

6J-6

HICATS/JEのポストエディット支援環境

白井孝雄、坂本浩一、沢田覚（日立ソフトウェアエンジニアリング ㈱）
吉村紀久雄、梶博行（㈱ 日立製作所）

1. はじめに

日立日英翻訳支援システムHICATS/JE (Hitachi Computer Aided Translation System/Japanese to English)は、1986年4月に発売された統合型の翻訳システムであり、意味トランスファを含む高度な機械翻訳機能だけでなく、人間の高度な判断をも上手に取り入れるための支援機能（プリ/ポストエディット）や効率よく翻訳を行うための使い勝手のよい運用機能（大量一括翻訳、小量対話翻訳、辞書保守機能）をサポートしている。

本稿では、HICATS/JEのポストエディット支援について報告する。

2. 翻訳・校正用日英対照エディタ

現状の機械翻訳の技術レベルでは対処しきれない問題として、文脈・知識レベルを考慮しなければならない翻訳がある。HICATS/JEでは、このような翻訳に対処するため人間の判断を簡単に取り込めるエディタを採用した。このエディタは、図1に示すようにHICATS/JEの文書（ページ、段落、文といった論理単位で構成されている）を自由に編集できるだけでなく、論理単位の翻訳や翻訳結果の自動編集もサポートしている。ユーザは、このエディタを使用することにより最良の翻訳結果を得ることができる。

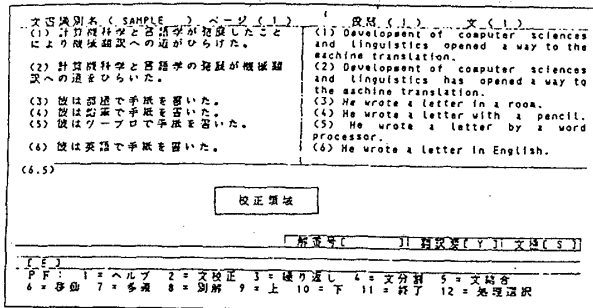


図1 日英対照エディタ

機械翻訳部とエディタとのインターフェイスには、メモリまたはファイルを用いている。1つの日本語文に対応する英語文、校正情報は図2に示すようになっており、これにより自動ポストエディット支援を実現している。

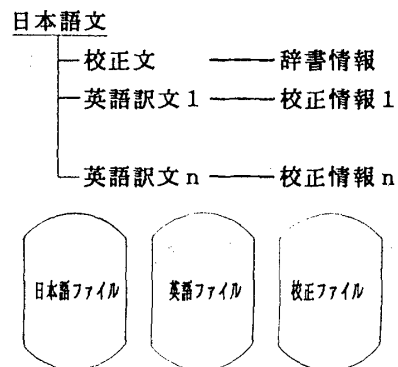


図2 原文/訳文/校正文の対応

3. 自動ポストエディット支援

HICATS/JEは、人間支援型機械翻訳 (Human Aided Machine Translation:HAMT)を採用している。すなわち、機械翻訳の結果を人間の判断をいれて校正するのであるが、さらに機械がその翻訳結果の校正作業をサポートする機能もある。この機能を自動ポストエディット支援と呼び、HICATS/JEでは、現在、次の2とおりの機能を提供している。

(1) 別英文表示選択機能

一つの日本語文には、一般に複数の英語文の対応が考えられる。

例

- ① 能動態 - I wrote a letter.
受動態 - A letter was written by me.
- ② 句 - at the program execution
節 - when the program executed

Post-editing Environment Supported by HICATS/JE

Takao USUI¹, Kouichi SAKAMOTO¹, Satoru SAWADA¹, Kikuo YOSHIMURA², Hiroyuki KAJI

1: HITACHI SOFTWARE ENGINEERING, Co., Ltd. 2: HITACHI, Ltd.

機械翻訳部の英語句構造生成¹は、日本語解析の結果得られた「概念依存図式²」から最適の英語文を作成する(本解)が、他の生成ルールによって別の英語文が作成できる場合(別解)これらを退避しておく。(この別解の生成数の上限はユーザが標準値パラメータで指定できる。)

ユーザが英語文を校正領域に移したとき解番号 n/m が表示され、別英語訳文の選択が可能となる。図3に別解表示/選択画面を示す。

複数の別解を校正領域に表示することにより、本解との比較が容易にできるという長所があるが、別解どうしの比較ができないという短所がある。また別解生成に対しての制御をユーザができないといった問題点がある。

(2)別訳語表示選択機能

一つの日本語単語に対して、一般に複数の英単語が考えられる。(訳語の多義という)

例 道 - way, road, street, path
作成する - generate, form, make

機械翻訳部では、動詞や名詞との共起及び語の性質による訳語選択等で、訳語を決定する処理が行われているが、決定が難しい場合も少なくない。そこで英語訳文中のこのような単語(英語品詞として名詞、形容詞、副詞、動詞の4種類を対象とする)を同時に表示する機能と、中の一語について別訳語を表示、選択できる機能を提供した。

ユーザは多義キーにより校正領域にある英語文中の多義語存在がすぐわかり、別訳語も選択できる。選択した際、形態素の合成が再度行われるの

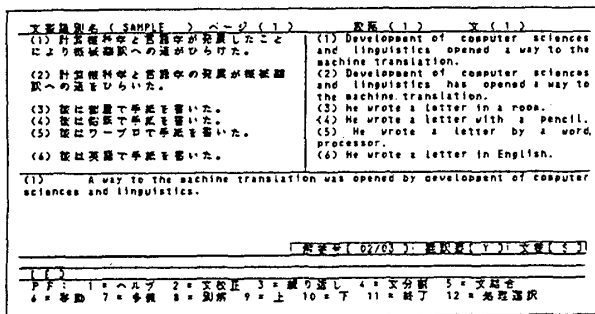


図3 別解表示/選択画面

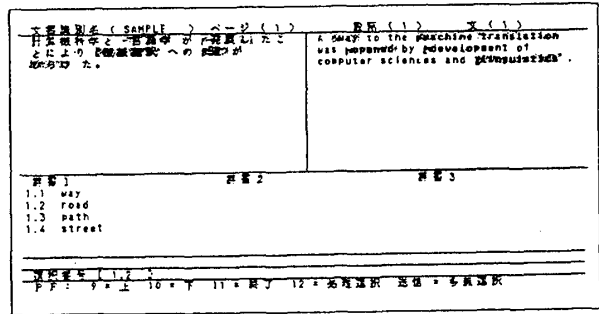


図4 多義選択画面

で、文中での変化形の変化も自動的に行われる。図4に多義選択画面を示す。

多義選択はユーザの判断を優先させ、形態素合成のみを行っているので、文体として正しくない英語が生成されることがあるが、訳語の多義を表示することを第1に考えているので、あまり問題とはなっていない。

4. まとめ

自動ポストエディット支援として

- 1)別英語文表示選択機能
- 2)別訳語表示選択機能

について述べた。これ以外にポストエディット支援の自動化として

- 1)電子辞書検索(用例検索)
- 2)前置詞・冠詞・数の自動変換 etc

を検討中である。

参考文献

[1] 梶、新田：概念依存図式からの英文生成、第28回情報処理全国大会5L-3 (1984-3)

[2] 梶、伊佐津：日英機械翻訳のための日本語文の依存構造解析、第30回情報処理全国大会7G-1 (1985-3)