

(様式2)

## 学位論文の概要及び要旨

氏 名 山田 節夫 印

題 目 機械翻訳用対訳知識の獲得に関する研究

学位論文の概要及び要旨

### 概要

本論文では、機械翻訳の訳質に大きな影響を与える対訳知識(原言語と目的言語の関係を表す知識)の獲得方法を提案する。対訳知識は、機械翻訳手法の違いによって実現方式が異なる。また、機械翻訳への入力文形式によっても対訳知識の内容が変わる。そのため、実現方式やその内容の違いによって、対訳知識の獲得における問題が異なる。以下に、対訳知識の実現方式及び内容に応じた4つの問題とその解決方法を挙げる。

- (1) ルールベースとパターンベースの翻訳手法に対して、対訳知識の登録方法の問題、すなわち、同じ原言語の表現に対して複数訳を登録する際に、どのような訳し分け条件を付与するのが望ましいかである。この問題の解決策として、既存の機械翻訳をその分野に適応するように対訳知識を構築する「分野適応型翻訳機構」を提案する。
- (2) ルールベースとパターンベースの翻訳手法に対して、対訳知識の拡充方法の問題、すなわち、全体としていかに翻訳精度を下げずに大量の対訳知識を効率的に拡充するかである。この問題の解決策として、原言語コーパスから得られる原言語の言語現象と、既存の翻訳システムで出力される目的言語の表現を機械的に抽出し、それを人が利用することによって、翻訳規則を効率的に拡充する「翻訳規則拡充支援法」を提案する。
- (3) 統計ベースの翻訳手法に対して、対訳知識の精度向上の問題、すなわち、翻訳モデルの精度をいかに向上するかである。この問題の解決策として、翻訳モデルを構築する時に用いた対訳コーパスを使って再度学習することによって、既存の翻訳モデルの精度が向上する「再学習による翻訳モデルの向上方法」を提案する。
- (4) 音声翻訳の翻訳手法に対して、対訳知識の登録方法の問題、すなわち、入力文が同じでも対話者の社会的役割に応じて訳出内容が異なる場合があり、その場合においていかに適切に訳出するかである。この問題の解決策として、対話者の社会的役割を対訳知識に取り入れることにより音声翻訳の訳質が向上する「対話者の社会的役割を利用した訳し分け手法」を提案する。

以上、4つの解決策により、従来に比べて効率良く効果的に機械翻訳用対訳知識が獲得できることを示す。

## 要旨

本論文では、機械翻訳の手法に応じた対訳知識（原言語と目的言語の関係を表す知識）獲得法を提案する。

機械翻訳として、良い翻訳の精度を目指してルールベース、パターンベース、統計ベースなど、様々な手法が研究・開発されている。しかし、どの手法においても、現段階での翻訳の質はまだ十分とは言えない。これら手法では、原言語と目的言語の関係を表した対訳知識が使用されている。対訳知識は、原言語を目的言語に翻訳するためには必須で、翻訳精度を直接左右する非常に重要な知識である。しかし、機械翻訳の手法によって対訳知識の実現方式が異なり、機械翻訳の手法に合わせて対訳知識を獲得しなければならない。また、機械翻訳へ入力する文形式によって、必要な対訳知識の内容が異なる。したがって、対訳知識は実現方式や内容によって、その獲得方法を考えなければならず、違う対処が必要である。そこで、本論文では、対訳知識の実現方式や内容に応じた次の4つの問題に対してその解決方法を提案する。

ルールベースとパターンベースの翻訳手法に対する対訳知識の登録方法の問題については、翻訳対象分野の事例（原言語と目的言語からなる対訳コーパス）を用いて、実用的ルールベース機械翻訳システムを自動的に分野適応させる枠組みである「分野適応型翻訳機構」を提案する。プロトタイプを構築し、技術マニュアル文で実験、評価したところ、分野適応型翻訳機構の有効性が確認できた。

ルールベースやパターンベース翻訳手法に対する対訳知識の拡充方法の問題については、単言語コーパスに含まれる言語現象を基に、人が効率的に翻訳規則を拡充する方法「翻訳規則拡充支援法」を提案する。提案手法では、翻訳規則の構成要素である原言語知識と目的言語知識に分けて翻訳規則を拡充する。天気予報のコーパスを用いて、翻訳規則拡充の実験を行った。この実験結果より、提案手法は従来手法に比べて、翻訳の質を下げずに効率的に翻訳規則を拡充できることが実証された。

統計ベース翻訳手法に対する翻訳モデルの向上の問題については、同じ対訳コーパスから非対称な2つの学習（翻訳方向が違う学習）によって構築された2つの翻訳モデルを利用することで、単語アライメントという尺度で、翻訳モデルを向上させる方法「再学習による翻訳モデルの向上方法」を提案する。辞書例文、科学技術記事、新聞記事の日英対訳コーパスを用いて実験を行った。その結果、翻訳モデルの向上が確認できた。

音声翻訳手法に対する対訳知識の登録方法の問題については、「対話者の社会的役割を利用した訳し分け手法」を提案する。社会的役割や性別は原言語を解析するだけでは取得困難であるが、音声翻訳の中では容易に得ることができる。この訳し分け手法を用いて、翻訳システムにとってオープンである対話を対象に評価したところ、効果があることが確認できた。この結果から、音声翻訳を使ってより自然な対話を行うことができる。

以上、実現方式や内容に応じて異なる、本論文で提案する4つの対訳知識獲得方法を利用することにより、従来に比べて効率良く効果的に対訳知識を獲得できる。したがって、本論文の手法は、機械翻訳の訳質を従来に比べて短時間で向上させることができると言える。