

対話式日英機械翻訳における 意味的なあいまい性の提示方法

1 F - 3

芦沢 実, 梶 博行
(株)日立製作所 関西システムラボラトリ

1. はじめに

現在の機械翻訳システムには、解析・生成におけるあいまい性を解決するための十分な知識がない。これに対する解答の一つとして、我々は、技術文書を対象に、利用者に対する問い合わせであいまい性を解決する対話式機械翻訳システムを検討している。この対話式システムは、文書執筆者自身が操作して問い合わせに答えることを想定している。このため、あいまい性をわかり易く提示することが重要である。形態素解析、構文解析におけるあいまい性については、前編集支援システム [1] [2] で開発した方法が適用できる。本稿では、日本語意味解析のあいまい性の提示方法について報告する。

2. 意味解析におけるあいまい性

(1) 名詞・動詞など内容語の語義

異なる概念が一つの単語で表わされる場合がある。例えば「電線を埋める」の場合と「すきまを埋める」の場合では「埋める」の語義が異なる。

(2) 語と語の関係

助詞は複数の意味を表わすものが多い。埋め込み文、複合語は、語と語の関係を表わす語が隠れているためにあいまいである。また、文のあいまい接続も、語と語の関係として扱う。

(3) 格パターン

格パターンは格助詞「が」「を」などの表わす格を決めるが、複数の格パターンを持つ動詞語義では格助詞が表わす格があいまいになる。例えば「(すきまを)埋める」の語義には次の3つの格パターンがある。

- (動作主)が, (場所)を, (材料)で
- (動作主)が, (場所)に, (材料)を
- (材料)が, (場所)を

このため、「排出物が埋める」という文では、「排出物」が(動作主)か(材料)かあいまいである。

(4) 格の省略

日本語では主語や目的語が省略されることが多い。

これらを省略したままでは適切な文体に翻訳できないことも多い。

3. 提示内容

主な利用者である文書執筆者には、文法の知識や目標言語の知識をあまり要求できない。従って、システムは、解釈の各々についてあいまい性がないように日本語で言い替えたものを選択肢として提示することとする。

(1) 内容語の語義の選択

和英辞典で語義を分類する小見出し程度の言い替えを提示する(図1)。

語義の決定だけでは類義語の範囲で訳語選択のあいまい性が残る。しかし、日英間のようにギャップが大きい言語間で訳語の違いを原言語だけで説明しようとしても[4]、同義語の微妙な違いを簡潔に表現することは難しい。また、類義語の範囲で訳語選択を誤っても致命的な誤訳にはならない。

(2) 語と語の関係

複数の意味を表わす助詞は、その表わす格が明確になるように言い替える(図2)。埋め込み文、複合語は、係り受け関係にある語の意味的關係を示す助詞を補って語順や品詞を変更して言い替える(図3, 図4)。文のあいまい接続は、並列接続の解釈に対しては文を分割して接続詞を補い、その他の解釈には関係を明確にする接続助詞を補う(図5)。

(3) 格パターン

格があいまいにならない助詞を使って言い替える。例えば、「(すきまを)埋める」の格パターンには「(材料)で」「(材料)を」はあるが「(動作主)で」「(動作主)を」がないので、「排出物が埋める」における「排出物」の格を(材料)とする解釈に対しては「で」または「を」によりあいまいなく言い替えられる。しかし、(動作主)とする解釈に対してはあいまいでない言い替えができない。このため、<言い替え不可>を選択肢とする(図6)。

(4) 格の省略

省略された語の入力を促す文を言い替えを使って作成する(図7)。利用者はその語の文字を入力し、システムは入力された語を補って翻訳する。

4. 書き直し

英語に比べて日本語では、あいまいな表現が多い。特に文のあいまい接続のように、意味を明確にするために文を書き直した方がよい場合が多い。このため、<書き直し>という選択肢を設ける。本稿の図では、この選択肢は省略した。

5. おわりに

対話式日英翻訳システムにおける意味的なあいまい性の提示方法を述べた。

助詞の用法、文のあいまい接続のあいまい性は、原文を構造を変えずに作り直すための言い替えで解決できた[3]。本稿で述べたその他のあいまい性は、そのような言い替えでは解決できなかった。

テンス、アスペクト、モーダルのあいまい性は技術文書では少ないと判断して、また、訳語選択、前置詞の選択、冠詞・数の選択は英文生成のあいまい性であるとして、対話の項目から除外した。これらに対しては、後編集が必要である。

プロトタイプによる、効果、限界の評価が必要であるが、提案方法によって後編集コストが削減できると考える。

[参考文献]

- [1] 平井, 他: 日英機械翻訳用前編集支援システム(1), 情報処理学会第36回全国大会, 2U-2(1988)
 [2] 芦沢, 他: 日英機械翻訳用前編集支援システム(2), 情報処理学会第36回全国大会, 2U-3(1988)
 [3] 長尾, 田中: 制限文法に基づく文書作成援助システム, 情報処理学会自然言語処理研究会資料44-5(1984.7)
 [4] Wood et al.: Machine Translation for Monolinguals, COLING, pp. 760-763(1988.8)

装置が漏出口を凝固剤で 埋め た
 (土の中に) 埋める
 (すきまを) 埋める
 (損失を) 埋める
 ぬるくする

図1 「埋める」の語義の提示

気流で加速する

気流において加速する
 気流を使って加速する
 気流のために加速する

図2 「で」の解釈の提示

制御するシステム

システムが制御する
 システムを制御する

図3 埋め込み文の解釈の提示

構造計算

構造を計算する
 構造が計算する
 構造と計算
 構造的な計算

図4 複合語の解釈の提示

バルブを閉めないで自然に乾燥させた。

…閉めない。および、…乾燥させた。

…閉めないので…乾燥させた。

…閉めないことによって…乾燥させた。

…閉めないならば…乾燥させた。

…閉めないけれども…乾燥させた。

図5 文のあいまい接続の解釈の提示

排出物が埋める

排出物で埋める
 <言い替え不可>

図6 選択肢に<言い替え不可>を含む提示

バルブを閉めないで自然に乾燥させた。

「『何を』乾燥させた」のですか。

図7 省略句の問い合わせ