

共同翻訳のためのプロトコルの開発

森田 大翼^{†a)} 石田 亨^{†b)}

Designing Protocols for Collaborative Translation

Daisuke MORITA^{†a)} and Toru ISHIDA^{†b)}

あらまし この論文では、バイリンガルでない異なる言語を用いる 2 人の人間が機械翻訳を利用して協力して文書の翻訳作業を行う方式、共同翻訳のプロトコルを提案する。翻訳機によって翻訳された文章は一般に正確ではなく、また流暢でないため、バイリンガルでなければ翻訳文が原文と同じ意味を表しているかどうかの判断ができない。そこで、この論文では、バイリンガルのいない状態で、(1) 機械翻訳の品質を確実に向上させ、(2) 正確で流暢な翻訳が得られた場合にそのプロトコルを終了する、共同翻訳プロトコルが存在することを示す。基本的なアイデアは翻訳元言語を理解する人（原文側）と翻訳先言語を理解する人（翻訳文側）が果たすべき役割を明確にすることである。すなわち、翻訳文側が機械翻訳文の流暢さについて改善を行い、原文側が翻訳の正確さを評価する。しかし、共同翻訳が成功した場合においても人為的な問題や言語の性質により、100% 正確な結果とならないことがある。我々は、この共同翻訳プロトコルの成功した場合の信頼度を、初期的な実験で評価した。その結果、共同翻訳プロトコルが成功した場合に、その翻訳品質が完全であると評価された文の割合は、日英翻訳で約 83%、日中翻訳で約 91% になるという結果が得られた。更に、共同翻訳の品質を向上させるプロトコルの拡張例を示し、今後の共同翻訳プロトコルの発展が期待できることを示す。

キーワード 機械翻訳、インタラクションプロトコル、分散協調作業、コンピュータ媒介コミュニケーション

1. ま え が き

国際化やインターネットの普及等により、我々は多くの言語を見聞きする状況に遭遇している。それに伴い、メンバーの母語が様々であるコミュニティも数多く発生している。従来、そのようなコミュニティはある一つの共通言語を利用しており、多くの場合は英語であった。しかし、非母語の理解やそれを用いたコミュニケーションは困難となる傾向がある [1] ~ [3]。

そのようなコミュニティにおいて機械翻訳は有用なツールである。機械翻訳を用いることで母語で自分の意思を伝えることができる。実際、機械翻訳を用いて活動する多くのコミュニティが存在している。

一般に機械翻訳は誤訳を多く生成する。翻訳品質が完全でなくても、いくつかの単語から意図をくみ取ることができ、アウェアネスの向上につながる [4] とい

う点では、機械翻訳は有用である。しかし、機械翻訳を用いたコラボレーションでは、その誤訳は双方に誤解を生じさせる。また、機械翻訳では訳語関係が必ずしも 1 対 1 に対応しないという非対称性の性質があり、誤訳の特定は困難である [5]。

この論文は、バイリンガルでない異なる言語を用いる 2 人の人間が機械翻訳を利用して協力して文書の翻訳作業を行う方式、共同翻訳を提案する。従来の翻訳では、翻訳文が原文の意味を表しているかどうかはバイリンガルのみが判断可能であった。しかし、この論文では、バイリンガルのいない状態で、(1) 機械翻訳の品質を確実に向上させ、(2) 正確で流暢な翻訳が得られた場合にそのプロトコルを終了する、共同翻訳プロトコルの存在を示す。

上記の問題を解決するアイデアとして、翻訳元言語を理解する人（原文側）と翻訳先言語を理解する人（翻訳文側）の果たすべき役割を明確にすることを考える。すなわち、翻訳先言語を知っている翻訳文側が機械翻訳文の流暢さについて改善を行い、原文の意味を知る原文側が翻訳の正確さを評価する。これらの事実に基づき、プロトコルを適切に設計する。しかし、

[†] 京都大学大学院情報学研究所社会情報学専攻, 京都市
Department of Social Informatics, Kyoto University,
Yoshida-Honmachi, Sakyo-ku, Kyoto-shi, 606-8501 Japan

a) E-mail: morita@ai.soc.i.kyoto-u.ac.jp

b) E-mail: ishida@i.kyoto-u.ac.jp

共同翻訳が成功した場合においても人為的な問題や言語的性質により、100%正確な結果とならない可能性がある。本研究では、初期的な実験によりプロトコルの成功率と信頼度を評価する。また、共同翻訳の品質を向上させるプロトコルの拡張例を示し、今後の共同翻訳プロトコルの発展が期待できることを示す。

2. 人間の補助を含む機械翻訳方式

2.1 異文化コラボレーション現場での実践

実際に、多くの異文化コラボレーションの現場ではコミュニケーションツールとして機械翻訳を利用している。国際的に活動している日本のNPO団体について例を挙げる。この団体は、韓国、オーストリア、ケニアの団体と連携して活動している。それらの団体のメンバーの母語は日本語、韓国語、ドイツ語、英語と様々である。

多言語の母語話者で構成されるコミュニティでは、共通言語としてしばしば英語が用いられる。しかし、英語を堪能に使うことができない人間が存在し、そのような人が情報を共有することが困難になる傾向があるという問題が発生する。その問題を解決し、国際間で情報の共有を促進して議論を活性化させるために、これらの団体間では機械翻訳を用いたWeb掲示板システムを独自に開発して利用している。図1にこのWeb掲示板システムのイメージを、日本語記事の投稿を例として示す。このWeb掲示板システムでは、各人は自分の母語で記事を編集する。その記事はシステム内部で翻訳され、各人はその記事を母語で閲覧する。しかし、機械翻訳の品質はしばしば低く、そのことがメンバー間での情報共有を困難にさせる場合がある。そこで、このシステムでは機械翻訳された文を手

によって自然な表現になるように修正する操作を可能にしている。また、文章中の文脈やコミュニティ内での共通知識を用いて文意を推測して修正することで、翻訳品質自体が改善される場合がある。これにより、団体内での適切な情報共有を可能にしている。例1に、翻訳文の修正による翻訳品質の改善の様子を示す。

(例1) 修正による翻訳品質の向上

“その絵を見た子供たちはみんな驚きました”という文を機械翻訳すると“Everyone was surprised at the children who saw the picture.”と翻訳される。この文の意味は原文とは異なるが、修正者は文脈やコミュニティの背景から意味を推測し、“Children were surprised to look at the picture.”と修正した。この修正された文は原文の意味と全く正しくなっている。

2.2 翻訳文の修正における問題点

翻訳文の修正で、文意が明確になり、団体内での情報共有が容易になる。この点で、異文化コラボレーションの現場で有用な方式であるといえる。しかし、この人の補助を含む機械翻訳方式の単純な実装には、主に二つ問題がある。以下にその問題を明らかにする。

(例2) 修正者の誤解

“彼は風邪をなおすのに1週間必要でした”という文を機械翻訳すると、“He was necessary to correct a cold for 1 week.”というように翻訳される。この文は文法、語法に誤りがあり、自然な表現に修正されることになるが、修正者はこれを“He should recover from a cold within 1 week.”と修正した。これは、“彼は1週間以内に風邪をなおすべきです”という意味になる。

(例3) 文意の理解不能

“彼はお腹が出ている”という文を機械翻訳すると、“A stomach has gone out to him.”と翻訳される。この文は修正者にとって理解不可能な文であるため、この文は修正されないまま残った。

人の補助を含む機械翻訳方式において、

- 修正された翻訳文が翻訳前の文の意味と正しいかどうか判定できない
- 翻訳文の意味が理解できない場合、修正されず品質の低いままになる

という問題がある。この方式は、翻訳文の品質を改善して情報を共有する手法として機械翻訳を用いたコラボレーションの現場で有用な手法であるが、これら二つの問題を解決できない限り、正しい情報共有をし、誤解を防ぐという課題を根本的に解決できない。

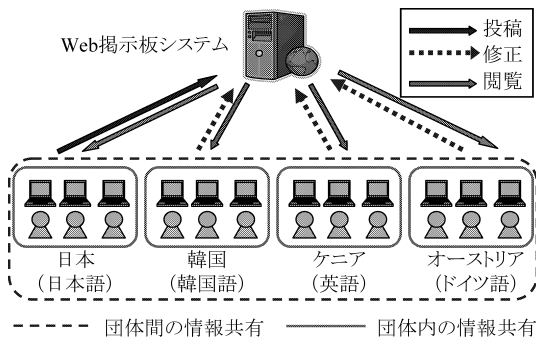


図1 Web 掲示板システムのイメージ

Fig. 1 Illustration of Web BBS system.

3. 共同翻訳

3.1 定義

参加者はバイリンガルでない2人の人間であり、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）に分かれる。その2言語間の翻訳は翻訳機が行う。作業者は端末を用い、ネットワークを介して作業する。共同翻訳は文書翻訳を対象とする。ここで、もとの原文自体には手は加えられないが、原文側は原文の参考のために代替案を提示できる。

また、原文側と翻訳文側の役割を明確にする。翻訳文側では翻訳文と原文の意味的類似度を評価できないが、翻訳文の流暢さは評価できる。翻訳文側は翻訳文から文意を推測し、修正して翻訳文を流暢にする。また、人が修正した文は常に流暢であるとする。翻訳文側と同様、原文側は翻訳文と原文の意味的類似度を直接評価できない。しかし、翻訳文側の修正翻訳文を、原文の言語に翻訳し直した文と原文が意味的に類似しているかは評価できる。共同翻訳では、その情報を利用し、翻訳文が原文の意味内容を正しく反映しているかを確認する。

図2に、以上の定義を表現した図を示す。本研究では上記に定義された共同翻訳を適切に実行するためのプロトコルを設計する。

3.2 基本プロトコル

効果的なコラボレーションのためには、相互理解の構築が不可欠である[6]。相互理解の構築過程は、情報の提供を行う提案フェーズと、その情報を受け手が理解できたかを確認し合う受取フェーズからなる[7]。共同翻訳では、提案フェーズは原文を翻訳して送信することであり、受取フェーズは、翻訳文の受信以降の、その翻訳文の正しさの確認までの処理全般を意味する。

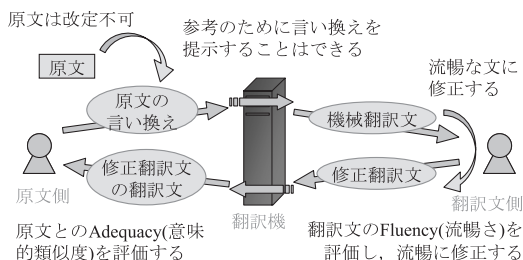


図2 共同翻訳のイメージ
Fig. 2 Basic concept of collaborative translation.

この受取フェーズにおいて翻訳文の正しさが適切に確認できるようにプロトコルを設計する必要がある。

図3に共同翻訳を実現するための最小のプロトコルである基本プロトコルの形式的記述を示す。

このプロトコルでは、上記の定義に加え、機械翻訳文の可読性判定を行っている。翻訳文側の流暢さの評価、原文側の意味的類似度の評価はもとの機械翻訳文が可読でない限り評価することができないためである。

また、プロトコルが停止することを保証するため、原文側及び翻訳文側の修正回数の上限 P, Q が定義されている。プロトコルの処理を繰り返しても翻訳品質の改善が見られない場合は、共同翻訳での品質の改善

原文を s 、文の意味 (content) が理解できる範囲か否かの真偽値を返す関数を c 、原文 s との正確さ (accuracy) が許容される範囲か否かの真偽値を返す関数を a 、文の流暢さ (fluency) が許容される範囲か否かの真偽値を返す関数を f とする。

また、人間による文の修正を示す関数を m 、翻訳機による文の翻訳を示す関数を t とする。例えば原文 s の翻訳文の修正文は $m(t(s))$ のように表す。

原文側の修正回数を p その上限を P とする。また、翻訳側の修正回数を q その上限を Q とする。

- 1) [原文側：原文の送信]
原文側の修正回数 $p := 0$ とする。
原文 s を翻訳文側に送信する。
- 2) [翻訳文側：翻訳文の可読性判定]
 s の機械翻訳結果を $t(s)$ とする。
もし、 $c(t(s))$ が偽なら：
原文側に原文 s の修正を依頼し 6) へ。
もし、 $c(t(s))$ が真なら：
翻訳文側の修正回数 $q := 0$ とする。
- 3) [翻訳文側：翻訳文の修正]
翻訳文側の修正回数 $q := q + 1$ とする。
もし、 $q \geq Q$ なら：
プロトコルを終了し失敗を返す。
もし、 $q < Q$ なら：
流暢性 $f(m(t(s)))$ を満たすよう翻訳文 $t(s)$ を修正した結果を $m(t(s))$ とし原文側に送信する。
- 4) [原文側：折り返し翻訳文の可読性判定]
 $m(t(s))$ の機械翻訳結果を $t(m(t(s)))$ とする。
もし、 $c(t(m(t(s))))$ が偽なら：
翻訳文側に $m(t(s))$ の修正を依頼し 3) へ。
もし、 $c(t(m(t(s))))$ が真なら：5) へ。
- 5) [原文側：折り返し翻訳文の正確性判定]
もし、 $a(t(m(t(s))))$ が真なら：
プロトコルを終了し成功を返す。
もし、 $a(t(m(t(s))))$ が偽なら：6) へ。
- 6) [原文側：原文の修正]
原文側の修正回数 $p := p + 1$ とする。
もし、 $p \geq P$ なら：
プロトコルを終了し失敗を返す。
もし、 $p < P$ なら：
原文を修正した結果を s とし 2) へ。

図3 基本プロトコルの形式的記述
Fig. 3 The formal statement of the basic protocol.

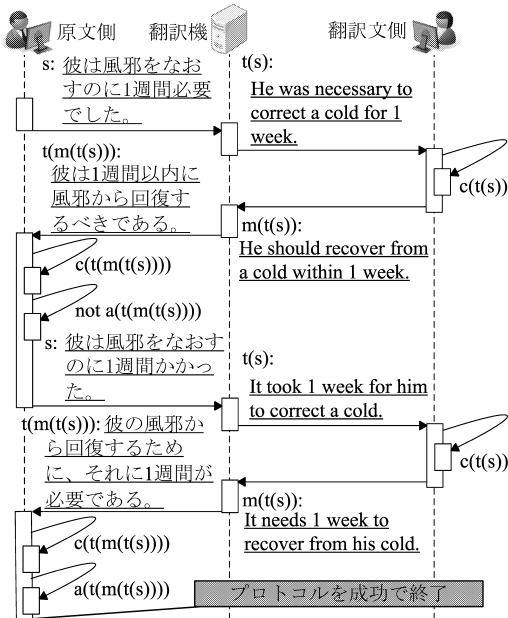


図4 2.2の例2にプロトコルを適用した場合
Fig.4 Application of the basic protocol to the example 2 in Sect. 2.2.

が不可能と判断し、適当な段階で処理を打ち切る。また、実際には人間が思いつくことができる文章の言い換え数は限られる。異文化コラボレーション実験 [4] では、機械翻訳の翻訳品質を改善するために原文を書き直す動作（自己主導型リペア）を最も行ったユーザの平均修正回数は8回であったという報告がある。しかし、これはまれなケースであり、熱心なユーザであっても4, 5回が限度であると考えられる。上限値の設定方針として、「上限値までは改善の可能性があるので努力せよ」と、「上限値以上の修正は時間の無駄になるので努力しない」という2通りの方針が考えられる。異文化コラボレーション実験の結果より、上限値 P, Q は前者の方針では3程度、後者の方針では5程度が妥当であると考えられる。

3.3 有効性

図4に2.2の例2にプロトコルを適用した場合に起こる一連の処理を示す。翻訳文側は最初の機械翻訳文を誤解して解釈し、修正している。原文側はその修正文の折返し翻訳文は原文の意味を表していないと判断する。そこで、原文を修正して再度翻訳文側に送信する。翻訳文側は前回とは異なる解釈をし、翻訳文を修正する。原文側はその文の折返し翻訳文は、十分に原文の意味を表していると判断する。プロトコルは成

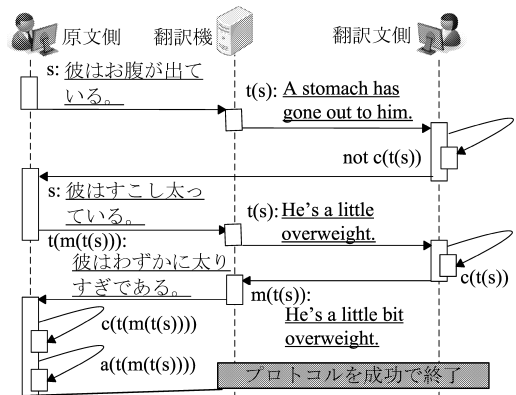


図5 2.2の例3にプロトコルを適用した場合
Fig.5 Application of the basic protocol to the example 3 in Sect. 2.2.

功で終了する。

図5に2.2の例3にプロトコルを適用した場合に起こる一連の処理を示す。原文が入力され翻訳文側に送信される。翻訳文側は最初の翻訳文の意味が理解できないため、原文側に原文の修正を依頼する。原文側によって原文が修正された後、再度得られた機械翻訳文は理解可能であるため、翻訳文側は翻訳文が流暢になるように修正する。原文側はその文の折返し翻訳文は十分に原文の意味を表していると判断する。プロトコルは成功で終了する。

共同翻訳の基本プロトコルでは、作業者は修正翻訳文が原文の意味と正しいかどうかを判断することができる。また、機械翻訳文が理解できない場合に、他方の作業者に翻訳元の文を再度修正させることができる。このプロトコルは人の補助を含む機械翻訳方式における二つの問題点を解決する。

しかし、語の多義性や構文解析のあいまい性等の言語学的な要因に伴う機械翻訳の性能の不完全さが存在する。更に、判断の誤りや、「理解できる」の判断基準が人によって異なるという人的要因も存在する。それらの性質より、共同翻訳が100%正確な結果とならない可能性がある。プロトコルの実験を行い、プロトコルの有効性を実験的に確認する。

4. 初期的実験

4.1 実験設定

基本プロトコルの評価実験を行った。実験に用いた機械翻訳のモジュールとして、独立行政法人情報通信研究機構 (NICT)、京都大学を中心として研究開発、

及び運営されている言語グリッドプロジェクト [8] にて提供される Web サービスを用いた。

NTT のコミュニケーション科学基礎研究所で提供されている機械翻訳機能試験文 [9] 3,718 文から日本語特有の固有名詞を含む文を除き、日本語 100 文を無作為抽出して、実験に用いた。翻訳先言語は英語と中国語とした。また、人的な影響を緩和するため、実験の参加者はそれぞれの言語対について 3 組で行われた。また、本実験では原文側及び翻訳文側の修正回数の上限 P , Q はそれぞれ 3 と設定された。

また、異文化コラボレーションの現場で機械翻訳を改善する手法として広く用いられている折返し翻訳についても同様に実験が行われた。共同翻訳と同様に修正回数の上限は 3 回と設定し、上限回数以内の修正で翻訳文に改善が見られない場合は失敗とみなす。

それぞれの方式で得られた翻訳文は、対となる 2 言語のバイリンガルによって、それぞれ意味の正確さに関して 5 段階 (All, Most, Much, Little, None) で評価が行われた^(注1)。評価は 3 人の評価者によって行われ、その中央値を最終的な評価値として用いた。人間によって修正された文章は完全に流暢であると仮定しているため、流暢さに関する評価は行わない。

また、共同翻訳、及び折返し翻訳による翻訳文の修正作業にかかる時間を計測し、比較する。翻訳に要する時間等のオーバーヘッドは考慮しないとす。共同翻訳、あるいは折返し翻訳を失敗で終了した作業の時間は考慮に入れるものとする。

4.2 評価結果

4.2.1 改善例

“この木の実を食べてもよい” という文の機械翻訳文は “These nuts may be eaten.” である。翻訳文側はこれを誤解して解釈し、 “These nuts may be already eaten.” と修正した。しかし、その折返し翻訳文は “これらのナットはすでに食べられているかもしれない” となり、原文側は翻訳文側の誤解を発見できた。そこで、原文側は “この木の実を食べることは問題ない” と原文を修正した。その翻訳文は “It’s no problem to eat these nuts.” となり、翻訳文側は文意を推測して、 “It’s safe to eat these nuts.” と修正した。その日本語訳は “これらのナットを食べることは安全である” となり、原文と同じ意味であることが確認できた。この翻訳では、評価が “Little” であった翻訳が共同翻訳で翻訳文側の誤解を解消でき、 “All” に改善された。

“交渉は次官クラスで行なわれる” という文の機械

翻訳文は “It’s negotiated for in the undersecretary class.” である。この文は翻訳文側には理解できない。そこで原文側は “交渉は次官クラスのメンバーで行なわれる” と原文を修正した。その翻訳文は “It’s negotiated for by a member in an undersecretary class.” となり、翻訳文側は文意を推測して “The negotiation was done by a member of undersecretary class.” と修正した。その日本語訳は “交渉は次官クラスのメンバーによってされた” となり、原文と同じ意味であることが確認できた。この翻訳では、評価が “None” であった翻訳が共同翻訳で意味の分かる翻訳文に修正でき、 “All” に改善された。

4.2.2 成功率

日英翻訳では折返し翻訳の成功率は約 67%、共同翻訳の成功率は約 70% であった。一方、日中翻訳では折返し翻訳の成功率は約 78%、共同翻訳の成功率は約 62% であった。折返し翻訳と共同翻訳の成功率の間に相関は見られず、ほぼ同程度の成功率であった。

4.2.3 翻訳品質

図 6 は日英と日中の翻訳について、機械翻訳、折

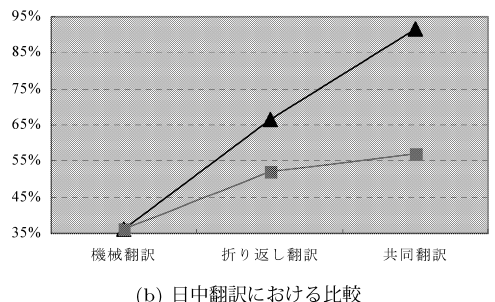
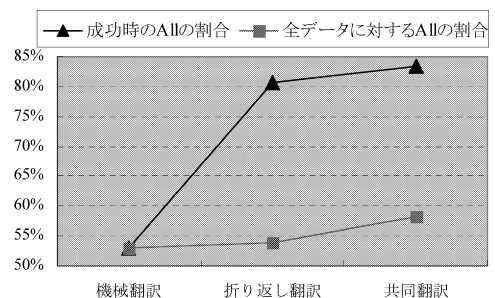


図 6 機械翻訳、折返し翻訳、共同翻訳の評価結果の比較
Fig.6 Results of collaborative translation, back-translation, and pure MT.

(注1): <http://projects.ldc.upenn.edu/TIDES/Translation/TransAssess02.pdf>

返し翻訳，共同翻訳の結果の比較であり，それぞれの方式が成功した場合にその翻訳結果の評価が“All”となった割合と，失敗した場合も含む全データに対して翻訳結果の評価が“All”となった割合を示している．共同翻訳が成功した場合，日英で約83%，日中で約91%が完全な翻訳品質となった．共同翻訳プロトコルが成功していれば，その文の正しさに関して高い信頼をおけるといえる．前章に，理論的にはプロトコルが成功しても正しい翻訳文が生成されない可能性があることを示したが，本実験によりその可能性は小さいことが確認された．また，折返し翻訳と比較した場合，翻訳品質が完全となる割合は，成功時のみのデータ，失敗した場合を含めた全データのそれぞれと比較して共同翻訳が大きいことが示されている．

表1は機械翻訳結果に対する共同翻訳結果の評価値の推移を示している．3組の被験者により実験が行われたため，300文の共同翻訳による結果が得られている．機械翻訳より評価値が下がった文が日英で5%，日中で1%程度と若干見られたが，プロトコルが成功で終了している場合，その翻訳結果の評価は機械翻訳よりおおむね高くなっている．この結果は共同翻訳は機械翻訳の品質を低下させることがほとんどなく，確実に翻訳品質を向上させていることを示している．

4.2.4 作業時間

折返し翻訳を用いたときの文章の修正に要する作業時間は1文当たり平均，日英で40.2秒，日中で29.1秒であった．一方，共同翻訳では，原文側が1文当たり

の作業に要する時間が日英で48.1秒，日中で56.2秒であり，翻訳文側が日英で30.4秒，日中で42.9秒であった．共同翻訳では1文ごとに並行してプロトコルを実行できるため，共同翻訳に要する時間は原文側と翻訳文側の作業時間の和ではなく，最大値である．機械翻訳に要する時間等，システムのオーバーヘッドを考慮すると，本実験で用いられた文に対して，共同翻訳は1文当たり平均1分前後の処理時間を要した．

機械翻訳のみの場合や，折返し翻訳を用いた場合と比較すると共同翻訳は処理が多く，多くの時間を要する．しかし，上述の実験結果より共同翻訳はその品質を大きく向上させることが示されており，翻訳文は人手で修正されるため，流暢さに関しては特に優れている．異文化コラボレーション現場での情報共有手段として有効であるといえる．作業時間は大きくはなるが，その有用性のすべてが失われるわけではない．

5. 基本プロトコルの改善

5.1 基本プロトコルの限界

4.より，基本プロトコルが成功で終了した場合，その翻訳文は正しいことがほぼ保証される．しかし，正しい翻訳文を生成するとき，常にプロトコルが成功終了することは保証しない．以下にその例を示す．

(例4) 修正翻訳文の翻訳文の文意の理解不能

“私はプリンを作る食材を買ってきた？”という原文は，“I have bought the ingredients which make pudding.”と翻訳される．翻訳文側はこれを“I have bought ingredients for a pudding.”と修正した．この修正翻訳文は原文と正しい意味をしているが，日本語訳は“私はプディングに成分を買った？”となり，意味が理解不能な文である．翻訳文側は再度修正翻訳文の作成を要求された．しかし，修正を繰り返しても原文側は翻訳文を理解できず，修正回数の上限を超えプロトコルは失敗で終了してしまった．

翻訳文の意味が伝わらなければ，適切な翻訳文の作成はできない．しかし，折返し翻訳文の意味が理解できない場合でも，その翻訳文が正しいという可能性がある．この点に着目し，プロトコルの拡張を行う．

5.2 拡張プロトコル

例4の場合，原文側は翻訳文の正しさの評価ができない．よって，翻訳文側に翻訳文の正しさの評価をさせる方法を考える．拡張プロトコルでは，同じ意味をもつ異なる表現の二つの原文に対し，翻訳文を得，翻訳文側がそれらを比較する．二つの翻訳文の意味が一

表1 機械翻訳結果に対する共同翻訳結果の評価値の推移
Table 1 Changes in adequacy scores of results of collaborative translation from scores of pure MT.

(a) 日英翻訳における推移

		共同翻訳					
		5	4	3	2	1	失敗
機械 翻訳	5	117	10	0	2	3	27
	4	21	5	0	0	0	22
	3	15	5	4	0	0	18
	2	11	1	2	0	0	16
	1	11	1	1	0	1	7

(b) 日中翻訳における推移

		共同翻訳					
		5	4	3	2	1	失敗
機械 翻訳	5	84	2	1	0	0	21
	4	30	6	0	0	0	18
	3	39	1	1	0	0	22
	2	16	2	1	2	0	51
	1	2	0	0	0	0	1

致すれば、原文の意味を表した翻訳文であるとみなす。

図 7 に拡張プロトコルを示す。基本プロトコルを継承しているため、図 7 にはその差分を示す。

図 8 に例 4 に拡張プロトコルを適用した場合の処理を示す。翻訳文側は、修正回数の上限の Q 回の翻訳文の修正を繰り返しても、その折返し翻訳文の意味が原文側に伝わらなかった。そこで、原文側は原文の言

い換え s' を提示する。翻訳文 $t(s')$ の意味は、前の翻訳文 $t(s)$ と同じ意味であるため、修正翻訳文 $m(t(s))$ を翻訳結果としてプロトコルを成功で終了した。このとき、翻訳結果は原文の意味を正しく表している。

基本プロトコルで解決できない問題が、プロトコルの拡張により解決できることが示された。共同翻訳はその成功確率を向上させる工夫が存在する。

5.3 評価

基本プロトコルの評価実験において、翻訳文側の修正回数 q が上限 Q 以上となった場合の文について拡張プロトコルを適用し、その成功率と信頼度を評価した。

表 2 は基本プロトコルと拡張プロトコルの実験結果を比較したものである。日英、日中ともに、拡張プロトコルの成功率は 10% 程度基本プロトコルより上昇した。日英翻訳でプロトコル成功時に評価値が “All” となった割合が減少しているが、全体として翻訳品質が完全になる文の数は増えており、翻訳品質改善の有効性が示されている。

実際の実験で拡張プロトコルによる改善が見られなかった文の例を示す。“以上の考察から、次の式が得られる。”という原文に対し、原文側は“前述したように考察した結果、次の式を得た。”という言い換えを提示した。それぞれ機械翻訳文は“The above mentioned consideration gives us the next system.”と、“After it was considered as it was above-mentioned, I got the next system.”となった。翻訳文側はそれらの意味が同じであると判断し、“The above mentioned description gives us the system.”が翻訳結果となった。しかし、英文の “system” は「方式、流儀」という意味であるのに対し、原文の「式」は「数式」等の意味であるため、評価値は “Much” と評価された。

二つの表現が異なる原文を作成した場合に、それぞれに同じ誤訳が生じた場合、翻訳文側がそれを誤訳と判定することは困難である。拡張プロトコルでは、このような問題は解決不可能である。これが原因で、翻

- c を文の意味 (content) が理解できる範囲ならその意味を返し、そうでないなら偽を返す関数と再定義する。
 原文側の言い換え回数を r その上限を R とする。
 基本プロトコルのステップ 3) の一部を以下のように上書きする。もし、翻訳文側の修正回数 q がその上限の Q 以上となるなら、 r を 0 に初期化し、ステップ 7) に移動する。
- 7) [原文側: 原文の言い換え]
 原文側の言い換え回数 $r := r + 1$ とする。
 もし、 $r \geq R$ なら:
 プロトコルを終了し失敗を返す。
 もし、 $r < R$ なら:
 原文を修正した結果を s' とし、翻訳文側に送信する
 - 8) [翻訳文側: 翻訳文の一致判定]
 s' の機械翻訳結果を $t(s')$ とする。
 もし、 $c(t(s))$ が偽なら:
 原文側に原文の言い換えを依頼し、7) へ。
 もし、 $c(t(s)) \neq c(t(s'))$ なら:
 原文側に原文の言い換えを依頼し、7) へ。
 もし、 $c(t(s)) = c(t(s'))$ なら:
 3) で得られた $m(t(s))$ を翻訳結果とする。
 プロトコルを終了し成功を返す。

図 7 拡張プロトコルの形式的記述

Fig. 7 Formal statement of the extended protocol.

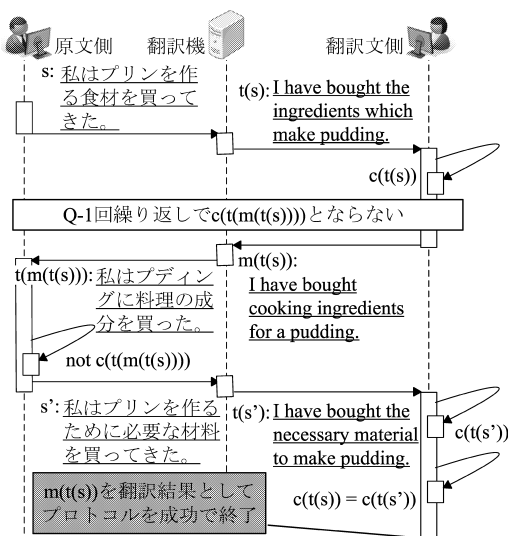


図 8 例 4 に拡張プロトコルを適用した場合

Fig. 8 Application of the extended protocol to the example 4.

表 2 基本プロトコルと拡張プロトコルの実験結果の比較
 Table 2 The experimental results of the basic protocol and the extended protocol.

	日英翻訳		日中翻訳	
	基本プロトコル	拡張プロトコル	基本プロトコル	拡張プロトコル
成功率	70%	80%	62%	71%
全データに対するAllの割合	58%	62%	57%	65%
プロトコル成功時のAllの割合	83%	77%	91%	92%

訳品質が低下する場合がある。しかし、一定の訳品質は保証され、かつ成功率を高めることができる。この点で、このような拡張方法は有効であるといえる。

6. む す び

近年、多くのコミュニティで機械翻訳を用いたコラボレーションが行われている。品質の低い機械翻訳を利用する上で、機械翻訳文を修正することは、情報共有のために有用である。しかし、その単純な実装では、修正翻訳文の正しさを保証できない欠点がある。

本論文の貢献はパイリンガルのいない状態で、(1) 機械翻訳の品質を確実に向上させ、(2) 正確で流暢な翻訳が得られた場合にプロトコルを終了する、共同翻訳プロトコルの存在を示したことである。翻訳文側では、機械翻訳文から意味を推測して流暢な翻訳文に修正できることを明らかにした。原文側では、その折返し翻訳文が、原文の意味を表しているかを評価し、翻訳の正しさを確認できることを明らかにした。

機械翻訳のみを適用した場合に、完全に正確な意味に翻訳された文の割合は日英翻訳で 53%、日中翻訳で 36%であったにもかかわらず、共同翻訳を適用してプロトコルが成功で終了すれば、その割合は日英翻訳で約 83%、日中翻訳では約 91%にまで改善されたことが、プロトコルの初期的評価より示された。すなわち、プロトコルが成功した場合の翻訳文の信頼度がかなり高くなることが実験的に示された。

原文側で修正翻訳文の折返し翻訳文の意味が理解できないために、プロトコルが失敗で終了する場合が存在する。本論文ではその場合に同一の意味をもつ、異なる表現で表された二つの原文に対する二つの翻訳文の意味を比較することで、翻訳文の正しさを翻訳文側で評価させるプロトコルの拡張の例を示し、プロトコルの成功確率を向上させる工夫が存在することを示した。それにより、今後様々な手法により共同翻訳プロトコルの発展が期待できることを示した。

謝辞 本研究は、日本学術振興会科学研究費基盤研究(A)(21240014, 平成 21 年度~23 年度)、京都大学グローバル COE プログラム「知識循環社会のための情報学教育研究拠点」並びに総務省戦略的情報通信研究開発推進制度から助成を受けた。

文 献

- [1] Y. Takano and A. Noda, "A temporary decline of thinking ability during foreign language processing," *J. Cross-Cultural Psychology*, vol.24, no.4, pp.445-462, 1993.
- [2] M. Aiken, C. Hwang, J. Paolillo, and L. Lu, "A group decision support system for the asian pacific rim," *J. International Information Management*, vol.3, no.2, pp.1-13, 1994.
- [3] K.J. Kim and C.J. Bonk, "Cross-cultural comparisons of online collaboration," *J. Computer Mediated Communication*, vol.8, no.1, Oct. 2002.
- [4] S. Nomura, T. Ishida, N. Yamashita, M. Yasuoka, and K. Funakoshi, "Open source software development with your mother language: Intercultural collaboration experiment 2002," *Proc. International Conference on Human-Computer Interaction (HCI-03)*, vol.4, pp.1163-1167, 2003.
- [5] N. Yamashita and T. Ishida, "Effects of machine translation on collaborative work," *International Conference on Computer Supported Cooperative Work (CSCW-06)*, pp.512-523, Nov. 2006.
- [6] H.H. Clark and D. Wilkes-Gibbs, "Referring as a collaborative process," *Cognition*, vol.22, pp.1-39, 1986.
- [7] H.H. Clark and S.E. Brennan, "Grounding in communication," in *Perspectives on Socially Shared Cognition*, ed. L.B. Resnick, J.M. Levine, and S.D. Teasley, pp.127-149, APA Books, Washington, 1991.
- [8] T. Ishida, "Language grid: An infrastructure for intercultural collaboration," *IEEE/IPSJ Symposium on Applications and the Internet (SAINT-06)*, pp.96-100, 2006.
- [9] 池原 悟, 白井 諭, 小倉健太郎, "言語表現体系の違いに着目した日英機械翻訳機能試験項目の構成," *人工知能誌*, vol.9, no.4, pp.569-579, 1994.

(平成 20 年 8 月 11 日受付, 12 月 12 日再受付)



森田 大翼

2008 京大・工・情報学卒, 現在, 同大大学院情報学研究所修士課程に在学中。主として異文化コラボレーションに関する研究に従事。



石田 亨 (正員)

1976 京大・工・情報工学卒, 1978 同大大学院修士課程了。同年日本電信電話公社電気通信研究所入所。ミュンヘン工科大学, パリ第六大学, メリーランド大学客員教授など経験。工博, IEEE フェロー, 情報処理学会フェロー, 現在, 京都大学大学院情報学研究所社会情報学専攻教授, 上海交通大学客員教授。自律エージェントとマルチエージェントシステム, セマンティック Web 技術に取り組む。デジタルシティ, 異文化コラボレーションプロジェクトを推進。