

4G-8

SL-TRANS における変換・生成手法

飯田 仁 長谷川 敏郎 上田 良寛 相沢 輝昭

ATR自動翻訳電話研究所

1. はじめに

自動翻訳電話実現に向けた基盤技術として、話し言葉翻訳技術の検討を進めた。その技術に要求されることは、書き言葉の翻訳技術に加え、断片的な文や文脈依存度の高い慣用的な句や文を翻訳する技術である。それらの技術について、対話文翻訳実験システムNADINE(1),(2)を使った音声言語日英翻訳実験システムSL-TRANS(3)を通して検討しているが、本稿ではその変換・生成の手法とその実験について報告する。

2. 対話翻訳に要求される技術(4)

文章を文単位に翻訳することは、文章中の文脈が目標言語側でも保存されていることが仮定されている。断片的な発話では、なおさらその内容が文脈に強く依存しているため、その文脈に適した目標言語側の適切な言語表現を生成しなければならない。また、対話のやり取りの滑らかさを保つ待遇表現などを目標言語においても表現しなければならない。翻訳の深さのレベルについては、解析内容の程度に合わせて、目的言語による表現がその単文自体で意味が確定する程度の記述とする。例えば、「登録用紙をお願いします」に対して、“registration form, too, please”では、文脈を無視するとその意味が確定しないので、“please send a registration form, too”、“please return the registration form, too”、あるいは、“please fill out the registration form, too”などの可能性を考慮することが望まれる。

問題となる主な処理対象として、(1)断片的な文(例:失礼します、わかりました)、(2)省略部を含む文(特に、ダ文、主題の省略)、(3)談話実体の表現(特に、冠詞や数詞)、(4)言語運用上の時制・アスペクト、(5)待遇的な表現、および発話意図の表現、などが挙げられる。

発話意図は、発話を介して聞き手の心的な状態に影響を及ぼすことを目的とした表現であり、対象領域の性質に依存した表現が多く現れる。例えば、「登録用紙をこちらにお送りいただけますか」では、疑問提示型の発話意図

(REQUEST S H
(INFORMIF H S
(POSSIBLE H
(RECEIVE-FAVOR S
Proposition
(SEND H S form))))

Transfer and Generation Method in SL-TRANS
Hitoshi HIDA, Toshiro HASEGAWA, Yoshihiro
UEDA and Teruaki AIZAWA

ATR Interpreting Telephony Research Laboratories

が得られるが、これを目標言語において適切に表現することが必要である。

3. SL-TRANS における変換手法

命題内容を目標言語に変換して、それに発話意図の記述を加味した表層表現を作ることで、滑らかさを保った発話の生成を実現することを、変換・生成過程の基本的な方針とする。従って、図1に示すように、変換過程では、命題内容をその対象として目標言語への変換を行う。発話意図の記述は中間言語として表現し、記述の書き換えを行わない。命題内容は、述語とその格要素を使って記述し、格構造の変換を基本にする。その際、下記の構造変換のための規則を語彙駆動で適用する。

(a) 補助的述語の表現の違い

解析の結果(5)得られる命題内容では、その意味記述が中間言語表現として扱えるほどに形式化が整っていない状況において、補助的述語の表現(例:(と)思う、(で)結構です)などに対して、画一的な格関係記述を与えることになる。そこで、「参加申し込みに関する問合せ」などの事務局と客の対話コーパスにおいては、補助的述語が多く現れ、それに関わる記述の目標言語への構造変換が重要な役割を果たす。この変換では、問題となる補助的述語などの語彙に関する構造変換規則を用意して、これを適用する。

(b) 利益追求主導の表現の違い

上記対話コーパスは、事務局が問合せ側に利益を与え、かつ問合せ側が何らかの利益を得るための情報交換であり、述語の使い方に特徴がある。日本語の受給表現は、英語においては表層上で対応する概念をもたないので、一般的には変換過程での書き換え規則を適用する。それに対し、SL-TRANSではその受給表現は発話意図として記述し、生成過程でその表現を処理する。

一方、結果的に利益が得られるような命題(例:

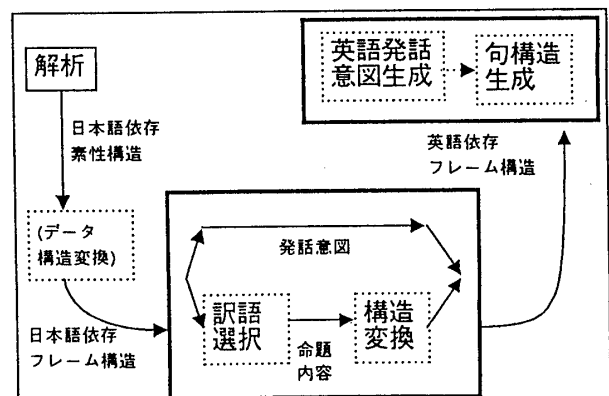


図1 変換・生成過程の概要

割引がある、割引をおこなう)では、英語において受益行為に関する動詞(例: give, offer, etc.)を使う。日本語の対象格を中心にした記述を、それらの動詞を中心にした記述に書き換える。

(c)言語運用上の時制とアスペクトの日英の違い

日本語の滑らかな対話を作り出すために、論理的に冗長となるアスペクト句が付加される場合が多い(例えば、無料です→無料となっています)。滑らかな英文生成の点で、日英のアスペクトの違いの扱いは重要である。そこで、書き換え規則を使って、動詞固有の性質を考慮しながら、日本語のアスペクトのタイプ(progressive, resultive など)を英語に適したタイプに変換する。

その他、下記に示す一般的な変換規則がある。

▶名詞句から単一名詞への変換

(例: 詳しい内容; details, 分からない点; question)

▶本動詞スルの動詞句からサ変動詞相当句への変換

(例: 手続をする, 発表をする)

▶冗長な句の削除

(例: 用紙のようなもの; 用紙, 口座に振り込むということ; 口座に振り込む)

▶慣用表現、固定的な応答文、熟語などの変換

以上、変換過程の構造変換規則を中心に述べたが、格要素などを手掛かりとする訳語選択の規則などを備えている。

4. SL-TRANSにおける生成手法

生成過程では、変換の結果得られる格構造を基本にした命題内容と、中間表現的に扱う発話意図(6)とを融合して、対話に適した発話文を作る。その過程の概略を次に示す。

- ①発話意図の記述に対する英語の目標構造を選択する。このとき、命題内容部分の変数とする。
- ②命題内容部分の欠落している必須格要素を決定する。但し、現状は文脈を考慮した決定はできず、デフォルト値を使い便宜的に処理している。
- ③その目標構造の命題内容部分に、変換結果の値を結合し、主語の統一などを行う。
- ④表層の語列を作り、その形態を満足する。

以下、扱っている発話意図表現の例を示す。

▶依頼/要求

例: お聞かせ願えますか; could you give / tell me, させてもらいたい; would like to, (て)ください; please

▶行為/態度拘束

例: なくてはならない(丁寧表現); must (need to) させたい; want to have

▶使役

例: させる; have

その他、句構造生成規則を用意して、形容詞句などを作る。また、解析の結果得られる発話意図の記述は、英語の表層表現に応じた変形を行う。

5. 対話文変換/生成の実験

解析モジュールでは、言語的かつ音声的な評価尺度を使って解析結果内から最適候補を一つ選択する。変換モジュールはこれを入力とする。

発話意図の記述については、言語不変性を仮定しているが、実験を通してその妥当性を判断する。模範的な対話例を翻訳した実験結果を図2に示す(音声認識の誤りのあるものも含む)。

生成過程では多くの課題を残している。冠詞の扱い、ダ文、主題の省略などについて文脈的な情報の扱いは導入していない。冠詞処理の現状は、不定冠詞を使うことを原則とし、話題領域内で固定的な対象について定冠詞を使っている。ダ文は翻訳例にあるように固定的な表現内のものだけを扱っている。解析モジュールでは、聞き手や話し手の省略補完がある程度解決されるが、主題の省略などは直前に現れた名詞句などを参照する便宜的な方法で対処している。

6. 問題と今後の課題

発話意図の変換手法、素性構造で統一した翻訳手法、文脈情報を使った決定的翻訳手法、自動翻訳電話技術に成り得る翻訳処理速度などを明らかにすべく各手法の見直しと実験を進める。

謝辞: 日頃ご指導戴くATR自動翻訳電話研究所 樽松明社長には有益なご意見を戴き深謝します。また、SL-TRANSの各部の担当者、およびNADINEの作成に携わってきた言語処理研究室の研究員諸氏に感謝します。

参考文献

- [1] 飯田、小暮、野垣内、有田、相沢: 「端末間対話通訳の実験システムの構成」、情処学会第36回大会、1988.
- [2] Iida H., et al., "An Experimental Spoken Natural Dialogue Translation System Using a Lexicon-Driven Grammar", EuroSpeech89, 1989.
- [3] 森元、小倉、相沢、樽松: 「音声言語日英翻訳実験システム(SL-TRANS)の概要」、情処学会第39回大会、1989.
- [4] 飯田: 「異言語間対話を目指す端末間通訳モデル」、認知科学の発展第1巻、認知科学会編、1988.
- [5] Kogure K., et al., "A method of analyzing Japanese speech act types", in Proc. of 2nd MT Conference, 1988.
- [6] Kume M., et al., "A Descriptive Framework for Translating Speaker's Meaning", the 4th EACL, 1989.

[翻訳例]

もしもし (入力文) ▶ Hello (翻訳結果) — 以下同様
そちらは会議事務局ですか
▶ Is that the office for the conference?
はい ▶ Yes
そうです ▶ I see.
会議に申し込みたいのですが
▶ I would like to make a registration for the conference?
登録用紙は既にお持ちでしょうか
▶ Do you already have a registration form?
いいえ ▶ No.
まだです ▶ Not yet.
分かりました ▶ I see.
それでは登録用紙をお送り致します
▶ Then, I send you a registration form.
ご住所とお名前をお願いします
▶ Could you give me an address and a name?
;; your name などと決定する一般的な手法を扱っていない
住所は大阪市北区茶屋町二十三です
▶ An address is 23 Chaya-machi, Kita-ku, Osaka.
名前は鈴木真弓です ▶ A name is Mayumi Suzuki.
在りました ▶ is there?
;; 音声認識の誤り(正しくは、「わかりました」)
登録用紙を至急送らせていただきます
▶ I send you a registration form immediately.
分からない点がございましたらいつでもお聞き下さい
▶ If there is a question, please ask me at any time.

図2 SL-TRANSの翻訳例