

# WHY 型 QA システムにおける回答抽出方法の改良

大友謙一 村上仁一 徳久雅人 池原悟

鳥取大学 工学部 知能情報工学科

{s052014,murakami,tokuhisa,ikehara}@ike.tottori-u.ac.jp

## 1 はじめに

### 1.1 情報検索と質問応答技術

現在、検索 KW を用いた情報検索が主流となっている。しかし初心者にとって、欲しい情報を得るために検索 KW を作成し、情報を抽出する事は困難である。この問題を解決するため、自然言語で質疑応答を行う技術が研究されている [1]。本技術は、自然言語で記した質問を用いて検索 KW とフィルタを作成し、検索エンジンより得た情報に対しフィルタリングを行い回答を抽出するものである。

### 1.2 質問応答技術に対する研究の流れ

質問応答の分野において質問形態は二つに分類される。名称や日付け・数値など事実に基づく回答を求める factoid 型質問と、理由や事象の説明に基づく回答を求める non-factoid 型質問である。質問応答技術を評価するワークショップとして、国立情報学研究所主催の NTCIR(NII Test Collection for Information Retrieval and Text Processing) の QAC(Question Answer Challenge) タスクがある。当初 QAC タスクは factoid 型質問を対象とした。QAC4(2006.4-2007.5) にて why や how といった non-factoid 型質問を研究対象とした。その後 non-factoid 型質問の QA システムの研究が行われている [1]。

### 1.3 本研究の目的

先行研究にて本研究室は全文検索を用いた WHY 型質問に対する質問応答システムを作成した [2]。本システムでは WEB 上のドキュメントをソースとし、検索エンジンとして Google を使用している。特定の質問に対して精度の高い返答を行うが、質問によって、回答を出力しない場合や、正答とは異なる回答を出力する場合がある。

本研究では以上の問題を踏まえ、先行研究のシステムに対し回答出力の過少の改善と正答出力の精度の向上を行う。

## 2 Why 型質問応答システム

本研究における Why 型質問応答システムとは、自然言語で記述された「なぜ」または「どうして」を含む質問に対し適切な返答を出力するシステムである。本システムは質問応答技術として代表的な「質問文解析」、「情報検索」、「回答候補抽出」、「回答候補ランキング」という 4 つのモジュールで構成している。動作の行程を図 1 に示す。

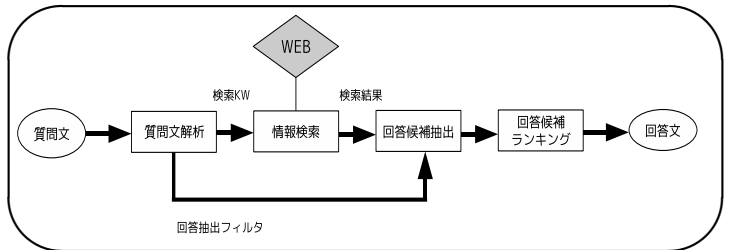


図 1 システム概要図

### 2.1 質問文解析

与えられた質問文に対し、解析、言い替えを行い、情報検索に用いる「検索 KW」を作成する。同時に情報検索で得た文章をフィルタリングし回答を抽出する「回答抽出フィルタ」を作成する。作成例を図 2 に示す。(詳細は先行研究 [2] を参照。)

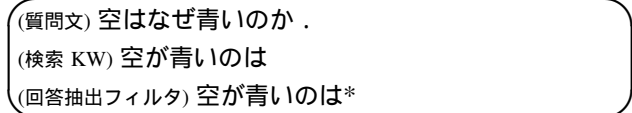


図 2 検索 KW 及び回答抽出フィルタ作成

### 2.2 情報検索

質問文解析モジュールで作成した検索 KW を検索エンジンに与え、情報検索を行う。

### 2.3 回答候補抽出

情報検索モジュールで得た文字情報に対し、質問文解析モジュールで作成した回答抽出フィルタを用いてフィルタリングを行い回答候補を抽出する。

### 2.4 回答候補ランキング

回答候補抽出モジュールによって複数の回答を得る。抽出後、「検索結果には質問に回答する情報が多く含まれている」という仮説の元、名詞の出現頻度を指標としたランキングを作成する。

## 3 改良点

### 3.1 「全文検索」から「類似検索」への変更

先行研究において検索エンジンの検索方式は「全文検索」である。「全文検索」は検索 KW と一致する文章のみ抽出する。そのため抽出可能な回答が限られる。そこで検索方式を「類似検索」へと変更する。「類似検索」は検索 KW に似た文章を持つ記事も抽出する検索方式である。

### 3.2 回答抽出フィルタの改良

#### 3.2.1 先行研究の回答抽出フィルタ

先行研究において回答抽出フィルタは文節の形態をとる。図3にフィルタの一例を示す。

(質問文) 携帯電話の料金はなぜあんなに高いのか  
(フィルタ) 料金があんなに高いのは\*

図3 回答抽出フィルタの一例

#### 3.2.2 先行研究の回答抽出フィルタの問題

先行研究にて用いられた回答抽出フィルタには様々な問題がある。以下に問題を示す。

##### A. 助詞の差異

文節の形式をとる回答抽出フィルタでは、助詞の違い等で単語が1文字でも異なる文章は抽出できない。図4に「に」の不一致による回答抽出失敗の例を示す。

(先行研究の回答抽出 PT) 料金があんなに高いのは\*  
(回答) 電話料金があんな高いのは基地局維持のため。

図4 助詞の不一致による抽出失敗例

##### B. 類語等の単語の差異や単語の欠落

先行研究のシステムでは、類語等の単語の差異や、単語の欠落で回答候補抽出モジュールに抽出されなかった文章でも、回答候補として成り立つ事がある。図5に例を示す。

(先行研究の回答抽出フィルタ) 雲が目に見えるのは\*  
(回答) 雲が見えるのは光の拡散のため。

図5 回答抽出 PT の回答抽出失敗例

したがって、回答抽出フィルタが文を抽出する際に、質問文中の単語をそれぞれ任意に参照するならば、単語の違いによる回答抽出失敗の対策になる。

##### C. 質問文中の単語を用いた任意抽出

任意の単語を抽出する回答抽出フィルタは、膨大な数の文を回答候補として抽出する。一方、理由を示す文には理由を示す文字パターンが含まれやすい事が挙げられる。そこで理由を示す文字パターンを含む文を抽出する事で、膨大な数の回答候補から正答文を抽出出来ると考える。

正答文に「のは」という文字パターンが含まれる事が多い。そこで「のは」という文字パターンが必ず含まれる文を抽出する。

#### 3.2.3 本研究において提案する回答抽出フィルタ

本研究では質問文に含まれる名詞、形容詞及び動詞を任意に含み、「のは」という文字パターンを含む文章を抽出する回答抽出フィルタを使用する。回答抽出フィルタの例を図6に示す。

(質問文) 携帯電話の料金はなぜあんなに高いのか  
(回答抽出フィルタ) (携帯電話 or 料金 or 高い)+”のは”

図6 回答抽出フィルタ改良例

## 4 実験

### 4.1 実験条件

実験には学研サイエンスキッズ [3] の「科学なぜなぜ110番」の自然ジャンルを参考に25件、雑学のすゝめ [4] の「素朴な疑問」を参考に75件、計100件の質問文を用いる。また質問文作成の参考にしたサイトから正解を含む部分を正解文として引用する。形態解析器として MeCab[5] を用いる。

### 4.2 評価

#### 4.2.1 回答取得成功率

与えられた質問文100件に対し、システムが回答を得た割合を示す。

#### 4.2.2 人手評価

本システムの精度を確認するために人手による評価を行う。ランキング1位の回答候補に対して“A”, “B”, “C”, “D”で評価する。評価基準を図7に示す。本研究では評価“A”, “B”に当たる回答を正答とする。

- A : 回答が正答を完全に網羅している。
- B : 回答が正答をほぼ網羅している。
- C : 回答が質問内容に関する記述を含む。
- D : 回答が質問内容よりかけはなれている。

図7 人手評価

#### 4.2.3 F値

人手評価をもとにF値 [6] を求める。式を図8に示す。

$$Recall = \frac{\text{正答数}}{\text{質問数}} \quad Precision = \frac{\text{正答数}}{\text{出力数}}$$
$$F \text{ 値} = \frac{2 * Recall * Precision}{Recall + Precision}$$

図8 F値

#### 4.2.4 検索成功率

本指標は回答ランキング上位10位以内に正答を含む目標の達成割合を表す。

#### 4.2.5 Mean Average Precision

システムの精度を示す指標として MeanAveragePrecision(MAP)[7] を用いる。MAPはシステムが正答数を網羅している割合を示す。

$$AP_n = \frac{k_n}{RANK}$$

(RANK...正答数,  $k_n$ ...出力された正答数)

$$MAP = \frac{1}{n} \sum AP_n$$

図9 MeanAveragePrecision(MAP)

### 4.3 実験結果

表 1 に改良前及び改良後システムの実験結果を示す。

表 1 改良前及び改良後システムの実験結果

	改良前	改良後
回答取得成功率	24%	96%
F 値	0.080	0.177
検索成功率	8%	39%
MAP	0.03	0.05

改良後は回答取得成功率が 70% 以上上昇しており多くの回答を抽出出来る事が分かる。また F 値及び検索成功率の値が改良前を上回っている。以上の事から、提案手法の有効性を確認した。

### 4.4 改善例

回答抽出フィルタの改良により回答出力が改善した例を図 10、図 11 に示す。

回答抽出フィルタ改良による回答未出力解決の例

質問:飛行機雲はどうしてできるの  
 回答例:排気ガスが核となりこのまわりに水分が集まり水滴となるから。  
 (改良前システムの出力):出力無し  
 (改良後システムの出力):飛行機雲 飛行機の後ろに飛行機雲が出るのは、エンジンからの排気ガスや水蒸気が核となって、機体周辺の気圧変化から水滴が出るため。

図 10 回答出力改善例 その 1

回答抽出フィルタ改良による回答内容向上の例

質問:酸性雨はどうして降るの  
 回答例:二酸化窒素を空気中の水分が吸収し、それが雨が降る際に一緒に降るから。  
 (改良前システムの出力):酸性雨が降るのはそのためです。  
 (改良後システムの出力):酸性雨というのは、空気中の水分が冷えて雲ができるときや、雲から雨が降ってくる途中で空気の中にある二酸化硫黄や二酸化窒素といった物質が溶けこんで酸性になった雨のことです。

図 11 回答出力改善例 その 2

### 4.5 結果の分析及び改善点

4.3 節、表 1 の改良後システムの検索成功率は 39% と低い。この原因として、正答が回答抽出モジュールで抽出されなかったからと考えている。検索成功率の低さを改善するため「より多くの正答を抽出する工夫」を行う。

## 5 より多くの正答を抽出する工夫

「より多くの正答を抽出する工夫」として「理由を示す他の文章パターンの利用」を行う。

### 5.1 理由を示す他の文字パターン「から」「ため」

3.2.3 節では、理由を示す文字パターン「のは」を含む文を抽出対象とする回答抽出フィルタを用いた。本節では「のは」の他に「から」と「ため」が理由を示す文字パターンであると考えられる。

### 5.2 「から」「ため」の有効性調査実験結果

理由を示す文字パターン「から」もしくは「ため」を単独で用い、質問文 100 件を用いて回答抽出を行った。

結果を表 2 に示す。

表 2 「から」もしくは「ため」を用いた回答抽出の実験結果

	「のは」	「から」	「ため」
F 値	0.177	0.071	0.114
検索成功率	39%	23%	27%
MAP	0.05	0.04	0.06

「から」や「ため」を用いた場合の F 値や検索成功率が、「のは」を用いた場合の値より低い。しかし「のは」で得られなかった正答が得られた。具体的な数を表 3 に示す。

表 3 「のは」では正答が得られなかった質問数

から	12
ため	7

表 3 の結果より、理由を示す文字パターンごとに得られる回答が異なる事が分かった。

### 5.3 理由を示す文字パターン複数使用有効性調査実験

5.2 節の、理由を示す文字パターンごとに得られる回答が異なるという結果から、「のは」・「から」・「ため」の 3 つの文字パターン全て用いて回答抽出を行った場合に精度が向上すると考える。3 つの文字パターン全てを用いて実験を行った結果を表 4 に示す。

表 4 「のは」「ため」「から」を用いた回答抽出の実験結果

F 値	0.09
検索成功率	43%
MAP	0.06

各評価値を 4.3 節の表 1 の値と比べると F 値が 0.09 に下がったが検索成功率は 5% 上昇した。この結果から、正答がランキング 1 位にある質問数は減ったが、ランキング 10 位以内に正答がある質問数は増加した事が分かる。つまり理由を示す文字パターンを複数用いる事により、ランキングの精度は下がるが正答抽出数は増加する事が分かった。

## 6 考察

### 6.1 言い替えによる質問文の再構築

正答の得られない質問に対し、言い替えを行った場合正答数が増加すると考える。図 12 に例を示す。

(質問文) オーロラができるのはなぜ  
 (言い替えた質問文) オーロラが発生するのはなぜ  
 (正答一例) オーロラがこの領域でよく発生するのは、オーロラ発光の原因であるプラズマ粒子がほぼ磁力線に沿って動くという性質を持っているから

図 12 言い替え例

#### 6.1.1 再構築の方法

似たような意味を持つ語を用いて、質問文の言い替えを手で行う。変更内容を図 13 に示し、具体例を図 14 に示す。

- 似たような意味を持つ動詞を用いて言い替え．
- 似たような意味を持つ名詞を用いて言い替え．
- 動詞を，似たような意味の名詞に変更．
- 漢字を正しいものに訂正．
- 代名詞など不明確な表記を明確な表記に変更．

図 13 言い替え方法一覧

(質問文) どうして海には引き潮と満ち潮があるの  
 (再構築後質問文) どうして海では引き潮と満ち潮が起ころの

図 14 言い替えによる質問文の再構築の例

### 6.1.2 言い替えによる質問文の再構築の実験結果

質問文 100 件の中で 6.1.1 節に挙げた言い替えを行う事の出来る質問文は 55 件存在した．言い替えを行った後，再びシステムを用いて回答を求めた．この結果，55 件のうち 5 件，新たに正答の出力を確認した．表 5 に手法別の件数を示す．

表 5 出力を確認した言い替え手法

似たような意味を持つ動詞に言い替え	1
動詞を同義の名詞に変更	1
漢字を正しいものに交換	1
不明確な表記を明確な表記に変更	2

以上の結果から，言い替えにより正答抽出数が増加する事が分かった．本節では手動にて質問の言い替えを行った．今後，言い替えを自動で行う手法を確立する事で QA システムの精度向上に貢献すると考えている．

### 6.2 正解を抽出したサイトの分布

今日，Wikipedia[8] や QA サイト [9] といった，個人の疑問に対する回答を記載するサイトがある．そこで，4.3 節の表 1 の改良後のシステムの回答抽出先の調査を行う．分類は，HP (個人，団体問わず)・QA サイト・ブログ・掲示板とする．結果を表 6 に示す．

表 6 正解を抽出したサイトの分布

HP(個人・団体問わず)	QA サイト	ブログ	掲示板
18	8	11	1

表 6 より QA サイトだけでなく多くのサイトから正答を抽出している事が分かる．

一般的に，回答を記載するサイトからの抽出が多い場合，検索対象サイトを限定する事により精度の向上を計ることが出来ると考えられている．しかしこの結果より，現在では対象サイトを限定することは出来ない事が示された．

### 6.3 質問文から作成する回答抽出フィルタの限界

本研究において提案した改善手法を用いても検索成功率は 5 割弱である．5 割強の質問においてランキング内に正答が存在しない．質問文より作成した回答抽出フィルタでは抽出できない正答が存在する．本研究で用いる回答抽出フィルタで回答が抽出出来ない例を図 15 に示す．

(質問文)  
 3 ポイントシュートができたのはどうして  
 (回答抽出フィルタ)  
 (3 ポイントシュート or でき)+” のは or から or ため”  
 (正答)  
 理由についてはいろいろですが、インサイドばかりでは・・・というのとプロのリーグでは現在の NBA と前記の ABL や CBA (NBA の対抗リーグ) はあまり関係が良くなって差別化を図ると同時にショー的要素を強めるという理由で導入されたというのもあるみたいですね。

図 15 本研究の回答抽出フィルタで回答が抽出出来ない例

この例では質問文中の単語が一切含まれず，理由を示す文字パターンも含まないため正答を抽出できない．

以上より質問文を回答抽出フィルタに変換して回答を抽出する手法には限界があると考えられる．質問文を変換した回答抽出フィルタは回答が記載されているページを特定する事に用いて，特定したページの文章構造から回答を抽出する手法を確立する必要があると考えられる．

## 7 おわりに

先行研究 [2] で作成したシステムに対する改良を行い，F 値が 0.01 上昇し，検索成功率は 35% 上昇した．回答抽出フィルタとして，質問文中の任意の数の単語と理由を示す文字パターンを用いることは回答抽出に対し有効である．だが正答に質問文の単語が含まれない事もあり，質問文を用いた回答抽出 PT には限界があると考えられる．今後，質問文の単語に左右されない回答抽出法を考える必要がある．

## 参考文献

- [1] 諸岡 心ら:非 Factoid 型質問に対応した質問応答システム, 言語処理学会第 13 回年次大会, pp.958-961, (2007)
- [2] 田村 元秀ら:web 検索エンジンを用いた Why 型質問応答システムに関する研究, 言語処理学会第 14 回年次大会, pp.1021-1024, (2008)
- [3] 学研サイエンスキッズ  
<http://kids.gakken.co.jp/kagaku/index.html>
- [4] 雑学のすゝめ  
<http://homepage1.nifty.com/tadahiko/INDEX.HTML>
- [5] MeCab:Yet Another Part-of-Speech and Morphological Analyzer  
<http://www.chasen.org/taku/software/mecab/>
- [6] ウィキペディア—情報検索  
<http://ja.wikipedia.org/wiki/%E6%83%85%E5%A0%B1%E6%A4%9C%E7%B4%A2>
- [7] 検索メインアック!—地図ではない MAP  
<http://voice.fresheye.com/sakai/2007/04/map.html>
- [8] ウィキペディア フリー百科辞典  
<http://ja.wikipedia.org/wiki/>
- [9] OKWave etc..  
<http://okwave.jp/>