

かな漢字変換評価技術の開発

恒川昌昭、高尾直弥、角史生
松下電器産業(株) 情報通信関西研究所

かな漢字変換の変換性能の定量的な把握と性能向上のための材料の抽出、および信頼性の検証の自動化を実現する技術を開発した。本技術では、大量の評価例文を使って自動的に変換実験を行い、その結果を集計するとともにかな漢字変換の改良のための材料を得ることにより、改良作業の迅速化が可能になった。また、今まで人手でしか実現できなかったかな漢字変換の信頼性の検証を自動的にかつ短期間で行うことができ、ソフトウェアの信頼性検証を飛躍的に効率化することが可能になった。なお、本技術を使ってかな漢字変換を開発することにより、約1年半の間に8%の変換率の向上を得た。

Technology for evaluation of Kana-kanji conversion

Masaaki Tsunekawa, Naoya Takao, Fumio Sumi

Kansai Information and Communications Research Laboratory
Matsushita Electric Industrial Co., Ltd.

We have developed a system for evaluation of Kana-kanji conversion system. This system declares the performance of conversion quantitatively, collects the stuff for improvement, and automates the testing of Kana-kanji conversion system. Totalizing the results of conversion on large volumes of sample data automatically and collecting the stuff for improvement, this system enabled us to improve the performance of conversion in a much shorter period. Furthermore, verifying the reliability of system automatically in a short period, which had to be executed by many hands, this system enabled us to verify it much more efficiently. With this system, we've got the improvement of 8% on the performance of conversion for one and a half years.

1. はじめに

かな漢字変換技術は、日本語ワードプロセッサを中心としてさまざまなコンピュータシステムへ適用される基本技術となっており、依然として変換性能の向上が強く望まれている。また製品のシリーズ展開の必要性から、かな漢字変換の開発期間の短縮、開発コストの低減などの効率化が要求されている。さらに、変換精度を向上させるために、形態素解析技術だけではなく構文解析技術、意味解析技術の利用も試みられるなど、かな漢字変換がますます複雑化してきており、システムの信頼性の確保が重要となってきた。

また、今までかな漢字変換を定量的に評価する技術が確立されておらず、変換性能の向上を定量的に裏付けることができていなかった。

そこで今回、次に示す機能をもったかな漢字変換評価システムの開発を行った。

(1) 変換性能把握の自動化

かな漢字変換の変換性能を定量的に把握するために大量の評価例文を用いた変換実験とその結果の集計を自動的に行うとともに、かな漢字変換の改良のための材料を提示する。

(2) 信頼性検証の自動化

ユーザモデルの導入により、今まで人手でしか実現できていなかったかな漢字変換の信頼性の検証を自動化する。

本稿では、この評価システムについて報告する。

2. 評価システム

本システムの構成を図1に示す。

図1に示すように、本システムは評価例文と変換率算出・改良の材料提示・信頼性検証プログラムからなる。

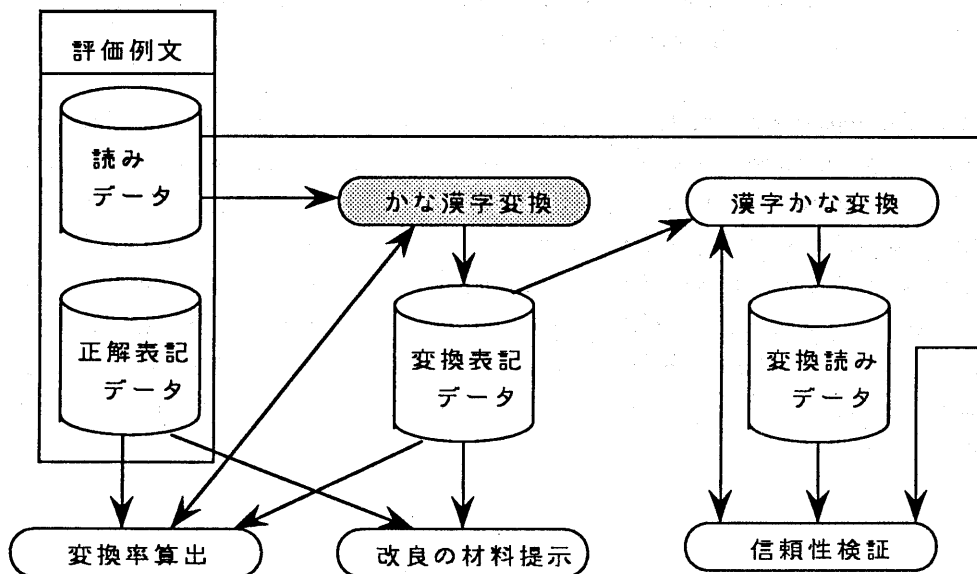


図1 評価システム構成図
《2》

2. 1 評価例文

電子化された評価例文の総数は、約20万文節（約500文例）である。

2. 1. 1 評価例文の内容

評価例文は、読みデータと正解表記データの2つから構成される。

表1にその例を示す。

(1) 読みデータ

- ・かな漢字変換を行うための読み入力データであり、ひらがなおよび記号類を含み、漢字は含まない。

(2) 正解表記データ

- ・原文の漢字かな交じり文字列の表記データである。
- ・文節の区切りごとに、区切り記号「/」が挿入されている。

読みデータ
はいけいしましますごせいえいのこととおよろこびもうしあげます。 さて、 3がつ15にちづけをもちまして「しんがたわーくすてーしょん」の みつもりしょをうけとりました。
正解表記データ
拝啓／貴社／ますます／ご／清栄のこととお／喜び／申し上げます／。 さて／、 3／月／15／日／付けを／もちまして／「／新型／ワークステーション／」の／ 見積書を／受け取りました／。

表1 評価例文の内容

2. 1. 2 評価例文の種別

評価例文の種別として次の2種類を用意することにより、例文に片寄らない客観的な変換率の評価ができるようにした。

(1) 解析用例文（10万文節）

かな漢字変換による誤変換例を解析して、かな漢字変換の改良策を得るための例文である。

(2) 検証用例文（10万文節）

かな漢字変換の変換率を確かめるためだけに使う例文である。

2. 1. 3 評価例文の分野

評価例文の分野としては、日本語ワードプロセッサを使って作成される可能性のあるできるだけ広い

分野を対象とした。そこで、表2に示す分類に基づき、評価例文を収集した。

①一般大衆対象の 文書	販売促進	PR誌、DM、カタログ、チラシ
	教養を深める	文庫本、生活図書、教養書
	レクリエーション	趣味・娯楽、旅行用図書
	教育関連	教科書、参考書、学習書
	ニュース	新聞、月刊誌、週刊誌
	学問	専門書、研究図書、ビジネス書
	商品添付資料	保証書、取扱説明書、マニュアル
②組織で作成する 文書	一般企業	通達、案内状、招待状、挨拶状、辞令、 論文、決裁書、契約書、特許
	官公庁	白書、法律
③プライベート	手紙類	手紙、招待状、年賀状、通知状
④情報ファイル	固有名詞	住所録、出席簿、名刺
	その他	メニュー、辞典、統計資料

表2 評価例文の分類

2. 2 機能

次に、本システムが持つ機能について述べる。
大きく分けて、次の3つの機能を持っている。

- (1) 変換率算出
- (2) 改良の材料提示
- (3) 信頼性検証

2. 2. 1 変換率算出

2. 1で説明した評価例文の読みデータを用いてかな漢字変換させ、そこで得られた変換結果の表記と正解表記データとを比較することにより、自動的に変換率を算出する。

現在、当社製ワークステーション(ソルボーン5/601)では、約20万文節の評価例文を用いた変換実験、結果の集計、変換率の算出を約1.5時間で実現している。

変換率は一般に、次のように定義される。

$$\text{変換率} = \frac{\text{正解文節数}}{\text{全文節数}} \times 100 (\%)$$

我々は変換率として、次の3種類を定義している。

- (1) 第1変換率：文節区切り位置が正解と一致し、かつ第1候補が正解と一致する割合
- (2) 第2変換率：文節区切り位置が正解と一致し、かつ候補中に正解と一致するものがある割合
- (3) 文節変換率：文節区切り位置が正解と一致する割合

2. 2. 2 改良の材料提示

この機能では、改良に先立ってそのためのポイントを見つけるための材料と、改良後にその改良の是非を判断する材料を提示する。

2. 2. 2. 1 改良のポイント

評価例文のうちの解析用例文についてのみ、その読みデータを用いてかな漢字変換させた結果の表記と正解表記データとを比較し、その中から以下に示す2通りの誤った変換例（以下誤変換例と呼ぶ）を抽出する。その誤変換例を解析することにより、改良すべきポイントを得ることができる。

(1) 文節区切り誤り

表3に示すように、文節の区切り位置が間違っている文節のみを抽出したものである。

(2) 同音語誤り

表4に示すように、文節の区切り位置はあっているが、第1候補が正解と一致していない文節を抽出したものである。

変換表記データ	正解表記データ
国内で／配膳	国内では／以前
関連し／多極を	関連した／曲を

表3 文節区切り誤り

(誤) この／分布／要領は
(正) この／分布／容量は
(誤) モータを／開店させます／。
(正) モータを／回転させます／。

表4 同音語誤り

2. 2. 2. 2 改良の是非

評価例文のうちの解析用例文について、改良前のかな漢字変換を用いて変換させた結果の表記データと改良後のかな漢字変換を用いて変換させた結果の表記データとを比較し、良くなったもの（以下改善例と呼ぶ）と悪くなったもの（以下改悪例と呼ぶ）を抽出する。この改善例と改悪例の2つの内容を見ることにより、改良が有効であるかどうかを判断することができる。

表5と表6に改善例、改悪例の一例を示す。

(新) OAが一般化されればされるほど、
(旧) OAが一般かされればされるほど、
(新) 不正経理が暴露されると、
(旧) 不/生計/利が暴露されると、
(新) 日本/支部の/事務局では、
(旧) 日本史/部の/事務局では、

表5 改善例

(新) 発明/者/および/特許/検車/共々、
(旧) 発明/者/および/特許権/者/共々、

表6 改悪例

2. 2. 3 信頼性検証

かな漢字変換の最終テストとして、例えば人間による約3人月の打ち込みテストを行うとすると、3人の入力テスト者を用意するとしても、約1ヵ月の期間がかかり、開発期間の短縮を行う上での大きなネックとなっていた。また、この程度の打ち込みテストで検出できる誤動作には自ずと限界があり、ソフトウェアの信頼性を確保するためには、人間によるテストでは到底できないほどの時間が必要となるという問題があった。

そこで、今回、これらの問題点を解決するために、この人間による打ち込みテストの完全自動化を目指した。

2. 2. 3. 1 動作検証のユーザモデル化

まず、人間による打ち込みテストの手順を分析し、どのような手順で誤動作の検出を行っているのかをモデル化する。次に、そのモデルを具体的に計算機に取り込むことにより、打ち込みテストの自動化を図る。以下に、これまで行われていた打ち込みテストの手順を示す。

[打ち込みテスト手順]

- (1) テスト用例文を用意し、変換/次候補等の操作を行うことにより目的の表記を得る過程で、誤動作*1が起らないことを確認する。
- (2) 入力可能な最大文字数、入力可能な最大文節数等、かな漢字変換の仕様の限界にあたる例文を使って変換させる過程で、誤動作が起らないことを確認する。
- (3) 通常行われないであろうアットランダムな操作を行う過程で、誤動作が起らないことを確認する。

*1 誤動作には、暴走、正解が得られない、学習が行われていない、入力した読みからは考えられない変換結果が得られる等がある。

上記の3つの手順に対応して、今回、次に示す3つのモデル化を行った。

- (1) 標準ユーザモデル

一般的なユーザが変換操作を行う手順をモデル化したもので、評価例文を用いて正解表記に近づける動作を自動化している。

(2) 限界ユーザモデル

ソフトウェアの仕様限界の動作を検証する手順をモデル化したもので、入力可能な最大文字数の例文入力や、可能な限り文節長を短くする等の動作を自動化している。

(3) ランダムユーザモデル

ユーザが誤って行う操作の手順をモデル化したもので、アットランダムに例文や変換操作を発生させる動作を自動化している。

現在は、上記の(1)、(2)の2つのモデルが完成しており、(3)のランダムユーザモデルに関しては、まもなく完成する予定である。

2. 2. 3. 2 漢字かな変換技術の導入

打ち込みテストで人間が検出する誤動作には、以下に示す4つのものがある。

- (1) 暴走または停止する。
- (2) 目的の正解表記が得られない。
- (3) 学習が正常に行われない。
- (4) 入力した読みからは考えられない変換結果が得られる。

ところが、このうちの(1)~(3)の3つを自動的に検出することは比較的容易であるが、(4)を検出するには、テストで使う評価例文の読みデータに対して考えられるすべての表記データを電子化する必要がある。かな漢字変換には同音語が多いことから、非常に多くのコストが必要となる。

そこで、漢字かな変換技術を導入することにより、従来極めて検出が困難であった(4)の誤動作を、評価例文の読みデータと正解表記データのみを用いて検出することを可能とした。

3. かな漢字変換開発への適用

本システムを使ってかな漢字変換の開発を行う手順を以下に示す。

[開発手順]

- (1) まず、現在の交換性能を把握するために、評価例文全体を使い、交換率を算出する。
- (2) 次に、交換結果の中から、誤変換例を抽出し、改良すべきポイントを得る。
- (3) そのポイントに応じて、かな漢字変換を改良する。
- (4) 解析用例文を用いた実験により、改良前と改良後の交換の違い(改善例と改悪例)を解析し、この改良の効果を検討する。
- (5) 検証用例文を用いた実験により、改良の前後の交換率の変化を見て、改良の妥当性をより客観的に検討する。
- (6) 最後に、信頼性検証機能を使って、改良したかな漢字変換が誤動作しないかどうかをテストし、誤動作がないことをもって改良作業を終了する。

4. 適用結果および効果

本システムを使ってかな漢字変換の開発を行った結果について述べる。

4. 1 変換率の向上

本システムを使って開発したかな漢字変換技術の変換率の推移を図2に示す。

この図に記述してあるV0～V3は、かな漢字変換のバージョンであり、それぞれ次のようなものである。なお、V0～V3へと開発の時期は新しくなっている。

V0：初期のバージョン

V1：本システムを使わずに開発したバージョン

V2：本システムを使って開発した1つ目のバージョン

V3：本システムを使って開発した2つ目のバージョン

結果として、本システムを使って開発したV2の第1変換率のV1からの向上分は、本システムを使わずに開発したV1の第1変換率のV0からの向上分の3倍以上であった。

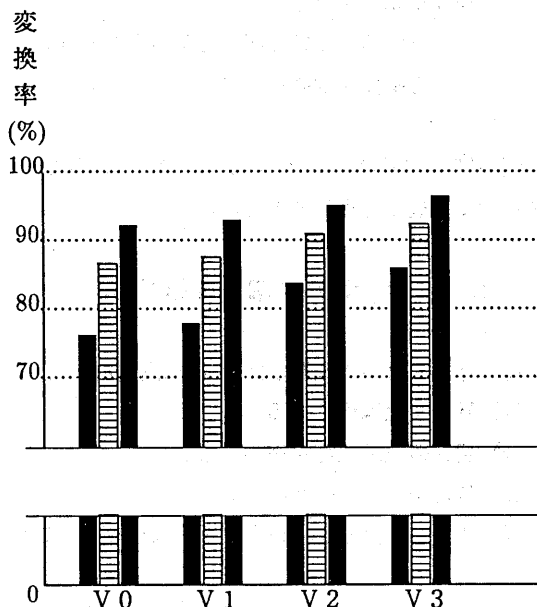
4. 2 テスト工数の削減

本システムの信頼性検証機能により、今まで人間により行われていた3人月の打ち込みテストが計算機による約10日間に短縮され、かな漢字変換の開発期間の大幅な短縮が可能になった。と同時に次に示す信頼性の向上もあった。

4. 3 信頼性の向上

本システムの信頼性検証機能により、人間による3人月の打ち込みテストではできない非常に広範囲な変換動作の検証を自動的に行うことができるので、信頼性の高いソフトウェアの供給が可能になった。

実際には、バージョンV3のかな漢字変換プログラムを本システムの信頼性検証機能で検証したところ、新たな誤動作を2つ検出することができた。このうちの1つの誤動作は、バージョンV2の開発時に発生したものであり、バージョンV2とV3の開発における合計6人月の打ち込みテストで検出できなかった誤動作を検出できたことになる。



(それぞれ左から順番に、第1変換率、第2変換率、文節変換率を表す。)

	V0	V1	V2	V3
第1変換率	76.36	78.10 (+1.74)	83.91 (+5.81)	86.15 (+2.24)
第2変換率	86.68	87.59	91.02	92.47
文節変換率	92.43	93.19	95.31	96.65

図2 変換率の推移

5. おわりに

現行の評価システムは、信頼性検証機能のうちの、ランダムユーザモデルが完成していないので、この機能を早急に完成させることが必要である。

今後の課題として、信頼性検証機能をより充実させるため、検証範囲を拡大させ、使用する例文の数をさらに増強するなど、かな漢字変換の評価技術としてのさらなる強化を図っていき、本システムを使ったかな漢字変換の開発の効率化を引き続き行っていく。

また、かな漢字変換技術の開発効率化を目的とした本評価技術を、他の日本語処理技術の評価へも適用していきたいと考えている。