

FEP 型英文作成支援ツールにおける対話的曖昧性解消方式

2B-6 山端 潔 土井 伸一 亀井 真一郎 佐藤 研治 田村 真子 村木 一至

NEC 情報メディア研究所

e-mail: {yamabana, doi, kamei, sato, shinko, k-muraki}@hum.cl.nec.co.jp

1 はじめに

前回、FEP 型英文作成支援ツールの試作について報告した [1, 2, 3]。このツールは、単語および熟語の日英辞書引き機能の上に、対話的な日英構文変換機能を有し、ユーザの英文作成を支援する。構文変換部は一種の対話型機械翻訳システムであり、ユーザは訳語選択と単文範囲の変換をインタラクティブに繰り返して英語表現を得る。

対話型翻訳システムとしてのこのシステムの最大の特徴は、対話が常にユーザによって主導される点にある [4]。このことは、過剰な対話を抑制し操作性を高める利点を持つが、対話をシステムが主導するアプローチ [5, 6] とは異なり、ユーザ操作が直接的な曖昧性解消指示ではないため、システムがユーザ操作の意味を適切に解釈する必要がある。さらに、対話は常にユーザから開始するため、システムによる選択肢のデフォルト推定の精度が重要になる。

本稿では、試作したシステムにおける、曖昧性解消のためのユーザ操作の解釈と選択肢のデフォルトの決め方について述べる。さらに、対話がユーザ主導であるという特徴を中心に、ユーザとシステムの協調、自動翻訳との整合性について論じる。

2 日英変換操作の概要

対話の日英変換の手順を示す。原文 (a) は入力と同時に形態素解析され、自立語が英語に変換される (b)。システムは変換範囲と訳語を提示するので、ユーザはこれらを確認したのち構文変換を指示する。構文変換部が日英の変換を行い、結果を返す (c)。複文、重文は単文区間の変換を多重に繰り返して結果を得る (d)。

- (a) 私は彼が本を読むのを助ける
- (b) I は he が book を read の を help
- (c) I は he reads a book を help
- (d) I help him to read a book

ユーザにとって可能な操作は、訳語選択と変換範囲の指定の二種類である。ある時点で変更可能な選択肢 (訳語および変換範囲) は、システムによる推測とともに常に画面に表示されており、ユーザはこれを任意に変更す

Interactive Disambiguation Method of FEP-type English Writing Support Tool

Kiyoshi YAMABANA, Shinichi DOI, Shin-ichiro KAMEI,
Kenji SATOH, Shinko TAMURA, Kazunori MURAKI
NEC Corporation

ることも、システムの推定を受け入れて次に進むことも自由である。この意味で、対話は完全にユーザが主導する。何も選択しなければ自動翻訳と同様の処理となるが、その場合のステップ数はほぼ入力文中の動詞の数と一致する。

3 曖昧性解消のための戦略

システムは、訳語と変換範囲の選択状態をもとに解析・生成における曖昧性を解消していく。解消できない曖昧性が残った場合、原則として優先度付きの複数解を生成しユーザによる選択に任せる。

3.1 係受けの曖昧性解消

変換範囲の指定はまとまりの指定であり、直接的に係受けの指定となる [2, 4]。

呼応の副詞など、統語的な非交差条件を破りうる語彙対は共起熟語として登録し、形態素解析の段階で共起関係を認定する。必要な素性の移動は共起認定と同時に行われるが、訳語選択操作により解除できる。すなわち、係受けの有無の指示は訳語選択操作に還元できる。

3.2 訳語選択

訳語選択の処理は、自立語と付属語とで趣が異なる。自立語の場合、原則として必ずユーザまたはシステムが訳語を推定し、構文変換部は必ずそれを採用する。何も選択せずに、日本語を訳語として送ることもできるが、その場合、システムは前回の選択結果を優先して採用する。構文変換後も訳語は自由に変更できるが、格フレーム等の構文情報は反映されない。

付属語の場合、システムは訳語の優先順位を推定するが、構文変換部にはこの優先順位のついた訳語候補の全体を送る。ただし、ユーザが明示的に訳語の一つを選択した場合はこの限りでない。付属語の用法決定にシステムの文法知識を活用するためである。

これらの処理は、語の辞書中に、特殊な英語エントリーであるダミーブロックが存在するかどうかで区別される。ダミーブロックは原語自身を訳語とし、訳語選択の際に無条件で優先される。ダミーブロックが選択されると、構文変換部は他の訳語をすべてロードし、文法規則を適用して選択を試みる。ダミーブロックは単なる一訳語であり、ユーザは他の訳語を自由に選択することができるため、ユーザの選択の自由を損なうことなく、本来システムに訳語選択を任せるべき語を明示することができる。

3.3 付属語の訳語選択と用法の決定

付属語は、原則としてシステムが訳語/用法の選択を行う。しかし、ユーザが明示的に訳語選択した場合はこの限りでない。例えば、「で」には助詞と助動詞の曖昧性があるが、ユーザは訳語“be”を選択し助動詞解釈を指定することができる。何も選択しなければ、構文変換部には日本語のまま送られて、専用の処理フェーズが用法の曖昧性解消を行う。

機能語の用法決定フェーズでは、既存の機械翻訳システムと同様の処理を行う。ただし、変換範囲の情報を有用に用いることができる。例えば、変換範囲は基本的に単文なので、変換範囲末尾または英語接続詞の直前の場合には助動詞、それ以外では助詞を優先するのが基本となる。

日本語または英語で機能語として振舞う語の訳語選択が生成処理においても有用であることは別に示した通りである [2, 4]。

4 システムによるデフォルト推定

対話型翻訳においてユーザの操作数を削減するには、システムによるデフォルト推定の精度を高めることが重要である。デフォルト推定が正しければ、ユーザは対話型であることを意識するまでもなく正しい結果を得ることができる。本システムでは、(1) 訳語 (2) 訳語の品詞 (3) 変換範囲 (単文範囲) の3つの種類の推定を構文変換前に行う。

4.1 訳語の推定

訳語推定には、(1) 品詞 (2) 学習 (3) 共起 (熟語) の3つの要因が関わる。システムは、語の前後の並びを見て訳語の品詞を決定する (下記参照)。訳語はこの品詞を持つものの中から、以前学習された訳語を優先して選ばれる。ただし、共起辞書 (熟語辞書) に登録された語のペアが現れたらそちらを優先する。

4.2 訳語品詞の決定

訳語選択により訳文の構造をコントロールするため訳語は品詞込みで選択する。そのため、システムは訳語推定と同じタイミングで英語品詞を推定する。推定は、日本語側の品詞と周囲の語の並びから行う。以下に推定規則の一部を示す。

日本語品詞	条件	英語品詞優先順位
サ変	直後がサ変語尾または助動詞、ただし助動詞「だ」を除く	Verb, Adj, Noun
サ変	上記以外	Noun, Verb, Adj
名詞	Prepを持つ訳語があり、以下の条件を満たす ・直前が連体形か体言連続 ・直後が読点、助詞、体言連続	Prep, Noun
名詞	その他	Noun

表1: 日本語並びからの英語品詞推定規則 (一部)

4.3 変換範囲の決定

デフォルトの変換範囲として、動詞を一つ含む単文区間を推定する。原則として、形容詞・形容動詞以外の動詞を一つだけ含み、接続詞を含まない最大の区間を変換範囲とする。品詞は訳語のものを使用する。結果として、変換範囲は、訳語選択結果に応じて動的に変化する。

5 ユーザ主導の対話型翻訳

本方式の最大の特徴は、対話のタイミングと内容を完全にユーザがコントロールするという意味の完全ユーザ主導型の対話型翻訳にある。システムは変更可能な選択肢と現在の選択を常に表示しているが、ユーザは任意の順序で任意の操作を行うことができる。そのため、ユーザ操作は本質的にシステムの判断を向上させるために用いられる。この意味でユーザとシステムは協調的である。

システム主導型の対話型翻訳方式 [5, 6] では、曖昧性が発生したらシステムはユーザに問い合わせを行う。その場合、不要な問い合わせが発生する、特別に訓練を受けたオペレータが必要になる、等の問題が発生することが報告されている。我々のシステムでは対話をユーザが主導することにより前者の問題を回避するとともに、オペレーションを訳語選択と変換範囲推定という簡単な操作に限定することにより後者の問題に対処している。

本方式では、システムは常に受動的であるため、単独で用いるだけでなく、例えば自動翻訳システムと組み合わせ後編集用のインターフェースとして用いることも容易である。自動翻訳結果を順次キャンセルして、単なる辞書引きから、部分の構文変換に至るまでの様々なレベルの支援を利用することができる。

6 まとめ

対話型日英変換における曖昧性解消方式について述べた。さらに、従来提案されている対話型翻訳との比較において、ユーザ主導であることの利点を示した。今後は、システムに用例検索機能を組み合わせ、構文変換機能と用例検索機能のための共通プラットフォームとして拡張することを検討したい。

参考文献

- [1] 土井他: 「FEP 型英文作成支援ツール — 訳語選択のユーザインタフェースと辞書記述 —」 第51回情処全大 5H-1 (1995).
- [2] 山端他: 「FEP 型英文作成支援ツール — 日英構文変換部 —」 第51回情処全大 5H-2 (1995).
- [3] 亀井他: 「FEP 型英文作成支援ツール — 外国語情報発信の効果的インタフェース —」 第51回情処全大 5H-3 (1995).
- [4] Kiyoshi YAMABANA et al. “Interactive Machine-Aided Translation Reconsidered — Interactive Disambiguation in TWP—”. Proceedings of NLP-95 (1995).
- [5] Alan K. Melby et al. “ITS: Interactive Translation System”. Proceedings of COLING-80 (1980).
- [6] Masaru Tomita. “Disambiguating Gramatically Ambiguous Sentences by Asking”. Proceedings of COLING-84 (1984).