

かな漢字変換における文節切り直し学習機能

3J-2

*横田 浩 **中里 茂美 **大黒 和夫 ***堤 義直
 *東芝ソフトウェアエンジニアリング株式会社
 **株式会社 東芝 情報通信システム技術研究所
 ***株式会社 東芝 青梅工場 ワープロ部

1. はじめに

最近の日本語ワードプロセッサ等で用いられているかな漢字変換技術は向上し、変換率も高くなり、操作性も良くなっている。操作性の向上については、全文まるごと変換方式によるところが大きいが、全文まるごと変換方式の問題点としては、機械が判断した文節が誤っている場合に正しく変換されず、オペレータが思い通りの文章に変換するためには、変換結果を正しい文節に区切り直ししなければならない。この正しい文節に区切り直す機能が文節切り直し機能である。

このため、最終変換率の向上をはかる上で文節切り直し機能は欠かすことの出来ない機能となっている。ただし、現行の文節切り直し機能ではいくつかの問題点がある。本報告では、これらの問題を取り上げ、その対策を述べ、より有効な文節切り直し学習機能について報告する。

2. 文節切り直し学習機能の必要性

全文まるごと変換方式の処理において、オペレータが入力する読みがな列に対し、変換システムは分割単位を検出し自動ブロック分割を行っている。しかし、例1のように自動ブロック分割で失敗した場合には、オペレータの意図するかな漢字変換結果が得られないことがある。このような場合、誤りを訂正しオペレータが意図するかな漢字変換結果が得られるようにするには、細分して再変換しなくてはならない。これを解決するために、オペレータが文節を切り直すことにより、簡単に意図する変換結果を得られるようにした機能が文節切り直し機能である。

<例1>

(原文) 自己像の持ち方がいかに大切か
 (誤変換結果) 自己像の持ちか／互いかに大切か

では着物を脱ぐ。」という変換結果を望んでいる場合の文節切り直しの処理を示している。ここで文節切り直しを行った位置の前の文節(第1文節)の読みと後の文節(第2文節)の読みを文節切り直し学習辞書に登録する。このときの文節切り直し学習辞書の内容を図1に示す。次に同じ文章が入力されたとき、文節切り直し学習辞書を参照し、入力文字列の読みと文節切り直し学習辞書の内容が一致すれば文節切り直し位置をかな漢字変換部に指示し、オペレータの望む文節群を含んだ変換結果が得られるようかな漢字変換を行う。

<例2>

入力文字列「ここではきものをぬぐ。」
 ↓
 変換結果「ここで履物を脱ぐ。」
 (同音語に「ここでは着物を」が含まれている)

↓
 切り直しするために読みに戻す
 「ここではきものを脱ぐ。」
 ↓
 カーソル等で切り直し位置を指定し再変換する。
 「ここではきものを脱ぐ。」
 ↓
 再変換結果「ここでは／着物を脱ぐ。」
 (同音語が「ここでは」と「着物を」に分かれている)

第1文節の読み 第2文節の読み

こ	こ	は	は	き	き	も	も	の	の	を	を
こ	こ	は	は	き	き	も	も	の	の	を	を
こ	こ	は	は	き	き	も	も	の	の	を	を

図1 文節切り直し学習辞書の内容

3. 従来の文節切り直し学習機能の概要および問題点

1) 概要

文節切り直しを実行し、再度同じ文字列が入力された際、オペレータの手間を省くことを目的として文節切り直しの学習を行う。例2は、オペレータが「こ

2) 問題点

従来の文節切り直し学習機能は、学習参照において入力文字列と文節切り直し学習辞書の内容を比較し、一致したときは無条件に切り直し位置で文節を区切り変換するようにしていたため、同音語が区切られてし

The sentence clause cutting repair study function in Kana to Kanji conversion.
 *Hiroshi YOKOTA **Shigemi NAKASATO **Kazuo OOGURO ***Yoshinao TSUTSUMI
 *TOSHIBA SOFTWARE ENGINEERING Corp.
 **TOSHIBA Information and Communication Systems Laboratory.
 ***TOSHIBA Corp. OME WORKS.

まい、自動ブロック分割の特長である同音語の中に異構造を含ませることが生かされない。

また、文節を切り直す位置を指定する際、使い慣れていないオペレータにとってはどの位置で文節を区切れば良いのか理解しづらい面がある。そのため誤った位置で文節を切り直してしまう可能性があり、この場合、また初めから文節切り直しを実行しなければならない。

4. 今回の改良点

1) 学習参照

自動ブロック分割を有効に活用するために、文節切り直し学習機能の適用に条件をつける。

[条件1]

文節切り直し学習辞書の内容と一致した入力文字列の範囲内に自動ブロック分割による区切りが含まれていないときは、文節切り直し学習機能を適用しない。

たとえば、例3の場合、文節切り直し学習辞書と一致する範囲「ここではきものを」の中に自動ブロック分割による区切りが含まれていないので文節切り直し学習機能は適用しない。

<例3>

入力文字列「……ここではきものを……。」

自動ブロック分割による区切り

「……／ここではきものを／……。」

—学習辞書と一致する範囲

／—自動ブロック分割による区切り

[条件2]

文節切り直し学習辞書の内容と一致した入力文字列の範囲内に自動ブロック分割による区切りが含まれているときは、文節切り直し学習機能を適用する。

たとえば、例4の場合、例5のような文節切り直し学習辞書と一致する範囲「もちかたがいかに」の中に自動ブロック分割による区切りが含まれているので、自動ブロック分割による区切り位置を補正するために文節切り直し学習機能を適用する。

<例4>

入力文字列

「……もちかたがいかにたいせつ……。」

自動ブロック分割によるブロック切り

「……／もちか／たがいかに／……。」

—学習辞書と一致する範囲

／—自動ブロック分割による区切り

<例5>

変換結果 「自己像の持ちか互いかに大切に。」

切り直し処理

↓

変換結果 「自己像の持ち方がいかに大切に。」

文節切り直し学習辞書の内容

第1文節の読み 第2文節の読み

もちかたが	いかに
：	：

2) 取り消し可能文節切り直し機能

文節を間違った位置で区切った場合、1回のキー操作で、文節を切り直す位置を指定できる状態に戻すことができれば使いやすい機能になると考えた。そこで変換キーのほかに取り消し可能変換キーを押下すると指定された区切り位置で変換を行い、次に左右カーソルキーが押下されると変換を行う前の状態、すなわち文節を切り直す位置を指定できる状態に戻るよう改良した。取り消し可能文節切り直し機能の処理を例6に示す。

<例6>

自己像の持ちか／互いかに大切に。

↓

切り直しキーを押下し区切り位置を指定

↓

自己像のもちかたがいかに大切に。

↓

取り消し可能変換キー

↓

自己像の持ち方／外貨に大切に。

↓

右カーソル押下

↓

自己像のもちかたがいかに大切に。

↓

変換キー押下

↓

自己像の持ち方がいかに大切に。

取り消し可能変換キーを押下した直後に左右カーソルキーが押下された場合、それぞれのキーに従って、取り消し可能変換キーを押下する前の区切り位置より左右にひとつカーソルが移動する。

5. おわりに

本報告では、従来の文節切り直し学習機能における問題点とその改善策について述べた。今回の改善策によって、よりユーザフレンドリなマンマシンインターフェースを実現することができた。