

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	スクランブリング自動生成の各機械翻訳自動評価への適用と比較					
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	教授	自然言語処理	発案、研究指導	
	分担者	門田 悠一郎	岡山県立大学大学院M1		実装、実験	
調査研究実績の概要	<p>翻訳ソフトを改良する時、訳の品質の採点が重要である。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、2 磯崎の提案した RIBES 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。「太郎は大阪でお好み焼きを食べた。」を「大阪で太郎はお好み焼きを食べた。」と言ってもよい。RIBES は語順を重視しているが、このような言い換え（スクランブリング）に対応しておらず、良い訳でも、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、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、その効果について2014～2015年に国際会議で発表した。しかし、翻訳自動評価法には他にもいろいろある。そこで今年度は、WERとIMPACTという翻訳自動評価法に適用した。その結果、これらの手法でも、文レベル相関係数が改善することが観測された。</p>					

<p>調査研究実績 の概要</p>	<p>RIBESは、翻訳ソフトが訳した「機械訳」と、人間が訳した理想的な訳である「参照訳」の語順を比較して、語順が似ていれば高い点数をつけ、語順が大きく違えば、悪い点をつける。</p> <p>したがって、「大阪で太郎はお好み焼きを食べた。」や「太郎はお好み焼きを大阪で食べた。」などの機械訳を参照訳「太郎は大阪でお好み焼きを食べた。」を基準として採点すると、少し悪い点数になる。でも、日本語がちゃんとわかっている人であれば、これらは同じことを言っていて、自然な日本語なので、満点をつけるであろう。</p> <p>そこで、機械訳と参照訳の係り受け木が、同じ形の木になればよい、という方法が考えられる。しかし、この手法には2つの問題がある。一つは、文全体が正しく係り受け解析できる精度はまだそれほど高くない、という問題である。もう一つは、機械訳が係り受け解析できるほど良い日本語ではないことがある、という問題である。参照訳の係り受け解析も失敗する可能性はあるが、あらかじめチェックしておく。</p> <p>そこで2015年の国際会議で、参照訳を係り受け解析して、その係り受け木をpost-orderでたどることによって生成される全ての語順の文をあらかじめ自動生成し、生成された文を係り受け解析した結果が、元の参照訳と同じ係り受け木になる場合に、参照訳として追加しておく、という手法を考案した。こうして増えた参照訳のそれぞれに対してRIBESで採点を行い、その中で一番点数の高いものを、最終的なスコアとする。機械訳が新たに追加された文と完全に一致すれば、満点が与えられる。</p> <p>NTCIR-9の特許翻訳タスクの英日翻訳に参加した17システムの機械訳で実験したところ、この改良により、17の翻訳ソフト全てにおいて、RIBESと人手評価の文レベル相関が向上した。これは符号検定で統計的有意差があった。</p> <p>今年度の研究は、この手法を他の翻訳自動評価法に適用するとどうなるかを調査することである。欧米で用いられている翻訳自動評価法の多くは、BLEUと同じように、語順の違いを軽視する傾向があることがわかっており、これらはもともと、英日翻訳では人手評価との相関が低い。そこで、欧米の研究者が語順に厳しいというWERという古い手法と、RIBES同様、日本語と英語の翻訳の自動評価のために考案されたIMPACTという手法で試すことにした。</p> <p>実験によると、WERは約70%、IMPACTは約65%の翻訳ソフトにおいて、人手評価との文レベル相関が向上したが、どちらも、符号検定では統計的有意差が出なかった。</p> <p>この違いは以下のように考えられる。RIBESは語順を直接計測しているが、WERやIMPACTは、間接的にしか計測していない。したがって、RIBESは語順に関わる改良が直接、人手評価との相関の向上として観測される。しかし、WERやIMPACTでは、そこまで明確な差としては観測されない。</p>
-----------------------	---

<p>成果資料目録</p>	<p>門田雄一郎・磯崎秀樹：和訳の自動評価のための係り受け木の比較、情報処理学会自然言語処理研究会、2016.</p>
---------------	---