

研究課題別事後評価結果

1. 研究課題名： 談話構造に基づく教師なし生成型要約
2. 個人研究者名
磯沼 大（東京大学大学院工学系研究科 特任助教）
3. 事後評価結果

本研究は、長い文章を主要な意味を保ったまま短い文章に変換する要約生成技術に関する研究であり、教師データを必要としない教師なし要約手法の確立を目指している。教師データを用意する必要がないため、本手法は多様なドメインや言語の要約生成に応用できる可能性を秘めている。コアとなるアイデアとして、文章のトピック構造を木構造で表す潜在表現に着目し、木構造のノードからトピック潜在表現を適切に選択し、各トピックについて文生成を行うことで、要約文の生成を行う。

ACT-X においては、木構造トピックモデルを大規模文章から学習する手法や、木構造トピックモデルから要約生成を行う手法を提案した。これらの研究は、自然言語処理分野のトップカンファレンス ACL およびトップジャーナル TACL に採択され、また国内学会・研究会で3件の受賞対象となっており、高く評価されている。今後は、本手法の応用範囲をさらに拡大する研究や、要約とトピックの関係に関する理論的な貢献が期待できる。

（加速フェーズ）

上記の評価を受けて研究実施期間を1年間延長し、加速フェーズを実施した。

加速フェーズでは、これらの成果の応用として新たにサーベイ論文自動生成の研究に取り組んだ。サーベイ論文生成は実用性が高いがこれまで要約研究の対象とはなっていなかった。まず、時系列上でトピックの分岐・統合をモデル化する時系列構造化トピックモデルを提案し、既存モデルでは捉えられないトピックの変化が得られることを示した。さらに、サーベイ論文生成の大規模教師データを作成し、これを用いてサーベイ論文生成モデルの学習・評価を行った。これらの成果は、これまでの研究成果を新たな応用分野へ展開しようとする意欲的なものである。サーベイ論文生成手法や自動要約の理論的定式化の研究は発展途上であるが、今後の成果が期待される。