

論文間の参照情報を考慮した関連論文の組織化

難波英嗣[†] 神門典子^{††} 奥村学[†]

我々は、サーベイ論文自動作成の実現を目指している。論文はいくつかの他の論文を参照しており、また、論文中には参照先論文について言及している箇所(参照箇所)がある。参照箇所には、参照元論文の著者の観点から見た参照先論文の要点や、参照・被参照関係にある論文に関する記述がある。特定分野の参照箇所を集め、著者の観点毎に参照箇所を分類・整理したものをユーザに提示すれば、その分野の研究動向を把握することができると考えられる。本研究では、論文間の参照・被参照関係に着目し、参照関係を用いて参照箇所を自動的に組織化するシステムを構築した。

Classification of Research Papers using Citation Links and Citation Types:

HIDETSUGU NANBA,[†] NORIKO KANDO^{††} and MANABU OKUMURA[†]

We are studying towards automatic generation of a survey(or a review) article in a specific domain. In a research paper, there are fragments where the author describes the essence of a cited paper and the differences between the current paper and the cited paper(we call them citing areas). These fragments can be considered as a kind of summary of the cited paper from the current author's viewpoint. We can know the state of the art in a specific domain from the collection of citing areas. Further, if the citing areas are properly classified and organized, they can play a role of a kind of review article. We implement such a system using citation links and citation types.

・参照箇所の組織化手法

本研究における参照箇所の組織化とは、論文間の参照情報をを用い、参照箇所を分類・整理することである。ここで参照情報とは、論文間の参照・被参照関係、論文から抽出された参照箇所と参照箇所を解析して明らかにされた参照の理由を指す。本稿では、参照情報を考慮し、書誌結合¹⁾という引用分析の手法に基づいて、参照箇所を組織化する手法を提案する。

引用分析とは、論文間の参照・被参照関係を用いて、論文間の関連度を測る手法である。書誌結合は、引用分析の代表的な手法の1つであり、論文間の関連度を測る時に、2論文間でどれだけ同じ論文を引用しているか、という基準に基づいている。

ここで、従来の書誌結合ではすべての引用を等価に扱っているが、実際には色々な参照の理由が存在する。従って、関連論文をより正確に組織化するためには、単純な参照・被参照関係だけでなく、より豊富な参照情報を考慮することが必要不可欠であると考えられる。

過去の我々の研究²⁾では、参照の理由を参照タイプとして3種類に分類し、自動的に参照タイプを決定する手法を提案した。この手法を用い、本研究では、2論文間で同一論文を共に参照しており、かつそれらの参照タイプが一致している結合のみを数えるという方法(BCCT: Bibliographic Coupling using Citation Types)で、2論文間の類似度を測る。

ある論文を複数の他の論文が参照する場合、それらを参照タイプ毎に組織化することは可能であった²⁾。本研究では、さらにBCCTを参照箇所の組織化に適用することにより、トピック毎の組織化も同時に行えるようになった。

提案手法はPerlで実装し、またCGIを用いることでWorld Wide Web上からの利用が可能となっている^{*}。

参考文献

- 1) Kessler, M.M. *Bibliographic Coupling between Scientific Papers*. American Documentation, Vol.14, No.1, pages 10-25, 1963.
- 2) 難波 英嗣, 奥村 学. “論文間の参照情報を考慮したサーベイ論文作成支援システムの開発”. 自然言語処理 Vol.6, No.5, pp. 43-62, (1999)

[†] 北陸先端科学技術大学院大学情報科学研究科
School of Information Science, Japan Advanced Institute of Science and Technology

^{††} 国立情報学研究所
National Institute of Informatics

^{*} <http://galaga.jaist.ac.jp:8000/pub/tools/sum>