

# 相関ルールを用いた組織内における文書校正支援

上田芳弘\* 古川真士\*\* 加藤直孝\* 林克明\* 木村春彦\*\*

## 研究の背景

一般に組織内では情報周知をはじめ、知識の共有や組織の意思表示のために文書が作成され、文書作成者とは別の校正者が様々な側面から校正を行っている。しかし、この校正作業は紙面上で行われるために校正結果の情報は組織内で蓄積、共有されていない。そこで、組織内の文書校正過程で得られる校正データに相関ルールを応用して、文書校正に役立つルールの抽出手法を確立する。この校正データには、組織固有の情報が内在しているので、表記や文法レベルのほか、組織情報に基づいた校正支援が抽出したルールにより可能になるものと考えられる。

## 研究内容

工業試験場では、官公庁における公文書作成を対象とした文書決裁支援システムを開発し、運用している。本システムは、業務担当者が作成した文書を組織内で決裁するとともに、文書作成者や承認者が文書内容を様々な角度から検討して、文書を校正する機能を有する。本システムを運用する中で、文書校正内容を表すデータ（校正データ）を、校正前後の文書ファイルの差分から生成できるようにした。この校正データに相関ルールの手法を応用して、単語の置換、挿入及び削除についてのルールを抽出した。このとき、校正の頻度は少なくとも、重要なルールを見逃さないために適用度と呼ぶ指標を導入した。さらに、文書や業務の種類（ドメイン）によって変化するルールの使用確率に対応するために、校正が完了した文書を用いて、ドメインごとに適用度を変更した。すなわち、ドメインが予め設定できる場合は、一つのルールに対してドメインごとに適用度を計算することにより、ドメインごとに異なった校正支援が可能となる。また、提案手法で抽出したルールによる校正結果と人間の校正者が行った結果を比較して、人間の校正結果に対してヒット率が高く、かつ誤り校正率が低くなるような適用度のしきい値を求め、これによってルールの有効性を決定することとした。

## 研究成果

上記の研究結果から以下の3点が明らかになった。

- (1) 提案手法により表記や文法、意味レベルの校正と組織固有の情報に基づいた校正を表すルールが実際に抽出できた。
- (2) ルールの適用度をドメインごとに変更させた結果、ルールに含まれる用語の割合から適用度の変更は適切に行われたものと推測された。
- (3) 図のとおり、校正者による校正の80%がルールにより校正可能であり、誤り校正率は適用度のしきい値を適切な値(0.7)に設定することにより低下可能であることが分かった。

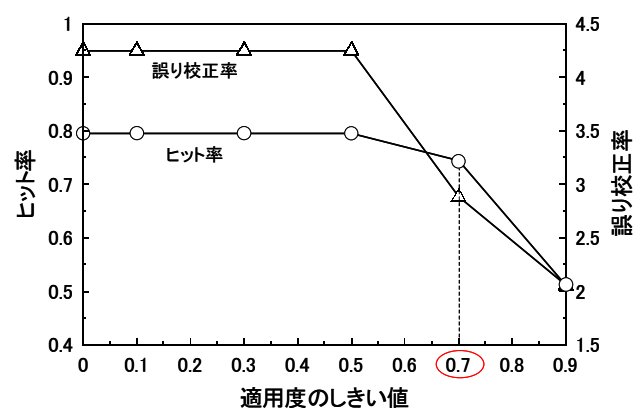


図 校正者とルールによる校正の比較

## 論文投稿

電子情報通信学会論文誌. Vol.J85-D- , No.7, 2002, p.681-690.

\*製品科学部 \*\*金沢大学