

WHY型QAシステムにおける 回答抽出方法の改良

作成者：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)

発表者：三沢博章

WHY型質問応答システムとは？

「なぜ」、「どうして」を含む質問



検索エンジンより得た情報に対しフィルタリングを行う



適切な返答をする

具体例

質問文: 空はなぜ青いのか

↓ 入力

本システムにおいて質問文の回答を抽出する

↓ 出力

回答文: 太陽からの光が、大気中の酸素、窒素、水蒸気などの分子や、光の波長よりも小さい微粒子にぶつかり散乱するとき、青のように波長の短い光が強く散乱されるので、空の色は青く見える。

使用されている質疑応答技術

- 質問文解析
 - 与えられた質問文に対し、解析、言い換えを行い情報検索に用いる「検索KW」を作成
 - 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

使用されている質疑応答技術(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

- 回答候補抽出

- 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

- 回答候補ランキング

- 回答候補の中で名詞の出現頻度を指標としたランキングを作成する

本研究の目的

質問の内容によって・・・

- 回答を出力しない
- 正答とは異なる回答を出力する

- ・回答出力の過少の改善
- ・正答出力の精度の向上

改良点

- 「全文検索」から「類似検索」への変更

全文検索・・・検索KWと一致する文章のみ抽出

類似検索・・・似た記事も抽出

改良点(2)

- 回答抽出フィルタの改良

正答文に「のは」という文字が含まれることが多い

↓そこで……

「質問文の名詞、動詞、形容詞」+「のは」

(例)

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”

実験

- 実験条件

計100件の質問文を用いる

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

4種類の評価方法を行う

評価方法

- ・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

- ・F値

$$\text{Recall} = \frac{\text{正答数}}{\text{質問数}} \quad \text{Precision} = \frac{\text{正答数}}{\text{出力数}}$$

$$\text{F値} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

F値が高ければ性能が良い

評価方法(2)

- 検索成功率

- 回答ランキング上位10位以内に正答を含む割合

- MAP

- システムが正答数を網羅している割合

実験結果

	改良前	改良後
回答取得正答率	24%	96%
F値	0.08	0.18
検索成功率	8%	39%
MAP	0.03	0.05

具体的な改善例

質問：飛行機雲はどうしてできるの

(改良前)：出力なし

(改良後)：飛行機の後ろに飛行機雲が出来るのは、エンジンからの排気ガスや水蒸気が核となって、期待周辺の気圧変化から水滴ができるため

より多くの正答を出す工夫

- ・他の理由を示す文字パターンの利用

「のは」の他に「から」と「ため」を利用

以上の3つの文字パターンを用いて、
再度実験を行うと、、、

実験結果(2)

F値	0.09	▼0.09
検索成功率	43%	△4%
MAP	0.06	△0.01

ランキングの精度は下がる

正答抽出率は上がる

考察

正答の得られない質問に対し、似たような意味をもつ語を用いて言い換えを行った場合、正答数が増加するのではないか？

(例)

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
起こるの

質問文の再構築の実験結果

質問文100件の中で、言い換えを行える質問文は55件存在

↓ 実験後...

55件のうち5件、新たに正答の出力を確認した

おわりに

WHY型質疑応答システムの改良は成功

しかし・・・

正答の中には、

質問文の単語が含まれないものも存在