

共同作業中の対話に潜在する情報ニーズの類型化

Classification of Information Needs Latent to Dialogue in Collaborative Work

学籍番号：201521619

氏名：志賀 奏介

Sosuke SHIGA

近年、音声対話システムがブームとなっている(河原, 2013)。2016年の報告によると、アプリやAndroidでのクエリの20%が音声を利用した検索であることが報告されており(Greg, 2016)、情報探索における音声対話システムの重要性が増していることを示唆している。しかし、今までの音声対話システムではあらかじめ決められた質問文やコマンドを扱っており、曖昧な情報ニーズの表現には上手く対応できていない。一方、情報検索分野ではTaylor(1968)が情報ニーズを不明確性の4つのレベルに分類したモデルを提案しており、これらの実データにおける振る舞いを調査することは、曖昧に表現された情報ニーズに対応した音声情報検索システムの実現につながると考えた。

本研究では3つのRQを設定した。RQ1:対話中の情報ニーズはどのように類型化できるのか。RQ2:類型化した情報ニーズにはどのような特性が存在するか。RQ3:類型化した情報ニーズを同定するにはどのような素性が有用か。RQ1ではTaylorのモデルを元にJarvelinら(1995)のモデルを適用し曖昧な情報ニーズ、明確な情報ニーズの下に10カテゴリを作成し対話的情報ニーズのモデルを作成した。さらにペアで行われた旅行計画タスクのコーパス32950文に対しクラウドソーシングを用いて情報ニーズの10カテゴリのアノテーションを行なった。分析の結果、RQ2に対しては対話中の情報ニーズは16%程度存在しており、タスクが進行するにつれて減少する傾向が見られた。また、曖昧な情報ニーズと明確な情報ニーズ両方でも同じように減少する傾向が見られた。状態遷移確率の分析では曖昧な情報ニーズから明確な情報ニーズに遷移する際に直接遷移するのではなく、それ以外の対話を介して遷移する確率が高いことが示唆された。また、RQ3に対してはRandomForestを用いて素性の重要度を測定した結果、意味的素性の他に時系列的素性、言語的素性、統計的素性、対話的素性などが情報ニーズを同定するのに有用であることが示唆された。今後の展望として、曖昧な情報ニーズを含むダイアログを活用したクエリの自動生成や拡張などが考えられる。

研究指導教員：眞榮城 哲也

副研究指導教員：高久 雅生