2D-5

複数翻訳機の利用による折り返し翻訳の精度不一致防止効果の検証

宮部　真衣 †　　吉野　孝 ‡,††
† 和歌山大学大学院システム工学研究科　　‡ 和歌山大学システム工学部
††（独）情報通信研究機構言語グリッドプロジェクト

1　はじめに

世界規模のインターネットの普及により，ネットワークを介した多言語間コミュニケーションの需要が高まっている．しかし，一般に多言語を十分に習得することは容易ではない．母語でのコミュニケーションを支援するために，機械翻訳を用いた支援が行われている [1].

機械翻訳技術は急速に進展してきているものの，高精度な翻訳を行うことは困難である．機械翻訳を介したコミュニケーションでは，翻訳精度が低い場合，十分な相互理解ができず，思い違いが発生する [2]．このような思い違いを回避するためには，自分の発言がどのように伝わっているのかを把握する必要がある．母語のみを用いた対象言語の翻訳精度の把握は，折り返し翻訳（対象言語翻訳結果の原言語への再翻訳）を利用することにより実現可能である．

我々はこれまでに，折り返し翻訳の精度確認手法としての妥当性の検証を行った [3]．検証の結果，折り返し翻訳文と対象言語翻訳文の精度には正の相関がみられた．一方で，対象言語翻訳文が不正確であるにもかかわらず，折り返し翻訳文が正確であるという状況（第 1 種の精度不一致）が発生する場合があった．

そこで，本稿では，第 1 種の精度不一致を防止するための仕組みを提案し，提案手法の効果について述べる．

2　折り返し翻訳とその課題

折り返し翻訳とは，対象言語へと翻訳した結果を，原言語へと再翻訳することである．折り返し翻訳の流れを図 1 に示す．対象言語に関する知識がない場合でも，折り返し翻訳文を確認することにより，対象言語翻訳文の翻訳精度を確認することができる [3]．また，翻訳自動評価において，対象言語の参照訳を用意せずに，翻訳精度を算出することができる [4, 5]．しかし，原言語への再翻訳によって得られる折り返し翻訳文は,「原言語から対象言語への翻訳」および「対象言語から原言語への翻訳」という，2 回の翻訳を介している．そのため,「対象言語から原言語への翻訳」を行うことにより，対象言語翻訳文の意味と折り返し翻訳文の意味が同一でなくなる可能性がある．

我々はこれまでに，折り返し翻訳の精度確認手法としての妥当性の検証を行った [3]．検証にあたり，以下の 2 種類の精度不一致を定義した．

［第 1 種の精度不一致］：折り返し翻訳文の精度が高いが，対象言語翻訳文の精度が低い

［第 2 種の精度不一致］：折り返し翻訳文の精度が低いが，対象言語翻訳文の精度が高い

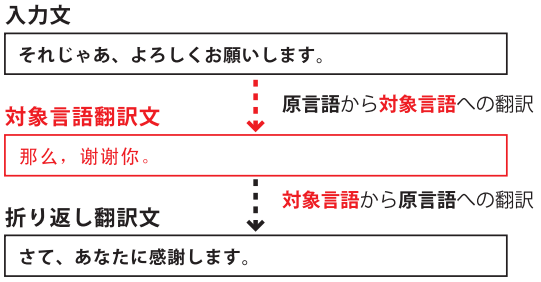

図 1: 折り返し翻訳の流れ

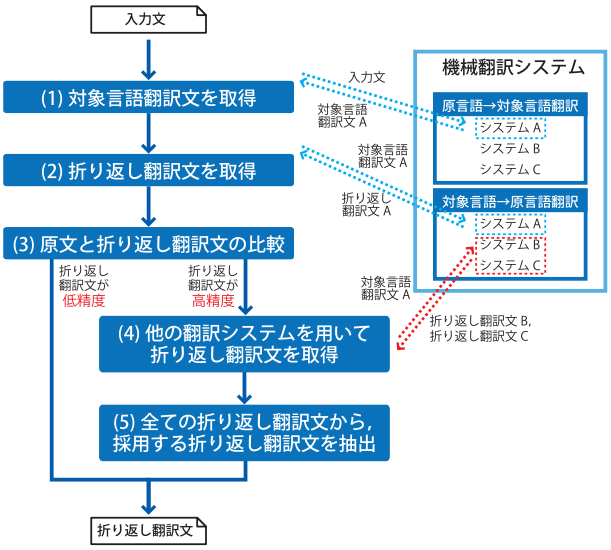

図 2: 提案手法の流れ

第 1 種の精度不一致が発生すると，入力者は伝わったと判断した内容が，相手の言語では正しく伝わらない．一方，第 2 種の精度不一致が発生すると，実際は修正しなくても伝わる文を，伝わらないと判断してしまう可能性がある．この場合，ユーザは本来不要な修正作業等を行う可能性があるが，第 1 種の精度不一致のような，意思疎通等の問題の発生にはつながらない．

検証の結果，第 1 種の精度不一致の発生率は低いものの，0%ではないことがわかった．第 1 種の精度不一致の発生は，意思疎通の阻害などを引き起こす可能性が高い．第 1 種の精度不一致の発生を防止し，正しい精度をユーザに提示できるようにする必要がある．

3　第 1 種の精度不一致の防止手法

本稿では，第 1 種の精度不一致を防止するための手法を検討する．折り返し翻訳文と対象言語翻訳文の精度検証実験 [3] においては，原言語から対象言語への翻訳および対象言語から原言語への翻訳を行う際に，同じ会社の翻訳システムを利用し，精度不一致の発生について検証を行った．しかし，第 1 種の精度不一致は，両方向（「原言語から対象言語」および「対象言語から原言語」）の翻訳が同じ手法が採用されている場合や，同じ言語資源から作られている場合などに発生しやす

Effect of Integrated Evaluation of Multiple Translation Systems for Preventing Accuracy Mismatch between Back-translated Sentence and Target-translated Sentence
Mai MIYABE†　Takashi YOSHINO‡,††
†Graduate School of Systems Engineering, Wakayama University
‡Faculty of Systems Engineering, Wakayama University
††Language Grid Project, National Institute of Information and Communications Technology

Copyright ©2011 Information Processing Society of Japan.
All Rights Reserved.

表 1: 実験に用いたテキストの一部

(1) 彼は駅からの距離を計った。
(2) 彼は私の顔をつぶした。
(3) とてもお財布にやさしいですね。
(4) チョコレート菓子ではなくてチョコレートですか？

い可能性がある．折り返し翻訳文の生成において，別の翻訳システムを用いることにより，精度不一致を防止できる可能性があるのではないかと考えた．

そこで，複数翻訳機を用いた折り返し翻訳の精度不一致防止手法を提案する．提案手法を用いた折り返し翻訳提示の流れを図 2 に示す．生成した折り返し翻訳文の翻訳精度が高精度である場合，複数翻訳機を用いて折り返し翻訳文の再生成を行う（図 2 手順 (4)）．それらの翻訳精度から総合的に精度を判定し，ユーザに提示する折り返し翻訳を選択することにより（図 2 手順 (5)），第 1 種の精度不一致の防止を目指す．

4 検証実験

提案手法による第 1 種精度不一致の防止効果を検証するために，実験を行った．本実験では，折り返し翻訳文が高精度であるが，対象言語翻訳文が低精度である文（精度不一致文）を評価テキストとして利用し，複数翻訳機を用いた場合の折り返し翻訳文の精度を評価する．また，複数の折り返し翻訳文に関する評価結果を考慮した場合に，第 1 種の精度不一致を防止できたかどうかを検証する．

4.1 精度評価方法

折り返し翻訳文，対象言語翻訳文の主観評価は，Walker らの適合性評価（5 段階評価）[6] により行う．評価指標は，5：同じ意味，4：文法などに多少問題があるが，大体同じ意味，3：意味は何となく掴める，2：雰囲気は残っているが，もとの意味はわからない，1：全く違う意味，となっている．上記の評価基準を用いて，2 つの文（原文および折り返し翻訳文）の意味の比較を行う．評価者は，日本人大学生 3 名である．

なお，本稿では，上記の評価基準において，3 未満の場合は「意味が理解できない」，3 以上の場合は「意味が理解できる」と分類することとし，対象言語翻訳文の評価結果が 3 未満である場合に第 1 種の精度不一致が発生しているとみなす．

4.2 評価テキスト

Walker らの適合性評価による折り返し翻訳文の評価結果が「**5（同じ意味）**」である場合は，原文と折り返し翻訳文の見た目にほとんど違いがない場合や，原文と折り返し翻訳文が同一である場合が多い．このような場合，ユーザは第 1 種の精度不一致の発生に気付くことができないため，多言語コミュニケーションにおける問題を引き起こす可能性が極めて高い．そのため，これらのテキストに関しては，特に第 1 種の精度不一致を防ぐ必要性が高い．

そこで，評価テキストとして，折り返し翻訳の精度検証実験 [3] において用いたテキストのうち，折り返し翻訳文の評価結果が「5」であり，第 1 種の精度不一致が発生していた 19 文の文章を用いた．評価テキストの一部を表 1 に示す．評価テキストには，2 種類の文（機械翻訳試験文[*1] およびチャットにおける発言[*2]）が含まれている．次節で述べる 3 種類のシステムを用いて，これらのテキストの対象言語翻訳文の折り返し翻訳文を生成し，その精度の評価を行う．

表 2: 第 1 種の精度不一致の防止数

総文数	防止数（中央値）	防止数（最小値）
19 文	5 文（26.3%）	11 文（57.9%）

4.3 翻訳システム

本実験では，3 種類の翻訳システム[*3 *4 *5] を用いて，評価テキストの折り返し翻訳文の生成（図 2 における手順（4））を行った．なお，各翻訳システムは言語グリッド [7] を介して利用した．

5 実験結果

今回の実験においては，3 つの翻訳システムによる翻訳結果の精度から，第 1 種の精度不一致を防止できるかどうかを判断する．折り返し翻訳文 3 文それぞれに対する評価結果から，中間値および最小値を代表値とした場合の防止率について検証する．提案手法による第 1 種の精度不一致の防止率を表 2 に示す．

表 2 より，複数翻訳機の翻訳結果に関して，中間値を代表値とした場合の防止率は 19 文中 5 文（26.3%）であった．また，最小値を代表値とした場合の防止率は 19 文中 11 文（57.9%）であった．最小値を代表値にすることにより，半数程度は第 1 種の精度不一致を防止できることがわかった．

6 おわりに

折り返し翻訳における第 1 種の精度不一致（対象言語翻訳文が不正確であるが，折り返し翻訳文が正確であるという状況）を防止するために，本稿では，複数翻訳機を用いた折り返し翻訳の精度不一致防止手法を提案した．提案手法の効果を検証するために，折り返し翻訳の精度検証実験を行った結果，複数翻訳機による翻訳結果の精度に関して，最小値を代表値とした場合，第 1 種の精度不一致の防止率は 57.9%であった．

今後は，精度不一致が発生していない文に対して本提案手法を適用した場合の誤検出率について検証する．

謝辞

本研究の一部は，日本学術振興会科学研究費 基盤研究 (B)(22300044) の補助を受けた．

参考文献

[1] Inaba, R.: Usability of Multilingual Communication Tools, Proceedings, LNCS 4560, pp.91-97 (2007).
[2] Yamashita, N. et al.: Automatic prediction of misconceptions in multilingual computer-mediated communication, Proc. the 11th international conference on Intelligent user interfaces, pp.62-69 (2006).
[3] 宮部真衣 ほか：機械翻訳を介したコミュニケーションのための折り返し翻訳の妥当性の検証，電子情報通信学会技術報告，人工知能と知識処理，AI2009-41，pp.65-70（2010）.
[4] Uchimoto, K., et al.: Automatic Rating of Machine Translatability, 10th Machine Translation Summit (MT Summit X), pp.235-242 (2005).
[5] Rapp, R.: The Back-translation Score: Automatic MT Evaluation at the Sentence Level without Reference Translations, In Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, pp.133-136 (2009).
[6] Walker, K., et al.: Multiple-Translation Arabic (MTA) Part 1, Linguistic Data Consortium, Philadelphia (2003).
[7] Ishida, T.: Language Grid: An Infrastructure for Intercultural Collaboration, SAINT-06, pp.96-100(2006).

[*1] NTT Natural Language Research Group, http://www.kecl.ntt.co.jp/icl/mtg/resources/index.php

[*2] チャットにおける発言とは,「好きなもの・嫌いなもの」というテーマでのチャットにおける対話文である．

[*3] J-Server（高電社，http://www.kodensha.jp/）

[*4] Google 翻訳（Google，http://translate.google.co.jp/）

[*5] WEB-Transer（クロスランゲージ，http://www.crosslanguage.co.jp/）

Copyright ©2011 Information Processing Society of Japan.
All Rights Reserved.