

卒業論文

コグニティブサービスを用いた 翻訳エージェント

指導教官 村上 陽平 准教授

立命館大学 情報理工学部
先端社会デザインコース 4 回生
2600170397-7

松岡 勇樹

2020 年度（秋 Semester）卒業研究 3（CH）
令和 2 年 2 月 1 日

コグニティブサービスを用いた翻訳エージェント

松岡 勇樹

内容梗概

近年、IT 技術の発達と共に世界のグローバル化が広がり、多くの外国人が来日し、日本にいても外国人を目にすることが当たり前となってきた。その結果、外国人とコミュニケーションを取る機会が増加している。さらに、機械翻訳の発展によって、外国語を話せない人も機械翻訳を媒介することで外国人とコミュニケーションを行うことが可能となった。しかしながら、ユーザの意図とは異なる翻訳結果が出力された場合、ユーザ間でコミュニケーションの齟齬が生じる。この問題は、現在の機械翻訳媒介コミュニケーションが、ユーザの言葉を機械翻訳に投入し、翻訳結果を対話相手に出力するという一方向の通信チャネルであるために引き起こされる。

そこで、本研究ではユーザとのインタラクションを行いながら翻訳結果を伝える翻訳エージェントを提案する。具体的には、チャットシステムを介した多言語間コミュニケーションを想定し、そのチャットシステムにユーザの視線と表情の情報を元に翻訳不適格箇所を分析し、翻訳結果を修正する翻訳エージェントを組み込む。本手法の実現にあたり、取り組むべき課題は以下の2点である。

翻訳不適格箇所の検知

翻訳不適格箇所を検知するためには、メッセージの受信者が理解できなかった翻訳文の箇所を翻訳エージェントが受信者とのインタラクションを通して判定する必要がある。

翻訳不適格箇所の修正

翻訳不適格箇所を修正して受信者が理解できる翻訳文を作成するために、検出された翻訳不適格箇所を同じ内容を表す別の表現で言い換えて再度翻訳する必要がある。

翻訳不適格箇所を検出するために、コグニティブサービスである表情分析サービスと視線分析サービスを用いる。表情分析サービスを用いた事前実験により、誤訳ではユーザがニュートラルな表情からそれ以外表情を示す傾向があることが分かったため、誤訳の有無をニュートラルな表情からの変化で行う。また、視線分析サービスを用いた事前実験により、ユーザは訳文中の誤訳箇所を繰り返し読み直す傾向があったことから、視線の再訪回数を計測し、複数回以上の再訪があった箇所を翻訳不適格箇所として推定する。次に、翻訳不適格箇所を修正

するために、検出された翻訳不適格箇所を類義語で置き換える。翻訳不適格箇所の類義語は Wikipedia で学習させた word2vec で獲得する。本研究の貢献は以下の通りである。

翻訳不適格箇所の検知

翻訳不適格箇所を検知するために、コグニティブサービスを用いて、メッセージの受信者の表情データと視線データを利用する方法を提案した。表情データと視線データの計測方法を変更することで、翻訳不適格箇所の検出の精度を向上できると考えられる。

翻訳不適格箇所の修正

翻訳不適格箇所を修正するために、Wikipedia のデータを用いて学習した word2vec によって翻訳不適格箇所を類義語で置き換える手法の提案した。BERT を用いて、メッセージ中で最も適切であると判定された類義語で置き換えることで、翻訳不適格箇所の修正の精度を向上できると考えられる。

Translation agent using cognitive services

Yuki Matsuoka

Abstract

In recent years, with the development of IT technology, the globalization of the world has spread, and many foreigners have come to Japan, and it has become commonplace to see foreigners even in Japan. As a result, the opportunities to communicate with foreigners are increasing. Furthermore, the development of machine translation has made it possible for people who can't speak foreign languages to communicate with foreigners by mediating machine translation. However, if the translation result which is different from the user's intention is output, there will be a communication discrepancy between users. This problem is caused by the one-way communication channel. The one-way communication channel means inputting the user's words into machine translation and outputting the translation results to the conversation partner.

Therefore, in this research, we propose the translation agent that conveys translation results while interacting with users. Specifically, we assume multilingual communication using a chat system. We incorporate the translation agent into the chat system that analyzes mistranslated parts based on the user's line-of-sight and facial expression information and corrects the translation results. In realizing this method, the following two issues should be addressed.

Detection of mistranslated parts

In order to detect mistranslated parts, the translation agent needs to determine the part of the translated text that the recipient of the message didn't understand through interaction with the recipient.

Correction of mistranslated parts

We correct translation ineligibility and create a translation that the recipient can understand. Therefore, it is necessary to paraphrase the detected mistranslated parts with another expression that expresses the same content and translate it again.

We use to detect mistranslated parts in order to facial expression analysis service and line-of-sight analysis service of cognitive service. Preliminary experiments using facial expression analysis services shows that users tend to

show facial expressions other than neutral in mistranslation. Therefore, the presence or absence of mistranslation is performed by changing from a neutral facial expression. In addition, due to preliminary experiments using the line-of-sight analysis service, users tend to repeatedly read mistranslated parts in the translated text. Therefore, we measure the number of revisits of the line-of-sight, and estimate the parts where there were multiple revisits as mistranslated parts. Next, we replace detected mistranslated parts with synonyms in order to correct mistranslated parts. We obtain synonyms for mistranslated parts from word2vec learned on Wikipedia. The contributions of this research are as follows.

Detection of mistranslated parts

We proposed the method of using facial expression data and line-of-sight data of message recipients using cognitive services to detect mistranslated parts. It is considered that the accuracy of detection of mistranslated parts can be improved by changing the measurement method of facial expression data and line-of-sight data.

Correction of mistranslated parts

We proposed a method to replace mistranslated parts with a synonym by word2vec learned using the data of Wikipedia in order to correct mistranslated parts. We can replace it with synonyms that is determined to be the most appropriate in the message by using BERT. Therefore, it is thought that the accuracy of correction of mistranslated parts can be improved.

コグニティブサービスを用いた翻訳エージェント

目次

第 1 章 はじめに	1
第 2 章 多言語コミュニケーション	3
2.1 機械翻訳による対話	3
2.1.1 対話方法	3
2.1.2 問題点	4
2.2 翻訳エージェントによる対話	6
2.2.1 翻訳エージェント	6
2.2.2 対話方法	8
第 3 章 コグニティブサービスによる認知分析	12
3.1 表情認識サービス	12
3.2 視線トラッキングサービス	12
第 4 章 システムアーキテクチャ	14
4.1 翻訳不適合箇所の検知	15
4.1.1 翻訳不適合箇所の有無の判定	15
4.1.2 翻訳不適合箇所の検出	17
4.2 翻訳不適合箇所の修正	17
第 5 章 考察	18
第 6 章 おわりに	20
謝辞	22
参考文献	23

第1章 はじめに

IT 技術はここ数十年で劇的な発展をとげ、世界のグローバル化を促してきた。Facebook や Twitter といった SNS の出現により、インターネットを通じて場所や時間に制御されることなく、他人と交流を行える。身近な人間とは限らず、直接会ったことのない人とも交流ができ、国境を超えた外国人とのコミュニケーションが容易に行える。また、日本の文化がインターネットによって発信されたことで、日本の文化に関心を持つ外国人が増加した。それに伴い、来日する外国人が増加し、日本の日常生活の中で外国人を目にすることが当たり前となった。それらの結果、日本にいても外国人と交流する機会が増加している。ただ、誰もが外国語を話せるわけではなく、そのような人は外国人その問題を解決するのが機械翻訳である。近年、Google 翻訳やみらい翻訳のような高品質の機械翻訳が登場した。統計的機械翻訳を用いた機械翻訳は精度の問題から実際に使われる機会が少なかったが、ニューラル翻訳の登場によって機械翻訳の精度が大幅に向上した。機械翻訳技術の発展により、外国語を話せない人も機械翻訳を媒介することで外国人とコミュニケーションを行うことができる。

しかし、機械翻訳によって高品質な翻訳結果が得られるとしても、ユーザの意図とは異なる翻訳結果が出力される場合がある。ユーザの意図を正しく反映した翻訳結果でなければ、ユーザ間でコミュニケーション齟齬が発生する。この問題は、従来の機械翻訳媒介コミュニケーションにおいて、発信側のユーザから入力された内容を機械翻訳が翻訳し、機械翻訳から翻訳結果を受信側のユーザへ出力されるといったユーザ間の一方向の通信チャネルであるために引き起こされる。

そこで、本研究ではユーザとのインタラクションを行いながら翻訳結果を伝える翻訳エージェントを提案する。具体的には、チャットシステムを介した多言語間コミュニケーションを想定し、そのチャットシステムにユーザの視線と表情の情報を元に翻訳不適合箇所を分析し、翻訳結果を修正する翻訳エージェントを組み込む。本手法の実現に辺り、取り組むべき課題は以下の 2 点である。

翻訳不適合箇所の検知

翻訳不適合箇所を検知するためには、メッセージの受信者が理解できなかった翻訳文の箇所を翻訳エージェントが受信者とのインタラクションを通して判定する必要がある。

翻訳不適合箇所の修正

翻訳不適合箇所を修正して受信者が理解できる翻訳文を作成するために、検出された翻訳不適合箇所を同じ内容を表す別の表現で言い換えて再度翻訳する必要がある。

以下本論文では、第 2 章で従来の機械翻訳による対話と翻訳エージェントによる対話翻訳について述べる。続いて、第 3 章で表情認識サービスと視線トラッキングサービスの 2 つのコグニティブサービスについて述べ、第 4 章で翻訳エージェントの実装方法を説明する。その後、第 5 章で考察を述べ、第 6 章で今後の展望や課題について述べ結論とする。

第2章 多言語コミュニケーション

本章では、機械翻訳、または翻訳エージェントを用いた多言語コミュニケーションについて説明する。大別して、従来の機械翻訳による対話と翻訳エージェントによる対話について、順に説明する。

2.1 機械翻訳による対話

従来の機械翻訳媒介コミュニケーションについて説明する。機械翻訳を媒介した多言語コミュニケーションの対話方法と問題点を述べる。

2.1.1 対話方法

一般に、異なる言語話者間で対話を行うとき、お互いに相手の言語の知識がなければ、機械翻訳を媒介する。メッセージを送る側を発信者、受け取る側を受信者とする。例として、それぞれ異なる言語を話すユーザ 1 とユーザ 2 を想定し、ユーザ 1 がユーザ 2 へメッセージを発信する場合の対話のステップを以下に示す。(図 1)

1. ユーザ 1 が母国語でメッセージを機械翻訳に入力する。
2. 機械翻訳が入力されたメッセージをユーザ 2 の言語に翻訳する。
3. 機械翻訳が翻訳結果をユーザ 2 に出力する。

ユーザ 2 は出力された翻訳結果を元にユーザ 1 の伝えたい内容を理解しようとする。ユーザ 2 がユーザ 1 へメッセージを発信する場合は、発信者をユーザ 2 受信者をユーザ 1 とし、上記の役割を交代して行う。対話が連続して行われるとき、以上のステップが繰り返される。



図 1: 機械翻訳による対話

2.1.2 問題点

機械翻訳を媒介することで生じる問題がある。それは、機械翻訳が発信者であるユーザ 1 の意図を反映していない翻訳結果を出力する点である。発信者の意図を反映していない翻訳結果とは、発信者の伝えたい内容とずれが生じている翻訳結果を指す。具体例として、日本語話者と英語話者の対話を挙げる。日本語話者が英語話者に対し、「昨日、本屋でブラックさんを見かけた」ことを伝えたいとする。その場合、日本語話者はメッセージとして、「昨日、本屋でブラックを見た」と機械翻訳に入力する。機械翻訳は入力された日本語を”I saw black at the bookstore yesterday”と英語に翻訳し、英語話者に出力する。ここで、翻訳結果を読んだ英語話者は以下 3 つのような解釈をされると考えられる。(図 2)

1. 昨日、本屋でブラックさんを見かけた。
2. 昨日、本屋でブラックさんと会った。
3. 昨日、本屋で黒絵の具を見た。

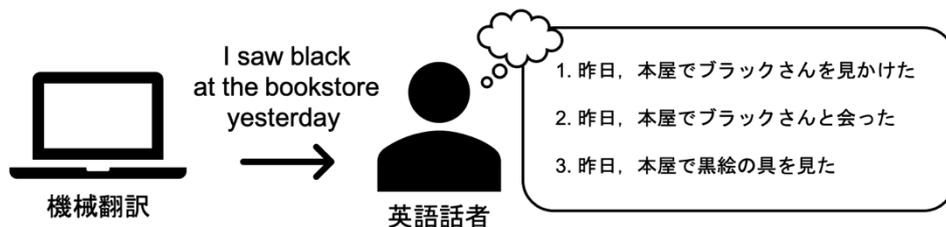


図 2: 英語話者が行う 3 つの解釈

1 つ目は、英語話者が”saw”を”会った”，”black”を”ブラックさん”と認識した解釈である。これは、日本語話者の伝えたい内容と一致しており、この対話における正解である。英語話者がこの解釈をすれば、日本語話者の意図は英語話者に正しく伝わったことになる。2 つ目は、”saw”を”会った”，”black”を”ブラックさん”と認識した解釈である。これは、”saw”が多義語であり、”見た”以外に”会った”という意味も考えられることが原因である。日本語話者は”見た”のに対し、英語話者は”会った”と判断したため、ブラックさんに対して何をしたかの認識に違いが生じている。3 つ目は、”saw”を”見た”，”black”を”黒絵の具”と認識した解釈である。これは、人名である”ブラック”が”Black”ではなく、”black”と誤訳されたことが原因である。日本語話者が”ブラックさん”を見たのに対し、英語話者は”

黒絵の具”を見た”と判断したため、何を見たかの認識の違いが生じている。英語話者が 2 つ目や 3 つ目の解釈をした場合、日本語話者の意図が英語話者に伝わっておらず、日本語話者と英語話者間で認識のずれが生じたことになる。

次に、英語話者は日本語話者のメッセージに対し返信を行うとする。英語話者は返信メッセージを機械翻訳に投入し、翻訳結果を日本語話者へ出力する。返信メッセージの内容は、日本語話者から送られてきたメッセージの解釈によって変化する。先に示した解釈ごとの返信の例を以下に挙げる。

1. Did you talk to Mr. Black?
2. Why did you meet Mr. Black yesterday?
3. Were you looking for black paint?

1 つ目の”Did you talk to Mr. Black?”の日本語訳は、”あなたはブラックさんと話しましたか”となる。これは、日本語話者が英語話者に対して伝えたい内容の「昨日、本屋でブラックさんを見かけた」ことに対する返信として適切である。この場合、日本語話者と英語話者の認識は一致しているため、コミュニケーションは成功していると考えられる。(図 3) 2 つ目の”Why did you meet Mr. Black yesterday?”と 3 つ目の”Were you looking for black paint”の日本語訳は、それぞれ”なぜあなたは昨日ブラックさんと会ったのですか”、”黒いペンキを探していたのですか”となる。これらは、「昨日、本屋でブラックさんを見かけた」ことに対する返信としては不適切である。この場合、日本語話者と英語話者間で認識のずれによるコミュニケーション齟齬が発生している。

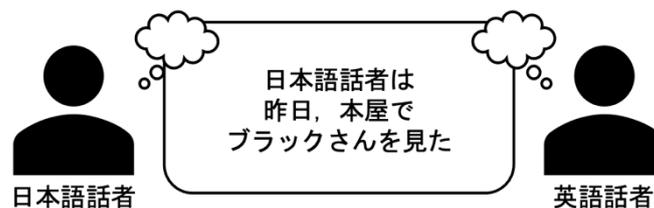


図 3: 日本語話者と英語話者の認識が一致

以上のように、翻訳結果から複数の解釈が考えられる場合があり、受信者が発信者の意図を汲み取った解釈をすることは限らない。また、受信者が発信者の意図とは異なる解釈をした場合、両者間で認識のずれが生じる。これは、機械翻訳

が発信者の意図を反映していない翻訳結果を出力したことが理由として挙げられる。例の場合、日本語話者の意図に反して、“見た”は複数の意味を持つ多義語“saw”で出力され、人名の“ブラック”は“Black”ではなく、“black”で出力された。機械翻訳が“昨日、本屋でブラックを見た”を“ I saw black at the bookstore yesterday”と翻訳したのは、単語や文法レベルでは相応しいと考えられる。しかし、発信者の意図が翻訳結果に反映されていないことは、多言語コミュニケーションにおいて、ユーザ間の相互理解を妨げる原因となる。したがって、翻訳結果は機械翻訳によって一意に決められるものではなく、ユーザが翻訳にインタラクティブに介入できる必要がある。

従来の機械翻訳媒介コミュニケーションは、発信者がメッセージを機械翻訳に入力し、翻訳結果を受信者に出力するという一方向の通信チャネルで行われる。そのため、ユーザは両者間で認識のずれが生じたときに、何が原因で発生したかを判断することができない。そこで、ユーザとのインタラクションを通して翻訳結果を伝える翻訳エージェントを提案する。

2.2 翻訳エージェントによる対話

機械翻訳媒介コミュニケーションで発生する問題を解決するために、翻訳エージェントを用いた対話を提案する。まず、翻訳エージェントの概要を示し、翻訳エージェントを媒介した多言語コミュニケーションの対話方法を述べる。

2.2.1 翻訳エージェント

翻訳エージェントは、林田らが機械翻訳システムをエージェントとして捉えることで、ユーザと機械翻訳のインタラクティブ性を向上させることを目的として提案したシステムである[2]。インタラクティブ性とは、どれだけ相手に伝わっているかという相互作用性を指す。ICE2002において、機械翻訳を用いた多言語コミュニケーションが行われたが、当時の機械翻訳は翻訳品質が低かったために、ユーザが翻訳結果を改善しようと、原文を修正する行為が見られた。原文は発信者が機械翻訳に入力しようとするメッセージを指す。当時の機械翻訳の問題点として以下の3つを挙げている。

- 送信メッセージを修正する上で、どの部分が原因で正しく翻訳されないのか、利用者は知ることができない。
- 利用者には、機械翻訳システムがどのように翻訳処理しているのか見当がつかず、機械翻訳システムのメンタルモデルが適切に形成されない。

- 機械翻訳システム自身が、どの程度翻訳できているのかを認識できないため、利用者に対する問い合わせなどが行えない。

林田らは、これらの問題点から、機械翻訳の翻訳精度だけではなく、ユーザと機械翻訳間のインタラクティビティも重要であるとし、翻訳品質とインタラクティビティの関係性を示し（図 4）、機械翻訳媒介コミュニケーションの支援を行う方法を提案した。

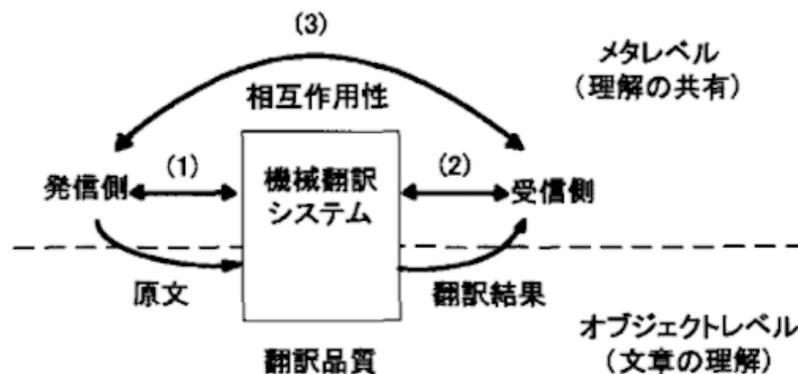


図 4: インタラクティビティと翻訳品質

出典: [2]

また、インタラクティビティの尺度の適用する場面として、以下の 3 つを挙げている。

- (ア) 発信側利用者と機械翻訳間での理解の共有。
- (イ) 機械翻訳と受信側利用者間の理解の共有。
- (ウ) 発信側・受信側利用者間での理解の共有。

特に、(ア)において、ICE2002 で観察されたユーザと機械翻訳間のインタラクティビティの不足として以下の 2 点を挙げている。

- 発信者利用者は、入力した文章中のどの部分が翻訳に適していないのを知ることができない。
- 発信者利用者は、どのように文章を書けば正しく翻訳されるのを知ることができない。

このようなインタラクティビティの不足を解消するために、翻訳エージェントは、機械翻訳のインタラクティビティ向上支援と異文化コミュニケーション支援を行う必要があるとした。翻訳エージェントのアーキテクチャを次のように示している。（図 5）林田らは、翻訳エージェントを用いて、(ア)の発信者と

機械翻訳間での理解の共有に着目し、多言語コミュニケーションの支援を行った。

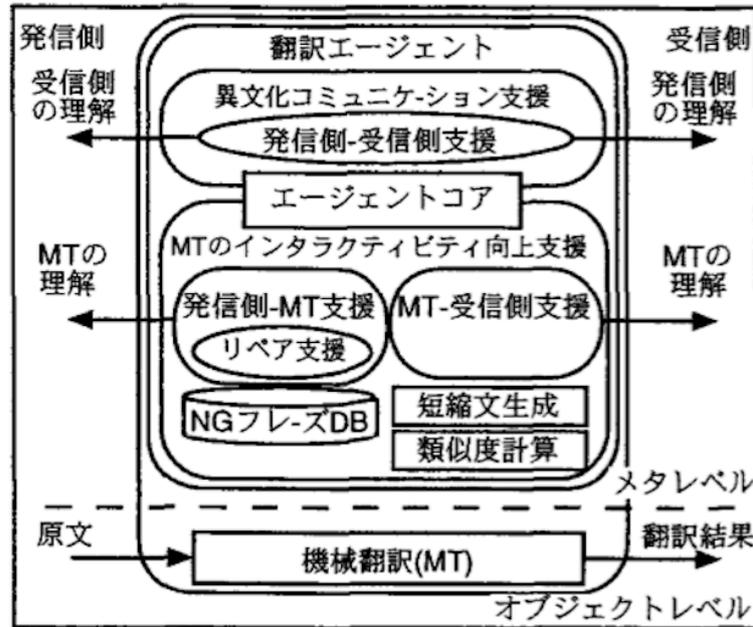


図 5: 翻訳エージェントのアーキテクチャ

出典:[2]

本研究では、(イ)の機械翻訳と受信者間の理解の共有に着目する。機械翻訳と受信者間のインタラクティビティを向上させる方法として、コグニティブサービスを用いた翻訳エージェントによって受信者の表情データと視線データの分析を行う。表情データは、機械翻訳から出力されたメッセージ中の翻訳不適格箇所の有無の判定に使用し、視線データは翻訳不適格箇所の検出に使用する。コグニティブサービスを用いた翻訳エージェントの実装方法は、第 4 章で説明する。

2.2.2 対話方法

次に、翻訳エージェントを媒介した他言語話者間の対話方法について説明する。機械翻訳による対話と同様に、それぞれ異なる言語を話すユーザ 1 とユーザ 2 を想定する。ユーザ 1 からユーザ 2 へのメッセージの発信は、機械翻訳による対話と同様のステップで行われる。(図 6)

1. ユーザ 1 が母国語でメッセージを翻訳エージェントに入力する。
2. 翻訳エージェントが入力されたメッセージをユーザ 2 の言語に翻訳する。

3. 翻訳エージェントが翻訳結果をユーザ 2 に出力する

翻訳エージェントには機械翻訳が組み込まれており，翻訳自体はそれによって行われる．ユーザ 2 がユーザ 1 へメッセージを発信する場合も，機械翻訳による対話と同様に行われる．



図 6: 翻訳エージェントによる対話

次に，翻訳エージェントの機能について説明する．翻訳エージェントは，受信者であるユーザ 2 の表情データと視線データを分析し，翻訳結果の解釈がスムーズに行われているかを判定する．以下のステップで翻訳エージェントはユーザ 2 とインタラクションを行う．

1. ユーザ 2 の表情データから翻訳不適格箇所の有無を判定する．(図 7)
2. 1 で翻訳不適格箇所が発生していると判定した場合，ユーザ 2 の視線データから翻訳不適格箇所を検出する．(図 8)
3. 翻訳不適格箇所を修正し，修正後の翻訳結果を出力する．(図 9)

翻訳エージェントの動作を，機械翻訳による対話で挙げた日本語話者と英語話者間の対話の例で説明する．日本語話者は「昨日，本屋でブラックさんを見かけた」ことを英語話者に伝えるために，“昨日，本屋でブラックを見た”と翻訳エージェントに入力する．翻訳エージェントは入力された日本語を”I saw black at the bookstore yesterday”と英語に翻訳し，英語話者に出力する．このとき，英語話者は翻訳結果の解釈に悩み，表情が硬く，視線が”saw”と”black”に集中しているとする．すると，翻訳エージェントは，英語話者の表情データから翻訳結果から翻訳不適格箇所が発生していると判断する．次に，視線データから翻訳不適格箇所が”saw”と”black”で発生していると予測する．そして，翻訳エージェントは”saw”を類義語の”looked at”，”black”を人名の”Black”に修正し，“I looked at Black at the bookstore yesterday”を出力する．この修正後の

翻訳結果から英語話者は、”saw”は”見た”，”black”は”ブラックさん”を表していたことが分かり，同時に日本語話者の意図を理解できる．



図 7: 表情データから翻訳不適格箇所の有無を判定



図 8: 視線データからの翻訳不適格箇所を検出



図 9: 翻訳不適格箇所を修正し，ユーザ 2 へ出力

以上のように，翻訳エージェントによる対話は，受信者による翻訳結果の解釈の支援を行う．これにより，複数の解釈が考えられるような状況でも，受信者はどの解釈をすべきかを判断できる．また，表情データと視線データを

用いるため、受信者自信が翻訳結果の修正を行う必要がなく、それに伴う負担が受信者にかからない。したがって、コミュニケーションを円滑に行うことができる。

第3章 コグニティブサービスによる認知分析

本章では、コグニティブサービスについて説明する。大別して、表情認識サービスと視線トラッキングサービスについて、順に説明する。

コグニティブサービスとは、人の認知をする能力をサポートすることを目的としたサービスを指す。数値やテキストのような構造化データだけではなく、自然言語や画像といったデータも利用することで、人間には知覚できない情報を含めた認知情報を提供する。また、無意識に知覚した情報をコンピュータで分析することで、人間の知覚範囲の拡張を行うことができる。本研究では、翻訳結果の翻訳不適格箇所を検出するために、コグニティブサービスとして表情認識サービスと視線トラッキングサービスを利用した。それぞれのサービスについて説明する。

3.1 表情認識サービス

受信者の表情データを読み取るために、表情認識サービスとして、Vincent Mühler によって作成された JavaScript API の `face-api.js` を利用した。`face-api.js` は、画像に写っている人の顔を自動で認識、検出、推定するサービスである。特徴として、顔認識、ランドマーク検出、表情の認識、年齢推定と性別認識ができる。また、ウェブカメラを用いて、リアルタイムで表情の認識ができる他、`neutral` (通常) , `happy` (喜び) , `sad` (悲しみ) , `angry` (怒り) , `fearful` (恐怖) , `disgusted` (嫌悪) , `surprised` (驚き) といった7種類の表情を認識できる。また、表情認識の精度が高く、本研究では、表情認識の機能を用いて、受信者の表情データを取得した。

3.2 視線トラッキングサービス

受信者の視線データを読み取るために、視線トラッキングサービスとして、Alexandra Papoutsaki らによって開発された `WebGazer.js` を利用した。`WebGazer.js` は、ウェブカメラを用いて、ユーザのページ上の視線位置をリアルタイムで推測するサービスである。特徴として、JavaScript で記述されているため、個人のウェブページへの組み込むコストが低く、クライアントのブラウザ上で実行できるメリットがある。しかし、視線認識の精度が低いデメリットがあった。デフォルトでは、ユーザの視線位置はブラウザ上に赤色のポインタで表示さ

れる．本研究では，この視線トラッキング機能を用いて，受信者の視線データを取得した．

第4章 システムアーキテクチャ

第3章で説明したコグニティブサービスを用いた翻訳エージェントの構成を示す。大別して、コグニティブサービスを利用した翻訳不適格箇所の検出と修正について、順に説明する。

初めに、翻訳エージェントの概要を示す。チャットシステムを介した多言語コミュニケーションを想定し、そのチャットシステムにユーザの視線と表情の情報を元に翻訳不適格箇所を分析し、翻訳結果を修正する翻訳エージェントを組み込む。以下に、翻訳エージェントのシステムアーキテクチャを示す。(図10)

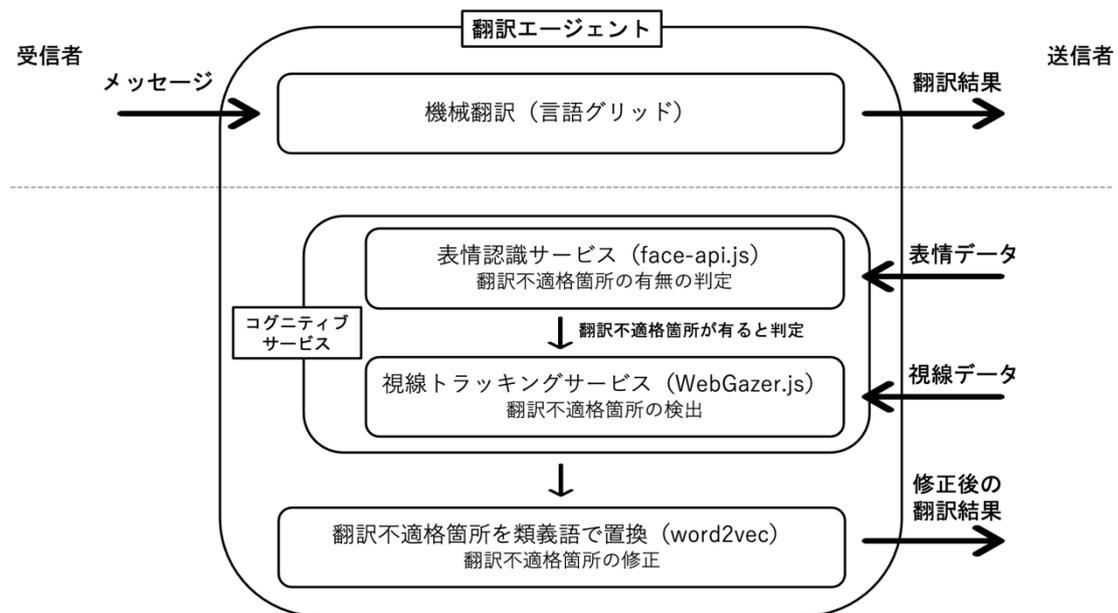


図10: 翻訳エージェントのアーキテクチャ

まず、翻訳エージェントの機能について説明する。翻訳エージェントには、多言語コミュニケーションで翻訳を行うための機械翻訳を搭載する。機械翻訳には言語グリッドを使用する。言語グリッドとは、インターネット上の言語資源や機械翻訳等の言語処理機能を自由に組み合わせて使うことができるサービスを指す。翻訳エージェントは、言語グリッドで送信者のメッセージを翻訳する。それと同時に、表情認識サービスを用いて、ユーザの表情データを分析する。対話中は、常にユーザの表情データを取得する。表情データから、翻訳結果に翻訳不適格箇所があると判定された場合、視線トラッキングサービスを用いて、ユーザの視線データを分析する。視線データの分析によって、翻訳不適格箇所を検出し、

翻訳不適格箇所の修正を行う。修正は、翻訳不適格箇所を類義語で置き換える。そして、修正後の翻訳結果をユーザに出力する。以上の翻訳エージェントの機能は、対話中に繰り返し行う。次に、翻訳不適格箇所の検出と修正の実装方法を説明する。

4.1 翻訳不適格箇所の検知

翻訳不適格箇所を検出する前に、翻訳結果に翻訳不適格箇所が含まれているかを判定する必要がある。よって、翻訳不適格箇所の有無の判定と、翻訳結果に翻訳不適格箇所があると判定されたときの翻訳不適格箇所の検出方法について述べる。

4.1.1 翻訳不適格箇所の有無の判定

翻訳不適格箇所の有無の判定には、表情認識サービスの `face-api.js` を利用する。まず、翻訳結果中の翻訳不適格箇所の有無によるユーザの表情の変化についての研究を示す[4]。向井らは、専門家が翻訳した訳文と評価したい訳文を比較したときに誤訳が生じているかを被験者に判断させる実験を行なった。実験に使用した 50 文のうち、誤訳と感じなかった場合の表情の時系列データ 10 文分のデータをグラフにプロットしたものを図 11 に、誤訳と感じた場合の表情の時系列データ 10 文分のものを図 12 に示している。それぞれのグラフは、横軸が訳文を表示してからの経過時間（秒）、縦軸がその際の表情データ値を示している。図 11 と図 12 から誤訳と感じなかった場合の表情データは、表情値の変化が少ないのに対し、誤訳と感じた場合の表情データは、`neutral` の数値が一度下がってから元に戻るという変化が生じたことを示した。これは、誤訳を見た被験者の表情が大きく動くことを意味している、と結論づけた。

以上から、翻訳結果に翻訳不適格箇所が生じている場合、ユーザの表情が大きく動くと考えられる。`face-api.js` は `neutral`（通常）、`happy`（喜び）、`sad`（悲しみ）、`angry`（怒り）、`fearful`（恐怖）、`disgusted`（嫌悪）、`surprised`（驚き）の 7 種類の表情を認識する。翻訳不適格箇所が生じていない場合は、受信者が `neutral` の表情を示し、翻訳不適格箇所が生じている場合は、一時的に `neutral` 以外の表情を示すと定義することで、表情データから翻訳不適格箇所を検出することが可能となる。

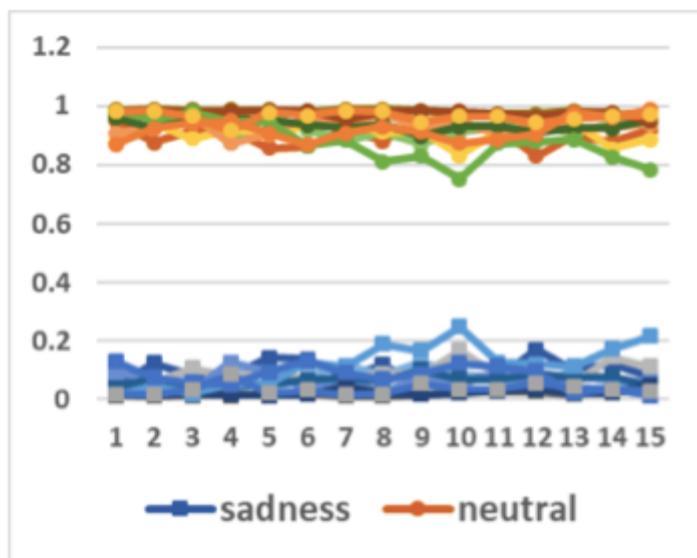


図 11: 誤訳と感じなかった場合の表情の時系列データ
出展:[4]

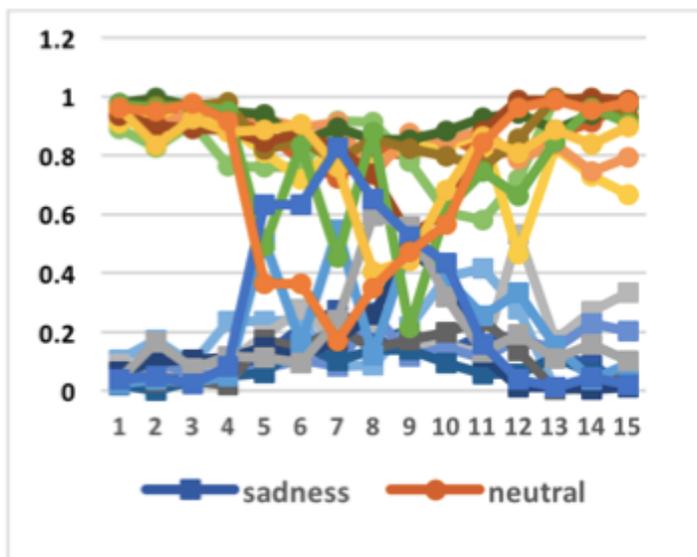


図 12: 誤訳と感じたときの表情の時系列データ
出展:[4]

4.1.2 翻訳不適合箇所の検出

翻訳不適合箇所の検出には、視線トラッキングサービスの **WebGazer.js** を利用する。検出方法として、受信者は翻訳不適合箇所があった場合、そこに視線が集中するとし、翻訳不適合箇所に対する視線の再訪回数が複数回行われるとする。また、翻訳不適合箇所は翻訳結果の形態素ごとに検出する。まず、送信者がチャットに入力したメッセージを、言語グリッドで翻訳する。次に、翻訳結果を形態素解析し、形態素ごとに **html** タグで表示する。そして、視線座標がタグの範囲内に入った時に、クリックイベントを発生させ、発生したクリックイベントの回数を計測する。これにより、視線データを用いて、翻訳不適合箇所の検出を行うことができる。

4.2 翻訳不適合箇所の修正

翻訳不適合箇所を修正するために、**word2vec** を用いて翻訳不適合箇所の類義語を獲得する。**word2vec** とは、Tomas Mikolov らによって提案された、単語をベクトルで表現する手法である。翻訳不適合箇所を **word2vec** によって獲得した類似語で置き換え、修正後の翻訳結果を出力する。以上の方法で、翻訳不適合箇所の修正を行うことができる。

第5章 考察

前章において、コグニティブサービスを組み込んだ翻訳エージェントの構成を説明した。図 13 と図 14 は本研究で実装したチャットシステムである。図 13 は日本語話者である「タロウ」のチャットシステムを、図 14 は英語話者である「John」チャットシステムの画面を示している。チャット内容がそれぞれの言語で翻訳されており、ユーザは自身の言語を用いてコミュニケーションを行うことができる。

しかし、翻訳エージェントの機能のうち、翻訳不適格箇所の検出を正確に行うことはできないと考えられる。視線トラッキングサービスとして、WebGazer.js を用いたが、実際の視線と WebGazer.js の認識した視線データが一致しないことが多かった。改善案として、チャットの文字を大きくする方法が考えられるが、その場合、チャットシステムのユーザビリティを考慮する必要がある。他に、翻訳不適格箇所の検出に視線の再訪回数の計測が有効であるかを評価する必要がある。

また、翻訳不適格箇所の修正に、Wikipedia で学習させた word2vec から類義語を獲得する手法を提案したが、BERT 等を用いて、翻訳不適格箇所の前後の単語との関係を考慮した同義語を獲得できる。したがって、これらを用いることで、メッセージ中で最も適切であると判定された同義語で置き換えられ、翻訳不適格箇所の修正の精度を向上できると考えられる。

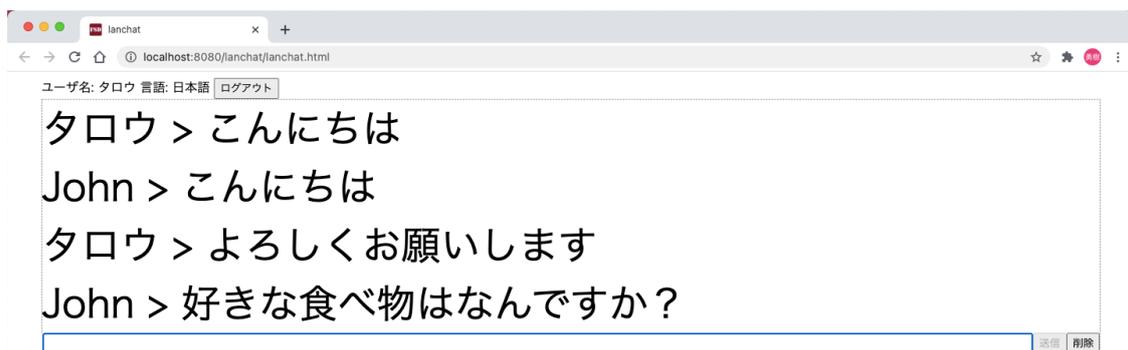


図 13: 日本語話者「タロウ」に対応したチャットシステム

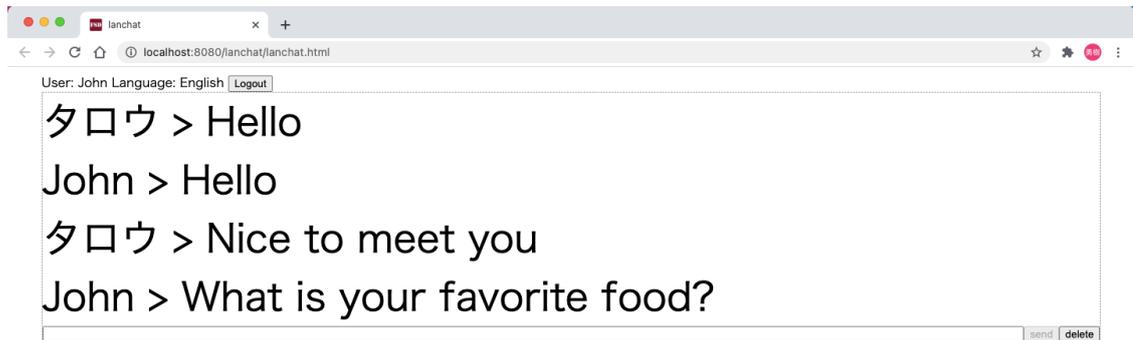


図 14: 英語話者「John」に対応したチャットシステム

第6章 おわりに

多言語コミュニケーションの支援のために、機械翻訳が用いられる。機械翻訳の翻訳精度は大きく向上したが、機械翻訳を使うことで生じる問題もある。

本研究では、機械翻訳媒介コミュニケーションにおける問題点を挙げ、多言語コミュニケーションを支援する方法として、ユーザとのインタラクションを行いながら翻訳結果を伝える翻訳エージェントを提案した。また、コグニティブサービスを組み込んだ翻訳エージェントの実装方法を説明した。

本研究の貢献は以下の2点である。

翻訳不適格箇所の検出

翻訳不適格箇所を検出するために、コグニティブサービスを用いて、メッセージの受信者の表情データと視線データを利用する方法を提案した。表情データと視線データの計測方法を変更することで、翻訳不適格箇所の検出の精度を向上できると考えられる。

翻訳不適格箇所の修正

翻訳不適格箇所を修正するために、Wikipedia のデータを用いて学習した word2vec によって翻訳不適格箇所を類義語で置き換える手法の提案した。BERT を用いて、メッセージ中で最も適切であると判定された類義語で置き換えることで、翻訳不適格箇所の修正の精度を向上できると考えられる。

本研究では、機械翻訳と受信者間のインタラクションを行う翻訳エージェントを提案したが、送信者と機械翻訳間、送信者と受信者間のインタラクションも行う翻訳エージェントであれば、対話者間のインタラクティブリティをより高められると考えられる。

また、コグニティブサービスとして、表情認識サービスと視線トラッキングサービスを提案したが、インタラクティブリティ向上のために使用できるコグニティブサービスはそれら以外もあると考えられる。よって、コグニティブサービスの検証が必要である。

最後に、チャットシステム上で行われる対話を想定したが、多言語コミュニケーションは、そのようなテキストコメントベース以外に、日常会話のように会話ベースでも行われる。コミュニケーションの円滑性を重視した場合、表情認識サービスや視線分析サービスを以外のコグニティブサービスを考える必要がある。コグニティブサービスはコミュニケーションが行われる環境に適したもの

を利用すべきである.

謝辞

本研究を行うにあたり，熱心なご指導，ご助言を賜りました村上陽平准教授に深謝申し上げます。また，普段からお世話になっている社会知能研究室の皆様
に心より感謝申し上げます。

参考文献

- [1] 瀬上和典: 機械翻訳の限界と人間による翻訳の可能性, *AGLOS: Journal of Area-Based Global Studies* pp.1-23(2018).
- [2] 林田尚子, 石田亨: 翻訳エージェントによる自己主導型リペア支援の性能予測, *電子情報通信学会論文誌, D-I, Vol.J88-D-I, No.9*, pp.1459-1466(2005).
- [3] Chunai SHI, Toru ISHIDA, Donahui LIN: Translation Agent: A New Metaphor for Machine Translation, *New Generation Computing*, pp.163-186(2014).
- [4] 向井涼, 田中高志, 大谷雅之 (近畿大学): 機械翻訳における人の振る舞いを用いた誤訳検出.
- [5] 石田亨, 村上陽平, 稲葉利江子, 林冬恵, 田中正弘: 言語グリッド: サービス指向の多言語基盤, *電子情報通信学会論文誌 D, Vol.J95-D, No.1*, pp.2-10(2012).