

同一議題の議論文書に対するアノテーション

Annotation Method for Argumentation Records of The Same Subject

大竹圭彦*1 岡田将吾*1 新田克己*1
Yoshihiko Ootake Shogo Okada Katsumi Nitta

*1東京工業大学大学院総合理工学研究科
Tokyo Institute of Technology

Though logical tags attached to argumentation records helps to understand the records, annotation takes much time. In this research, we propose an annotation method for argumentation records of the same subject. First of all, each argumentation record is divided into some segments using the word clustering method. Then, for each segment, topic is assigned. And by comparing and mapping segment sequences of more than one records, we improve the precision of topic estimation.

1. はじめに

長時間にわたる複雑なテーマの議論文書を解析することは大変、手間のかかる作業である。解析を助けるために発言の論理構造（理由、結論など）を表すタグを議論文書につけることができれば、解析の効率を高めることが期待できる。

われわれは、交渉の発言記録から議論スキルの評価を行う研究の一環として、議論文書にタグをつけるためのアノテーションのエディタを開発し（[Kubosawa2012]）、多くの議論文書の解析を行ってきた。

しかしながら、アノテーションは大変時間がかかる作業であり、しかも人により、アノテーションの結果に大きな違いがあることがある。安定したアノテーションを効率よく行うことができれば、交渉記録の解析の質が大きく向上する。

大学の法学部における議論スキルの教育の場では、模擬裁判や模擬交渉のトレーニングや、交渉のコンテストが行われている。このような場では、同じ交渉課題に関して、異なる学生が複数の交渉を行い、教員がそれらを比較して交渉スキルの説明を行うこともなされている。そのような交渉記録は、発言中のトピックの推移が類似していることが予想され、それを利用することによって、アノテーションを効率化できる可能性がある。

そこで、われわれは、複数の交渉記録を比較しながらアノテーションを行う方式について予備的な検討を行った。

2. 大学対抗交渉コンペティション

本研究で検討した議論文書は大学対抗交渉コンペティション（以後、交渉コンペと呼ぶ）のものである。交渉コンペは大学別約40のチームが約20の会場に分かれ、事前に与えられた共通の紛争事例に関して対立する立場に分かれ、模擬仲裁とも模擬交渉の二形式の議論を約3時間ずつ行い、それぞれの会場で法律の専門家による採点を行い、その点数を比較して各チームの議論能力の優劣を競い合う。

このような交渉コンペの交渉課題は、複数の論点が含まれており、しかもその論点間に依存関係があることもある複雑な問題である。しかし、主な論点はあらかじめ抽出することがで

き、大きな論点の流れは、どの議論文書にも共通するものと予想できる。

2.1 交渉課題の例

今回取り扱った交渉コンペの交渉課題は以下のとおりである。

「電子部品メーカーのR社と電子機器メーカーのB社は携帯電話事業の強化のため、協力することになった。その方法は、R社の集積回路部門をB社に譲渡する、両社が互いに研究員や出資金を出し合い新たな合弁会社を設立する、両社が研究員や出資金を出し合い長期研究協力を行う、の3つの形が考えられる。

事業譲渡の場合、売却価格や、売却後の同一部品に関する競争禁止期間や、譲渡後の社員の労働環境の保証など、決めなくてはいけないサブ論点が複数存在する。合弁会社の設立の場合、社長や社名、出資金の割、海外進出方針などのサブ論点があり、長期研究協力の場合も、交流形態に関するサブ論点が存在する。

2.2 解析に必要なアノテーション

議論の解析に必要なアノテーションは発言のタイプと争点である。

発言のタイプは、「主張」、「提案」、「確認」、「合意」、「否定」、「説明（論証）」などであり、争点は「譲渡価格」、「男女平等」、「雇用保証」など、課題からあらかじめ読み取れるものである。

3. アノテーション支援手法の概要

3.1 手法の概要

トピック推定に関する支援手法の流れは図1のとおりである。

1. 準備 サンプル抽出した交渉記録に関して、人手でアノテーションを行う。呂が開発したssLDA手法を用い[呂2013]、トピック（サブ論点）ごとにそのトピックを識別するための特徴語の抽出を行う。
2. 記録内アノテーション アノテーション付けたい交渉記録に関して、まず単語間の共起度を利用した単語クラスタリングによる文書分割を行い、それぞれの区間に関して、頻出単語と特徴語を比較して、トピック推定を行う。
3. 記録間アノテーション サンプル抽出した交渉記録と、そうでない交渉記録の区間を比較し、交渉記録中のほぼ

同じ位置に存在する区間同士で単語ベクトルを利用したマッピングを行う。マッピングされた区間同士は同じトピックになるため、トピックが異なっていた場合は書き換えを行う。

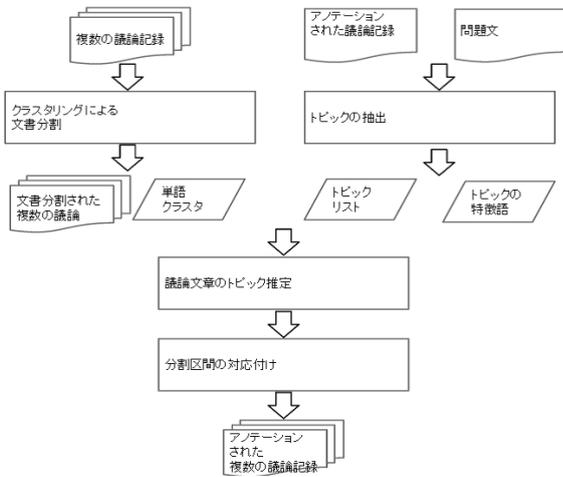


図 1: システムの流れ

3.2 トピック推定の予備評価

本手法によるトピック推定の予備の評価を行った。評価に用いた3つの交渉記録のトピックの推移は図2に示すとおりである。上から下の時系列になっており、区間の長さは発言数に比例しており、それぞれにトピックが割り当てられている。議論Aは3つの案(譲渡、合併、長期協力)のすべての案について満遍なく議論したが、その結果、時間不足になり、合意形成ができなかった事例である。議論BとCは事業譲渡についてメインに交渉が行われ、合併会社の設立や長期研究協力についてはあまり議論されず、結局は事業譲渡で決着がついた事例である。

A	B	C
事業譲渡確認	事業譲渡確認	デュエリレンジョス期間
労働環境	デュエリレンジョス期間	番付
就業禁止	雇用保証	雇用保証
金銭	労働環境	労働環境
デュエリレンジョス期間	借付	雇用保証
雇用保証	関係会社	研究員人数
長期研究 確認	将来性	借
合併会社 確認	合併確認	研究員人数
有限責任	会社	就業禁止
社長	国際化	有限責任
出資金	関係会社	LSI
国際化	確認	デュエリレンジョス期間
出資金	借付	番付
国際化	雇用保証	確認
有限責任	就業禁止	社長
出資金	デュエリレンジョス	方針
国際化	労働環境	出資金
期間	合併確認	就業禁止
研究員人数	方針	雇用保証
期間	借	研究員人数
金銭	デュエリレンジョス期間	有限責任
役員	借付	番付
	研究員人数	デュエリレンジョス期間
	借付	雇用保証

図 2: 人出による議論のアノテーションの結果

記録内のトピック推定と記録間のトピック推定の結果を表1に示す。この表は、記録Bをサンプル記録として、記録Aと

記録Cのトピック推定を行った結果である。いずれも交渉記録間の比較によって、トピック推定の精度が高まっている。

表 1: トピック推定の精度

議論	記録内での推定	記録間の推定
A	10	12
B	13	no count
C	10	16

4. まとめ

交渉記録へのアノテーションを支援するため、複数の記録を比較してトピック推定をする方式の概略を説明し、その予備的な評価を行った。その結果、トピック推定に関しては有効性が認められることが示された。次の段階としては、トピックのより詳細な推測を行う必要がある。たとえば、同じトピックであっても、その前半は提案や説得が行われ、後半は合意形成が行われることはどの交渉記録でも同じである。詳細レベルのトピック推定が行えれば、アノテーションからの負担が一層軽減できることが期待される。

参考文献

[Kubosawa2012] Shumpei Kubosawa, Youwei Lu, Shogo Okada, Katsumi Nitta: Argument Analysis with Factor Annotation Tool. JURIX2012:61-70

[Huruta2003] 古田一雄, 前原基芳, 高島亮祐, 中田圭一, 知的支援機能を備えた電子会議システム, 社会技術研究論文集, Vol.1, 299-306, 2003

[Tanaka07] 田中貴紘:事例に基づいたオンライン調停支援システムの研究, 博士論文, 東京工業大学大学院,2007.

[Sakahara2014] Makoto Sakahara, Shogo Okada, Katsuki Nitta, Domain-Independent Unsupervised Text Segmentation for Data Management, Data Mining Workshop (ICDMW), 2014 IEEE International Conference on

[Rubner1998] Y. Rubner, C. Tomasi, and L. J. Guibas, "A metric for distributions with applications to image databases," in Proceedings of the Sixth International Conference on Computer Vision, ser. ICCV '98. Washington, DC, USA: IEEE Computer Society, 1998, pp. 59?. [Online]. Available: <http://dl.acm.org/citation.cfm?id=938978.939133>

[呂 2013] Lu, Youwei, Shogo Okada, and Katsumi Nitta. "Semi-supervised Latent Dirichlet Allocation for Multi-label Text Classification." Recent Trends in Applied Artificial Intelligence. Springer Berlin Heidelberg, 2013. 351-360.