

# 回答の根拠を提示するヘルプデスク型質問応答システム\*

佐々木智 (学籍番号 200721534)

研究指導教員：藤井敦

## 1. はじめに

インターネットの普及に伴い、多種多様な情報が Web に発信されるようになった。大量の Web 文書から有益な情報を抽出する代表的な手法としてキーワード検索がある。しかし、「情報要求によってはキーワードで表現しにくい」、「文書中のどこに欲しい情報があるか分かりにくい」という問題があり、これらを解決する手段として質問応答(QA)がある。

QA には、組織化された情報を用いて回答を推論する「推論型」と、組織化されていない文書集合から回答を抽出する「抽出型」がある。推論型 QA は情報の組織化が高価であり、回答できる分野が限定される。そこで、近年は抽出型 QA が注目されている。抽出型 QA は、対象とする回答の種類によって手法を分類することができる。名称や数値などの客観的事実を答える「factoid 型」、行動や手順を答える「how 型」、原因や根拠を答える「why 型」などがある。本研究は how 型 QA に焦点を当てる。

how 型 QA の研究事例として、三原ら [1] の「ヘルプデスク型 QA」は、行動を問う質問に対して、名詞句と動詞からなる「行動表現」を回答として出力する。例えば、「蜂に刺されたらどうすればいい?」という質問には、「針を抜く」や「流水で洗う」などの行動表現が回答となる。

しかし、三原らのヘルプデスク型 QA は、なぜその行動を取るべきか根拠を提示しないため、ユーザは行動の適否を判断することが難しい。渡辺ら [2] や Moriら [3] も取るべき行動を回答する QA の手法を提案している。しかし、根拠も共に答える手法は存在しない。以上より、本研究は、取るべき行動を根拠と共に答える QA システムを提案する。

## 2. 提案する QA システム

「蜂に刺されたらどうすればいい?」という how 型の質問をした人が「流水で洗う」ことの根拠を知りた

い場合、「なぜ蜂に刺されたら流水で洗うのか?」といった why 型の質問をするだろう。そこで、how 型と why 型の QA を組み合わせることで、取るべき行動を根拠と共に答える QA システムを提案する。当該システムの構成を図 1 に示す。ユーザが質問を入力すると、「how 型 QA」、「根拠を問う質問文の生成」、「why 型 QA」、「回答の統合」の順に処理を行う。最終的に、行動表現と根拠を組み合わせられた回答が出力される。

### 2.1 how 型 QA

ユーザが「蜂に刺されたらどうすればいい?」といった行動を問う質問を入力すると、その質問をクエリとして Web を検索する。収集された Web 文書から「患部を洗う」や「針を抜く」など「名詞+助詞+動詞」という係り受け表現を抽出し、回答の候補として出力する。回答の候補にスコアを付けて順位付けするために、ヘルプデスク型 QA [1] の手法を用い、さらに以下の拡張を行った。

- 「他人に聞く」や「情報を集める」などの名詞が 2 文字以下である行動表現はユーザの情報要求を満たす可能性が低いいため、スコアを下げる。
- 「トップに戻る」などの Web に頻出する表現は誤答であるにも拘らず抽出されやすい。そこで、Web に頻出する表現のリストを人手

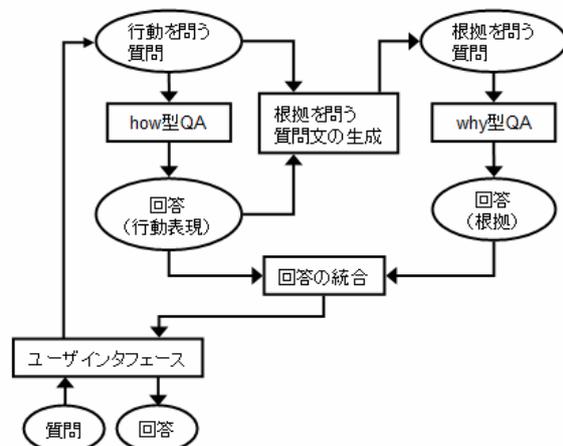


図 1: 本研究で提案する QA システムの構成

\* “A Helpdesk-style Question Answering System Presenting Grounds for Answers” by Satoru SASAKI

で作成、登録された表現はスコアを下げる。

- 助詞が「が」や「は」である行動表現は誤答である場合が多いため、スコアを下げる。

## 2.2 根拠を問う質問文の生成

how 型 QA で得られた行動表現に対し、その行動の根拠を問う why 型の質問文を生成する。例えば、how 型 QA において、「蜂に刺されたらどうすればいい?」という質問が入力され、「患部を洗う」という回答が出力された場合、「なぜ蜂に刺されたら患部を洗うのか?」という why 型質問文を生成する。

## 2.3 why 型 QA

why 型質問文が生成されると、その質問文をクエリとして Web を検索する。収集された Web 文書から、「患部を洗うと毒を洗い流せる」といった根拠を抽出し、回答の候補として出力する。why 型 QA は、渋沢ら [4] の手法を拡張して用いた。

## 2.4 回答の統合

how 型 QA で得られた行動表現と、why 型 QA で得られた根拠を一つの回答として統合する。ここで、根拠が得られなかった行動表現は不適切な回答としてスコアを下げる。

## 3. 実験評価

「2.1 節で提案した how 型 QA の拡張」と「how 型と why 型の QA を統合」によって行動表現の順位付けがどの程度改善されたのかを評価した。評価尺度は、以下の式で計算される精度を用いた。

$$\text{精度} = \frac{\text{上位 } N \text{ 件の正解数}}{N}$$

回答数  $N$  による精度の推移を調べるために、出力する行動表現を段階的に増やしていき、 $N$  の値ごとに精度を計算した。結果を図 2 に示す。グラフの横軸は行動表現の件数、縦軸は精度を表す。図 2 中の系列 A、B、C は以下の手法に対応する。

A: 三原ら [1] が提案した手法

B: 2.1 節に基づいて A を拡張した手法

C: A に why 型 QA を統合した手法

図 2 の A と B を比較すると、全体的に B が優位である。2.1 節で提案した拡張手法は有効であった。しかし、A と C を比較すると、全体的に C は精度が低い。正解の行動表現に対して、「Web 検索において、根拠が書かれた Web 文書が得られない」、

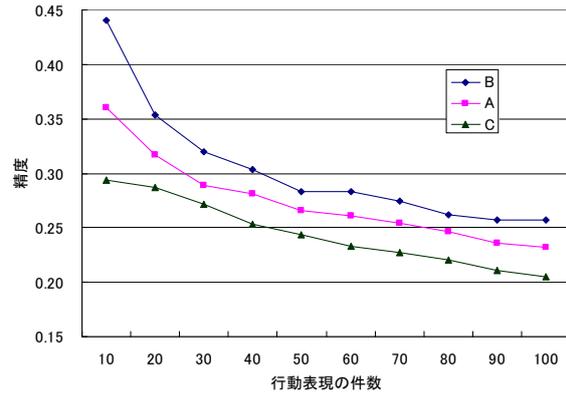


図 2: 異なるスコア付け手法の比較

「抽出された根拠のスコアが低い」という why 型 QA 固有の誤りによって、精度が下がった。

## 4. おわりに

従来の how 型 QA には、取るべき行動を回答してもその行動を取る根拠を提示しないという問題があった。そこで本研究は、how 型 QA と why 型 QA を統合し、取るべき行動を根拠と共に答える QA システムを提案した。また、提案手法の有効性を実験によって部分的に評価した。残された課題は、システムを様々な観点から評価することである。

## 文献

- [1] 三原英理, 藤井敦, 石川徹也. Web を用いたヘルプデスク指向の質問応答システム. NLP2005 発表論文集, pp. 1096-1099, 2005.
- [2] 渡辺靖彦, 園和也, 岡田至弘. メーリングリストを利用した質問応答システムのための知識獲得. 情報処理学会研究報告, 2004-NL-162, pp. 131-138, 2004.
- [3] Mori Tatsunori, Okubo Takuya, and Ishioroshi Madoka. A QA system that can answer any class of Japanese non-factoid questions and its application to CCLQA EN-JA task. *Proceedings of the 7th NTCIR Workshop Meeting*, pp.41-48, 2008.
- [4] 渋沢潮, 林貴宏, 尾内理紀夫. Why 型質問の回答文を Web ページから抽出するシステム RE:Why の試作. *コンピュータソフトウェア*, Vol.24, No.3, pp.20-28, 2007.