

ホスト/パソコン機能分散型機械翻訳システム (2)

4B-8

武田 公人*

長谷部 浩一*

中島 美也子*

正木 康夫**

天野 真家*

(株) 東芝*

東芝CAEシステムズ(株)**

1. はじめに

当社は、ワークステーション (AS3000シリーズ) で動作する機械翻訳システム "AS-TRANSA C" を開発した。現在、多くのユーザで利用中である。この機械翻訳システムを、パソコンを端末として手軽に利用したいとする要求が強い。

本原稿は、ホスト計算機とパソコンをネットワークで接続する、分散処理型の (英日) 機械翻訳システムについて報告する。

2. システム構成

2.1 ハードウェア構成

センターのホスト計算機と利用者の手元のパソコンをネットワークで接続する分散処理型のハードウェア構成を図1に示す。パソコン側にOCRとプリンタが接続可能である。

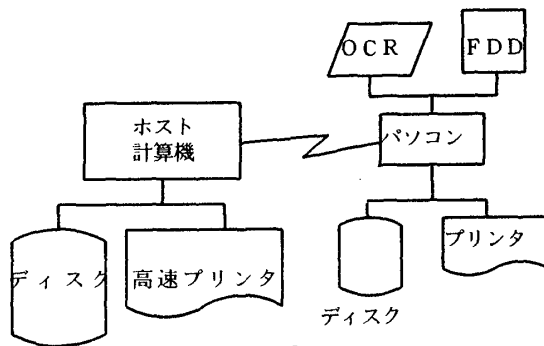


図1

(a) 入力機器: フロッピー・ディスク、磁気テープ、CD-ROM、OCR等

原文の入力は種々のメディアで提供される。例えば、英文マニュアル等の電子化された文書は、磁気媒体で入手することが多い。また、電子化されていない文書は、OCRで読込むと効率的である。

(b) 計算機: センターの大型機とパソコンをネットワークで接続する。

(c) 出力機器: 少量の印刷は、パソコンに接続されたプリンタ、大量の印刷はセンターのプリンタを利用する。

(d) その他: マウスを接続利用する。

2.2 ソフトウェア構成

AS-TRANSA Cの運用で翻訳機能以外に原文編集と2か国語編集の必要性が再確認された。

(a) 原文編集

原文編集の機能は、つぎの機能を持つ。

- ・ワードプロセッサ機能
- ・翻訳特有の編集機能

当社の翻訳システムは前処理のような手間のかかる処理を利用者にさせないという基本思想がある。したがって、原文を機械にかかりやすく書きかえるのではなく、原文編集は原文中のスペリング・ミスに対処する機能が中心である。

例えば、OCRで読込んだ原文は、文書の印字品質、OCRの性能によるが、一般には数%の読み誤りを含む。この状態の文書をそのまま翻訳しても、辞書にない語が頻繁に発生する。とくに、動詞、前置詞等に誤りがあると、文の翻訳を正しくすることは困難である。このため、スペリング・ミスの修正が必須となる。

(b) 2か国語編集

一括翻訳で翻訳された訳文を校正するエディタで、つぎの機能を持つ。

- ・日本語ワードプロセッサの基本機能
- ・英文ワードプロセッサの基本機能
- ・翻訳特有の校正、編集機能

各機能の操作は、AS-TRANSA Cで得られた使用頻度の高い機能を操作しやすい位置に配置した。

(c) 翻訳機能

対話翻訳と一括翻訳の両方の機能を備えることにした。対話翻訳は、2か国語編集集中に文単位で翻訳する。

一括翻訳は、指定された文書をまとめて翻訳する。

(d) 文書管理機能

計算機の知識を持たない利用者を想定し、ホスト計算機のファイル名等を意識せずに、利用できるようにする。利用者がつけた文書識別名をマウスで指示するだけで、原文編集、翻訳、2か国語編集を実行する。

3. 機能分担と翻訳の流れ

分散処理型の機械翻訳システムの利用形態、均一な応答速度、操作性の良さ、利用者の負担等を考慮して、つぎの機能分担とした。

- (a) ホスト計算機は、翻訳、翻訳辞書の管理、翻訳済文書の管理を行う。
- (b) 入力、及び全ての編集処理はパソコンで行う。原文文書の入力に関しては、OCRで原稿を読取って、直ぐに読取り誤り（スペリング・ミス）チェックをする必要があるため、原文文書とスペルチェックの辞書はパソコン側に持つこととした。

これらの機能分担を持つ翻訳の流れを図2に示す。つぎに、翻訳の流れを簡単に説明する。

- (a) OCRから読取った原文は、パソコンのファイルに記憶される。英文マニュアル等の電子化された文書はフロッピーから読込む。ホスト計算機にある文書は、スペルチェックを必要とする場合のみ、パソコンに送る。
- (b) 原文編集の終わった文書は、ホスト計算機へ送り、一括翻訳を行う。少量の文書の場合は、直接、対話翻訳を行うこともできる。
- (c) 一括翻訳の終わった対訳結果に赤字訂正後、対話翻訳で文単位の校正をする。
- (d) 文単位の校正が終了したら、清書機能で印刷形式を整えて印刷する。

4. 機能

分散処理型の機械翻訳システムは、AS-TRANSACと機能的に同一である。

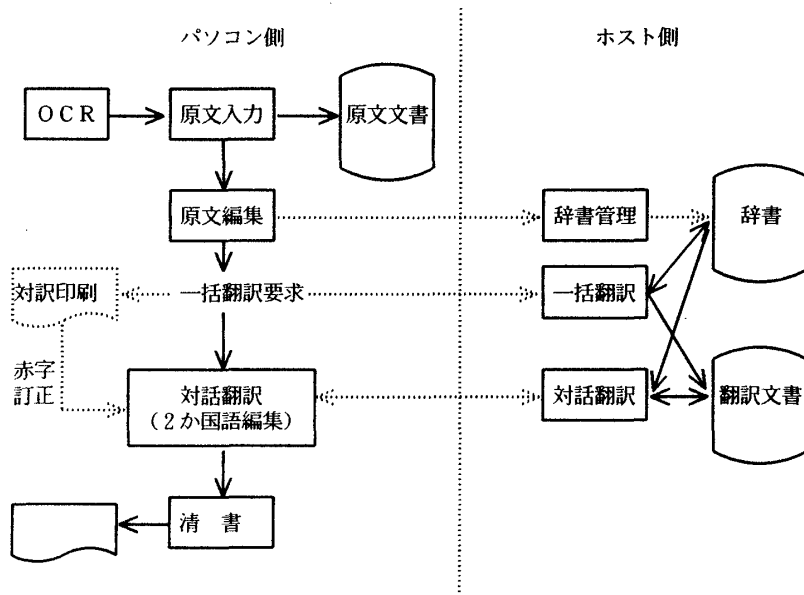


図2

つぎの機能は、ホスト計算機とパソコン間でデータ交信をしながら処理する。

- ・原文編集：辞書登録／削除、優先度変更
- ・一括翻訳：翻訳情況
- ・対話翻訳：再翻訳、辞書登録／削除、辞書参照、優先度変更、係受け変更、別解釈語の学習
- ・文書管理：文書名一覧表、文書の登録、削除

一例として、優先度変更のデータ交信手順を示す。

- ① PC側 単語を指定する
- ② PC→ホスト 単語を送る
- ③ ホスト側 辞書を引き
- ④ PC←ホスト 品詞情報を返す
- ⑤ PC側 品詞を画面に表示する
利用者が指定する
- ⑥ PC→ホスト 選択された品詞を送る
- ⑦ ホスト側 優先度を変更する
- ⑧ PC←ホスト 結果を返す

5. おわりに

ワークステーション（AS-3000シリーズ）で動作する機械翻訳システム“AS-TRANSAC”をベースに、マンマシーン・インターフェースをシステム構築の中心にした分散処理型の（英日）機械翻訳システムについて述べた。現在、開発中であるがAS-TRANSACと同程度の使い易さが実現可能である。

<参考文献>

- (1) 正木 他 【ホスト／パソコン機能分散型機械翻訳システム（1）情報処理学会第37回全国大会 1988】
- (2) 天野 他 【ワークステーションで実用速度を実現した翻訳システム】 archive No. 8
- (3) 安達 他 【日英相互機械翻訳システムのエディタ方式】情報処理学会第34回全国大会 6W-4 1987