

## 質問応答システムにおける回答候補の選定手法

丹羽達洋 福本淳一

立命館大学大学院理工学研究科

{t\_niwa, fukumoto}@nlp.cs.ritsumei.ac.jp

## 1 はじめに

質問応答(QA)とは、自然文で記述された質問文に対して、大量な文書データからの確かな答えのみをユーザに返す技術である。例えば「青色発光ダイオードを開発した人は誰ですか」という完全な形の質問文を入力とし、大量な文書データの中から「中村修二」を特定し回答として返す。現在のQAでは、与えられた質問から文書を検索するためのキーワード抽出や質問文解析、文書検索、回答出力という処理過程が一般的となっている。

回答出力の段階では、検索された文書に固有表現抽出の技術を適用することが多く、認識される固有名の種類は、IREX[1]における固有表現抽出課題で示された8種(人名,地名,組織名,固有名,日付表現,時間表現,金額表現,割合表現)がよく用いられる。しかしながら、分野限定をしない多種多様な質問を扱うQAでは、対応できないものがある。

日本語質問応答技術の評価プロジェクト Question Answering Challenge (QAC-1)[2]で実際に使用された質問文を分析したところ、物の長さや重さ、場所の面積、人物の年齢、人数など多様な数値表現の回答が求められている。数値表現を回答に持つ質問に答えるために、山下ら[3]は、文章から数値、物事、属性の関係を取得する手法を提案している。しかし、数値情報構造では対応できないものが存在するため、十分な精度が得られていない。

山下らの研究でも述べられているように、数値は対象としている物事との関係があって有益な情報となっている。したがって、質問文においても、単に回答が数値表現であるということだけでなく、「何についての数値なのか」という対象を考慮することが正確な回答を得る手掛りとなる。

本研究では、数値に付随する助数詞などの表現と共に起する単語情報と、質問内容を表す語との照合により回答候補を絞り込む手法を提案する。

## 2 数値表現の分類

数値表現を数詞または助数詞を伴う場合はそれも含む文字列と定義する。

数値表現がどのような対象を表すのに用いられているかを分類するため、3年分の新聞記事の中に現れる助数詞を調査し、18種類の数値表現に分類した。その結果を表1に示す。

表1: 数値表現の分類

数値表現の種類	対応する助数詞	例
金額 (MONEY)	円, ドル, ユーロなど	82 兆 200 億円
長さ (LENGTH)	メートル, km など	1000km
重さ (WEIGHT)	グラム, トンなど	5 トン
面積 (AREA)	ヘクタール, 坪など	61 ヘクタール
体積 (VOLUME)	リットル, cc など	50cc
容量 (CAPACITY)	GB, メガバイトなど	12GB
年齢 (AGE)	歳, 才	78.07 歳
人数 (NPERSON)	人, 名	一億二千万人
割合 (PERCENT)	割, %, パーセント	0.001 %
世紀 (CENTURY)	世紀	21 世紀
年 (YEAR)	年, 年度, 年代	2002 年
月 (MONTH)	月	12 月
日 (DAY)	日	27 日
時 (HOUR)	時	20 時
分 (MINUTES)	分	39 分
秒 (SECOND)	秒	33 秒
期間 (DURATION)	ヵ年, か月, 日間など	10 日間
特殊 (OPTIONAL)	キロ	40 キロ
その他 (NUMBER)	数詞のみまたは その他の助数詞	123, 50 本, 200 回, 3 度

表中の特殊 (OPTIONAL) は、助数詞「キロ」のような「長さ」や「重さ」のどちらに対しても表現出来る可能性があるため設定したものである。18種類以外の数値表現は、その他 (NUMBER) に分類した。

## 3 質問文の分析

数値表現を問う質問に答えるためには、疑問詞に付随している助数詞を用いることで適切な回答を絞り込むことができる。そこで、我々は質問文における助数詞の有無に注目して、QAC用に作成された質問のうち、回答が数値表現である192問の分析を行った。

### 1. 疑問詞に助数詞が付属している質問

助数詞が付属している質問は67問存在した。以下に例を示す。

- (a) 「日本人男性の平均寿命は何歳ですか。」  
 この場合、回答として求められるものは、質問で使われている助数詞「歳」を含んだ数値表現であり、対象文書内の「日本人の平均寿命は男が77・19歳」という個所から数値表現「77・19歳」を回答候補として特定できると考える。
- (b) 「インドネシアのスハルト大統領が辞任した時、彼は何歳でしたか。」  
 この場合、回答として求められるものは同じく「歳」を含んだ数値表現であるが、対象文書には「スハルト大統領(76)」のように数詞のみで表現されてものも存在した。

### 2. 疑問詞に助数詞が付属していない質問

助数詞が付属していない質問は125問存在した。以下に例を示す。

- (a) 「地球の人口はどれくらいですか。」
- (b) 「1998年の日本の経常黒字はいくらですか。」  
 助数詞が含まれていない質問の場合、疑問詞から回答が数値表現あることは判断できても、1(a)の例のように具体的な助数詞の形態的特徴まで特定することはできないため、回答として考えられる候補を絞り込むのは困難である。

2の場合に使用される疑問詞のパターンを調べると「いつ」が62問、「いくつ」が17問、「いくら」が23問、「どれ(どの)くらい(ぐらい)、どれだけ」が23問存在した。これらの疑問詞と数値表現の種類を、回答になる可能性があるかという観点で対応付けたものを表2に示す。

この表から、「いつ」以外の疑問詞は、回答としていくつかの数値表現の種類を取りえることがわかる。これは、疑問詞だけでどのような助数詞を含んだ数値表現が回答候補として適切であるか判断できないためである。

表2: 疑問詞と数値表現の種類との対応表

数値表現の種類	疑問詞のパターン			
	A	B	C	D
金額 (MONEY)		○	○	○
長さ (LENGTH)		○	○	○
重さ (WEIGHT)		○	○	○
面積 (AREA)		○	○	○
体積 (VOLUME)		○	○	○
容量 (CAPACITY)		○	○	○
年齢 (AGE)		○	○	○
人数 (NPERSON)		○	○	○
割合 (PERCENT)		○	○	○
世紀 (CENTURY)	○			
年 (YEAR)	○			
月 (MONTH)	○			
日 (DAY)	○			
時 (HOUR)	○		○	
分 (MINUTES)	○		○	
秒 (SECOND)	○		○	
期間 (DURATION)			○	○
特殊 (OPTIONAL)		○	○	○
その他 (NUMBER)	○	○	○	○

A:いつ B:いくつ C:いくら  
 D:どれ(どの)くらい(ぐらい)、どれだけ

## 4 回答候補の選定手法

疑問詞だけで答えの候補を十分に絞りきれないケースに対応するため、質問内容を最も示している語が、どのような助数詞と関係しているかということを手掛りとして用いることにした。そのため、助数詞と単語の共起関係を新聞記事から統計的に求め、その情報を利用することで、回答の数値表現の種類を判断する手法を提案する。

### 4.1 主題語の決定手順

質問内容を最も示している語を質問の主題語とし、その決定手順を以下に示す。

1. 質問文中の係助詞「は」の前にある最も近い「一般名詞」または「形容詞」
2. 係助詞「は」がない場合、疑問詞に最も近い「一般名詞」または「形容詞」

「形容詞」に関しては、例えば「長さはどれくらい」といった「形容詞+接辞語(さ)」が重要語になる可能性があるため設定した。

## 4.2 共起情報の作成

共起情報の作成のため、1997年から2001年の新聞記事5年分中の「助数詞-一般名詞」と「助数詞-形容詞」の共起頻度を求めた。共起の範囲については、助数詞を基準にウィンドウサイズを形態素数10個とした。作成した共起情報から、実際に質問文で使われた主題語と共起頻度が高い助数詞を示した例を表3に示す。

表3: 主題語と共起頻度が高い助数詞の例

主題語	助数詞
犯人	人
値段	円
本数	本
盲導犬	頭
人口	人
全長	メートル
広さ	平方メートル

## 4.3 質問対象の決定

共起情報から得られる助数詞を、そのまま回答候補の選定に用いるには制約が強い。例えば主題語が「全長」であるからといって、回答は「メートル」含む数値表現であるとは限らない。そこで、表1を用いて、選ばれた助数詞が属する数値表現の種類を回答候補の範囲とした。すなわち主題語が「全長」の場合、長さ(LENGTH)に属する助数詞(メートル, センチ, kmなど)を含んだ数値表現が回答の候補となる。

## 5 実験と評価

質問対象を特定する手法の精度を調べるため、3節で述べた疑問詞だけでは答えが絞れない「いくつ」「いくら」「どれ(どの)くらい(ぐらい)、どれだけ」の計63個の質問を用いて実験を行った。4節で述べた手法に基づき、質問の対象となっている数値表現の種類を特定するシステムを作成した。このシステムでは、選ばれた主題語と助数詞の共起頻度順に数値表現の種類が出力される。同種類が出力された場合は、上位のみを正解した。システムが出力した数値表現の種類と、正解が属する数値表現の種類とが一致しているかという点で評価を行った。その結果を表4に示す。

表4から、システムが出力した数値表現の種類で1位のみを採用した場合、正確に特定することができた

表4: 質問対象の正解率

	正解数/質問数	正解率
1位を正解	46/63	73.0%
2位までを正解	54/63	85.7%
3位までを正解	58/63	92.1%

のは全体の73.0%であった。

次に、1位で正しく質問対象が特定できた質問を使って、どれくらい回答候補の数が絞り込まれているか調べるため、実際に我々が構築した既存のQAシステム[4]に適用して実験を行った。対象記事は1998, 1999年の毎日新聞であり、このシステムは文書の検索にNamazu2.0.9[5]を用いている。評価に用いた質問文とその質問対象を表5、それぞれの結果を表6に示す。

表6: 質問ごとの結果

質問番号	絞込率	正解/候補1	正解/候補2	記事数
1	16.7%	1/1	1/6	1
2	14.3%	1/1	1/7	1
3	2.4%	1/2	1/85	1
4	45.2%	1/14	1/31	2
5	7.2%	2/37	2/514	2
6	26.3%	6/46	6/175	7
7	21.4%	2/6	2/28	2
8	23.2%	2/13	2/56	3
9	33.3%	1/4	1/12	1
10	13.6%	1/8	1/59	3
11	13.0%	3/30	3/230	6
12	3.7%	1/4	1/108	3
13	16.7%	1/1	1/6	1
平均	12.6%	1.8/12.8	1.8/101.3	2.5

表6において、候補1は質問対象として特定された数値表現の種類に属する候補数、候補2は検索記事に全て出現する数値表現の数、正解は各候補中に存在する正解の数である。絞込率は(候補1÷候補2)から算出した。この数値より回答候補となる数値表現をどれだけ絞ることができたかということがわかる。その絞込率は2.4%~45.2%で平均12.6%という結果を得た。

## 6 考察

質問対象を正確に特定できた例として、「沖縄本島の広さはどのくらいですか。」という質問では、主題語を「広さ」とし、質問対象を面積(AREA)とすることができた。一方、誤った質問対象が選ばれた例と

表 5: 評価に用いた質問文と質問対象

質問番号	質問文	質問対象
1	ユニーの社長の父はいくつで亡くなりましたか。	年齢 (AGE)
2	プルトニウムの原子番号はいくつですか。	その他 (NUMBER)
3	日銀新潟支店長の住宅の建ぺい率はいくつですか。	割合 (PERCENT)
4	ゲームソフト「ドラゴンクエスト6」の出荷本数はいくらですか。	その他 (NUMBER)
5	98年の、日本人女性の平均寿命はいくらですか。	年齢 (AGE)
6	NTT移動通信網の初値はいくらですか。	金額 (MONEY)
7	ソニーはジャストシステムの第三者割当増資にいくら出資しますか。	金額 (MONEY)
8	伊藤忠商事はいくらでファミリーマートを買収しましたか。	金額 (MONEY)
9	最も遠い位置にある銀河「RD1」は地球からどのくらい離れていますか。	長さ (LENGTH)
10	サッカーくじで全試合的中する確率はどのくらいですか。	割合 (PERCENT)
11	2006年の冬季五輪が開かれるイタリアのトリノの人口はどのくらいですか。	人数 (NPERSON)
12	日本とブラジルの時差はどれだけありますか。	期間 (DURATION)
13	アンドレ・ザ・ジャイアントの身長はどれくらいですか。	長さ (LENGTH)

して、「沢木耕太郎はいくつの時に大宅壮一ノンフィクション賞をとりましたか。」という質問で、選ばれた主題語が「ノンフィクション」、質問対象がその他 (NUMBER) となった。この場合、質問は人の年齢を尋ねているため、質問対象は年齢 (AGE) となるべきである。

回答候補の絞り込みに関しては、平均絞込率 12.6% という結果から 9 割近い余分な回答候補を排除することができた。特に、質問番号 1, 2, 13 は質問対象を特定した後、選ばれた回答候補そのものが正解であった。これは、数値表現を詳細に 18 種分類したのが効果的であったと考えられる。また、各質問において対象記事内の正解数が減少せず、回答候補の絞り込みを行うことができた。

## 7 おわりに

本論文では、数値に付属する助数詞などの表現と共に起する単語情報と、質問内容を表す語との照合により回答候補を絞る手法を提案した。実験には、「いくつ」「いくら」「どれ(どの)くらい(ぐらい)、どれだけ」といった疑問詞だけでは答えを絞りきれない質問を対象に行った。その結果、質問対象を 1 位で正確に特定することができたものは質問全体の 73.0% であった。また、このような質問において、回答候補の数をどれだけ絞ることができたかを示す絞込率は平均 12.6% となった。

今後の課題は、質問対象を正確に特定する手段として、主題語以外の語に関しても分析を進めることと、数値表現の種類の詳細化があげられる。また、今回は

回答候補の絞り込みに焦点をあてたが、今後は QA システムの最終的な回答における精度向上を目指していく予定である。

## 参考文献

- [1] Information Retrieval and Extraction Exercise (IREX) Homepage, <http://cs.nyu.edu/cs/projects/protus/irex/>.
- [2] J. Fukumoto and T. Kato. "An Overview of Question and Answering Challenge (QAC) of the next NTCIR Workshop", Proceedings of the Second NTCIR Workshop Meeting, National Institute of Informatics, 2001.
- [3] 山下竜之, 藤畑勝之, 角田久直, 志賀正裕, 森辰則. "質問応答システムにおける数量表現の取り扱い", 言語処理学会 第 7 回年次大会 発表論文集, pp34-37, 2001.
- [4] J. Fukumoto, T. Endo, and T. Niwa. "RitsQA: Ritsumeikan question answering system used for QAC-1", Working Notes of the Third NTCIR Workshop Meeting, National Institute of Informatics, 2002.
- [5] 全文検索システム Namazu Homepage, <http://www.namazu.org/>.