

点訳・翻訳の統合システム

3J-6

十川 奈美子[†] 佐藤 浩史[†] 原 俊介[‡] 大武 信之[‡]
[†]お茶の水女子大学人間文化研究科 [‡]筑波技術短期大学

1 はじめに

コンピュータの普及に伴い数々の点訳ソフトウェアが利用されているが、我々はこれまでにいくつかの点訳及び翻訳システム [1-6] を開発してきた。

これらの点訳システムはそれぞれ数式点訳・英語点訳などといった利用目的に合わせて独立して提供されている。そのため、英語、日本語、数式、化学式など複数の文字体系が混在した文書を点訳する場合の自動化が完全でなく、視覚障害者を含む利用者に大きな負担を与えてしまう。また、点字に詳しくない晴眼者が点訳システムを利用して点字文書を作成した際に、点訳が正確か否かを判断することは困難である。そのため、点訳とは逆機能を果たす翻訳システムがあると便利である。翻訳システムを開発する上で問題となるのが、現行日本点字規則における多義性・曖昧性 (2章参照) である。統一日本語点字コード (UJBC: Unified Japanese Braille Code) [7] によればこれらの問題点が解消でき、翻訳システムの開発 [5] が可能となった。

本研究では、これまで開発されてきたそれぞれのシステムを見直し、墨字から点字、点字から墨字への双方向変換の自動化を可能にするための制御・統合システムを開発した。

2 現行点字と UBC, UJBC

現行日本点字規則が制定されたのは、1890年のことである。その後初等中等教育の充実とともに、1980年頃から、数学や各種分野の点字体系が整備され始め、同時に現行日本点字において様々な問題があることが明らかになった。この問題は、日本だけでなく英語圏においても生じている。特に点訳・翻訳を自動化する上で支障をきたす問題点を以下に示す。

- 分野ごとに点字コード体系が異なる。
- 墨字に対する未定義の点字コードがある。
- 各国の統一がとれていない。

通常、晴眼者が使用する墨字、英語、数式記号や化学式は、世界共通の文字・記号として標準的に用いられてい

る。しかし、点字ではこれが当てはまらない。

このような現行点字における問題点を解消するために、英語圏において統一英語点字コード (UBC: Unified English Braille Code) [8] が 1991 年に提案された。これは、世界標準の英語点字コード体系として各国で評価実験が行われ、2003 年の採用に向けて努力されている。日本もこれらの動きに合わせて UJBC が提案された。UJBC では仮名表記の部分は現行日本点字規則を、数式・英語の部分は UBC を用いている。

3 点訳・翻訳システムとその利用

晴眼者が使用する墨字を 6 点の凸点で表示した点字に書きかえることを点訳、点字から墨字への変換を翻訳と呼ぶ。これまでに開発された点訳システムは、日本語処理、数式処理、英語処理を、それぞれ、独立したシステムとして提供されている。また、翻訳システムも点訳システムとは独立に開発した。

数式点訳システムとして、MathBraille [1] と MathUJBC [3] の 2 つが開発されている。MathBraille が現行日本点字規則を基に開発されたのに対し、MathUJBC は新しい点字コード体系 UJBC に基づいて開発された。使用する点字規則に違いはあるが、両者とも L^AT_EX で表現された数式を含む文書を自動点訳できる。また、数式点訳の前処理として、`\newcommand` などのマクロ定義の展開を行う MacroEx [4] が開発されている。

点字における英語の記述方法は Grade I と Grade II の 2 種類がある。UBC では、Grade I は現行通りの規則を用い、Grade II は現行英語点字において曖昧性をもたらす 4 つの略字を廃止し、さらに単語間の空白文字を省略するという英語点字特有の規則も廃止した。英文を点訳する場合、Grade II を用いるのが標準である。この UBC に基づく英語点訳システム EngBraille [2] を開発した。

UJBC から墨字への翻訳システム BrtoPr [5] も開発された。英語、数式、日本語の混在した文書において、UJBC では日本語と英語・数式の切り分け記号が未定義である

ため、このシステムは日本語処理と英語・数式処理を切り離して処理を行っている。現在、曾根ら [6] は、このシステムを発展させている。

各システムは独立して利用されるが、日本語、英語、数式の混在する文書进行处理するには、システム間のファイルの受け渡しなどに手作業が必要となり使用者に負担がかかる。しかし、システム設計面では、各機能が独立していることで、拡張性の高いシステムの実現が可能である。

点訳の際には、市販のソフトウェア [9] を利用して、漢字から仮名分かち書きへの変換を行っているが、統合システムに市販のソフトウェアを組み込むのは問題がある。また、翻訳における仮名から漢字への変換にも困難が伴う。

以上のことから、英語、日本語、数式など複数の文字体系を含むテキストの点訳・翻訳は、これまでに開発したシステムに頼ってもまだ容易だとはいえない。

4 システムの概要

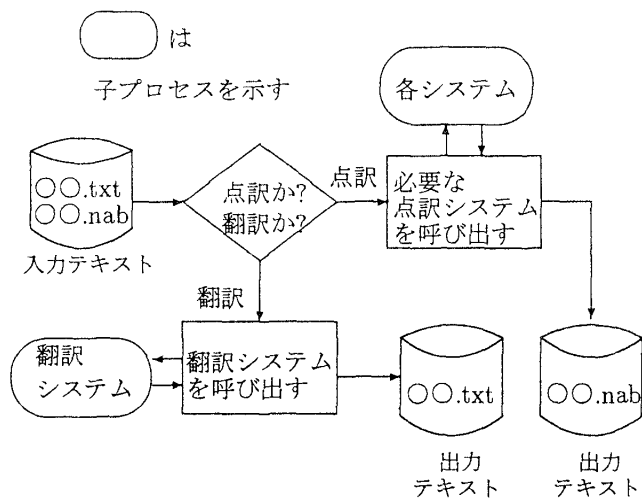


図 1: 本システムの流れ

点訳及び翻訳システムを統合する方法として、入力から出力までの全処理を一つのプロセスで処理する方法と、前節で述べたシステムを子プロセスとして利用する方法の 2 通りが考えられる。本システムは、各システムの修正および拡張が容易に行える後者を採用し、これまで開発された個々のシステムを制御・統合する点が大きな特徴である。このため、各システムの設計に影響を及ぼす部分は手を加えず、システム間の調和をとるための入出力部分の変更・修正のみを行った。また、漢字仮名分かち書き変換及び仮名漢字変換処理を付加した。これによ

り、本システムでは市販ソフトウェアは不要となった。

5 おわりに

本システムは、パソコン上で開発を行い、これまで UNIX 上で開発された各システムをパソコンに移植した。視覚障害者用のソフトウェアのほとんどは DOS/V または Windows 環境で開発され、視覚障害者およびボランティアの間では UNIX は普及していないため、高コスト・大規模なワークステーションより、パソコン上で開発されたシステムの方が、広範囲に利用される。

UBC および UJBC の採用にあたっては検証が必要である。本システムは、UBC、UJBC に基づいた点訳・翻訳システムであり、検証システムとしても機能する。

参考文献

- [1] 宮崎紀子, 樋口美佳, 佐藤浩史, 原俊介, 大武信之: 数式自動点訳システム, 電子情報通信学会, 信学技報, ET94-119, pp.9-16, 1995.
- [2] 十川奈美子, 佐藤浩史, 原俊介, 大武信之: UBC に基づく英語点訳システム, 電子情報通信学会, 信学技報, ET96-76, pp.1-8, 1996.
- [3] 安井民子, 佐藤浩史, 原俊介, 大武信之: UJBC に基づく数式点訳システム, 電子情報通信学会, 信学技報, ET96-77, pp.9-16, 1996.
- [4] 楠佳奈子, 佐藤浩史, 原俊介, 大武信之: 自動点訳システムのための L^AT_EX マクロ展開, 電子情報通信学会, 信学技報, ET97-82, pp.9-16, 1997.
- [5] 野辺純子, 藤山薫, 佐藤浩史, 原俊介, 大武信之: UJBC 文書の墨字翻訳システム, 電子情報通信学会, 信学技報, ET97-81, pp.1-8, 1997.
- [6] 曾根佐枝子, 佐藤浩史, 原俊介, 大武信之: UJBC 文書の墨字翻訳システム II, 電子情報通信学会, 信学技報, ET98-93, pp.1-8, 1998.
- [7] 藤芳衛, 石田透, 木塚泰弘, 澤崎陽彦, 山口雄二: 新しい統一日本語点字記号の開発, 電子情報通信学会, 信学技報, ET96-78, pp.17-26, 1996.
- [8] 点字科学記号検討資料編纂会訳: 統一英語点字コード, 日本点字委員会, 1995.
- [9] (例えば) 石川准: 日本語英語自動点訳プログラム「EXTRA for Windows」Ver 1.0 1998.