対話システムは、非常に限られた領域でしか実現できない。一方、Elizaに代表されるキーワード駆動の対話システムでは、対話の領域は制限されないが、システム発話は利用者に発話を促すための相槌などであり、利用者は対話によって新たな情報を得ることはできない。

本発表では、前者に比べて対話の領域が広く、しかも、ユーザ発話に関連した情報を提供する新たな応答文の生成手法を提案する。本手法では、対話領域に関連する Web 上の文書からシステム発話の候補文を収集しておき、この中からユーザーの発話文と表層的結束性および意味的関連性の高い文を検索し、システムの応答文とする。

我々はよく、自分でもはっきりしていない考えが他者との会話を通して明確になるという経験をする。また、会話を通して新しい見方や情報を得るということもある。本発話文生成手法は、会話による思索支援や明確に言語化されていない検索要求に対する情報の取得に利用可能ではないかと考えている。なお、本発表での試作システムは、映画に関する対話を行うことを想定して作成している。

## 2 対話システムの概要と対話例

システムとの対話は、ユーザ発話とシステム発話が交互になされる二者対話を想定している。

試作したシステムは、以下の3つのシステム発話モードを持つ。

**完全理解対話モード** ユーザ発話が質問文のとき試されるシステム発話のモード。バス路線案内システムなどと同じく、パターンマッチによりユーザ発話を理解し、知識ベースを検索してシステム発話を生成する。以下の対話例では  $S_5$  がこのモードによる出力である。

**情報検索発話モード** ユーザ発話が質問文でないか、完全理解対話モードでの発話に失敗した場合に試されるシステム発話のモード。このモードの詳細は次節で述べる。なお、以下の対話例では、 $S_2 - S_4$  がこのモードによる出力である。

**相槌モード** 情報検索発話モードでの発話に失敗した場合に行われる発話モード。ただし、試作システムでは、『へえー』を生成し、これにより相槌モードによる発話であることを示すだけに留めている。

以下に、試作したシステムとの対話例を示す。 $U_i$  がユーザ発話で、これに対するシステムの応答が  $S_i$  である。

$S_0$  : 映画の話をしましょう。

$U_1$  : 亡国のイージスを観ました。

$S_1$  : へえー。

$U_2$  : 自衛隊に協力してもらって、映画を作ったそうです。

$S_2$  : 何か、説教じみていて、映画にロマンを感じず、単に、自衛隊の宣伝映画に過ぎないという感じを受けました。

- $U_3$  : でも戦争万歳という映画ではないんです。
- $S_3$  : 原作で好きなエピソードもカットされていてショックでした。
- $U_4$  : 主人公の真田広之が好きでした。
- $S_4$  : 真田広之が熱く、寺尾聡が渋く、キャストも良かったです。
- $U_5$  : 寺尾聡が演じていた役は何でしたかね？
- $S_5$  : 《いそかぜ》副長・宮津弘隆<sup>2</sup>等海佐です。

### 3 情報検索発話モードによる発話

情報検索発話モードでは、WWW 上から求めたシステム発話の候補文集合（候補文コーパス）から、ユーザ発話に対する応答として、表層的な結束性を持ち、意味的に最も関連する文を求め、これをシステム発話として出力する。なお、ユーザ発話および候補文に対しては、形態素解析<sup>1</sup>のみを行っている。

#### 3.1 候補文コーパス

候補文コーパスは以下の手順で構築した。Web 上から、サーチエンジンで検索キーを“ $\alpha$  映画”として Web ページを収集する。 $\alpha$  は対話のメインテーマとなるもので、映画に関する対話を行う試作システムでは、 $\alpha$  として映画タイトルを用いた。検索キーに“映画”を含めるのは、映画とは直接関係ない文書を除くためである。こうして得られる Web ページには、その文書のメインテーマが“ $\alpha$  映画”ではないものも含まれる。そのため、得られた Web ページに“ $\langle \text{title} \rangle \alpha \langle \backslash \text{tilde} \rangle$ ”があるものだけを残し、他の Web ページは削除した。こうして収集した Web ページから文を抜き出し、 $\alpha$  ごとの候補文集合  $\Gamma(\alpha)$  を構築した。なお、 $\Gamma(\alpha)$  には、どの Web ページの何文目から得られた文かの情報も付与している。

#### 3.2 センタリング理論

表層的な結束性を満たすか否かの判定は、基本的にはセンタリング理論 [2] に基づいて行う。本研究では、中心性は、ゼロ代名詞『は』が後置する名詞句、『が』が後置する名詞句、『を』が後置する名詞句の順に高いものとする。発話にゼロ代名詞があるか否かを正確に求めるためには格フレームなどの知識を必要とするため、本研究では単純に、『は』『が』でマークされた名詞句がともにない場合にゼロ代名詞があるものとして扱う。また、代名詞の先行詞は、その前の発話中の最も中心性が高い名詞句（ゼロ代名詞の先行詞を含む）とする。

#### 3.3 情報検索発話モード概要

我々の対話システムとの対話は、比較的短いもので、対話の最初のユーザ発話の話題が対話のメインテーマになり、対話を通して、メインテーマの変更はないものと仮定する。

情報検索発話モードでは、以下の手順でシステム発話を生成する。

##### (1) メインテーマによるフィルタリング

対話が始まると、システムは第一ユーザ発話で最も中心性の高い名詞句を求める。すなわち、『は』が後置する名詞句、『が』が後置する名詞句、『を』が後置する名詞句の順で第一ユーザ発話中での名詞句を求め、最初に見つかった名詞句を対話のメインテーマとして設定する。前述のように、発話の候補文はメインテーマごとに収集されている。求めた対話のメインテーマを  $\alpha$  とすると、 $\alpha$  をメインテーマとする Web ページの文書からシステム発話の候補を求めるのが、より妥当性の高い応答文を生成できると考えられる。そこで、以降のシステム発話の候補文集合を  $\Gamma(\alpha)$  に制限する。

##### (2) システム発話の焦点とすべき名詞句 $f$ の決定

直前のユーザ発話  $U$  の中で最も中心性の高い名詞句を中心的な話題（本発表では焦点と呼ぶ）とする文をシステム発話とすると、話題の流れが自然であると考えられる。焦点  $f$  は、

<sup>1</sup> 形態素解析には茶筌を利用している。

3.2節で述べたことを考慮すると、以下のよう  
にして求まる。

- $U$  に『NP は』があれば、 $f$  は  $NP$  .
- $U$  に『は』が後置する名詞句がなく『NP  
が』があれば、 $f$  は  $NP$  .
- $U$  に『は』が後置する名詞句がなく『が』  
が後置する名詞句もなければ、 $U$  にはゼ  
ロ代名詞が存在し、その先行詞は  $U$  で  
最も中心性が高い名詞句である。これは  
前回のシステム発話の焦点であるから、  
 $f$  は前回のシステム発話の焦点（つまり  
焦点を継続）。

### (3) $\Gamma(\alpha)$ からの焦点を含む文の選出

$\Gamma(\alpha)$  中の文から、 $f$  を先行詞とするゼロ代名  
詞、『 $f$  は』、『 $f$  が』、あるいは『 $f$  を』を  
含む文を求める。なお、照応解析の精度を考慮  
し、『 $f$  は』を含む文に後続する文で、3.2節に  
述べた判定法でゼロ代名詞があると判定され  
る連続した最大  $m$  文を  $f$  を先行詞とするゼ  
ロ代名詞を含む文とする<sup>2</sup>。最終的に選択さ  
れた候補文からシステム発話を生成する場合  
は、焦点を削除しゼロ代名詞化する（たとえ  
ば、『 $f$  は』で検索された候補文の場合、『 $f$  は』  
を削除)<sup>3</sup>。センタリング理論に従うならば、  
このようにすることで、表層的結束性を満た  
す文をシステム発話として生成できる。

### (4) 名詞・述部の関連度によるフィルタリングと 優先順位付け

直前のユーザ発話  $U$  と、前記 (3) で求めた  
各候補文  $S$  との名詞・述部の関連度  $R(U, S; f)$   
を求め、これが  $U$  の述部のタイプ (3.3.2 節  
参照) ごとに設定されている閾値以上である  
場合、 $E(U, S; f)$  が最大のものを選択し、焦  
点の名詞句をゼロ代名詞化してシステム発話

とする。 $E(U, S; f)$  は次のように定義される。

$$\begin{aligned} R(U, S; f) \\ &= C_N \cdot R_N(U, S; f) \\ &\quad + C_P \cdot R_P(\text{Pred}(U), \text{Pred}(S)). \end{aligned}$$

ただし、 $f$  は (2) で設定した焦点であり、  
 $\text{Pred}(U)$  は  $U$  の述部である。 $R_N()$  は名詞  
の関連度、 $R_P()$  は述部の関連度であり、それ  
ぞれ 3.3.1 節と 3.3.2 節で述べる。また、 $C_N$   
および  $C_P$  は定数である<sup>4</sup>。

### (5) メインテーマを焦点とした発話の生成

焦点つまり中心的な話題がメインテーマに移  
ることは自然な話題の推移の一つであると考  
えられる。そこで、前記 (4) で名詞・述部の関  
連度が閾値を越える候補文が存在しない場合、  
(1) で求めたメインテーマ  $\alpha$  を焦点とするシ  
ステム発話の生成を試みる。つまり、 $f = \alpha$   
として、前記 (3)(4) と同様の操作を行う。た  
だし、焦点のゼロ代名詞化は行わない（ゼロ  
代名詞化すると、先行詞がわからなくなる恐  
れがあるため）。

#### 3.3.1 名詞の関連度

1 文中で共起しやすい 2 つの名詞は、意味的に関  
連が高いと考えられる。そこで、

$$r(n, n') = \log \frac{p(n, n')}{p(n)p(n')}$$

を名詞  $n$  と  $n'$  の関連度と定義する。ただし、  
 $p(n, n')$  は 1 文中に  $n$  と  $n'$  が発生する確率であ  
り、 $p(n) = \sum_{n'} p(n, n')$  である。共起確率は、映  
画に関する Web ページから求めた文書集合 (総文  
数 5,875,515) から最尤推定で求めた。

$N_S(U; f)$  を  $f$  を除く  $U$  中の全名詞から成る集  
合を表すものとする。焦点が  $f$  の場合の、ユーザ  
発話  $U$  とシステム発話候補文  $S$  との、名詞に関  
する文全体の関連度  $R_N(U, S; f)$  を  $N_S(U; f) \times$   
 $N_S(S; f)$  中の全ての  $(n, n')$  のうち、 $r(n, n')$  が

<sup>2</sup> 予備実験の結果より、試作システムでは  $m = 2$  として  
いる。

<sup>3</sup>  $f$  が連体修飾を受ける場合、『 $f$  は』等を削除すると、不  
自然な文になってしまうため、このような場合、 $f$  を適切な  
代名詞や『その人』のような連体詞付きの名詞句にすべきで  
ある。しかし、試作システムでは、このような場合、『 $f$  は』  
等を削除しないようにしている。

<sup>4</sup> 予備実験結果より、試作システムでは、 $C_N = C_P = 0.5$   
としている。

高い正の上位  $K$  個まで<sup>5</sup> の和と定義する．上位  $K$  に制限するのは，長い候補文が優位になること避けるためである．

### 3.3.2 述部のの関連度

述部は，動詞，形容(動)詞，その他の3つのタイプに分類し，ユーザ発話の述部  $p$ ，候補文の述部  $p'$  の関連度を

$$R_{p,p'} = \begin{cases} R_v(p, p') & ; p, p' \text{ 共に動詞,} \\ R_a(p, p') & ; p, p' \text{ 共に形容(動)詞,} \\ 0 & ; p \text{ がその他,} \\ \mu & ; \text{その他} \end{cases}$$

と定義する．試作システムでは  $\mu = -\infty$ <sup>6</sup> ．

#### [動詞間の関連度]

近くに出現しやすい2つの動詞は，因果関係や事象の全体部分関係，事象の時間的前後関係などの意味的な関連性があると考えられる．そこで，動詞  $v$  と  $v'$  の意味的な関連度  $R_v(v, v')$  を

$$R_v(v, v') = \begin{cases} \log \frac{p(v, v')}{p(v)p(v')} & ; v, v' \text{ が共に学} \\ & \text{習データに出現} \\ 0 & ; \text{その他} \end{cases}$$

と定義する．ただし， $p(v, v')$  は  $v$  と  $v'$  が2文以内の範囲で共起する確率である．

これらの確率を推定するために，前述の映画に関する Web ページの文書から2文以内で共起する動詞対の列を求め，これから低頻度(100未満)の動詞から成る対を削除して学習データを作成した(対の総数 29,819,802，対の異なり数 8,198,797，動詞異なり数 13,347)．この学習データは名詞の共起確率推定の学習データに比べスパースなため， $p(v, v')$  は PLSA モデル [1] により推定した<sup>7</sup> ．

#### [形容詞・形容動詞の関連度]

形容(動)詞が述部である発話は，何らかの評価をしており，対話において利用者がそのような文

<sup>5</sup> 予備実験の結果より，試作システムでは， $K = 3$  としている．

<sup>6</sup> 異なるタイプ同士の述語でも，共起情報量などを用いて，関連度を求めることも考えられる．これに関しては今後の課題である．

<sup>7</sup> 予備実験の結果より，試作システムでは latent variable の次元を 20 と推定したものを使用している．

を発話した場合，システムが自分の評価に同意することを利用者は期待すると考えられる．ここでは，形容(動)詞が表す評価を単純に，正，負，その他，に分けるものとする．したがって，ユーザ発話の述部の形容(動)詞  $a$  とシステム発話候補との述部の形容詞(動)詞  $a'$  との関連度  $R_a(a, a')$  は，

$$R_a(a, a') = \begin{cases} \lambda & ; a, a' \text{ 共に正または共に負,} \\ -\lambda' & ; a, a' \text{ の一方が正, 他方が負,} \\ 0 & ; a \text{ の評価がその他.} \end{cases}$$

とした<sup>8</sup> ．形容(動)詞の評価の極性は，前述の映画に関する全 5,875,515 文から，頻度の高い順に 1000 単語を選び，人手で付与した．なお，極性付与の自動化に関しては，[3] などの評価表現の自動抽出技術が参考になる．

## 4 おわりに

今回試作したシステムは，まだ開発段階であり，十分な評価実験をしたわけではないが，情報検索発話モードでシステムが生成した発話のうち，7割程度は妥当であった．今回の名詞の関連度，述部の関連度はときどき直観に合わない場合があり，定義および推定法を吟味する必要がある．また，ユーザ発話の名詞句が1つ以下の場合，情報検索発話モードは失敗し，相槌モードでの発話となってしまう．この点の改良も今後の課題である．

## 参考文献

- [1] Hofmann, T.: Probabilistic Latent Semantic Analysis, In: *Proc. of Uncertainty in Artificial Intelligence, UAI'99* (1999).
- [2] Walker, M., M. Iida, and S. Cote: Japanese discourse and the process of centering, *Computational Linguistics* **20**(2), pp.193–233 (1994).
- [3] 鈴木 泰裕, 高村大也, 奥村 学: Weblog を対象とした評価表現抽出, 人工知能学会 セマンティックウェブとオントロジー研究会, SIG-SWO-A401-02 (2004).

<sup>8</sup> 試作システムでは，予備実験の結果より， $\lambda, \lambda'$  共に 1 ．