

氏名	きよ た よう じ 清 田 陽 司
学位(専攻分野)	博 士 (情 報 学)
学位記番号	情 博 第 135 号
学位授与の日付	平成 16 年 11 月 24 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研究科・専攻	情 報 学 研 究 科 知 能 情 報 学 専 攻
学位論文題目	Dialog Navigator: A Navigation System from Vague Questions to Specific Answers based on Real-World Text Collections (ダイアログナビ: 実世界テキスト集合に基づく漠然とした質問から具体的な回答へのナビゲーションシステム)
論文調査委員	(主 査) 教 授 松 山 隆 司 教 授 河 原 達 也 助 教 授 佐 藤 理 史

論 文 内 容 の 要 旨

本論文は、大規模な電子テキストとして蓄積されている知識ベースに対して、ユーザが自然言語で検索・質問するタスクを想定し、ユーザに適切なテキストを提示するための方法に関する研究をまとめたものである。テキスト集合とユーザ質問の自然言語表現の間にどのようなギャップが存在するのかを明らかにし、それらのギャップに対処するための具体的な方法を提案している。さらに提案手法を、実世界のソフトウェアサポート用知識ベースを対象とした対話システム「ダイアログナビ」として実装し、公開運用の結果に基づいた総合的な評価を行っている。

第1章は序論で、ユーザの質問とテキストの間に存在する様々なギャップの問題と、その解決策について論じている。特に、従来の方法では対応できない、キーワードを越えるレベルの表現のギャップと、内容の具体性のギャップに対処することの必要性を強調している。また、人工知能的アプローチに基づく先行研究では、実世界テキストへの適用が困難であることを指摘している。

第2章では、自然言語処理技術を利用し、ユーザの質問とテキストの正確かつ柔軟なマッチングを行うためのアルゴリズムを述べている。特に、上記のギャップに対処するためには文の構造を重視したマッチングを行う必要があることを指摘し、同義フレーズ辞書の利用、日本語の係り受け関係を重視したスコア計算手法などを提案している。収集した質問文による評価を行い、提案手法の有用性を示している。

第3章では、換喩によって生じる文の構造のギャップに対処する方法を、第2章の方法の拡張として提案している。まず、大量のコーパスから換喩とその解釈の表現のペアを自動的に抽出する新規手法を提案し、得られたペアのうち80%以上は換喩の解釈として適切であることを示している。さらに、得られたペアを同義フレーズとして扱うことでマッチングの精度が向上することを評価実験によって示している。

第4章では、第2章・第3章で提案したマッチング手法を応用し、漠然とした質問をしたユーザを具体的な答えのテキストに導くための戦略を述べている。具体的には、ボトムアップ的手法として状況説明文抽出による聞き返し、トップダウン的手法として対話カードによる聞き返しを組み合わせる方法などを提案している。いずれの手法でも上記のマッチング手法が活用されており、様々な抽象度の質問に対して2つの手法が相補的に有効に働くことを示している。

第5章では、以上の提案手法を対話システム「ダイアログナビ」として実装し、Webサイト上で公開運用を行った結果について述べている。2年間の運用で約25,000回のアクセスを得た。収集した運用ログを、(1)システムはユーザに適切なテキストを提示できたか、(2)ユーザとシステムは対話中でどのようにふるまったか、(3)抽出された状況説明文は有用であったか、という3つの観点から評価し、それぞれ提案手法の有用性を示している。具体的には、(1)対話セッションの成功率は76%、(2)対話カードは短く漠然とした質問文に対して有効に働いている、(3)状況説明文のうち61%は聞き返しの選択肢として妥当、という結果を得ている。

第6章では結論を述べている。

論文審査の結果の要旨

本論文は、自然言語による情報検索の研究において、従来あまり扱われていなかった、「ユーザの質問（文）と知識ベース（テキスト集合）の間に存在する様々なギャップ」を解消するという問題に、実世界のソフトウェアサポート用の大規模なテキスト集合を具体的素材として本格的に取り組んだもので、得られた主な成果は以下の通りである。

1. 高度な自然言語処理技術、特に文構造の解析技術を利用することにより、ユーザの質問文とテキスト集合の間に存在する、(1)キーワードを越えるレベルの表現のギャップと(2)内容の具体性のギャップを埋め、両者の柔軟なマッチングを行う方法を提案し、表層的な単語の情報のみに基づく従来手法より有効であることを示した。
2. 重要な言語現象である換喩によって生じる文構造のギャップの問題を明確にした上で、大規模コーパスから同義フレーズを抽出することにより、換喩の解釈を行う方法を提案し、質問応答システムの高度化において有用であることを示した。
3. 上記の提案手法を組み合わせ、実世界のソフトウェアサポート用知識ベースを対象とした対話システム「ダイアログナビ」を実装し、Web サイト上で公開した。このシステムの運用ログを分析し、提案手法が実際に有効に機能することを実証した。このように自然言語で対話を行える大規模な情報検索システムを実世界で運用・評価した事例はほとんどない。

以上のように本論文は、自然言語による情報検索において、ユーザの質問（文）と知識ベース（テキスト）の間に存在する(1)キーワードを越えるレベルの表現のギャップと(2)内容の具体性のギャップを埋め、両者の柔軟なマッチングを行う方法を提案し、その有効性を実世界システムによって実証したもので、学術上、實際上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。

また、平成16年10月26日実施した論文内容とそれに関連した試問の結果合格と認めた。