

「なぜ」に答える質問応答システム NAZEQA

—原因表現の自動獲得と機械学習に基づく理由検索—

どんな研究?

- ・ユーザの「なぜ?」という理由を尋ねる質問に対して回答を出すシステム
- ・昨年のシステムから大幅に精度を改善
- ・新聞記事からだけではなく、ウェブ上の文書からも検索可能

もたらされる変革

- ・答えをすばり検索できる
- ・欲しい情報をすぐに届けるようなサービスの実現
- ・応用例：検索ポータルへの付加価値、社内のトラブルシューティング、ロボットの知識源の自動獲得

Q : 「なぜホタルは光るのか?」

構成図

文内の要素の意味関係(原因→結果など)が手動でタグ付けされた文集合(約20万文)

回答候補抽出

新聞記事・ウェブデータ

検索

原因表現の自動獲得(約700種)

意味情報付きコーパス

特徴量抽出

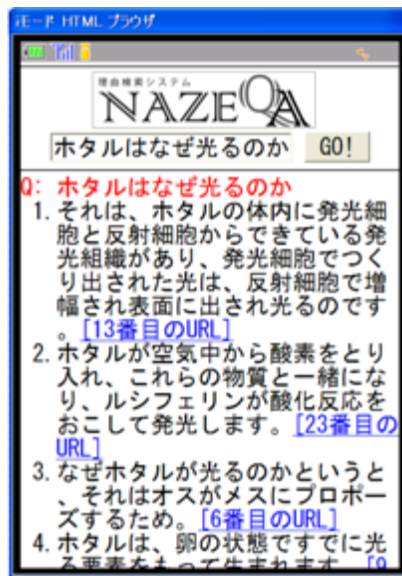
画面例

理由検索に適切な特徴量を導入することで、昨年度から性能が改善

質問応答コーパス

回答候補ランキング

1000問の質問応答事例から機械学習により採点関数を作成



A : それは、ホタルの体内に発光細胞と反射細胞からできている発光組織があり、発光細胞で作らだされた光は、反射細胞で増幅され表面に出され光ります。

関連文献

R. Higashinaka, H. Isozaki: Automatically Acquiring Causal Expression Patterns from Relation-annotated Corpora to Improve Question Answering for why-Questions, ACM Trans. on Asian Language Information Processing, Special Issue on NTCIR-6, Volume 7, Issue 2, Article 6 (29 pages), 2008.

連絡先: 東中竜一郎(Ryuichiro Higashinaka)

協創情報研究部 言語知能研究グループ

東中 磯崎

目黒