

対話型質問応答システムにおける曖昧な質問に対する問い返し文の生成

坂本 篤史 白井 清昭

北陸先端科学技術大学院大学 情報科学研究科

{asakamot, kshirai}@jaist.ac.jp

1 はじめに

我々は曖昧な質問を取り扱う質問応答システムに関する研究を行っている。本研究における「曖昧な質問」とは、質問文中のキーワードの意味が曖昧であるために解答を一つに絞ることができない質問を指す。例えば、「ワールドカップの優勝国はどこですか?」という質問は、ワールドカップにはサッカー、ラグビーなどの種類があるという意味で曖昧であり、その種類に応じて解答が複数存在する。我々は、このような質問に対してシステムの側から「ワールドカップの種目は何ですか?」といった問い返しを行い、これに対するユーザの返答に応じて適切な解答を取り出す対話型質問応答システムの構築を目指している。ユーザとの対話を伴う質問応答システムに関する先行研究はいくつかある [1, 2] が、その多くはドメインを限定してシステムを構築している。これに対し、本研究では、新聞記事を知識源としたオープンドメインの質問応答システムにユーザとの対話を取り入れることを試みる。

提案する質問応答システムの処理の流れを図 1 に示す。処理 [A]~[C] は通常の質問応答システムとほぼ同じである。ここまではユーザの質問に対する解答候補が複数得られる。処理 [D] では解答群を作成する。詳細は 2 節で述べるが、この処理は問い返し文の発話内容を決めることに相当する。処理 [E] では解答群から問い返し文の表層表現を生成する。詳細は 3 節で述べる。最後の処理 [F] は、ユーザからの返答を受けて質問の曖昧性を解消し、適切な解答を 1 つ選択してユーザに提示する。また、必要ならユーザに対して再度問い返しを行う。本論文では処理 [D] と処理 [E] について報告する。

2 解答群作成

2.1 概要

本節では解答群の作成について述べる。解答群の作成方法は松本らの手法 [3] を改良したものである¹。本論文ではその概要のみを述べる。詳細は [5] を参照していただきたい。

¹松本らは曖昧な質問に対して図 2 のような解答群を生成し、それをリスト表示する手法を提案している。ユーザとの対話を行うことは想定していないが、ユーザの質問の曖昧性を検出するという目的は同じであるため、彼らの手法は本システムにも適用できる。

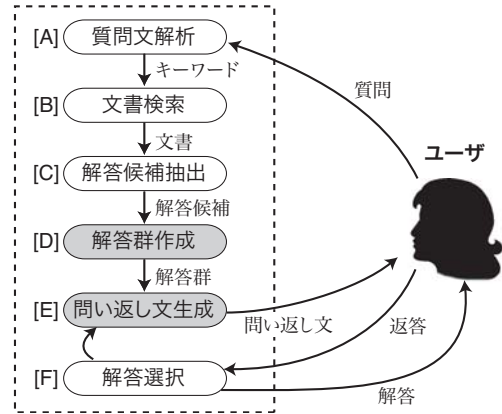


図 1: システムの処理の流れ

まず、複数の解答候補 a_i のそれぞれに対して、それを含む元のテキストから (a_i, k_j, s_k) という 3 つ組を抽出する。 k_j はユーザが入力した質問文内のキーワードであり、 s_k はキーワード k_j の意味を限定する限定表現である。 1 節で述べたワールドカップの質問の例では、「ワールドカップ」がキーワード、「サッカー」や「ラグビー」が限定表現となる。キーワードの限定表現は、あらかじめ用意されたいくつかの抽出パターンによって抽出する。例えば、キーワードの直前・直後にある名詞やキーワードと意味的関連度の高い名詞が限定表現として抽出される。次に、 (a_i, k_j, s_k) の集合から、 k_j が共通でかつ限定表現 s_k に何らかの共通の属性 ($attr$) が存在する部分集合を発見する。共通属性の例としては、末尾 N 文字が共通 ($N=1,2,3$)、意味クラスが共通、「数+助数詞」という表現であることが共通、などがある。このように発見された部分集合を $AG(k, attr) = \{ (a_i, s_i) \}$ という形式で表現したものが解答群である。 k は 3 つ組の部分集合に共通するキーワードであり、これはユーザの質問文に含まれる曖昧なキーワードに該当する。一方、 s_i は $(a_i$ の周辺に位置する) キーワード k の限定表現、 $attr$ は s_i が共通して持つ属性である。図 2 に解答群の例を示す。図 2 の解答群には「パルトロー」「デンチ」「スピルバーク」の 3 つの解答候補がある。また、質問文中の曖昧なキーワード k は「アカデミー賞」である。限定表現「主演女優賞」「助演女優賞」「監督賞」は、賞の部門という観点で「アカデミー賞」の意味を限定している。限定表

質問文=アカデミー賞を受賞したのは誰ですか

AG(アカデミー賞, 末尾1文字:賞) = {
 (パルトロー, 主演女優賞), (デンチ, 助演女優賞),
 (スピルバーグ, 監督賞) }

図 2: 解答群の例

現の共通属性 *attr*(=末尾1文字:賞) は, 末尾1文字が「賞」であることが全ての限定表現に共通していることを表わす。

一般に図 2 のような解答群は多数得られる。そのため, 様々な観点から解答群の候補に対してスコアを与え, 最大のスコアを持つ解答群をひとつ選択する。

解答群の作成は, 質問文中の曖昧なキーワードを検出するとともに, ユーザに対する問い返し文の内容を決める役割も持つ。すなわち, 曖昧なキーワードの意味をユーザに問う文, より具体的にはキーワードの限定表現をユーザから聞き出す文を生成できれば, 解答候補の中から適切な解答を選択することができる。図 2 の例では, 「アカデミー賞」の部門が「主演女優賞」「助演女優賞」「監督賞」のいずれかをユーザから聞き出せれば, ユーザが求める解答を特定できる。

2.2 解答群作成手法の改良

本研究では松本らの手法 [3] の問題点を見直し, 以下のような改良を行った。

- 限定表現の抽出パターンを追加した。具体的には, 記事の先頭にある名詞を限定表現として抽出するようにした。新聞記事の先頭には記事のトピックが書かれていることが多く, この名詞がキーワードの限定表現となることがあるためである。
- 解答群を作成する基準となる共通属性の種類を増やした。具体的には「直前」(全ての限定表現がキーワードの直前に出現することが共通), 「直後」, 「連体修飾」(全ての限定表現が「 s_i の k 」といったようにキーワードを連体修飾することが共通) などである。例えば, 「サリン事件はいつ起こりましたか」という質問で, 「サリン事件」には地下鉄サリン事件と松本サリン事件の2つがあるという曖昧性がある。このとき曖昧なキーワード k が「サリン事件」, 限定表現 s_i が「地下鉄」「松本」といった解答群が作成されるべきである。松本らの手法では「地下鉄」と「松本」をまとめる共通属性を用意していなかったため, このような解答群は得られなかったが, 「直前」という共通属性を導入したことで適切な解答群を作成できるようになった。

| | テンプレート | サ変 | 属性 | 順序 |
|---|-------------------------------|----|----|-----|
| A | k の s_1 ですか, s_2 ですか? | - | - | k |
| A | k したのは s_1 ですか s_2 ですか? | ○ | - | k |
| A | $s_1 k$ ですか $s_2 k$ ですか? | - | - | s |
| B | k の $wh+s$ ですか? | - | - | k |
| B | $wh+s$ の k ですか? | - | - | s |
| B | $wh+s$ で k したのですか? | ○ | - | s |
| C | 何 k ですか? | - | 直前 | - |
| C | 何の k ですか? | - | 連体 | - |

図 3: 問い返し文生成テンプレート (抜粋)

- 解答群の候補のスコア付け手法を見直した。例えば, 質問文中のキーワードと解答群の限定表現の意味的関連度が高い場合にはその解答群に高いスコアを与えるようにした。単語間の意味的関連度は新聞記事コーパスにおける共起度によって定量化した。また, 限定表現の抽出パターンに応じてスコアを調整した。限定表現はいくつかの抽出パターンによって抽出されるが, その抽出パターンが限定表現をどれだけ正確に抽出できるかといった信頼度を解答群のスコアに反映させた。

3 問い返し文生成

本節では, 解答群が与えられたとき, 曖昧なキーワードの意味をユーザに問う問い返し文を生成する手法について述べる。

3.1 処理の流れ

問い返し文は基本的にはテンプレートを用いて生成する。問い返し文の生成テンプレートの例を図 3 に示す。A,B,C はテンプレートのタイプである。一方, 「サ変」「属性」「順序」はテンプレートの適用条件を表わす。これらの詳細は 3.3 項で述べる。本研究で用意したテンプレートの数は 23 である。

タイプ A は二者択一の問い返し文を生成するテンプレートである。解答群中の限定表現が s_1 と s_2 の 2 種類しかないときにはこのテンプレートが適用される。

タイプ B は曖昧なキーワード k と問い返し主題 $wh+s$ を用いて生成するテンプレートである。ここで問い返し主題とは, 解答群に含まれる限定表現全体を指し示す名詞に疑問詞をつけた表現であり, ユーザに対する問い返し文において中心的な役割を果たす。例えば, 図 2 の解答群では, 3 つの限定表現は賞の種類を表わしている。このとき, 問い返し主題は「何賞」となり, タイプ B の 1 行目のテンプレートから「アカデミー賞の何賞ですか?」という問い返し文が生成される。

タイプ C は曖昧なキーワード k のみを用いて生成するテンプレートである。解答群によっては問い返し主題が

決まらない場合がある。例えば、「NHK 杯の優勝者は誰ですか」という質問は、「NHK 杯」にはフィギアスケート、体操、将棋などの種類があるという意味で曖昧である。このときはフィギアスケート、体操、将棋のいずれかを聞き出す問い返し文を生成すればいいが、これら全体を指すような「疑問詞+名詞」といった問い返し主題を生成するのは難しい。このようなときはキーワードのみで問い返し文を生成する。例えば、図 3 のタイプ C の 2 行目のテンプレートから「何の NHK 杯ですか?」といった問い返し文が生成される。

問い返し文の生成処理の流れを図 4 に示す。まず、解答群に含まれる限定表現の異なり数が 2 ならタイプ A のテンプレートで問い返し文の候補を生成する。それ以外はタイプ B とタイプ C の両方のテンプレートを用いる。タイプ B のテンプレートを適用する前には問い返し主題 $wh+s$ を生成する(詳細は 3.2 項を参照)。 $wh+s$ が生成できない場合にはタイプ C のテンプレートのみを用いる。また、テンプレートによる生成には各種の適用条件が考慮される(詳細は 3.3 項を参照)。最後に、生成された問い返し文の複数の候補に対して n-gram の頻度を基にしたスコアを与え、スコアが最大のものを最終的な問い返し文とする(詳細は 3.4 項を参照)。

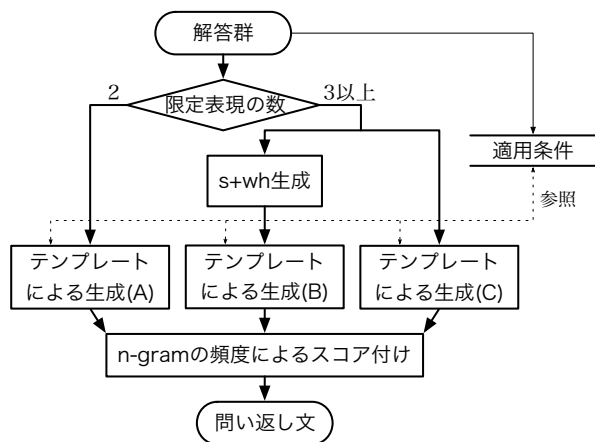


図 4: 問い返し文の生成処理

3.2 問い返し主題の生成

本項では問い返し主題 $wh+s$ の生成方法について述べる。まず、以下の手続きに従って問い返し主題の候補をいくつか生成する。

- 疑問詞 + 限定表現の共通末尾

解答群の限定表現 $\{s_i\}$ の末尾に共通する文字列を抽出する。この文字列の前に疑問詞「何」「何の」「どんな」をつけた表現を問い返し主題の候補とする。

(例) 図 2 の解答群の場合、限定表現の末尾には共通して「賞」という文字があるので、「何賞」「何の賞」「どんな賞」を生成する。

- 疑問詞 + 助数詞

解答群の限定表現 $\{s_i\}$ が全て「数+助数詞」という表現のとき、助数詞の前に「何」をつけた表現を問い返し主題の候補とする。また、助数詞が順序を表わす序数詞のときは「第何」を前につける。

(例) 限定表現が {20 回, 22 回, 25 回} のとき、「第何回」を生成する。

- 疑問詞 + 意味クラス

解答群の限定表現 $\{s_i\}$ の意味クラスが共通しているとき、意味クラス名に「何」「何の」「どんな」という疑問詞をつけた表現を問い返し主題の候補とする。意味クラスは角川類語辞典 [4] を参照した。

(例) 限定表現が {ソフトボール, サッカー, 野球} のとき、これらの意味クラスは 898d(球技) であるので、「何球技」「何の球技」「どんな球技」を生成する。

次に、生成された問い返し主題の候補に対して、ウェブ検索エンジンにおけるヒット数を調べ、最もヒット数が多い表現をひとつ選択する。最大ヒット数を持つ問い返し主題が複数ある場合はその全てをタイプ B のテンプレートにあてはめて問い返し文を生成する。ウェブ検索エンジンは Google API を用いた。

3.3 テンプレートの適用条件

生成テンプレートには以下に示す 3 種類の適用条件があり、これを満たすときに限り適用される。

曖昧なキーワード k がサ変名詞

図 3 の「サ変」の列が「○」であるテンプレートは k がサ変名詞であるときのみ適用される。

解答群の共通属性 $attr$

図 3 の「属性」の列は $attr$ に関する適用条件を表わす。例えば、タイプ C の 1 行目のテンプレートは $attr$ が「直前」、2 行目のテンプレートは $attr$ が「連体修飾」のときのみ適用される。

キーワードと限定表現の出現順序

図 3 の「順序」の列はキーワードと限定表現の出現順序に関する適用条件である。「 k 」はテンプレートにおいてキーワード k が限定表現 s_i (タイプ A のとき) または問い返し主題 $wh+s$ (タイプ B のとき) よりも前にあることを、「 s_i 」は逆に s_i や $wh+s$ が k の前にあることを示す。「順序」の制約があるテンプレートは、テンプレートにおける k と s_i (または $wh+s$) の位置関係が、解答群中の

解答候補 a_i を取り出した根拠文におけるキーワードと限定表現の位置関係と一致するときのみに適用可能である。ただし、解答候補毎に両者の出現順序が異なるときは多数決によって根拠文中における k と s の出現順序を一意に決定する。

この制約は、より自然な問い返し文を生成するために k と s_i (または $wh+s$) の順序を決定する働きをする。例えば、図 2 の解答群の場合、「アカデミー賞の何賞ですか?」と「何賞のアカデミー賞ですか?」という 2 通りの問い返し文が考えられるが、解答候補を取り出した根拠文の多くにおいてキーワードが限定表現の前に位置することから、前者の問い返し文の方が自然である。このとき、「順序」の制約から後者の問い返し文は生成されない。

3.4 n-gram の頻度によるスコア付け

テンプレートから生成された複数の問い返し文の候補に対し、式 (1) のスコアを計算し、それが最大となる文を最終的な問い返し文として選択する。

$$S(q_{sys}) = \frac{1}{|q_{sys}|} \sum_{n=1}^6 \left(w_n \cdot \sum_{c_1^n \in q_{sys}} O(c_1^n) \right) \quad (1)$$

q_{sys} は問い返し文を表わす。 c_1^n は q_{sys} に含まれる文字の n-gram, $O(c_1^n)$ はコーパスにおける c_1^n の出現頻度である。 w_n は n-gram の文字数 n に応じて与えられる重みである。すなわち、 $S(q_{sys})$ は問い返し文中に含まれる全ての文字 n-gram の頻度の重み付き和であり、問い返し文 q_{sys} の自然さを評価する指標である。すなわち、自然な問い返し文ほどコーパスにおける部分文字列の出現頻度が高いという考えに基づいている。

本研究では 1-gram から 6-gram の文字 n-gram の頻度を考慮した。重み w_n は $w_6 = 0.9$, $w_5 = 0.09$, $w_4 = 9 \cdot 10^{-3}$, $w_3 = 9 \cdot 10^{-4}$, $w_2 = 9 \cdot 10^{-5}$, $w_1 = 1 \cdot 10^{-5}$ とした。また、文字 n-gram の頻度を数えるコーパスとして毎日新聞の 13 年分の新聞記事を用いた。

4 予備実験

本節では提案手法を評価した予備実験について述べる。まず、曖昧な質問を 50 問用意した。これらの質問を我々の対話型質問応答システムの入力とし、解答群の作成ならびに問い返し文の生成を行った。

解答群の作成に関しては、50 問の質問のうち、適切な解答群が作成できた質問が 46 問、その順位が 1 位であった質問が 32 問であった。一方、松本らの手法 [3] では、前者が 35 問、後者が 17 問であった。この結果から、本研究で行った解答群生成手法の改良 (2.2 項) が大きな効果を挙げたことがわかる。

表 1: 実験結果 (問い返し文生成)

| | |
|-----------|----------|
| 問い返し文が適切 | 35 (76%) |
| 問い返し文が不適切 | 11 (24%) |

一方、問い返し文生成については、適切な解答群が得られた 46 個の質問に対して、その解答群を入力として問い返し文を生成し、それが適切かを人手で判定した。ここで適切な問い返し文とは、ユーザが問い返し文に答えることによって質問の曖昧性を解消でき、かつ文として自然であるという条件を満たす文とした。結果を表 1 に示す。実験の結果、76% の質問に対して適切な問い返し文を生成できることがわかった。問い返し文の生成に失敗した主な原因は、n-gram の頻度によるスコア付けが正しくない事例が 6 例、ウェブ検索エンジンのヒット数で選択した問い返し主題が不適切な事例が 1 例、キーワードと限定表現の位置関係による制約が適切な問い返し文の生成を妨げた事例が 1 例であった。これらの事例は現在のテンプレートでも適切な問い返し文を生成できていることから、問い返し文の候補のスコア付け手法などを改良することで、問い返し文生成の正解率の向上が期待できる。

5 おわりに

本論文では、対話型質問応答システムの構築を目標とし、その要素技術となるユーザへの問い返し文を生成する手法について述べた。実験では生成した問い返し文の内容や表層表現が適切であるかを評価したが、問い返し文を実際にユーザに提示し、ユーザから適切な返答を得て質問の曖昧性をうまく解消できるかについては検証していない。図 1 の処理 [F] の実装も含め、ユーザとの効果的なインタラクションを実現する手法を検討することが今後の課題である。

参考文献

- [1] 清田陽司, 黒橋禎夫, 木戸冬子. 大規模テキスト知識ベースに基づく自動質問応答-ダイアログナビ-. 自然言語処理, Vol. 10, No. 4, pp. 145-175, 2003.
- [2] Sadao Kurohashi and Wataru Higasa. Dialogue help-system based on flexible matching of user query with natural language knowledge base. In *Proceedings of the 1st ACL SIGdial Workshop on Discourse and Dialogue*, pp. 141-149, 2000.
- [3] 松本匡史, 白井清昭. 質問の曖昧性を検出し複数の解答を提示する質問応答システム. 言語処理学会第 12 回年次大会発表論文集, pp. 935-938, 2006.
- [4] 大野晋, 浜西正人. 角川類語辞典. 角川書店, 1981.
- [5] 坂本篤史. 対話型質問応答システムにおける問い返し文の生成に関する研究. Master's thesis, 北陸先端科学技術大学院大学, 3 2007.