

理工学専攻修士論文要旨

提出年度：2005年度

提出日：2006年1月30日

専修コース：知能情報コース

学生番号：35604107

学生氏名：竹原 一彰

研究指導教員：原田 実 教授

(論文題目)

質問応答における照合と回答抽出の高精度化

(内容の要旨)

近年、インターネットや辞書など電子化された膨大な知識にアクセスできる環境が整ってきた。しかし、情報量が充実していくにつれてその中から欲しい情報のみを獲得するのに多大な労力が必要になっている。この現状を踏まえ原田研究室では情報検索と情報抽出を融合した質問応答システムの研究に着手し2003年度にプロトタイプ、2004年度にはWeb検索により知識文を獲得する意味グラフベースの質問応答システム **Metis** を開発してきた。しかし、その精度にまだ問題があり、本年度は **Metis** の回答抽出精度の向上を図ることを主な研究テーマとした。さらに、**Metis** の要素技術であるグラフ類似度計算(文類似度計算)アルゴリズムを他の自然言語処理研究でも活用できるようにライブラリ化し提供することをもう一つの研究テーマとした。

従来 **Metis** の研究・開発から得られた主な問題点は以下のようなものである。入力された質問文を意味解析、質問箇所特定を行い、その語意(意味制約)を定めるが、その制約が弱く質問内容とそぐわない回答を抽出してしまうことが頻出していた。知識文検索では、質問文からのキーワード抽出精度が低く知識文の検索率が48%と非常に低かった。また、疑問や推量、仮定を表す文を知識文として採用してしまうことがあり誤回答を誘発していた。質問グラフと検索されてきた知識グラフの照合において、従来方式は2つのグラフ間を根から縦型探索しながら照合する方式であり、係り受け構造に強く依存するのでグラフ構造が異なるとうまく解を抽出できなかった。また、言い換え表現に対応するためノード飛び越しという語間の意味的接続関係を無視する対応付けを行っており誤対応が起こることがあった。

以上のような問題点を解決するために本研究では以下のような改善を行った。

1. 質問文解析の改善

質問箇所に与える意味制約決定ルールを昨年度以上に精密に定めた。

2. 知識文検索の改善

検索エンジンから知識文を獲得する際に用いるキーワードの抽出ルールを洗練し、効率的な検索ができるような組合せの優先順位付け手法を提案した。

3. 新しいグラフ照合方式の採用

新照合方式は、質問グラフと知識グラフのノードに対し、一定の閾値を越える照合ノードペアの組合せを生成する。全ての照合ノードペアの組合せについて、どの程度構造が類似しているかを示す照合ノード間のアーク類似度を求めることでグラフ全体の類似性を評価し、最大類似度を与えるものを照合結果として選出する。さらに、さまざまな言語的な言い換えに対応するために、意味グラフベース言い換えエンジン **Serpent** を用いた。

知識文検索率、正解抽出率の評価実験を行った。知識文検索率が48%から89%へ、正解抽出率が36%から78%と大幅に改善することができた。また、提供したグラフ照合アルゴリズムは、アンケート分類やテキストマイニングツールの開発研究で採用され、良好な結果が得られたとの報告がある。