

シリーズ型質問に回答する質問応答システム Metis2013

濱田大和[†] 太田雄介[†] 松村冬子[‡] 原田実[‡]
 青山学院大学理工学部情報テクノロジー学科^{†‡}

1. はじめに

近年の IT の急激な発展によって、大量の文書データからの知識の発掘（テキストマイニング）などの分野で、文章の意味解析への期待が高まっている。我々はこれまでに、意味解析システム Sage[1]を用いて意味グラフベースで質問文と知識文を照合し、類似度の高い知識文から回答抽出をする質問応答システム Metis[2,3]を開発している。しかし、対話における質問応答を想定して、文脈に沿って関連する内容を逐次的に連続して回答するシリーズ型質問には対応していなかった。シリーズ型質問とは、NTCIR などのコンテストで行われている質問で、「日本で一番高い山はなんですか。」の質問に続いて提示される、「その高さは何メートルですか。」のような前の質問やその回答などの文脈を考慮することで質問内容が完結する質問を指す。本研究ではこのシリーズ型質問に回答可能な Metis2013 の開発を目指す。

2. Metis2013 の概要

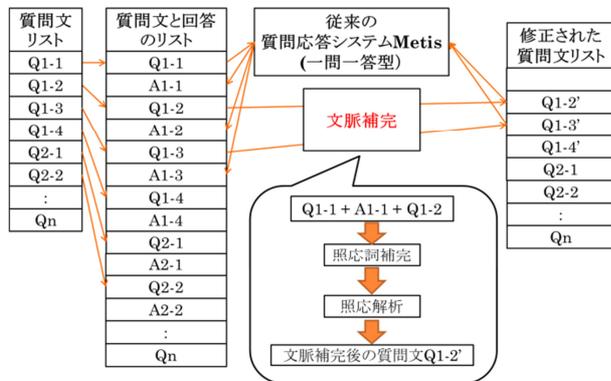


図1 Metis2013 のシステム構図

Metis2013 ではシリーズ型質問文を扱うので質問文をリストで扱っている。従来の Metis は 1 文で完結している質問にしか回答できないので連続している質問文は、入力する前に文脈補完を行う。文脈補完とは、与えられている質問文に適切な照応詞を補完し、それらを照応解析システム Anasys[4]に入力し、照応詞をその先行詞で置き換え

た質問文を作成する。照応解析とは、文中の照応詞(指示代名詞や人称代名詞など)の指示対象を推定するほか、省略された名詞節(ゼロ代名詞)を補完する解析を指す。文脈補完により作成された質問文を従来の Metis に入力し回答を抽出する。この作業をそれぞれの質問シリーズ毎に行う。例えば、Q1-1:「日本で一番高い山はなんですか。」、A1-1:「富士山」、Q1-2:「高さは何メートルですか。」のときに、これらの質問文と回答をリストに加える。次に Q1-1、A1-1、Q1-2 を文脈補完する。Q1-2 に照応詞「その」を補完することで、Q1-2':「その高さは何メートルですか。」となる。次に照応解析により「その」の先行詞が「富士山」となり、Q1-2':「富士山の高さは何メートルですか。」が作成される。この質問文を従来の Metis に入力し回答を得る。

3. 文脈補完

Metis では、まず各質問文から 4 つの特徴文節（質問ノード、主述語、主体、主題）を抽出する。質問ノードとは、質問文中での疑問詞を指す。主述語とは、質問中で問われている対象が行う行為を表す述語のことを指す。主体とは、主述語の主語となるものを指す。主題とは、質問ノードで問われている対象を表す。これらの概念をもとに、文脈補完を行う。シリーズ型質問では、連続して質問が出されるので前の質問文もしくは回答の内容を省略している質問文が多い。そこで、我々はこの省略している内容を補うことで従来の Metis が回答できるように以下のルールによって文脈を補完する。

3.1. 照応詞がある場合

質問文の形態素に「それ」、「その」、「この」、「これ」、「彼」、「彼女」などの照応詞がある場合は、連続した質問文と回答を照応解析システム Anasys に入力し照応詞に対応する先行詞を推定し補完する。例えば、Q1-1:「境港に入港し、抗議行動が起こった駆逐艦の名前は何か。」、A1-1:「クッシング」、Q1-2:「また、それはどこの船ですか。」のシリーズ質問がある。照応解析に Q1-1、A1-1、Q1-2 を入力し、照応詞「その」に対応する先行詞を推定する。ここでは「クッシング」が推定できる。その結果、Q1-2':「また、クッシングはどこの船ですか。」と前文の内容を補完した質問文が作成できる。

3.2. 照応詞がない場合

照応詞がない場合は、質問文中に適切な照応詞を加え、照応解析することで完結した文を作成する。

Question answering system Metis2013 revised for answering question series

Yamato Hamada[†], Yusuke Ota[†], Fuyuko Matsumura[‡] and Minoru Harada[†]

[†]Undergraduate school of Integrated Information Technology, Aoyama Gakuin University.

[‡]Department of Integrated Information Technology, Faculty of Science and Engineering, Aoyama Gakuin University.

A. 質問文に共起関係詞「は」があり、質問ノードの係り側ノードの文節末が「のは」の場合

1. 質問ノードの係り側ノードの品詞が「断定名詞節」: 質問ノードの係り側ノードの前に「その」を補完する。
2. 質問ノードの係り側ノードの品詞が「動名詞節」、「形容名詞節」、「形容動名詞節」で、質問ノードの係り側ノードにかかる名詞節があり、その名詞節にかかる修飾節が少ない:

質問ノードの係り側ノードにかかる名詞節の前に「その」を補完する。

3. 質問ノードの係り側ノードにかかる名詞節がなく、質問タイプが MONO、HITO:

文頭に「それを」を補完する。

4. 質問ノードの係り側ノードにかかる名詞節がなく、質問タイプが TOKI、BASYO、NANIX:

文頭に「それが」を補完する。

例えば、Q1-1:「ミュージカル「ライオンキング」の上演は何年に始まりましたか。」、A1-1:「1997年」、Q1-2:「台本を書いたのはだれですか。」のとき、質問ノード「だれですか。」の係り側ノード「書いたのは」は動名詞節で係り側ノードにかかる名詞節「台本を」があるので、Aの2つ目のルールに当てはまり、照応詞「その」を補完し、Q1-2:「その台本を書いたのはだれですか。」が作成できる。Q1-1、A1-1、Q1-2を照応解析にかけることで、Q1-2:「ライオンキングの台本を書いたのはだれですか。」に補完される。

B. 主述語の主格がない、または質問文に主格がない、または主述語にかかる共起関係詞「が」がある場合

文頭に「それは」を補完する。

例えば、Q1-1:「棋士・佐藤康光の生年月日はいつですか。」、A1-1:「1969年10月1日」、Q1-2:「どこで生まれましたか。」のとき、質問文に共起関係詞「は」がなく、質問ノードの係り側ノードの文節末が「のは」ではない。また、主述語「生まれましたか。」の主格がないので、照応詞「それは」を文頭に補完し、Q1-2:「それはどこで生まれましたか。」が作成できる。Q1-1、A1-1、Q1-2を照応解析にかけることで、Q1-2:「棋士・佐藤康光はどこで生まれましたか。」に補完される。

C. その他の場合

上の規則に当てはまらない質問文に関しては、文頭に「その」を補完して照応解析することで完結な質問文を作成する。

4. 実験及び評価

NTCIR-3(QAC1)の task3 は、2 個の質問からなるシリーズ型質問のセットが 40 個で 80 問、NTCIR-4(QAC2)の task3 は、5 から 10 問の質問からなるシリーズ型質問のセットが 36 個で 251 問である。これらの質問データに基づき、毎日新聞記事 1998 年から 1991 年までの 2 年分

を知識源として評価実験を行う。回答精度の尺度には情報検索システムの検索性能の評価を示す F 値を使う。F 値の計算方法は以下に示す。

$$\text{適合率 } Precision = \frac{\text{質問文より抽出された正しい回答の数}}{\text{質問文より抽出された回答の数}}$$

$$\text{再現率 } Recall = \frac{\text{質問文より抽出された正しい回答の数}}{\text{質問文に対する正しい回答の数}}$$

$$\text{F 値 } F = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

NTCIR-3(QAC1)、NTCIR-4(QAC2)のテストコレクションで実験を行った結果と過去のコンテストで成績の良かった 1 位、2 位の F 値を表 1 に示す。

表 1 評価実験の結果と他研究との比較

F 値	1 位	2 位	Metis2013
QAC1	0.23	0.19	0.22
QAC2	0.22	0.20	0.22

5. おわりに

質問応答システム Metis2013 ではシリーズ型質問に回答することで、より対話的な環境で質問応答できることを示した。今回の実験では、QAC2 で他研究の 1 位と同じ結果となったが、QAC1 では上回ることができなかった。

より回答精度の高いシステムにするためには、従来の Metis の回答精度の向上やシリーズ型質問がしっかりと補完できていないなどの問題点を見直していく必要がある。

参考文献

[1]原田実,水野 高宏:“EDR を用いた日本語意味解析システム SAGE”, 人工知能学会論文誌, Vol.16, No.1, pp.85-93 (2001.1).
 [2] Dongli HAN, Yuhei KATO, Kazuaki TAKEHARA, Tetsuya YAMAMOTO, Kazunori SUGIMURA, and Minoru HARADA: “QA System metis based on web searching and semantic graph matching”, Proc. of the 4th International Conference on Intelligent Information Processing, Adelaide, Australia, pp.123-133,(2006).
 [3]原田実, 西岡晋太郎: “深層格とオントロジーを用いた回答抽出による質問応答システム Metis の回答精度向上”,情報処理学会第 73 回全国大会論文集,1D-1, (2011.3).
 [4] 山田和正,松田源立,原田実: “照応解析システム ANASYS の精度向上-外界照応分離, 素性の洗練, ナイーブベイズ法の導入-”,情報処理学会研究報告, Vol.2010-NL-196, No.15, pp. 1-8, (2010.5).