

2C-2

問答式日英機械翻訳における
例文による深層格のあいまい性の提示方法

芦沢 実 , 梶 博行

(株)日立製作所 システム開発研究所)

1. 始めに

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、日本語原稿作成者自身を利用者として、利用者への問い合わせによってあいまい性を解決する問答式日英機械翻訳システムを検討している。

前回の報告[1]では、深層格、語義のあいまい性に対して、複数の解析結果の言い替えを選択肢として提示する問い合わせ方法を述べた。そこでは深層格のあいまい性は機能語の置き換えによって言い替えるとしたが、格パターンのあいまい性がある場合の言い替えが理解しにくいという問題があった。本稿では、深層格のあいまい性を例文によって提示する問い合わせ方法を報告する。

2. 深層格のあいまい性

深層格は、述語に対する格助詞などの役割を表わす。格助詞に対する深層格は、格パターンによって決まる場合と格パターンによらずに決まる場合がある。

格パターンは、述語ごとに格助詞に対する深層格を定義する。動詞の語義によっては複数の格パターンを持つので格パターンのあいまい性がある。例えば「埋める」の〈一杯にする〉という語義には、次の3つの格パターンがある。

- a. (動作主)が(対象)に(要素)を埋める
- b. (動作主)が(対象)を(要素)で埋める
- c. (要素)が(対象)を埋める

格パターンによらない場合、格助詞ごとに深層格との対応が定義される。格助詞によっては、次のように、深層格との対応を2つ以上もつ。

- d. (場所)で
- e. (道具)で

格助詞の解析における深層格のあいまい性は、これらの対応関係の組み合わせである。例えば、「編集者がカットで余白を埋める」という文では、「編集者が」の深層格は(動作主)か(要素)の2通りの解析が可能である。また、「カットで」は(要素)か(場所)か(道具)、「余白を」は(要素)か(対象)の解析が可能である。これらの可能な組み合わせを図1に示す。

3. 例文による提示

我々の問答式機械翻訳システムは、一般的な利用者に文法の知識や目標言語の知識をあまり要求しないことを前提としている。特に、深層格や格パターンのあいまい性を意識している利用者は少ないと考えられる。利用者に判断を仰ぐためには、本稿の場合、深層格のあいまい性があることを、利用者に文法の知識なしで理解できる提示を必要とする。

このためには、例文の提示が有効である。抽象的な例文ではなく、述語に対する格助詞の役割が明確で、具体的なイメージを利用者が想起しやすい例文が良い。例えば、前章で示した「埋める」の〈一杯にする〉の格パターンa~cは、次の例文で表現できる。

- a'. 作業員が 溝に アスファルトを 埋める
- b'. 作業員が 溝を アスファルトで 埋める
- c'. アスファルトが 溝を 埋める

また、格パターンによらずに対応が決まる場合も動詞の語義に応じて例文で表現できる。

- d'. 工事現場で 埋める
- e'. ショベルで 埋める

4. 文節逐次注目方法

システムの解析結果に対する判断を仰ぐためには、例文を提示した上で、例文の単語と入力文の単語の対応を示すことを必要とする。

そのための方法として、本稿では、文節逐次注目方法を提案する。これは、一つの問い合わせでは入力文の格助詞の一つとそれを含む文節に注目することとして、複数の問い合わせによってあいまい性を解決する。その問い合わせ例を図2に示す。図2の問い合わせ1は、入力文の「編集者が」に注目する問い合わせである。問い合わせ2は、問い合わせ1で「1」が正解とされたとした場合に「カットで」に注目する問い合わせである。

問い合わせにおいて、入力文と選択肢の例文の注目する文節を明確に示すために、注目文節を強調する。下線や分かち書きなどを使ったり、注目しない文節を括弧で囲んで選択肢の末尾に置いたりする。

注目する格助詞の解析結果が2つ以上の格パターンに共通する選択肢では、そのうちの格パターンの一つを仮定してその例文を提示する。例えば、「編集者が」を(動作主)とする格パターンは、a, bであるが、

問い合わせ1の第1選択肢とする例文は、a' だけである。注目する格助詞の解析結果の1つに対して2つ以上の例文を提示すると、利用者の注意が注目文節からそれらの例文の差違へそれしてしまうので、問い合わせの意図が伝わりにくい。一つの解析結果に対して例文を一つにすれば、そのようなことは起きない。

また、注目する格助詞の解析結果が格パターンによらない深層格である選択肢では、その解析結果が格パターンによらないことを明確にするために、その解析結果に対する例文を格パターンの例文と合成する。例えば、図2の問い合わせ2の第2選択肢、第3選択肢は、例文d'、e'をa'と合成したものである。

このようにすることで、別の格パターンが正解となる入力文に対してもあいまい性がない提示ができる。その例を図3に示す。

文節逐次注目方法では、一つの問い合わせごとにあいまい性を解決すれば、途中から自動的にあいまい性を解決することができる場合がある。図2の問い合わせ2において'1'が選択されれば格パターンbが正解であることが決まるので「余白を」は直ちに(対象)であることが決まる。

入力文の格助詞に注目する順序によっては、更に問い合わせの数を減らせる場合もあるが、利用者の選択次第で問い合わせの数が減らない可能性もある。減らない場合は注目する語が入力文の前後へ移動して問い合わせることになる。それよりは、注目する語は入力文の文頭から文末へ一定して移動した方が、利用者にとって分かりやすいと考える。

例文を利用して深層格のあいまい性を問い合わせる方法には、文節逐次注目方法の他に、例文の語順を入れ替えて入力文の語順と一致させる方法を考えることができる。つまり、例えば図1に示した組み合わせにおいて、カッコ内の深層格ラベルを前章の例文で使用した名詞に置き換えたものを提示する方法である。この方法は、1回の問い合わせでこの入力文の「埋める」に関する深層格のあいまい性を解決するが、利用者にとって入力文と選択肢で対応付けるべき点が多いので煩わしい。文節逐次注目方法には、この煩わしさが無い。

5. 語義のあいまい性との関係

提示する例文は、具体的なイメージを想起するものであるから、動詞の語義ごとに用意する。例えば、「埋める」に関しては、〈一杯にする〉、〈物の中にうずめる〉、〈補う〉の語義で、それぞれに適した例文を用意する。また、格パターンのあいまい性が、語義によって異なる。従って、例文によって深層格の問い合わせを提示する場合は、その問い合わせに先立って動詞の語義のあいまい性を問い合わせによって解決する。

6. まとめ

本稿は、深層格のあいまい性が生じる理由が利用者に入法の知識なしで理解でき、入力文と選択肢の対応を付けやすい、文節逐次注目方法について述べた。こ

れに対して機能語の置き換えによる言い替え方法[1]では、格パターンのあいまい性がない場合は充分であるが、格パターンのあいまい性がある場合に分かりやすく提示することが難しかった。

例文による方法と機能語を置き換える言い替え方法は、組み合わせて利用することができる。格パターンによらずに格助詞と深層格の対応が決まる場合は、機能語も置き換えた言い替えの方が明確である。例えば、「(場所)で」の例文は、「工事現場で」より「工事現場において」とした方が明確である。また、「(道具)で」の例文は、「ショベルで」よりも「ショベルを道具にして」の方が明確である。

□ 参考文献

[1] 芦沢実, 梶博行: 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, 情報処理学会第40回全国大会, 1F-3(1990)

入力文1: 編集者が カット で 余白を 埋める

1. (動作主)が (要素)で (対象)を 埋める
2. (動作主)が (場所)で (対象)を 埋める
3. (動作主)が (道具)で (対象)を 埋める
4. (動作主)が (場所)で (要素)を 埋める
5. (動作主)が (道具)で (要素)を 埋める
6. (要素)が (場所)で (対象)を 埋める

図1 深層格のあいまい性の組み合わせ

■問い合わせ1

入力文1: 編集者が カット で余白を 埋める

「編集者が」と「埋める」の関係と同じものは?

1. 作業員が 埋める (溝にアスファルトを)
2. アスファルトが 埋める (溝を)

↓ ('1' が選択された場合)

■問い合わせ2

入力文1: 編集者が カット で 余白を 埋める

「カットで」と「埋める」の関係と同じものは?

1. アスファルトで 埋める (作業員が溝を)
2. 工事現場で 埋める (作業員が溝にアスファルトを)
3. ショベルで 埋める (作業員が溝にアスファルトを)

図2 文節逐次注目方法による問い合わせ

入力文2: 記者が ホテルでロビーを埋める

「記者が」と「埋める」の関係と同じものは?

1. 作業員が 埋める (溝にアスファルトを)
2. アスファルトが 埋める (溝を)

図3 別の入力文に対する問い合わせ