

# 理工学専攻修士論文要旨

提出年度：2006年度  
提出日：2007年1月30日  
専修コース：知能情報コース  
学生番号：35605045  
学生氏名：加藤 裕平  
研究指導教員：原田 実 教授

## （論文題目）

複数文からの回答抽出による質問応答システム Metis の精度向上と質問内容の拡張

## （内容の要旨）

近年、自然文で書かれた質問文に対して膨大な量の文章より回答を抽出する質問応答システムの研究が盛んに行われており、NTCIR の Question Answering Challenge や Cross Language Question Answering 等の評価型ワークショップなどで研究成果が公表されている。現在行われている CAC-4 では以前までの人物、時間等といったものを回答とする factoid の質問だけでなく、理由や方法、定義といったものも回答とする質問も対象としている。原田研究室でも情報検索から情報抽出までを自動化する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質問応答システム Metis を開発し、2005年度ではシステム更なる精度向上を行っている。

2005年度までの Metis では人物、時間等といったものを回答とする factoid の質問のみを対象としており、理由や方法、定義といった質問には対応していない。また、文の照合を単文同士でしか行うことができず、複数文からの文照合は行うことができない。そこで本年度は Metis の対応する質問形式の拡張と複数の文にわたって記述された知識からの回答抽出を主な研究テーマとした。

以上の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. 照応関係を利用した複数の文からの回答抽出

複数の文に分かれて記述された知識を1つにまとめるため、知識を意味解析システム Sage によって解析して得られた照応関係を用いる。知識グラフに対し、他の文からの照応情報を付与することで複数の文からの回答抽出を実現した。

### 2. 対応質問形式の拡張

理由、方法、定義に関する質問に対してシステムを対応させた。回答抽出時には factoid の質問での質問箇所と照合したノードを回答として抽出するという方法ではなく、質問グラフ中の主題ノードと照合したノードより、質問の回答として相応しい深層格で結ばれているノードを根とする部分木を回答として抽出する。

知識文検索率、正解抽出率の評価実験を行った。昨年度との比較では知識文検索率が89%から92%へ、正解抽出率が72%から74%への改善がみられた。理由・方法・定義の質問を扱った NTCIR-QAC のテストコレクションでは正解抽出率は42%、factoid の質問を扱った NTCIR-CLQA のテストコレクションでは正解抽出率は33%であった。