

雑談システムにおけるトピックの関係性に着目した発話選択手法

福田 拓也

近年、人間と対話を行うシステムの研究が活発になされており、私達の社会に広がりつつある。雑談システムは特定のタスク達成を目的とせず、人間と自然な会話をすることを目指すものである。現在、システムの応答のために、Twitter 等の大規模データを用いた発話候補文を生成する研究がなされており、このとき問題となるのがユーザの発話に対してどのような発話候補文をシステム側が選択すべきかということである。

ここで特に重要な観点の 1 つは話題のつながりの自然さであると考えられるが、必ずしも類似する話題についての返答が適切なわけではなく、人間同士の会話には独立した一見関係のない話題同士が同時に出現する 경우가数多く存在する。Hobbs は、このような自然な話題の遷移を働きかける、異なる話題同士のシームレスな結合現象を話題のドリフトと呼んでいる。しかし、著者の知る限りでは、応答の際に話題の関係性を考慮して発話選択を行う対話システムは提案されていない。より人間らしい会話を行うシステムの実現のためには、話題を抽象的な枠組みで捉え、そのつながりに着目することにより、話題のドリフトをもたらす発話を選択する機構が必要であると考えられる。

本研究では話題のドリフトを定量的に表し、その度合いに基づいて発話を選択する手法を提案する。話題のドリフトの定量化とは話題間の距離を求めることであり、ドリフト度合いとは話題間の距離を表すものである。提案手法について、具体的には LDA により文書に含まれる話題と、話題を構成する単語群の学習、各単語トークンを生成した話題の推論をし、話題の共起度を計算することで話題の重み付き無向グラフを作成する。そして、2つの話題からそれぞれ想起される話題集合の類似度に従って話題のドリフト度合いを求める。

評価については、ドリフト度合いに従って選択された発話文の順序と、被験者によって判断される話題間の距離に従い選択された発話文の順序の相関について、順位相関係数を用いて評価した。

クラウドソーシングによる実験の結果、評価データの 70%に正の相関が見られ、30%が特に高い正の相関を示した。提案手法の応用としては話題の一貫性の判断への利用やユーザを楽しませる応答への利用が考えられる。

(指導教員 若林 啓)