

文書自動添削システムによる学生の文書改善履歴の調査

プロジェクトマネジメントコース 矢吹研究室 1442031 氏名 小山隆太郎

1. 背景

学生が行う研究では、研究だけではなく文書を作成する時間が多い。特に卒業論文は分量も多く、形式も指定されるため、文書校正にかかる労力は大きい。また、自分以外が読んでもわかりやすい文書を書く必要があり、文が長いほど理解が難しくなってしまう場合や、「だから」、「かなり」といった口語が混じり、文書の質が落ちてしまうことがある。

このような状況に RedPen[1] を執筆環境に導入することで、文書の質が向上することが期待されている。RedPen は技術文書をターゲットにした文書自動添削ツールであり、現在もコードの追加、改変が行われている。

2. 目的

RedPen は学校や会社等の組織のルールに対応できるように設定が柔軟に行える仕様になっている。マシンを用いた文書添削を繰り返し行い、論文向けの添削システムを確立し、文書の質の向上と、作成時間の短縮を図ることを目的とする。

3. 手法

添削システムに必要な要素を以下の手法で調査する。

1. 執筆中の文書の添削に、CI（継続的インテグレーション）サーバを導入する。GitHub に提出した文書の添削を自動化し、エラー内容を瞬時に確認できるようにする。文書を提出する度にエラー内容を集める [2]。
2. 集めた添削結果から添削システムに必要な要素を考察し、RedPen のコードを追加、改変する。

4. 想定される成果物

個人、複数人プロジェクトで活用できる文書添削システムを構築する。

5. 進捗状況

矢吹研究室に所属する3年生の課題文の添削を行い、エラー（文中の誤り）数が減る推移を図1にまとめた。RedPen は文長に関する添削機能や、表現に関する機能をいくつか提供している。

5.1 文長

SentenceLength（文字数）を最大120文字に設定した。また、CommaNumber 機能は、カンマが多すぎる文の場合にエラーを出力する。

5.2 文書表現

SuccessiveWord 機能は、同一単語が一文中に複数回使われていると、エラーを出力する。InvalidWord は「1 つ」や「一つ」等の数値表現が統一されていない場合にエラーを出力する。

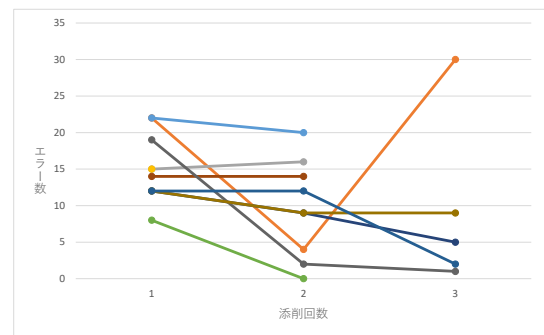


図1 添削項数の推移

添削結果から、エラー数が減っている文書は、執筆の度に文が短くなる特徴が得られた。

エラー数が減らない文書には、カンマや助詞が多いままの文書が見受けられた。また、添削システムの設定が不十分であることも要因である。

6. 今後の計画

論文に適切な文長や表現等の要素を考察し、添削機能の追加、改変を行い機能を実装する。文書作成に利用してもらう。

参考文献

- [1] TakahikoIto. Redpen1.9 ドキュメント.
http://redpen.cc/docs/latest/index_ja.html(2017.9.20 閲覧).
- [2] 伊藤敬彦. Redpen を使って技術文書を手軽に
校正しよう. [http://gihyo.jp/lifestyle/
serial/01/redpen](http://gihyo.jp/lifestyle/serial/01/redpen)(2017.9.20 閲覧).