

# 共同翻訳のためのプロトコルと支援システムの開発

## Designing Protocols and Support Systems for Collaborative Translation

森田 大翼†  
Daisuke Morita

石田 亨†  
Toru Ishida

### 1 はじめに

近年、インターネット等の通信基盤の整備は世界中でなされている。人々はPC等の通信端末を用いて容易に世界中の情報を受発信できる。しかし、実際には言語の壁の問題により取得可能な情報は制限される。

ここで、機械翻訳は有用なツールである。しかし、一般に機械翻訳の品質は実用に耐える程高くない。そこで、人間が機械翻訳による翻訳作業に参加する。参加者は翻訳機を介して相互に作用し、品質の高い翻訳文の作成を目指す。これが、共同翻訳の構想である。

共同翻訳を考える上で、以下のような課題がある。

#### 1) 作業者間のインタラクションの衝突

両者ともに同一の文章に手を加えることで、一方の作業が他方の作業に支障をきたす可能性がある。その結果、作業の競合が発生し、作業内容の一部あるいは全部が失われる。あるいは、他方の作業に支障をきたす可能性を極度に恐れ、作業が抑制される。その結果、作業効率が落ちる。また、作業者にストレスを与える。

#### 2) 文意の正確な伝達

機械翻訳で翻訳された文章が元の文の意味通りに翻訳されない可能性がある。また、翻訳文の解釈に誤解が生じた場合、翻訳文の修正によって翻訳文の文意が原文の文意から更に離れる可能性がある。

本研究の目的は上記の課題を解決し、共同翻訳により品質の高い翻訳文を生成するためのプロトコルの設計を行うことである。また、共同翻訳の支援システムを実装し、プロトコルの効果を検証する。

まず、2章にて共同翻訳の概要と、共同翻訳にプロトコルが必要であることを示す例を述べる。3章にてプロトコルの提案を行う。4章にて実験によるプロトコルの評価を行う。

### 2 共同翻訳

まず、共同翻訳の定義を示す。作業者は文書の翻訳元言語を理解する人間(原文側)と、翻訳先言語を理解する人間(翻訳文側)であり、それぞれはバイリンガルではない。その2言語間の翻訳は翻訳機が行う。また、作業者は翻訳する文書のドメインを理解している。作業者は端末を用いてネットワークを介して作業する。

共同翻訳は文書の翻訳を対象とし、チャットのような対話は扱わない。そのため、作業の同期性は緩和される。また、翻訳対象となる原文は完成されたものであり、手を加えられることはない。しかし、原文側は原文の参考のために言い換えを提示することはできる。

翻訳文側は翻訳文のFluency(流暢さ)を評価する。翻訳文側では翻訳文と原文の意味的な類似度を評価できない。機械

翻訳で翻訳された文章が完全に流暢であることは稀である。そのため、翻訳文側は流暢でない翻訳文を修正する作業を行う。ここで、人間が修正した文章は必ずFluentであると仮定する。

原文側は翻訳文側が修正した文の翻訳文と原文のAccuracy(意味内容の正しさ)を評価する。翻訳文側と同様、翻訳文と原文の意味的な類似度を評価できない。しかし、原文の言語に翻訳し直した文と原文との意味的な類似度は評価できる。その評価から、翻訳文側に原文の意味内容が正しく伝達しているかを確認する。

図1に、以上の定義を表現した図を示す。

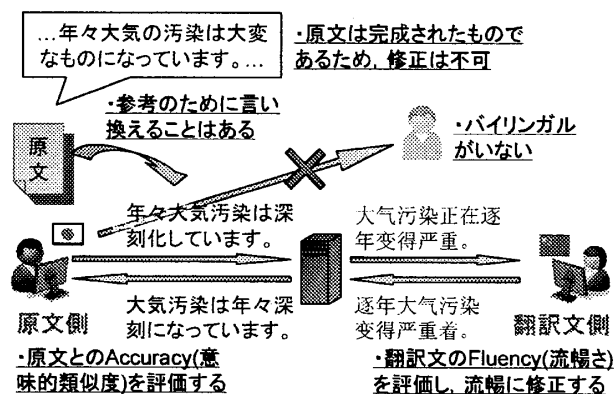


図1: 共同翻訳のイメージ

共同翻訳の適用例として、ある分野の専門家と翻訳機によってその分野の文書を多言語に翻訳する場合が想定される。翻訳機がある分野に特化した翻訳を行うことは不可能である。しかし、人間側はその分野の背景から意味を推測して翻訳文を修正することができる。即ち、ある特定の分野に精通した翻訳者を必要とすることなく、その分野の文書の翻訳を可能にする。

共同翻訳には、これらのプロセスの系列の適切な設計が必要である。即ち、プロトコルが必要である。

共同翻訳にプロトコルが無い場合の問題を例を用いて説明する。共同翻訳を行う最低限のインタフェースとして以下を仮定する。原文、翻訳文、折り返し翻訳文、翻訳文の折り返し翻訳文を表示する4つのテキストエリアがある。ここで折り返し翻訳とは、ある文章を対象言語に翻訳した文を、元の文章の言語に翻訳し直すことである。例えば、日本語の文を中国語に翻訳した後、もう一度日本語に翻訳し直すことである。折り返し翻訳文と原文の比較により、翻訳文の言語を理解していない場合にも、機械翻訳の精度をおおよそ把握できる。原文の中のある一文の修正によって、その他の文の対応する箇所に翻訳による変化が反映される。また、翻訳文のある一文の修正によって、原文以外の対応する箇所に翻訳による変化が反映される。2人の人間はそれぞれ同時に同じ画面を操作する。

† 京都大学大学院情報科学研究科, Graduate School of Informatics, Kyoto University

表 1: 翻訳文側の操作に原文側が支障をきたす例

| 原文                            | 翻訳文                                                         | 折り返し翻訳文                   | 翻訳文の折り返し翻訳文                                         |
|-------------------------------|-------------------------------------------------------------|---------------------------|-----------------------------------------------------|
| まず、自分たちの街を見て、その状況を把握する。       | First the town is seen and its situation is grasped.        | 第一に、町は見られて、その状況は把握される。    | First a town is seen, and its situation is grasped. |
| まず、自分たちの街を見ることによって、その状況を把握する。 | At first, we see the town and grasp the situation.(翻訳文側の修正) |                           |                                                     |
| まず、自分たちの街を見ることによって、その状況を把握する。 | First its situation is grasped by seeing the town.          | 最初のその状況は、町を見ることによって把握される。 | Its first situation is grasped to see a town.       |

表 1 に翻訳文側の操作に原文側が支障をきたす場合の例を示す。まず、1 行目の様に原文が入力され、それに対する翻訳文等が得られる。2 行目では翻訳文の修正作業が行われる。その修正にやや遅れて (ほぼ同時に)、修正が折り返し翻訳等に影響を及ぼす前に、3 行目の様に原文側の修正が行われる。3 行目は、その修正により、2 行目の修正が上書きされた結果を示している。2 行目の修正は適切であるが、作業の衝突により反映されない。即ち、翻訳文側の作業内容が失われている。

表 2: 折り返し翻訳文の品質が低い、翻訳文の品質が低い例

| 原文                      | 翻訳文                                                                | 折り返し翻訳文                          | 翻訳文の折り返し翻訳文                                                                              |
|-------------------------|--------------------------------------------------------------------|----------------------------------|------------------------------------------------------------------------------------------|
| この問題は、初期に見られる機会がない。     | There are no chances found in an early stage for this problem.     | この問題のために初期の段階において発見されたチャンスが全然ない。 | There are no chances found in the stage of the early stage for this problem at all.      |
| 初期に、この問題が発見されるような機会がない。 | There are no chances when this problem is found in an early stage. | この問題が初期の段階において発見される時に、チャンスが全然ない。 | When this problem is found in the stage of the early stage, there are no chances at all. |

表 2 に折り返し翻訳文の品質が低い、翻訳文の品質が低い場合の例を示す。1 行目の様に原文が入力され、それに対する翻訳文等が得られる。原文側は 1 行目の折り返し翻訳文の品質は不十分であると判断し、2 行目の様に修正を行う。しかし、1 行目の翻訳文は少々不自然ではあるが、原文の意味を十分に推測できる。このとき翻訳文側が修正した場合に良い翻訳文が得られる可能性が高い。しかし、原文側の修正により、その修正に対する新たな翻訳文に上書きされる。翻訳文側は前の翻訳文に対する修正を行えない。

上記の 2 例は共同翻訳ではプロトコルを適切に設計し、人間の行動を制御する必要があることを示す。

### 3 共同翻訳プロトコル

共同翻訳では、作業者は他方の様子を直接的に知ることができないため、それに関連する問題が生じる。また、文章の翻訳には機械翻訳を用いるため、それに関連する問題も生じ

る。その問題を以下に示す。

#### 1) 2 人の人間の作業の競合

同一文章に対して 2 人の人間が独立に作業を行うため、作業の競合が起こる可能性がある。プロトコルは操作権制御を行う必要がある [1]。

#### 2) 機械翻訳による文意の伝達の確認

機械翻訳による翻訳は、翻訳文が必ずしも原文と同じ文意を意味するとは限らない。文意が正しく伝達されているかを確認する必要がある。

上記の問題を考慮し、以下にプロトコルの基本の流れを示す。処理は一文単位で行われる。

- 1) 原文側は、原文の修正を行う。その文を翻訳し、翻訳文側に渡す
- 2) 翻訳文側は、受け取った翻訳文の Fluency を評価し、流暢に文を修正する。その文を原文側の言語に翻訳し直し、原文側に渡す
- 3) 原文側は、受け取った文の原文に対する Accuracy を評価する。正当な翻訳かどうかを確認する

上記のプロトコルの形式的記述を図 2 に示す。

原文を  $s$ 、意味内容が表現される文の accuracy の閾値を  $A$ 、意味内容が流暢に表現される文の fluency の閾値を  $F$ 、修正回数の上限を  $N$  とする。文章  $s$  の修正を  $m(s)$ 、翻訳を  $t(s)$ 、文章  $t$  との accuracy を  $a(s,t)$ 、fluency を  $f(s)$  と表す。例えば原文  $s$  の修正文の翻訳文は  $t(m(s))$  のように表す。

- 1) 原文側の処理に入る
- 2) 原文の修正文  $m(s)$  の翻訳文  $t(m(s))$  を作成し、翻訳文側の処理に入る
- 3) もし、 $f(t(m(s))) < F$  ならば、翻訳文の修正  $m(t(m(s)))$  の翻訳文  $t(m(t(m(s))))$  を作成して原文側の処理に入り、5)へ
- 4) もし、 $F \leq f(t(m(s)))$  ならば、 $t(t(m(s)))$  を作成して原文側の処理に入る
- 5) 原文側で受け取った文  $s'$  が、 $A \leq a(s',s)$  ならば、プロトコルを終了し、成功を返す
- 6) もし、 $N$  回目の原文側が修正をしていれば、プロトコルを終了し、失敗を返す
- 7) 2)へ

図 2: 基本プロトコルの形式的記述

翻訳文の品質が非常に低い時、翻訳文側が適切に修正することはほぼ不可能である。図 2 に示したプロトコルは、その場合に非効率となる。実際の共同翻訳プロトコルでは以下の改良を行う。

各ユーザが文章を修正する際、その翻訳文を即座に相手側に送信せず、折り返し翻訳文を得る。折り返し翻訳文と原文を比較し、折り返し翻訳文の Accuracy が低い場合、その Accuracy が十分高くなるまで文章の修正を繰り返す。各ユーザは翻訳文の品質を高めてから相手側に送信するという動作を行う。

また、文章を修正し続ける場合、その折り返し翻訳文の品質が必ずしも十分に高くなるとは限らない。しかし、その場合に、翻訳文の品質が十分に高い可能性がある。ここで、折り返し翻訳文の品質が原文の意味内容を推測できる程度高いのであれば、その翻訳文の品質は原文の意味内容を推測できる程度高いとみなす。

図 3 に改良したプロトコルの形式的記述を、図 4 にプロトコルのシーケンス図を示す。片側のユーザが一定回数以上

原文を  $s$ 、意味内容が推測できる文の accuracy の閾値を  $LA$ 、意味内容が表現される文の accuracy の閾値を  $UA$ 、意味内容が推測できる文の fluency の閾値を  $LF$ 、意味内容が流暢に表現される文の fluency の閾値を  $UF$ 、修正回数の上限を  $N$  とする。文章  $s$  の修正を  $m(s)$ 、翻訳を  $t(s)$ 、文章  $t$  との accuracy を  $a(s, t)$ 、fluency を  $f(s)$  と表す。

- 1) 原文側の処理に入る
- 2) もし、原文側の修正を  $N$  回行ったなら、プロトコルを終了し失敗を返す
- 3) もし、修正原文  $m(s)$  を作成して、 $a(t(t(m(s))), s) < LA$  ならば、2) へ
- 4) もし、 $LA \leq a(t(t(m(s))), s) < UA$  ならば、
  - a. 翻訳文側の処理に入る
  - b. もし、 $f(t(m(s))) < LF$  または  $UF \leq f(t(m(s)))$  ならば、1) へ
  - c. もし、修正翻訳文  $m(t(m(s)))$  を作成して、 $a(t(t(m(t(m(s)))), m(t(m(s)))) < LA$  ならば、
    - i. もし、翻訳文側の修正が  $N$  回目なら、1) へ
    - ii. c. へ
  - d. 原文側の処理に入る
  - e. もし、 $UA \leq a(t(m(t(m(s))), s) < UA$  ならば、プロトコルを終了し、成功を返す
  - f. 2) へ
- 5) 翻訳文側の処理に入る
- 6) もし、 $f(t(m(s))) < LF$  ならば、プロトコルを終了し、失敗を返す
- 7) もし、 $UF \leq f(t(m(s)))$  ならば、プロトコルを終了し、成功を返す
- 8) もし、修正翻訳文  $m(t(m(s)))$  を作成して、 $a(t(t(m(t(m(s)))), m(t(m(s)))) < LA$  ならば、
  - a. もし、翻訳文側の修正が  $N$  回目なら、プロトコルを終了し、失敗を返す
  - b. 8) へ
- 9) 原文側の処理に入る
- 10) もし、 $a(t(m(t(m(s))), s) \geq UA$  ならばプロトコルを終了し、成功を返す
- 11) もし、翻訳文側の修正が  $N$  回目なら、プロトコルを終了し、失敗を返す
- 12) 翻訳文側の処理に入り、8) へ、

図 3: 共同翻訳プロトコルの形式的記述

の修正を繰り返した時、「失敗でプロトコルを終了」する。即ち、共同翻訳プロトコルの適用で翻訳文が改善しなかった場合に「失敗」となる。

表 3 に上記表 1 の例にプロトコルを適用した場合を示す。1 行目で原文が入力され、その折り返し翻訳文を原文側は得る。原文側はその折り返し翻訳文の品質が原文の意味を推測できる程度高いと判断し、操作権は明確に翻訳文側に移る。翻訳文側は翻訳文を 2 行目の様に修正する。その文を日本語に翻訳し直した文を原文側が十分に Accurate であると判断し、共同翻訳を成功させる。プロトコルの適用により、表 3 で問題であった人間同士の作業の衝突を避けることが可能になる。

表 4 に上記表 2 の例にプロトコルを適用した場合を示す。この場合も同様に、1 行目の折り返し翻訳文の品質は原文の意味内容を完全に表していないが、推測できないほどではない。翻訳文側に修正をさせることができる。2 行目の様に、翻訳文側が修正を行った場合、その折り返し翻訳文は推測できる程度は意味内容を表しているので、原文側に評価を要求できる。原文側は折り返し翻訳文が原文の意味内容を十分に表しているので、成功でプロトコルを終了する。プロトコルの適用により、表 4 で問題であった、翻訳文の品質が良いが、

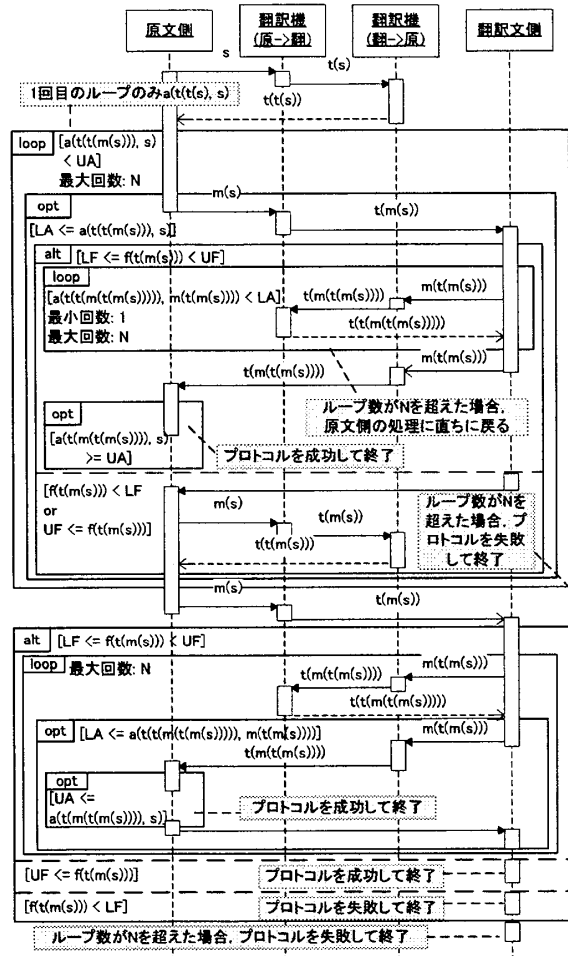


図 4: 共同翻訳プロトコルのシーケンス図

折り返し翻訳文の品質が悪いために翻訳文が棄却されるといふ問題を解決する。

#### 4 実装と初期的評価

共同翻訳プロトコルの評価のため、支援システムの実装を行う。システムは文書を文単位に分割し、独立にプロトコルを適用し、並行して作業を行う。即ち、2 人のユーザは互いの作業を待たない。処理可能な文から次々に可能な作業を行うことで、作業の効率化を図る。支援システムで用いる機械翻訳のモジュールとして、NICT、京都大学を中心とした言語グリッドプロジェクト [2] にて提供される Web サービスを用いる。その他共同翻訳プロトコルを実現する必要な機能を定義し、共同翻訳支援システムは実装された。

翻訳結果において原文の意味をどの程度理解できるか

- 5 完全に理解できる (All)
- 4 大部分を理解できる (Most)
- 3 大体理解できる (Much)
- 2 ほんの少ししか理解できない (Little)
- 1 全く理解できない (None)

図 5: 翻訳文の品質の評価基準

この支援システムを用いて共同翻訳の実験を行う。日本語の短い 10 文の文章を英語、中国語、韓国語、フランス語に

表 3: 表 1 の例にプロトコルを適用した場合

| 原文                      | 翻訳文                                                  | 折り返し翻訳文                | 翻訳文の折り返し翻訳文                                             |
|-------------------------|------------------------------------------------------|------------------------|---------------------------------------------------------|
| まず、自分たちの街を見て、その状況を把握する。 | First the town is seen and its situation is grasped. | 第一に、町は見られて、その状況は把握される。 |                                                         |
|                         | At first, we see the town and grasp the situation.   | 最初は、私達は町を見て、状況を把握する。   | In the beginning we see a town and grasp the situation. |

表 4: 表 2 の例にプロトコルを適用した場合

| 原文                   | 翻訳文                                                                | 折り返し翻訳文                          | 翻訳文の折り返し翻訳文                                                                                   |
|----------------------|--------------------------------------------------------------------|----------------------------------|-----------------------------------------------------------------------------------------------|
| この問題は、初期に発見される機会がない。 | There are no chances found in an early stage for this problem.     | この問題のために初期の段階において発見されたチャンスが全然ない。 |                                                                                               |
|                      | There are no chances that this problem is found in an early stage. | この問題が初期の段階において発見されるという可能性が全然ない。  | There are no possibilities that this problem is found in the stage of the early stage at all. |

翻訳することを実験のタスクとする。日英については3つの組、日中、日韓、日仏については1つの組が実験の参加者となった。また、共同翻訳実験により生成された翻訳文は、それぞれ Accuracy に関して5段階で評価を行う。各評価基準を図5に示す。

図6に各言語について共同翻訳実験を行った結果を示す。各言語の左側の帯は実験に用いた文に機械翻訳のみを適用した時の評価を示す。右側の帯は共同翻訳を適用した時の評価を示す。実験に用いた文書は文毎に5段階評価をする。即ち、日韓、日中、日仏については10文、日英については30文が評価対象となる。図6のグラフは評価値で文を分類し、全体の文の数に対する割合を示す。図6の「失敗」という評価は共同翻訳のプロトコルを失敗して終了したことを示す。

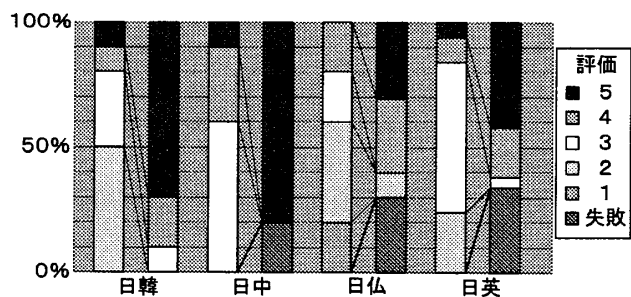


図 6: 機械翻訳のみの場合と共同翻訳を適用した場合の翻訳結果の評価の比較

図6より、機械翻訳のみを適用した場合、翻訳文が5と評価された割合が各言語1割以下であった。しかし、共同翻訳を適用した場合、評価値が5となる割合が機械翻訳のみの適用の場合と比較して大きい。また、評価値が5とならない場合においても、共同翻訳が成功した場合、翻訳精度が概ね向上した。共同翻訳のプロトコルにより、共同翻訳が成功したほとんどの場合は原文の意味が翻訳文に正しく伝わっている。

## 5 おわりに

本研究の貢献は以下の2点である。

- 1) 作業員間のインタラクションの衝突の防止  
プロトコルは、一方が作業している間には他方の作業を抑制するよう設計された。それにより、人間同士の作業の衝突を防止した。
- 2) 折り返し翻訳の導入による文意の正確な伝達の確認  
翻訳文側による翻訳文の修正文を、再度原文の言語に翻訳し直し、原文との Accuracy を比較する。それにより、翻訳文が原文の意味内容を正しく表しているかどうかを確認することが可能になった。また、修正文作成時に折り返し翻訳文を得ることによって、片側のユーザのみの処理で翻訳文の品質を高める。それにより、文意が正確かつ効率的に他方のユーザに伝達することを可能にした。

本研究では、共同翻訳の概念を提案し、そのプロトコルの設計を行った。共同翻訳プロトコルにより、機械翻訳を用いながらも、精度の高い翻訳を可能にすることが期待されることが実験により示された。

## 参考文献

- [1] Ellis, C. A., Gibbs, S. J. and Rein, G. L.: Groupware: some issues and experiences, *Communications of the ACM*, Vol. 34, No. 1, pp. 38-58 (1991).
- [2] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, *IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06)*, pp. 96-100 (2006).
- [3] Yamashita, N. and Ishida, T.: Effects of Machine Translation on Collaborative Work, *International Conference on Computer Supported Cooperative Work (CSCW-06)*, pp. 512-523 (2006).
- [4] Ellis, C. A. and Gibbs, S. J.: Concurrency Control in Groupware Systems, *ACM SIGMOD Record*, Vol. 18, No. 2, pp. 399-407 (1989).
- [5] Shen, H. and Sun, C.: Flexible Notification for Collaborative Systems, *International Conference on Computer Supported Cooperative Work (CSCW-02)*, pp. 77-86 (2002).
- [6] Sun, D., Xia, S., Sun, C. and Chen, D.: Operational Transformation for Collaborative Word Processing, *International Conference on Computer Supported Cooperative Work (CSCW-04)*, pp. 437-446 (2004).