

概要

論文において記載すべき情報が記載されていない場合が存在する。その場合、研究の内容が読者に伝わり難いという問題が発生する。本研究では、論文に記載すべき情報を「記載必要項目」と定義し、論文内で記載必要項目が欠落している論文を自動検出することで、論文の文章作成支援を行うことを目的とする。

多くの論文に出現する単語は記載必要項目である可能性が高いと考える。本研究では、全論文を出現論文数で割ることで出現率を算出し、出現率の高い単語を調査する。さらに、出現率の高い単語に類似している単語は記載必要項目の検出に役立つ単語であると考え、類似している単語を調査する。出現率の高い単語とその単語に類似している単語を参考に人手で検討し、記載必要項目とその項目の検出に役立つ単語を決定する。

決定した記載必要項目の検出に役立つ単語が一つもない論文を記載必要項目が欠落している論文であると判別する。このようなルールで判別し、ルールベースを利用して記載必要項目が欠落している論文を自動検出する。

本研究で決定した記載必要項目のうち「重要性」「新規性」以外の項目を検出することができた。さらに「比較」「問題」「目的」はF値が0.6から0.7で検出でき、「例」はF値が0.86で検出できた。それぞれの結果はベースラインよりもF値が高かった。

目次

第1章	はじめに	1
第2章	関連研究	3
2.1	論文の閲覧支援やサーベイ支援の研究	3
2.2	ルールベース手法を用いている研究	4
2.3	欠落個所の指摘を目的とした研究	4
2.4	文章作成支援の研究	5
2.5	論文作成支援の研究	5
第3章	研究の流れ	7
第4章	記載必要項目と検出に役立つ単語の決定	8
4.1	問題設定	8
4.2	決定手順	8
4.2.1	頻度調査	8
4.2.2	意味ソート	9
4.2.3	人手での検討	9
4.3	データ	9
4.4	決定結果	10
4.4.1	頻度調査の結果	10
4.4.2	意味ソートの結果	12
4.4.3	記載必要項目と検出に役立つ単語の決定結果	13
第5章	文章作成支援	14
5.1	問題設定	14
5.2	記載必要項目が欠落している論文の検出方法	14
5.3	データ	14

5.4	評価方法	15
5.4.1	評価の手順	15
5.4.2	F 値	15
5.4.3	人手での判別基準	16
5.4.4	カッパ係数による人手での判別基準の評価	16
5.5	実験結果	17
5.5.1	人手での判別基準とその基準の評価	17
5.5.2	文章作成支援の実験結果	18
5.6	考察	19
5.6.1	文章作成支援の実験考察	19
5.6.2	「重要」「異なる」などについての考察	19
5.7	具体例	21
5.7.1	「目的」についての具体例	21
5.7.2	「問題点」についての具体例	24
5.7.3	「比較」についての具体例	26
5.7.4	「例」についての具体例	26
第6章 おわりに		27

表 目 次

4.1	論文内で出現率が高い上位 100 単語	11
4.2	決定した記載必要項目と検出に役立つ単語	13
5.1	カッパ係数評価指標	16
5.2	各記載必要項目の判別基準	17
5.3	「比較」について文章作成支援の評価結果	18
5.4	「問題点」について文章作成支援の評価結果	18
5.5	「目的」などについて文章作成支援の評価結果	18
5.6	「例」について文章作成支援の評価結果	18

目 次

1.1	本研究の目的	1
4.1	意味ソートの結果(一部)	12
5.1	「目的」について正解であると人手で判別した例	21
5.2	「目的」について正解でないと人手で判別した例	23
5.3	「問題点」について正解であると人手で判別した例	24
5.4	「問題点」について正解でないと人手で判別した例	25
5.5	「比較」について正解でないと人手で判別した例	26

第1章 はじめに

論文において研究成果や研究の必要性・有効性などの記載すべき情報が記載されていない場合が存在する。その場合、研究の内容が読者に伝わり難いという問題が発生する。

本研究では、論文に記載すべき情報を「記載必要項目」と定義し、論文内で記載必要項目が欠落しているか否かを自動検出することで、論文の文章作成支援を行うことを目的とする。

文章作成支援の研究は既に先行研究が多数ある。「冗長な表現の改善」の研究としては文献 [6] が、「誤字の修正・適切な語の選択」の研究としては文献 [10][11][12] が、「語順の修正・語と語の係り受けの誤りおよび複雑性の修正」の研究としては文献 [10][13][14] がある。数多くの文章作成支援の研究があるが、論文の記載必要項目を利用して論文の文章作成支援を行う研究はないため本研究で扱うこととした。

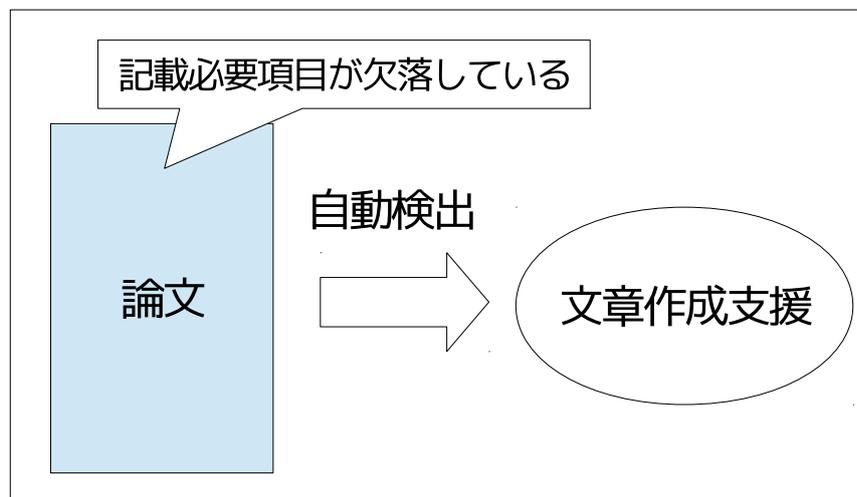


図 1.1: 本研究の目的

本論文の主な主張を以下に整理する.

- 論文内で記載必要項目が欠落しているか否かを自動判別し, 文章作成支援を行う.
- 論文の記載必要項目を利用して論文の文章作成支援を行う先行研究はないという新規性がある.
- 記載必要項目の検出に役立つ単語を決定し, それらの単語の有無によって記載必要項目の欠落を自動判別できる.
- 本研究では6つの記載必要項目を決定し, それらの項目が欠落している論文を検出した. その結果, 6つの項目のうち4つの項目が欠落している論文を0.6から0.8という比較的高いF値で検出できた.
- 全ての論文を文章作成支援の対象として検出した場合と比較した結果, 本研究の提案手法のほうが検出精度が高いことが確認された.

本論文の構成は以下の通りである. 第2章では, 本研究の関連研究について述べている. 第3章では, 本研究の全体の流れについて述べている. 第4章では, 記載必要項目とその項目の検出に役立つ単語について述べている. 第5章では, 文章作成支援について述べている.

第2章 関連研究

関連研究として以下の研究が挙げられる。

2.1 論文の閲覧支援やサーベイ支援の研究

村田ら [1] の研究では、論文の閲覧支援を目的として、論文アブストラクトから重要な情報を抽出し、その情報を表の形で可視化している。さらに重要な情報を抽出するための教師データを作成し、それをを用いて教師あり機械学習により重要な情報を抽出している。

檜本ら [2] の研究では、論文のサーベイを効率良く行うことを目的として、論文から表、図、脚注、参考文献の4つの論文構成要素をルール及び機械学習 (SVM) を用いて抽出を行っている。

難波ら [3] の研究では、論文の閲覧支援を目的として、研究分野の動向を概観するのに必要不可欠である研究動向情報を論文から自動的に抽出し、可視化する研究を行っている。特定の分野の論文を収集し、それらの論文から可視化に必要な情報を抽出している。抽出した情報を表にし、可視化を行う。

村田ら [1] の研究と檜本ら [2] の研究と難波ら [3] の研究は、論文データの可視化を行うことで、論文のサーベイに役立てることを目的としている。それらに比べて本研究では、論文データの可視化を目的にしているのではなく、論文内に記載必要項目の欠落している文章が存在しているかを確認し、論文作成の際の文章作成支援を行うことが目的である。

2.2 ルールベース手法を用いている研究

Martin ら [4] の研究では, ルールベースを用いた情報抽出システムを利用して半自動的に構造化データを取得するシステムのモデルを提案している. そのモデルにはテキスト文書からの情報抽出およびデータ取得のためのシステム (TEXT MARKER) を使用している.

Martin ら [4] の研究は, 本研究の提案手法でも利用しているルールベースを用いて情報抽出を行っており, 文章作成支援を行っていない. しかし, 本研究ではルールベースを用いて論文の文章作成支援を行っている.

2.3 欠落個所の指摘を目的とした研究

灘本ら [5] の研究では, SNS やブログのようなコミュニティ型コンテンツ内で議論が集中し, 視点が狭くなる可能性があるとして述べている. これにより議論におけるテーマを多面的に捉えられなくなる危険性を指摘している. 見落とされた視点をコンテンツホールと呼び, SNS やブログにおけるコミュニティ内の議論の履歴からコンテンツホールを抽出し, ユーザに提示している.

灘本ら [5] の研究と本研究を比較すると, 灘本ら [5] のユーザに見落された点を指摘するという目的と本研究の論文著者に欠落した文の存在を知らせ文章作成支援を行うという目的は類似していることがわかる. しかし, 灘本ら [5] はコミュニティ型コンテンツ内での研究であり, 本研究は論文内での研究であるという異なる点が存在している.

2.4 文章作成支援の研究

都藤ら [6] の研究では, 冗長な文章を自動検出することを目的としている. 冗長な文章の自動検出の提案手法として, 冗長度と機械学習を利用して実験を行っている. 冗長度を索性にすることにより機械学習だけを利用して検出するより精度が向上することが確認されている.

都藤ら [6] の研究と本研究を比較すると, どちらの研究も不適切な文章の検出を目的としているが, 都藤ら [6] は機械学習を用いて冗長な文章の検出を行っている. 一方, 本研究では冗長な文章ではなく記載必要項目が欠落している文章の検出が目的であり, 検出対象が異なっていることがわかる.

都藤ら [6] 以外にも数多くの研究で文章作成支援が行われている [10][11][12][13][14]. しかし, その数多くの文章作成支援の研究の中でも, 論文の記載必要項目を利用して文章作成支援を行っている研究はない.

2.5 論文作成支援の研究

Ptaszynski ら [7] の研究では科学論文の作成支援システムの提案をしている. システムが科学論文を執筆するのに必要な精度やスコアなどの実験データを自動的に計算することで科学論文の作成支援を行っている.

Liu ら [8] の研究では学生が自分の論文を改訂するための支援の形として自動質問生成 (AOG) ツールの開発を行っている. 「この論文の著者は先行研究の考えに反対ですか?」「あなたはこの論文を読んで著者が反対であるという証拠を見つけることができますか?」などの質問を自動作成する. その質問を学生が参考にすることで論文の改訂支援を行っている. 論文を読んで自動質問生成ツールでは, 引用から取得した構文情報や意味情報を基に質問を生成している.

Ptaszynski ら [7] の研究や Liu ら [8] の研究は論文を対象とした作成支援を目的に行っている. 論文の作成支援のなかでも, Ptaszynski ら [7] は実験データの自動計算システムにより論文の実験データ作成支援を行っており, Liu ら [8] は自動質問生成ツールにより論文作成支援を行っている. それらの研究に対し, 本研究は, 論文の記載必要項目の欠落しているか否かを自動判別することで論文の文章作成支援を行っている. Ptaszynski ら [7] の研究と本研究を比較すると, 論文の実験データの作成支援と論文の文章の作成支援といった作成支援の対象が違っている. Liu ら [8] の研究と本研究を比較すると, 自

動質問生成ツールを用いた論文作成支援と論文の記載必要項目を利用した論文作成支援といった論文作成支援へのアプローチ方法が違っている.

第3章 研究の流れ

本研究の全体の流れを以下に示す.

1. 記載必要項目とその項目の検出に役立つ単語を決定する.
2. 検出に役立つ単語とルールベースを利用して記載必要項目が欠落している論文を自動検出する.
3. 2の検出結果が文章作成支援に役立っているか否かを人手で判別し, 評価する.

本研究では, 初めに, 記載必要項目とその項目の検出に役立つ単語を決定する. 多くの論文に出現する単語は記載必要項目の検出に役立つ単語である可能性が高いと考える. そこで, 記載必要項目の検出に役立つ単語を決定するために, 論文内での単語の頻度調査を行う. 論文に出現する単語ごとの出現率を算出し, 出現率の高い単語を調べる. 出現率の高い単語から記載必要項目の検出に役立つ単語であるかを人手で検討し決定する.

さらに, 記載必要項目の検出に役立つ単語を増やすために, 記載必要項目の検出に役立つ単語に類似している単語を調査する. 類似している単語の調査には, 意味ソート [9] を利用して調査を行う.

次に, 記載必要項目の検出に役立つ単語が一つもない論文を記載必要項目が欠落している論文であると判別し, そういった論文を自動検出する. 本研究ではルールベースを用いて検出を行う. 最後に, ルールベースでの検出結果から文章作成支援に役立っているか否かを人手で判別し, 評価を行う. 本研究では記載必要項目を補う必要のある論文を文章作成支援に役立っている論文と判別し, F 値を算出して評価を行う.

第4章 記載必要項目と検出に役立つ単語の決定

4.1 問題設定

記載必要項目と記載必要項目の検出に役立つ単語を決定する。検出に役立つ単語が一つもない場合は記載必要項目が欠落している論文であると判別でき、記載必要項目が欠落している論文の検出に役立つと考える。

4.2 決定手順

記載必要項目とその項目の検出に役立つ単語の決定は以下の手順で行う。

1. 多くの論文に出現する単語を調査する (4.2.1 節)
2. 1の結果から意味ソート [9] を利用して意味の類似している単語をまとめて表示させる (4.2.2 節)
3. 2の結果を人手で検討して、記載必要項目とその項目の検出に役立つ単語を決定する (4.2.3 節)

手順の詳細を以下に示す。

4.2.1 頻度調査

多くの論文に出現する単語は論文の記載必要項目である傾向である可能性が高いと考えられる。単語の出現した論文数を全論文数で割ることで単語の出現率を算出する。例えば、全論文 300 件中 250 件の論文に単語「Z」が存在している場合、単語「Z」の出現率は $250/300$ となる。

4.2.2 意味ソート

記載必要項目の検出に役立つ単語に類似している単語も記載必要項目がの検出に役立つ単語である可能性があると考える。例えば「手法」という単語が記載必要項目の検出に役立つ単語である場合、その単語に類似している「方式」などの単語も記載必要項目の検出に役立つ単語である可能性がある。本研究では、記載必要項目の検出に役立つ単語に類似している単語を調査するために意味ソート [9] を利用する。意味ソート [9] は意味の類似している単語をまとめて表示させることができる。これにより出現率の低い単語も参考にでき、より詳細な記載必要項目とその項目の検出に役立つ単語が決定できると考える。

4.2.3 人手での検討

4.2.2 節の結果を参考にして、人手で記載必要項目とその項目の検出に役立つ単語を検討し決定する。

4.3 データ

記載必要項目の決定を行う際に使用した実験データは、1994年から2013年の言語処理学会論文誌 (393件) である。

4.4 決定結果

4.4.1 頻度調査の結果

本研究では, 4.2.1 節で挙げられた方法で頻度調査を行った. 全論文数は 393 件あり, その論文中出现する単語の総数は 19,234 単語であった. その内の出現率の高い上位 100 単語までの結果をまとめて表 4.1 に示す.

表 4.1: 論文内で出現率が高い上位 100 単語

単語	出現率	単語	出現率	単語	出現率
示す	1.000	従事	0.901	工学部	0.831
研究	1.000	実験	0.898	作成	0.831
場合	1.000	提案	0.898	有効	0.828
用いる	0.997	工学	0.898	分類	0.828
考える	0.997	基づく	0.898	今後	0.828
結果	0.997	比較	0.895	呼ぶ	0.826
言語	0.997	博士	0.895	科学	0.826
処理	0.997	同様	0.895	一方	0.826
情報	0.997	大きい	0.895	重要	0.811
必要	0.994	学会	0.895	定義	0.808
に対して	0.994	一般	0.895	論文	0.805
述べる	0.977	に対する	0.895	行う	0.802
得る	0.971	情報処理	0.895	抽出	0.797
可能	0.971	存在	0.890	次に	0.797
以下	0.965	与える	0.881	名詞	0.791
含む	0.962	に関する	0.881	出現	0.791
自然	0.959	表す	0.878	同年	0.788
対象	0.956	卒業	0.875	程度	0.788
関係	0.953	解析	0.875	会員	0.788
異なる	0.951	修了	0.875	条件	0.785
以上	0.951	複数	0.869	求める	0.782
方法	0.948	全体	0.869	適用	0.779
現在	0.948	手法	0.866	単語	0.779
大学院	0.942	日本語	0.863	構造	0.779
課程	0.942	例えば	0.858	目的	0.773
高い	0.936	対応	0.858	特に	0.770
利用	0.933	考慮	0.858	少ない	0.767
表現	0.933	見る	0.855	精度	0.765
多い	0.930	多く	0.852		
意味	0.930	計算	0.852		
持つ	0.922	以外	0.852		
問題	0.916	実際	0.849		
同じ	0.913	構成	0.843		
評価	0.907	種類	0.840		
システム	0.904	説明	0.834		
部分	0.901	データ	0.834		

4.4.2 意味ソートの結果

論文での出現率の高い上位 500 単語を意味ソート [9] を使ってソートし, 意味の類似している単語をまとめて表示させた. 意味ソート [9] の結果の一部を図 4.1 に示す.

(数量)

{ 量 } 出力 入力 総数 数値 頻度 番号 関数
{ 数 } 多く 多数 多い 大量 十分 少ない
{ 値・額 } 長い 短い 尺度 高い 低い 深い 近い 距離

(関係)

{ 因果 } 条件 有効 前提 原因 要因 結果 効果 影響
{ 理由・... } 理由 目的 実用
{ 異同 } 相対 相互 応じる 対応 相当 比べる 比較
{ 相対 } 同じ 似る 同様 類似 異なる 含む
含める 違い 区別
{ 有無 } 存在 既存
{ 出現 } 現れる 出現 実現 提案 提示 示す 出す

図 4.1: 意味ソートの結果 (一部)

4.4.3 記載必要項目と検出に役立つ単語の決定結果

4.4.1 節から研究の重要性・有用性を表す可能性がある「重要」や新規性を表す可能性がある「異なる」などの出現率が高いことがわかった。研究の重要性や新規性が存在しない論文は研究の内容が理解し難くなる可能性が高いので、「重要」「異なる」なども記載必要項目である可能性が高いと考えられる。

また、「問題」「目的」などが存在しない論文は何が問題で何を目的にしているかを理解できなくなる可能性が高いと考えられる。さらに、「例えば」などが存在しない論文でも理解しやすい具体例などが無い可能性があり、論文の内容の理解が難しくなる可能性があると考えられる。従って、「目的」「問題」「例えば」なども記載必要項目である可能性が高いと考えられる。

以上で記載必要項目である可能性が高いとされた単語と 4.2.2 節で述べた意味ソート [9] の結果を比べ、その単語に類似した単語を手で検討し、記載必要項目とその項目の検出に役立つ単語を決定した。結果を表 4.2 に示す。検出に役立つ単語が一つもない論文を記載必要項目が欠落している論文として自動検出でき、論文の文章作成支援に役立つ。

表 4.2: 決定した記載必要項目と検出に役立つ単語

項目名	検出に役立つ単語	記載必要項目の説明
重要性	必要 重要	研究の重要性
新規性	異なる 違う 違い	研究の新規性
比較	比較 比べる	先行研究との比較 精度の比較実験
問題点	問題	先行研究の問題点 研究の背景
目的	目的 目標 目指す	研究の目的
例	例えば 例 具体	具体例

第5章 文章作成支援

5.1 問題設定

表 4.2 の結果を基にルールベースで記載必要項目が欠落している論文を検出する。記載必要項目を補う必要がある論文が自動検出できれば、論文の文章作成支援に役立つとする。

5.2 記載必要項目が欠落している論文の検出方法

表 4.2 の検出に役立つ単語をルールとしてルールベースを利用し論文の検出を行う。表 4.2 の検出に役立つ単語が一つも出現していない論文を記載必要項目が欠落している論文であると判別し、検出する。

5.3 データ

文章作成支援の実験を行う際に、2011 年度の年次大会論文 (266 件) を学習データとして使用し、2012 年度の年次大会論文 (305 件) をテストデータとして使用した。また、学習データを 5.4.3 節の判別基準の設定に利用し、テストデータを評価に利用する。

5.4 評価方法

5.4.1 評価の手順

文章作成支援の評価は、以下の手順で行う。

1. 5.2 節で記載必要項目が欠落している論文であると判別され、ルールベースで検出されたものが文章作成支援に役立っている（その記載必要項目を補う必要がある）かを人手で判別する。
2. 1 の結果から提案手法であるルールベースの F 値を算出する。
3. 全ての論文をシステムの出力にした場合（全ての論文を記載必要項目が欠落している論文であると判別し、検出した場合）をベースラインとして、ベースラインの F 値を算出する。
4. 2 と 3 で算出した結果から提案手法であるルールベースとベースラインの F 値を比較し評価を行う。

5.4.2 F 値

本研究では文章作成支援の精度を再現率 (*recall*), 適合率 (*precision*), F 値 (*F-measure*) で評価する。再現率と適合率は以下の式で算出される。

$$\text{再現率} = \frac{\text{システムの正解数}}{\text{テストデータ中の正解数}} \quad (5.1)$$

$$\text{適合率} = \frac{\text{システムの正解数}}{\text{システムの出力数}} \quad (5.2)$$

本研究では文章作成支援に役立っている論文を正解として (5.1) と (5.2) を算出した。また、(5.1) と (5.2) の値の調和平均 (5.3) を求めることで F 値を算出できる。

$$F \text{ 値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}} \quad (5.3)$$

5.4.3 人手での判別基準

それぞれの項目の人手での判別でばらつきが生じないように項目ごとに基準を設定する。何故なら、曖昧な判別を行い、判別結果がばらつくとそれだけ再現率・適合率・F値が正確でなくなるからである。より正確な再現率・適合率・F値を求めるために2011年度の年次大会論文(266件)の学習データを使用し、提案手法で処理した結果を人手で評価する。そこでの評価を参考にして、人手での判別基準を設定する。

5.4.4 カッパ係数による人手での判別基準の評価

本研究では5.4.3節の人手での判断基準をカッパ係数を利用して評価する。カッパ係数とは、ある現象を観察者が観察した場合の結果がどの程度一致しているかを表す統計量である。カッパ係数は以下の数式で求められる。

$$= \frac{Pr(a) - Pr(e)}{1 - Pr(e)} \quad (5.4)$$

はカッパ係数, $Pr(a)$ は見掛け上の一致率で, $Pr(e)$ は偶然の一致率を示す。本研究では2人の観察者の一致率でカッパ係数を算出する。

表 5.1: カッパ係数評価指標

値	評価
0.00-0.20	わずかに一致
0.21-0.40	一致
0.41-0.60	中程度の一致
0.61-0.80	かなりの一致
0.80-1.00	ほぼ完全な一致

5.5 実験結果

5.5.1 人手での判別基準とその基準の評価

2011年度の年次大会論文(266件)の学習データ内の論文を人手で判別し、判別基準を設定する人物一人(人物Aとする)と人物Aとは別の人物一人(人物Bとする)が人物Aの設定した判別基準を参考に人手で判別を行った。人物Aと人物Bの合計二人の判別結果の一致率からカッパ係数を算出した。また、人物Aが学習データにおいて人手判別したのからランダムに文章作成支援に役立っているものと役立っていないものをそれぞれ12件ずつ取り出した。その合計24件の論文を人物Bが判別基準を参考にして人手判別した。本研究で設定した判別基準でのカッパ係数は0.67であった。

記載必要項目の判別基準を表5.2に示す。また、表5.2では、文章作成支援に役立つと判別したものは○、文章作成支援に役立たないと判別したものは×としている。

表 5.2: 各記載必要項目の判別基準

項目名	判別	基準
比較		先行研究との比較や実験結果の比較が行われていない論文
	×	「...という手法が先行研究で挙げられている。それに対して我々は...」などの文で先行研究との比較が行われている論文
問題点		世の中の問題(研究背景)や先行研究の問題点についての説明が不足している(説明が不明瞭な)論文
	×	研究背景や先行研究の問題点について詳しく説明されている(説明が明瞭な)論文
目的		詳しく読まないとな何を目的に研究を行っているのかが理解できない論文(一読するだけでは研究の目的が理解できない論文)
	×	詳しく読まなくても、一読すれば研究の目的が何であるかが理解できる論文
例		具体的な例がない論文(図中も含む)
	×	具体的な例がある論文(図中も含む)

5.5.2 文章作成支援の実験結果

2012年度の年次大会論文(305件)をテストデータとして実験を行った。結果を表5.3から表5.6に示す。

表 5.3: 「比較」について文章作成支援の評価結果

手法	再現率	適合率	F 値
ベースライン	1.00 (59/59)	0.19 (59/305)	0.32
ルールベース	0.58 (34/59)	0.60 (34/57)	0.59

表 5.4: 「問題点」について文章作成支援の評価結果

手法	再現率	適合率	F 値
ベースライン	1.00 (114/114)	0.37 (114/305)	0.54
ルールベース	0.61 (70/114)	0.80 (70/87)	0.69

表 5.5: 「目的」などについて文章作成支援の評価結果

手法	再現率	適合率	F 値
ベースライン	1.00 (94/94)	0.31 (94/305)	0.47
ルールベース	0.53 (50/94)	0.60 (50/84)	0.56

表 5.6: 「例」について文章作成支援の評価結果

手法	再現率	適合率	F 値
ベースライン	1.00 (9/9)	0.03 (9/305)	0.06
ルールベース	1.00 (9/9)	0.75 (9/12)	0.86

「重要」「異なる」などを学習データでルールベースで処理し、人手での判別基準を検討したが、ルールベースの性能が低いことがわかった。これは「重要」「異なる」などの人手の判別基準の設定が困難であったことが原因であると考えられる。これについては5.6.2節で考察する。

5.6 考察

5.6.1 文章作成支援の実験考察

表 5.3 から表 5.6 を見るとそれぞれベースラインよりもルールベースのほうが F 値が高いことがわかる。また、表 5.6 の「例」についての結果を見ると、ベースラインと提案手法であるルールベースの F 値の差は 0.80 もあることもわかる。さらに表 5.3 の「比較」についての結果もベースラインと提案手法であるルールベースの差が約 0.3 あることもわかる。しかし、表 5.4 と表 5.5 の「問題点」や「目的」についての結果は、ベースラインと提案手法であるルールベースの差が約 0.10 ほどであり、これは他の表 5.3 や表 5.6 の結果と比べると、ベースラインとの性能差があまりないように見える。これは「問題点」や「目的」などの項目のルールを詳細化することでベースラインとの性能差が生まれると考えられる。今後は本稿の提案手法だけでなく、更に新しい手法を取り入れ、記載必要項目を決定することでルールが詳細化され则认为。また、本稿では情報抽出の基本的な手法しか使っておらず、論文の記載必要項目という特性をほとんど利用できていなかった。したがって、更なる手法の改良が今後の課題であるとする。

5.6.2 「重要」「異なる」などについての考察

今回の実験で、「重要」「異なる」などが欠落しているだけでは記載必要項目が欠落しているという判別はできないということがわかった。「重要」「異なる」などが欠落している論文と欠落していない論文を見比べたところ、「重要」「異なる」などの有無に関係なく、研究の重要性や新規性が述べられている論文が散見された。さらに人手で論文を詳しく見た結果、記載必要項目「重要性」「新規性」の欠落している論文の検出漏れが多く存在していることがわかった。例えば「これらの重要な表現を抽出する。」「同じ形態素が異なる意味で使われるという変化は捉えられない。」などのような記載必要項目「重要性」「新規性」を表さない文章が誤って記載必要項目「重要性」「新規性」を表す文章であると自動判別された。その結果、実際は記載必要項目「重要性」「新規性」が欠落している論文のはずが欠落していない論文であるとして検出されなかった。この結果より、4.4.3 節で述べた「重要」「異なる」などが研究の重要性や新規性を表しているという考えが間違っている可能性があると考えられる。しかし、重要性や新規性などは研究にとって不可欠であり、それらの項目が存在しないと研究の詳細が理解しづらくなるのは容易に想像できるだろう。重要性や新規性の欠落は今回の手法で

検出するのは困難であると考えられるが、検出すべき項目であり、今後扱っていく課題である。

5.7 具体例

記載必要項目が欠落している論文として検出された結果のうち人手で正解であると判別した論文と人手で正解でないと判別した論文の一部を具体例として以下に示す。また、記載必要項目を補う必要がある論文を正解としている。

5.7.1 「目的」についての具体例

記載必要項目「目的」が欠落している論文として検出された結果のうち人手で正解であると判別した論文の一部を図 5.1 に示す。

... 近年、Web やブログの普及により大規模な日本語コーパスを手に入れることができるようになってきた。言語モデルをはじめとする統計的手法では、コーパスが大規模であるほどパラメータ推定の信頼性が上がったり、自由度の高いモデルを利用できたりするため、大規模コーパスの恩恵は大きい。

しかしながら、そのような大規模コーパスによる言語モデルには以下に挙げる 2 つの問題点がある。...

(中略)

... これらの問題はデータ量と性能とのトレードオフであり、完全に解決する方法は見つかっていない。そのため、言語モデルの応用を考える上でそのトレードオフについて知ることは重要である。

また、大規模なコーパスを扱える環境は現在では限られており、特に日本語のコーパスにおける研究は非常に少ない。そのため、大規模な日本語コーパスで言語モデルの振る舞いを定量的に確認することには意義がある。

そこで、我々は大規模な日本語コーパスの 1 つとしてブログコーパスを選択し、言語モデルの構築を行った。ブログコーパスを選択した理由は、入手や扱いが容易であることと、一般の人々によって書かれていることの 2 つがある。後者の性質によって、一般の人々が使う仮名漢字変換や音声認識などの入力システムにおいては Web 全体を使うよりも有利に働く可能性がある。...

図 5.1: 「目的」について正解であると人手で判別した例

この例示した論文において、読み手が想像する研究の目的として以下のものが挙げられる。

- 2つの問題点を解決することを目的にしている
- 完全に解決する方法が見つからないので、完全に解決する方法についての手法を提案することを目的にしている
- 問題点がトレードオフの関係であるということを知ることが目的にしている
- 大規模なコーパスを扱える環境を得ることを目的にしている
- 日本語のコーパスにおける研究を行うことを目的にしている
- 大規模な日本語コーパスで言語モデルの振る舞いを定量的に確認することを目的にしている

しかし、実際この例示した研究では、問題点がトレードオフの関係であることをを知ることが目的としていられる。また、大規模な日本語コーパスで言語モデルの振る舞いを定量的に確認することも目的としていられる。さらに、これら2つの目的には、この例で挙げられた2つの問題点を完全に解決する方法の提案に役立てるといふ目的も存在していると考えられる。

このような例の場合、本来の研究の目的が単語「目的」などを使用して明記しておらず、本来の研究の目的でない目的が多数想定されてしまう。その結果、どの目的がその研究における主目的なのか論文読者が理解し難くなる可能性が考えられる。そういったことが起らないために、単語「目的」などの表現を用いて研究の目的を明記する必要がある（記載必要項目を補う必要がある）と考え、文章作成支援に役立っていると判別した。

記載必要項目「目的」が欠落している論文として検出された結果のうち人手で正解でないと判別した論文の一部を図 5.2 に示す。

... 現在、機械翻訳の分野において、対訳文対から自動的に翻訳規則を生成し翻訳を行う統計翻訳が注目され、研究が盛んに行われている。統計翻訳では、イタリア語-英語など文法構造が類似する言語対において翻訳精度が高くなり、日本語-英語などの文法構造の異なる言語対においては翻訳精度が低くなる傾向がある。別の翻訳手法にパターン翻訳がある。パターン翻訳では文パターン辞書と単語辞書を用いて翻訳を行う。文パターンが有する大局的な文法情報を用いることで、翻訳文全体の構造を保持した翻訳精度の高い翻訳文を生成出来る利点がある。しかし、従来、文パターン辞書の作成は人手で行うため、開発にコストがかかる欠点がある。

そこで本研究では、文パターン辞書を対訳文対から自動的に作成する手法を検討する。文パターン辞書の自動作成により、開発にかかるコストの削減が可能となる。 ...

図 5.2: 「目的」について正解でないと人手で判別した例

この例示した論文において、読み手が想像する研究の目的として以下のものが挙げられる。

- 文パターン辞書の開発にコストがかかるという問題点を解消することを目的としている
- 自動生成された文パターン辞書と人手作成された文パターン辞書での翻訳精度調査を目的としている
- パターン翻訳の精度向上を目的としている

この例の研究では、文パターン辞書の開発にコストがかかるという問題点を解消することを目的としていると考えられる。この論文内に「X という問題点がある。そこで本研究では Y を行う」という文章が存在していることがわかる。このような文章は前に記述された問題点の解消を行うことを目的として研究を行っていることが容易に理解できる文章が使われている論文であると考えられる。

5.7.2 「問題点」についての具体例

記載必要項目「問題点」が欠落している論文として検出された結果のうち人手で正解であると判別した論文の一部を図 5.3 に示す。

... 近年, 評判分析の対象として, Twitter が注目されている。 は, 「:)」や「:(」などの emoticon を利用した訓練データの獲得と機械学習による極性判定 (肯定・否定判定) の手法について提案した。 × × は, Twitter 特有の機能である「リプライ」などに着目し, 極性判定する手法について提案している。 は, Twitter によく見られる「Cooool」のような繰り返し表現に着目し, その繰り返し表現の正規化や感情との関連性などについて検証している。

ここで, 我々は Twitter 上で用いられる文体に着目する。 Twitter 上にはさまざまな文体が存在し, その特性が異なる。例えば, 話し言葉に近い口語体では, 特徴的な文末表現や記号などによる顔文字の多用によってその感情を表現する傾向があるのに対し, 書き言葉的な文語体では, 言語表現そのものによってその感情が表されることが多い。このような文体の違いを考慮することは, さまざまな場面に有効であると考えられる。 ...

図 5.3: 「問題点」について正解であると人手で判別した例

記載必要項目「問題点」というのは先行研究の問題点や研究の背景を差している。この例では, 先行研究について述べられており, さらに研究の有効性も記述されている。しかし, 先行研究の手法の概要のみを記述しており, 先行研究で生じた問題についての記述が存在していない。仮に先行研究で問題が生じていなかったと考えても, その場合は先行研究の手法の概要と先行研究との明確な違いを記述する必要があると考える。

研究の背景として「近年, 評判分析の対象として, Twitter が注目されている。」とあるが, この例の文章であると何故 Twitter が評判分析の対象として注目されているかが理解し難いと思う。

次に、記載必要項目「問題点」が欠落している論文として検出された結果のうち人手で正解でないと判別した論文の一部を図 5.4 に示す。

... 近年, Facebook や Twitter などのマイクロブログが急速に普及し, ユーザによるマイクロブログを用いた情報発信が活発化している. 特に Twitter は, 140 文字という制限によりユーザの情報発信への敷居が大きく下がっており, 2011 年 3 月 11 日に発生した東日本大震災においては, リアルタイムに情報を伝える重要な情報インフラの 1 つとして活用された. しかし, 安否情報などの重要な情報の共有・伝搬が行われた一方で, 多くの流言も拡散された. 流言は適切な情報共有を阻害し, 特に災害発生時には, 流言が救命のための機会損失を生む場合もあるため, 流言の広がりにくい環境を作る必要がある. ...

(中略)

... そこで本研究では, 流言拡散を防ぐための仕組みとして, 流言情報クラウドを提案する. 流言情報クラウドは, リアルタイムに流言情報を蓄積し, その情報を提供することにより, 流言拡散を防止する. ...

図 5.4: 「問題点」について正解でないと人手で判別した例

この例の研究では, マイクロブログを用いた情報発信が活発化して様々な利点が生まれたが, その反面として多くの流言が発生してしまい, それが原因で適切な情報共有ができなくなるという問題点が説明されている. この論文では, 「ある事象 A があります. 事象 A により 〇〇のようなメリットが存在します. しかし, その反面で××というデメリットも存在してしまいます. 」という文章が使われていることがわかる. 研究の背景としてある事象を例として利点を挙げ, その後に欠点を記述することにより何が利点で何が欠点なのかが容易に理解できる文章が使われている論文であると考えられる.

5.7.3 「比較」についての具体例

記載必要項目「比較」が欠落している論文として検出された結果のうち人手で正解であると判別した論文は論文全体を読んでも先行研究の比較や比較実験を表す文章がなく、論文内のどこで先行研究との比較を行っているかが理解できない。また、論文全体を読む必要があるので、論文の一部を例として示すことができない。

記載必要項目「比較」が欠落している論文として検出された結果のうち人手で正解でないと判別した論文の一部を図 5.5 に示す。

... 対訳表現の抽出手法として、
らは、まず、対訳コーパスを形態素解析した後、
一定の頻度を超える自立語を辞書見出し語として抽出した。その後、拡張した Dice
係数を用いて日英の自立語間の類似度を計算し、類似度の高い対訳語ペアを抽出し
た。このように、辞書見出し語の決定と訳語抽出を逐次的に行い、高い精度での対
訳表現自動抽出に成功している。

それに対して我々は、それぞれを逐次的に行うのではなく、辞書見出し語の決定
と訳語抽出を同時に行う手法を提案する。...

図 5.5: 「比較」について正解でないと人手で判別した例

この例示した論文では「
という先行研究があり、A という手法を用いている。そ
れに対して（それに比べて）本研究では B という手法を用いている」という意味に捉
えることができ、読み手がこの論文ではここで先行研究との比較を行っていることが
容易に理解できる。

5.7.4 「例」についての具体例

記載必要項目「例」が欠落している論文として検出された結果のうち人手で正解で
あると判別した論文は論文全体を読んでも具体例がなく、理解が難しい複雑な手法な
どを読み手が理解できなくなる。また、論文全体を読む必要があるので、論文の一部を
例として示すことができない。

記載必要項目「例」が欠落している論文として検出された結果のうち人手で正解で
ないと判別した論文内には文章ではなく図として具体例を示している。

第6章 おわりに

本研究では、論文の文章作成支援を行うことを目的に、論文の記載必要項目を調査し、その結果を基にルールベースによって記載必要項目が欠落している論文を検出した。また、本研究で決定した記載必要項目のうち「重要性」「新規性」以外の項目を検出することができた。さらに「比較」「問題」「目的」はF値が0.6から0.7で検出でき、「例」はF値が0.86で検出できた。

今後の課題として、今回の実験で使用したルールベースの性能向上を考えている。そのために、更なるルールの詳細化が必要であると考えられる。それにより、本稿で検出できなかった論文の記載必要項目であろう研究の重要性や新規性を検出できるようになり、更なる論文の文章作成支援に繋がると考える。また、論文の記載必要項目という特性をうまく利用できるように手法を改良することも今後の課題である。

謝辞

本研究を進めるにあたり，終始に渡り研究の進め方や本論文の書き方など，細部に渡る御指導を頂きました，鳥取大学工学部知能情報工学科計算機工学C講座の村田真樹教授に心から御礼申し上げます．また，本研究を進めるにあたり，御指導，御助言を頂きました，村上仁一准教授，徳久雅人講師に心から御礼申し上げます．その他様々な場面で御助言を頂きました計算機工学C講座研究室の皆様方に感謝の意を表します．

参考文献

- [1] 村田真樹, Stijn De Saeger, 橋本力, 風間淳一, 山田一郎, 黒田航, 馬青, 相澤彰子, 鳥澤健太郎: “論文データからの重要情報の抽出と可視化”, 2009 年度人工知能学会全国大会 (第 23 回) 論文集, pp.1-4, 2009.
- [2] 樫本達矢, 太田学, 高須淳宏: “学術論文からの構成要素抽出の一手法”, 第 12 回日本データベース学会年次大会, C5-2, 2014.
- [3] 難波英嗣, 谷口裕子: “学術論文データベースからの研究動向情報の抽出と可視化”, 言語処理学会第 12 回年次大会発表論文集, pp.35-38, 2006.
- [4] Atzmueller Martin, Kluegl Peter, Puppe Frank: “Rule-Based Information Extraction for Structured Data Acquisition using TextMarker”, LWA, pp.1-7, 2008.
- [5] 灘本明代, 阿辺川武, 荒巻英治, 村上陽平: “コミュニティ型のコンテンツホール抽出手法の提案”, 日本データベース学会 Letters, 6 巻, 2 号, pp.29-32, 2007.
- [6] 都藤俊輔, 村田真樹, 徳久雅人, 馬青: “機械学習と冗長度を用いた冗長な文章の検出”, 言語処理学会第 20 回年次大会発表論文集, pp.939-942, 2014.
- [7] Michal Ptaszynski, Fumito Masui: “SPASS: A Scientific Paper Writing Support System”, ICIEIS2014, pp.1-10, 2014.
- [8] Ming Liu, Rafael A. Calvo: “An Automatic Question Generation Tool for support of Sourcing and Integration in Students Essays”, Australasian Document Computing Symposium (ADCS), pp.45-54, 2009.
- [9] 村田真樹, 神崎享子, 内元清貴, 馬青, 井佐原均: “意味ソート msort -意味的並べかえ手法による辞書の構築例とタグつきコーパスの作成例と情報提示システム例-”, 自然言語処理, Vol.7, No.1, pp.51-66, 2000.

- [10] 菅沼明, 牛島和夫: “テキスト処理による推敲支援情報の抽出”, 人工知能学会誌, 23巻, 1巻, pp.25-32, 2008.
- [11] Masaki Murata, Hitoshi Isahara: “Automatic detection of mis-spelled Japanese expressions using a new method for automatic extraction of negative examples based on positive examples”, IEICE Transactions, VOL.E85-D, No.9, pp.1416-1424, 2002.
- [12] 村田真樹, 井佐原均: “自動言い換え技術を利用した三つの英語学習支援システム”, 情報科学技術レターズ, 3巻, pp.85-88, 2004.
- [13] 内元清貴, 村田真樹, 馬青, 関根聡, 井佐原均: “コーパスからの語順の獲得”, 言語処理学会論文誌「自然言語処理」, Vol.7, No.4, pp.163-180, 2000.
- [14] 村田真樹, 馬青, 井佐原均, 内元清貴: “日本語文と英語文における統語構造認識とマジカルナンバー 7 ± 2 ”, 言語処理学会論文誌「自然言語処理」, Vol.6, No.7, pp.61-73, 1999.