

申請者	学科名	情報システム工学科	職名	教授	氏名	磯崎 秀樹
調査研究課題	翻訳ソフトの改良のための自動評価法の改良と翻訳の問題点の自動発見					
調査研究組織	氏名		所属・職	専門分野	役割分担	
	代表	磯崎 秀樹	情報システム工学科教授	自然言語処理	研究全般	
	分担者	河田 彰	情報システム工学科M2		翻訳の問題点の自動発見の研究	
調査研究実績の概要	<p>翻訳ソフトを改良する時、訳の品質を採点することが重要になる。人間が採点すると時間がかかるので、自動で採点するソフト（翻訳自動評価法）が用いられる。世界的には、IBMの考案したBLEUという翻訳自動評価法が用いられるが、英語と日本語のように、まったく違う語順の言語の間の翻訳では、BLEUの点数が人間の評価と合わないことが知られており、磯崎の提案したRIBES(ライビーズ)が用いられている。RIBESは語順を重視した採点を行う。</p> <p>しかし、日本語には語順に自由度がある。「太郎は大阪で好み焼きを食べた。」を「大阪で太郎は好み焼きを食べた。」と言ってもよい。RIBESはこのような言い換えに対応しておらず、低い点をつけてしまうという問題がある。</p> <p>この問題を解決するため、昨年度の独創的研究助成費では、日本語訳をRIBESで自動的に採点する際に利用する「参照訳（人間が作ったお手本の訳）」から他の語順の文を自動生成する方法を考案した。この方法を使うと、「太郎は大阪で好み焼きを食べた。」というお手本の訳から「大阪で太郎は好み焼きを食べた。」などの別の語順の文が自動生成できる。</p> <p>ただし、「太郎は花子から電話を受けた後にスマホを買った。」を「太郎は電話を受けた後に花子からスマホを買った。」と言い換えると、意味が変わってしまう。したがって、語順を入れ替える場合、意味が変わらないように注意しなければならない。</p> <p>この問題に対して、昨年度は、語順を入れ替えても問題がない範囲での入れ替え方法を考案し、機械翻訳ワークショップWMT-2014で磯崎が発表した。しかしこの手法は、制約が強すぎて、ごく一部の文にしか別の語順が生成できなかった。</p> <p>そこで今年度は、もっと色々な語順を生成できるように、まったく違う手法を試した。この新しい手法は、自動生成した新しい語順の文を係り受け解析して、元の文と違う係り受け関係になる文を除去するという方法である。</p> <p>これにより、昨年度の手法よりもずっと多くの参照訳に別の語順を生成することができた。去年は8割の文に対して別の語順を生成できなかったが、今年度は、8割の文に対して別の語順を生成できるようになった。その結果、人間の評価にさらに近づけることができた。これを磯崎がWMT-2015で発表した。</p>					
成果資料目録	Hideki Isozaki and Natsume Kouichi: Dependency Analysis of Scrambled References for Better Evaluation of Japanese Translation, In Proceedings of Workshop on Statistical Translation (WMT), 2015.					