

# 平成26年度 情報工学コース卒業研究報告要旨

長尾 研究室	氏 名	安 藤 聡 志
卒業研究題目	複数議事録を対象とした 議論のクラスタリングに関する研究	

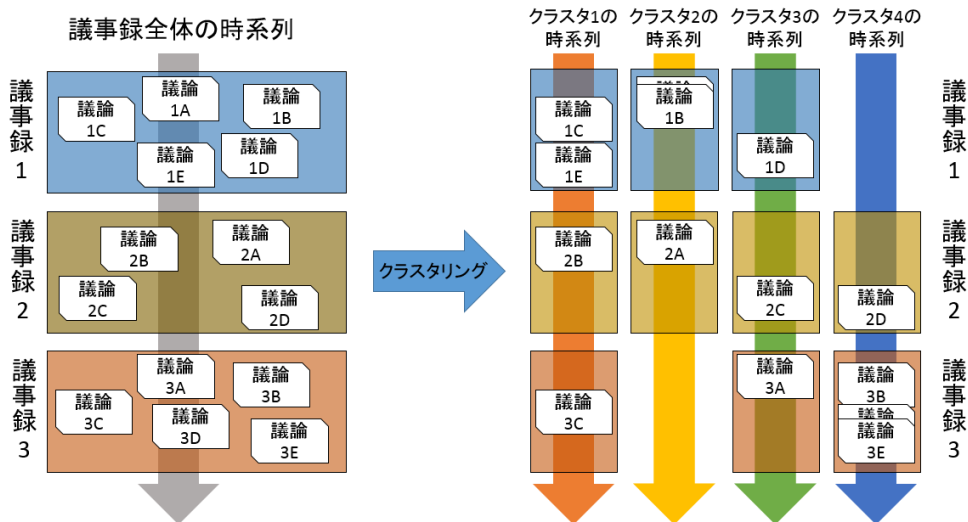
研究活動では議論が頻繁に行われ、その内容は議事録を閲覧することで想起できる。過去の議論を見返すことは、研究活動における新たな指針を見出すことにつながるため、議論を議事録として記録し、それを閲覧することにはメリットがある。また、複数回にわたって議論を行った場合、それらの議事録を時系列に沿って閲覧することで自身の研究活動についてより深い理解が得られる。

しかし、研究活動で行われる議論は、手法や実験などの複数の話題を含むため、複数の議事録にまたがる単一の内容を時系列に沿って閲覧することは容易ではない。この問題に対して、複数の議事録に含まれるそれぞれの話題の推移が分かりやすい形式で、議事録群を閲覧できることが望まれる。

本研究では、議事録を話題内容の観点から分割し、複数の議論を内容に応じてクラスタリングする手法を提案する(下図)。本手法では、議論セグメントをクラスタリングの対象とする。ここで、議論セグメントとは、構造化された議事録を構成する要素であり、個々の話題ごとのまとまりに対応する。議論セグメントをクラスタリングすることによって、話題内容ごとに分割された議論の閲覧を可能にする。

本研究では、議論セグメントのクラスタリングに文書クラスタリング手法を応用した。具体的には、Non-negative Matrix Factorization を用いた。クラスタリングで用いる議論セグメントの特徴情報としては、議論セグメントに含まれる情報を利用した。すなわち、書記が記録したテキスト情報、議論で参照されたスライド内のテキスト全文、及び、発言者が指示したスライド内の要素を使用した。スライドを用いる議論では、議論がスライドの内容に影響を受けるため、スライドの提示情報を活用することによって、議論セグメントの内容を反映したクラスタリングが可能となる。

提案手法のクラスタリング性能を評価するために実験を行った。実験では、人手で作成したクラスタリング結果を正解データとし、それと比較することにより評価した。実験の結果、議論セグメントが持つ議論内容の情報を特徴情報として用いることで、より正解に近いクラスタリングができることを確認した。



図：提案手法によるクラスタリングの概要