

機械学習による冠詞の用法と検索結果数に基づく 英文冠詞誤りの自動修正

尾崎 弘明^{1,a)} 新妻 弘崇^{1,b)} 太田 学^{1,c)}

概要: 英語を母語としない日本人には英語の冠詞の適切な使い分けは難しい。そこで我々は、類型化した冠詞の用法に重みを付与することで、冠詞を含むフレーズの検索結果数と冠詞の用法を組み合わせたスコアを算出し、それを基に冠詞を修正する手法を提案した。しかし、冠詞の用法と検索結果数を組み合わせることで適切に冠詞を判断することは難しく、改善の余地があった。そこで本稿では、検出した冠詞の用法と検索結果数を素性として機械学習により冠詞を判別する方法を提案する。また実験により、両手法による冠詞誤りの修正性能を比較する。

Automatic Article Error Correction by Machine Learning Using the Usage of Articles and the Number of Search Results

1. はじめに

我々日本人は、英語を第二言語として習得することが多く、また近年、母語ではなく第二言語を用いて文章を書く機会も増加している。しかし、英語を母語としない日本人にとって英語の習得は難しく、英語の用法を誤ることがしばしばある。日本語に存在しない冠詞には、日本人にあまり馴染みのない用法 [1] が多数あり、文献 [2] においても、日本人の英作文には冠詞の誤りが多いことが報告されている。そのため、このような冠詞の誤用を自動修正する様々な研究が行われている。例えば、検索エンジンを利用した冠詞誤りの検出 [3] や、名詞の可算/不可算の情報に着目した誤り検出手法 [4] が提案されている。しかしそれらの研究では、文脈や冠詞の例外的な用法が修正に反映されにくいという問題点がある。

そこで我々は、冠詞の用法と、検索エンジンを利用して得られる検索結果数を用いて英文の冠詞誤りを自動修正する手法を提案した [5]。これに用いる冠詞の用法は、理数系の学生や研究者を対象に書かれた技術英語のテキスト [1]

と、英文法に関する書籍 [6] で解説されている冠詞の用法に基づいている。手法 [5] では、対象となる名詞句にあてはまる冠詞の用法を検出し、それに重みを付与することで、用法と検索結果数を組み合わせた冠詞の種類毎のスコアを算出し、それを基に冠詞誤りの修正を行った。手法 [5] は、冠詞の用法のみでは判断できない場合に検索結果数を用いて修正する手法 [7] に比べ、評価実験において冠詞の用法と検索結果数の示す冠詞が異なっていた場合に起こる誤修正が減少した。そのため手法 [5] は手法 [7] に比べ、高い精度で冠詞誤りを修正することができた。しかし一方で、冠詞の用法が冠詞修正の判断に占める比重が減ったため、検出された冠詞の用法の中に正解の冠詞を示すものがあったとしても、それを採用せずに誤修正を行う例が増加した。よって、冠詞の用法と検索結果数の組み合わせ方に改善の余地があるという結論に至った。

そこで本稿では、検出した冠詞の用法と検索結果数を素性として、機械学習手法の人である SVM により冠詞を判別する方法を提案する。そして、評価実験により手法 [5] と冠詞誤りの修正性能を比較する。

2. 関連研究

英文誤りの自動検出・修正の研究は国内外で様々行われており、国内では 2012 年に英文の誤り検出・訂正ワークショップ (EDCW) が初めて開催された [8]。また、検索

¹ 岡山大学大学院自然科学研究科
Graduate School of Natural Science and Technology,
Okayama University

a) osaki@de.cs.okayama-u.ac.jp

b) niitsuma@suri.cs.okayama-u.ac.jp

c) ohta@de.cs.okayama-u.ac.jp

エンジンを用いた英文作成支援 [9] や用例文検索 [10] の研究なども行われている。このように英語学習者等の英作文を支援する研究は昨今関心が高いといえる。我々の研究室でも過去に [11], [12], [13], [14] のような検索エンジンを用いた英文修正に関する研究を行った。有富ら [11], 久保田ら [12] は前置詞誤り, 谷本ら [13], [14] は動詞やコロケーション誤りを対象に, 検索エンジンを用いて誤りを検出して修正する手法を提案した。

本稿で取り上げる冠詞誤りの検出と修正についても, 主としてコーパスに基づく誤り修正と検索エンジンを用いるものがあり, 以下でこれらの研究について簡単に説明する。

2.1 検索エンジンによる冠詞誤りの修正

冠詞誤りを検出して修正する手法の一つに検索エンジンを利用した手法 [3], [9] がある。これらは検索エンジンの検索結果数を比較して, 冠詞を修正する。平野らの手法 [3] では, 対象となる名詞句と前後の単語から, 検索フレーズを単数形と複数形両方の場合で生成する。その結果, 単数形と複数形の片方だけの検索結果数を用いる場合よりも, 精度が向上すること確認されている。また 2.2 節で説明するコーパスベースの手法と比較した結果, 検索エンジンを用いた方が良い結果となることが報告されている。しかし, 検索エンジンによる修正では, 使用頻度の高い語と一緒に用いられる冠詞の修正では精度が高いが, 使用頻度の低い語では精度が下がってしまうという問題点が挙げられる。

2.2 コーパスを用いた冠詞誤りの修正

冠詞誤りの検出・修正手法にはコーパスを利用した手法 [4], [15], [16], [17], [18], [19] も多い。乙武らの手法 [15], [16] では, ロイターニュース記事コーパスから名詞や周辺単語の出現状況を抽出し, それをルールとして獲得する。そして, そのルールを用いて冠詞の誤りを検出して修正する。コーパスを用いた冠詞誤り修正の問題点としては, コーパスの規模を大きくすることが困難である点と, 冠詞の例外的用法への対応が難しい点が挙げられる。

永田らは, 前置詞モデルを用いてコーパスから前置詞との生起確率を推定することで冠詞誤りを検出する手法 [17] と, 名詞のコーパスから学習し生成した可算/不可算判別モデルを誤り検出に用いる手法 [4] を提案した。どちらも冠詞誤り検出に有効であり, 検索エンジンと組み合わせることで性能が向上する可能性があるとして述べられている。また, 竹内らの提案した名詞の前方照応判定モデルを用いる手法 [19] では, 既出の名詞句がそれ以前と同じ内容の名詞句を指す場合に定冠詞が付く用法を考慮することで, 修正精度が向上することが示されている。

そこで我々は, 名詞の可算/不可算や前方照応などの冠詞の用法を検索結果数と組み合わせる手法を提案した [5]。

表 1 可算名詞と冠詞の組み合わせ

	a/an	the	φ (無冠詞)
単数形	○	○	×
複数形	×	○	○

表 2 不可算名詞と冠詞の組み合わせ

	a/an	the	φ (無冠詞)
単数形	×	○	○
複数形	-	-	-

3. 冠詞の用法

本節では文献 [1] の「冠詞の用法」を参考に, 提案手法において SVM の素性として用いる冠詞の用法について説明する。まず名詞が可算か不可算か, 単数形か複数形かに分けて冠詞の用法を述べ, 次に不定冠詞, 定冠詞, 無冠詞の用法について説明する。また, 下記で説明する用法以外に, 慣用句的な表現も素性に用いる。

3.1 可算/不可算名詞と冠詞の用法

名詞が可算か不可算かによって分類した冠詞の用法を表 1, 2 にまとめる。これらの表中の○は可能な組み合わせで, ×は通常不可能な組み合わせである。名詞が可算か不可算かという情報は冠詞の修正において有用なので, 冠詞修正の判断に利用する。

3.2 不定冠詞の用法

不定冠詞 (a/an) は基本的に複数あるものの一つを取り上げたことを表すのに用いられる。用法としては, 「一つ」という意味で one と同じ意味で使われたり, 多数存在するものの中から任意に選んだ一つという意味を表したりすることがある。システムが検出する不定冠詞の用法を以下にまとめる。

(1) 不定冠詞と組み合わせられる構文や表現で用いられる

例: There is **a** situation in which this approach is necessary.

この例では, “there is” の構文によって “situation” に不定冠詞がつく。他にも “here is” や “such as” などの構文がある。

(2) これまでに存在していなかった何かが新たに得られたり発生したりしたものを示す

例: From this, we can obtain **an** important theorem.

この場合, 生成, 発生, 発見, 取得, 導入, 定義を表す動詞と組み合わせられることが多い。例えば, “find”, “introduce”, “obtain” などがこれに該当する。また, これらの動詞の対象となる名詞として, “consequence”, “discovery” などがよく用いられる。ただし, 既に述べられたものを指す場合は当然, 定冠詞になる。本研究では, このような動詞との照合を (2-1), 名詞との照合を (2-2) のとして扱う。

- (3) 対象を個別化する形容詞に修飾されている
例: The result can be shown by **a** simple computation.
対象を個別化する形容詞 “simple” に名詞が修飾されているため不定冠詞がつく.
- (4) 数量形容詞 (few, little, good/great many) に修飾されている
例: **A** little learning is a dangerous thing.
- (5) (as, how, so, too) + 形容詞 + 不定冠詞 + 名詞
例: He is as great **a** musician as ever lived.

3.3 定冠詞の用法

定冠詞は基本的に対象のものが一つしか存在しないことを表す. 定冠詞の具体的な用法を以下にまとめる. 用法の番号は 3.2 節からの通し番号である.

- (6) 直前に述べたものを指す
(7) the ~ of ... の形で対象の性質や特徴を示す
例: **The** existence of such function is obvious.
- (8) その内容が that 以下に述べられていることを示す
例: This property is due to **the** assumption that the noise is Gaussian.
- (9) その個数だけ存在する基数を伴う名詞
例: When we view the line segment, **the** two endpoints may not be visible.
- (10) 普通名詞から転じた固有名詞
例: the United Nations, the White House など
- (11) 最上級の形容詞に修飾されている場合
例: It is **the** easiest way to prove this theorem.
- (12) 人名を冠した対象
例: the Fourier transform など
- (13) 読者が皆知っていると思われるものの場合で, 了解を意味する形容詞で済ませるとき
例: the standard ~, the famous ~ など
- (14) 定冠詞と組み合わせられると決まったもの
例: the same as ..., the past など
- (15) some や most など特定の形容詞+of+the+名詞の複数形
例: some of the ..., most of the ... など
- (16) by + the + 単位を表す名詞
例: We buy tea by **the** pound.

3.4 無冠詞の用法

名詞が無冠詞になるのは, 上記の不定冠詞や定冠詞の用法の範疇に入らない場合といえる. 具体的な無冠詞の用法を以下にまとめる. 番号は 3.3 節からの通し番号である.

- (17) 不特定の数えられない名詞
例: rain, snow など
- (18) アポストロフィ s を伴う人名の所有格
例: Green's theorem など

- (19) 列記された名詞が一まとまりの場合は, 冒頭の単語にのみ冠詞をつけて, 以降の名詞の冠詞を省略することがある
例: The symbols and **notations** used in this paper are as follows.
- (20) 冠詞相当語に修飾されている
例: his, those などに修飾されている
- (21) 記号を固有の対象として扱う時
例: Consider a line segment connecting **point P** and **point Q**.
- (22) 特定の数値を扱う時
例: Consider a circle of **radius 2**.
- (23) (kind, sort, manner, type) + of の後の名詞は無冠詞
例: This is a new type of **dictionary**.
- (24) every や each など特定の形容詞がついた場合
例: In every **case**, each **variable** assumes a different value.
- (25) 学問や専門分野の名前
例: mathematics, vector analysis など

4. SVMによる冠詞誤り自動修正

4.1 システム概要

冠詞修正システムの処理の流れを図 1 に示す. 入力された英文は montyTagger[20] によって品詞タグ付けされ, 英文解析部において名詞句が検出される. 次に検出された各名詞句に対して, あてはまる冠詞の用法を検出し, 検索エンジンにより検索結果数を取得する. そして, それらの結果に基づいて各名詞句の素性ベクトルを生成する.

学習データからこのようにして得られた素性ベクトルを用いて, SVM による識別モデルを生成する. 次にその識別モデルを用いて, テストデータの各名詞句の素性ベクトルに基づいてその名詞句の冠詞を判別する. そして, SVM の判別結果がテストデータの名詞句の冠詞と異なる場合, 冠詞誤りとして検出し修正案を出力する. ただし, 検出された冠詞の用法によって一意冠詞が定まる場合は, その用法に基づいて修正案を決定する. これらの処理について以下で詳しく説明する.

4.2 素性ベクトルの生成

提案手法では, SVM の素性として refrule 節で述べた冠詞の用法, 検索結果数の比と, 名詞句の特徴のいくつかを用いる (表 3). 本節では, 表 3 にまとめた素性ベクトルの生成方法について説明する.

4.2.1 検索エンジンによる検索結果数と比の算出

提案手法は入力文から名詞句を抽出し, その冠詞を不定冠詞, 定冠詞, 無冠詞として検索した検索結果数を取得する. そして, その検索結果数の比を素性に用いる.

具体的にはまず, 入力文から名詞句とその名詞句の一つ

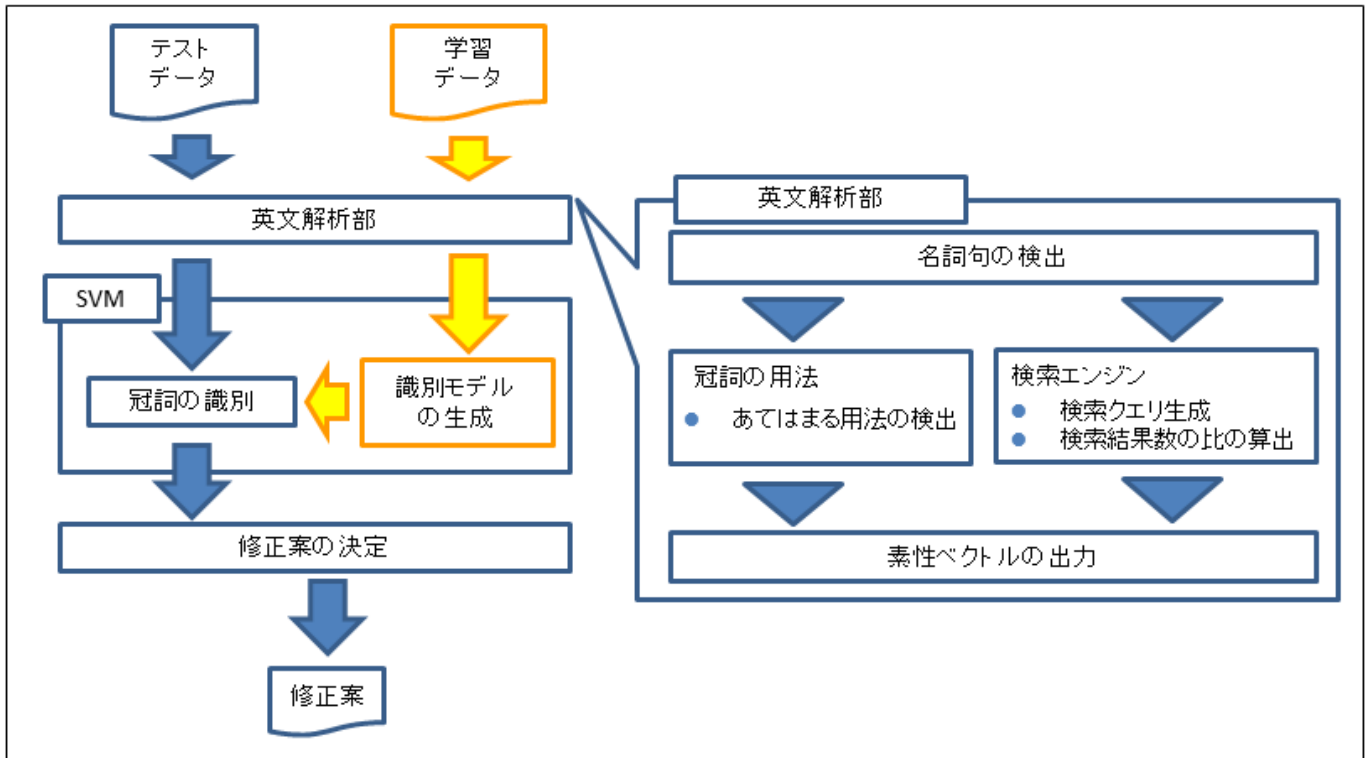


図 1 冠詞修正システムの処理の流れ

表 3 素性ベクトルの一覧

素性	次元数	内容
検索結果数	3	式 (1) で定義される各冠詞の出現比率
名詞の単数/複数形	2	対象の名詞句の最後の名詞が単数形か複数形かの判別
名詞の可算/不可算	3	対象の名詞句の最後の名詞が可算名詞、不可算名詞または両方の意味をもつかの判別
慣用句	3	照合した慣用句が、どの冠詞を含むか
冠詞の用法 (1) ~ (25)	26	各ルールに当てはまるか、当てはまらないかの 2 値 (ただし、用法 (2) は (2-1) と (2-2))
名詞句の特徴	1	対象の名詞句が文頭かどうか (2 値)
	1	対象の名詞句が序数に修飾されているかどうか (2 値)
	1	対象の名詞句の前の単語が前置詞かどうか (2 値)
	1	対象の名詞句が固有名詞に該当するかどうか (2 値)
	1	対象の名詞句が人名に該当するかどうか (2 値)
	1	対象の名詞句が曜日または月に該当するかどうか (2 値)
	1	対象の名詞句が地位や役職を表す名詞に該当するかどうか (2 値)
	1	対象の名詞句の名詞が、世界にひとつしかないものの名詞かどうか (2 値)

前の単語をつなげたものをフレーズとして抜き出す。そして、抜き出したフレーズの名詞句の冠詞 3 種類と、名詞句の最後の名詞の単数形と複数形 2 種類の合計 6 パターンの検索フレーズを生成し、検索する。例えば “This is a book.” の “book” を含む名詞句について生成される検索フレーズを表 4 に示す。しかし、この 6 パターン全ての検索結果数が 0 件の場合、検索フレーズ中の単語を一つ減らし、6 パターンの検索フレーズを新たに生成して検索する。単語の減らし方は、まず名詞句中の形容詞などの修飾語を先頭から減らす。修飾語がなくなり複合名詞が残った場合は、複合名詞の先頭の単語を削除する。このような再検索を、検

索結果数が得られるか、冠詞の後ろの名詞が一語になるまで繰り返す。このようにして、得られた検索結果数から各冠詞の比を算出する。冠詞 x の比 $Ratio(x)$ は式 (1) で定義する。

$$Ratio(x) = \frac{H(x)}{\sum_y H(y)} \quad (x, y \in \{a/an, the, \phi\}) \quad (1)$$

$H(y)$ は冠詞 y で取得した、単数形、複数形両方の検索結果数の和を示す。

4.2.2 冠詞の用法に関する素性

冠詞の用法 (1) ~ (25) の検出方法を表 5 にまとめる。まず、用法 (6) の前方照応は対象名詞句以前にその名詞

表 5 冠詞の各用法の照合方法一覧

照合手段	冠詞の用法
前方照応	(6)
オンライン辞書など Web 上のデータ	(10), (12), (18), (25)
検索エンジンを利用して作成したリスト	(2), (3)
形態素解析など	(1), (4), (5), (7) ~ (9), (11), (13) ~ (17), (19) ~ (24)

表 4 検索クエリの生成例

	a/an	the	ϕ (無冠詞)
単数形	is a book	is the book	is book
複数形	is a books	is the books	is books

が出現したかどうかで判断する。用法 (10), (12), (18), (25) はオンライン辞書や人名事典などの Web 上のデータから判断する。用法 (2), (3) は検索エンジンを用いて収集した単語のリストと照合して判断する。単語リストの収集は、文献 [1] の例文から収集対象の品詞をワイルドカードに変更した検索フレーズを生成し、検索結果のサマリよりワイルドカードに当てはまる単語を収集する [5]。その他の用法は、形態素解析の結果から用法のパターンと照合して判断する。ただし、用法 (21) は記号、(22) は数値が名詞句に含まれているかどうかで判断する。そして、それぞれの用法が検出されたかどうかの 2 値を素性として用いる。

4.2.3 その他の素性

その他の素性として、名詞の可算/不可算、単数/複数形、慣用句を利用する。名詞の可算/不可算、単数/複数形は montyTagger[20] の品詞タグ付結果から判断し、慣用句は Weblio 英和和英辞典 [21] から抽出した慣用句のリストとの照合によって検出する。

またそれ以外に、名詞句の特徴を示す 8 個の素性を使用した。表 3 に示した名詞の特徴を表す要素のうち、上から四つは形態素解析の結果から判断する。また、人名については Weblio 英和和英辞典、残り三つは書籍 [1], [6] からリストを作成して照合し、判断する。

4.3 SVM による冠詞誤りの検出と修正

冠詞誤りの検出と修正案の提示は、4.2 節で説明した素性ベクトルを用いて SVM により決定する。まず、学習データの英文の名詞句の素性ベクトルを用いて識別モデルを生成する。次にテストデータの各名詞句の素性ベクトルと識別モデルを用いて適切な冠詞を判断する。そして、SVM によって出力された冠詞とテストデータの名詞句の冠詞が異なる場合に誤りとして検出し、修正案を提示する。

ただし、用法 (20), (24), (25) が検出された場合、修正案を無冠詞に決定する。

5. 冠詞誤り修正システムの評価実験

評価実験では、以下の観点から提案手法を手法 [5] と比

較する。

(A) 冠詞の適切性

(B) 冠詞誤り修正性能

ここで、(A) は誤りを含まない英文における正しい冠詞を正しいと判別できるか、(B) はテストデータの冠詞の半数を無作為に別の冠詞（無冠詞を含む）に置き換えた英文に対して、その誤りを修正できるかを評価する。

また、本実験では libSVM[22] を使用し、線形カーネル関数と RBF カーネル関数を用いて比較した。本実験におけるカーネル関数のパラメータは、学習データを用いて事前に算出した値を用いる。

5.1 実験環境

提案する冠詞修正システムでは、検索エンジンには Yahoo!検索 WebAPI[23] を用い、冠詞の用法検出では Weblio 英和和英辞典 [21]、欧羅巴人名録 [24]、米国国勢調査局の Name Files[25] を利用した。ただし、Yahoo!検索 WebAPI は 2013 年 10 月現在サービスを停止している。実験では、2013 年 1 月と 2012 年 9 月に取得した検索結果数のログを用いた。

5.1.1 生成した単語リスト

3.2 節の用法 (2-2) で用いた生成、発生、発見、取得、導入、定義を表す動詞の対象になる名詞のリスト、用法 (3) の対象を個別化する形容詞のリストを、検索エンジン Yahoo!検索 WebAPI を利用して作成した。なお用法 (2-1) の生成、発生、発見、取得、導入、定義を表す動詞については Weblio 英和和英辞典を利用して、動詞を収集した。その結果、用法 (2-2) では 1,036 の名詞、用法 (3) では 3,619 の形容詞のリストが得られた。

5.1.2 使用した英文データ

本実験では [5] で使用した New York Times の記事 [26] と NCBI[27] にある英語論文の抄録をテストデータに用いた。この New York Times の記事は政治分野の記事で 51 文で構成されている。この記事中に冠詞判定の対象となる名詞句は 398 あり、また検索結果数は 2013 年 1 月 23 日の検索ログを用いた。一方、英語論文抄録は情報検索に関する論文五つの抄録であり、47 文で構成されている。この抄録中に対象となる名詞句は 334 あり、実験では New York Times の記事と同様に 2013 年 1 月 23 日の検索結果数の検索ログを使用した。

学習データにはテストデータとは異なる New York Times

表 6 冠詞の適切性評価実験 (New York Times)

	正答数	誤り	正解率
線形カーネル	344	54	0.864
RBF カーネル	348	50	0.874
手法 [5]	353	45	0.887

表 7 冠詞の適切性評価実験 (論文抄録)

	正答数	誤り	正解率
線形カーネル	255	79	0.763
RBF カーネル	270	64	0.808
手法 [5]	268	66	0.802

表 8 冠詞の判別結果の内訳 (New York Times)

	(i)	(ii)	(iii)	(iv)
線形カーネル	25	29	16	328
RBF カーネル	17	33	12	336

の記事を使用し、これは政治、科学、テクノロジー、ビジネスの分野から合計 6 記事 196 文からなる。そのうち対象となる名詞句は合計で 1,399 あり、検索結果数は 2012 年 9 月 10 日の検索ログから得た。

5.2 冠詞判別の適切性評価実験

誤りを含まない英文における正しい冠詞を正しいと判別できるか実験した結果を表 6, 7 に示す。これらの表より New York Times の記事では、手法 [5] の正解率が最も高く、続いて RBF カーネル、線形カーネルの順であることがわかる。一方論文抄録では、手法 [5] よりも RBF カーネルの方が正解率が高く、線形カーネルが最も低かった。

次に、冠詞の判別結果を (i) ~ (iv) のように分類すると、それぞれの件数は表 8, 9 のようになった。

(i) 手法 [5] では正しく判別できていたが、提案手法では誤って判別したもの

(ii) 手法 [5] と提案手法の両方で誤って判別したもの

(iii) 提案手法では正しく判別できていたが、手法 [5] では誤って判別したもの

(iv) 手法 [5] と提案手法の両方で正しく判別できたもの

表 8 の (iii) より、New York Times の記事では手法 [5] で誤判別した冠詞を線形カーネルでは 16 箇所、RBF カーネルでは 12 箇所正しく判別していた。しかし、(iii) の件数よりも、(i) の件数の方が多かったため、手法 [5] に若干及ばなかった。手法 [5] より提案手法が正しかった (iii) の例には次のようなものがあつた。例えば、“But he made it clear on Saturday that he was more focused on drawing attention to the issues of limited government, limited deficits and a restrained military than on winning.” の “the issues” は、検索結果数の比は無冠詞を示しているため、用法 (7) に当てはまるが、手法 [5] では無冠詞と判別した。しかし、提案手法の RBF カーネルでは正しく判別することができた。

表 9 冠詞の判別結果の内訳 (論文抄録)

	(i)	(ii)	(iii)	(iv)
線形カーネル	36	43	22	233
RBF カーネル	19	45	20	250

表 10 冠詞誤りの修正性能 (New York Times)

	再現率	適合率	F 値
線形カーネル	0.875	0.829	0.851
RBF カーネル	0.870	0.829	0.849
手法 [5]	0.885	0.855	0.870

表 11 冠詞誤りの修正性能 (論文抄録)

	再現率	適合率	F 値
線形カーネル	0.738	0.681	0.709
RBF カーネル	0.798	0.720	0.757
手法 [5]	0.798	0.724	0.759

5.3 冠詞修正性能の評価

本実験では、提案手法の冠詞誤りの修正性能を評価した。5.2 節と同じ学習データとテストデータを使用し、テストデータの冠詞の半数を無作為に別の冠詞 (無冠詞を含む) に置き換えた英文に対して、その誤りを修正できるか評価した。

5.3.1 評価尺度

冠詞誤り修正の評価尺度として以下の三つの指標を用いた。式 (4) の P は適合率、 R は再現率である。

$$\text{再現率} = \frac{\text{正しく修正した冠詞誤りの数}}{\text{実際の冠詞誤りの数}} \quad (2)$$

$$\text{適合率} = \frac{\text{正しく修正した冠詞誤りの数}}{\text{修正した冠詞誤りの数}} \quad (3)$$

$$\text{F 値} = \frac{2PR}{P + R} \quad (4)$$

5.3.2 冠詞修正性能

提案手法と手法 [5] による冠詞誤り修正の実験結果を表 10, 11 にまとめる。表 10 より New York Times 記事では、手法 [5] の F 値が最も高かった。しかし論文抄録では、RBF カーネルと手法 [5] の F 値は同程度であった。

6. まとめ

本稿では、冠詞の用法と検索結果数を素性とした SVM による冠詞誤りの修正手法を提案した。冠詞の適切性評価実験と冠詞誤り修正実験を行い、手法 [5] と提案手法を比較した。その結果、パラメータを手動で決定した手法 [5] とほぼ同程度の性能を得られることが分かった。

さらなる性能改善には、冠詞の用法や検索結果数に関する素性の追加、用法検出の精度向上などが考えられる。また、英語学習者支援の具体的な方法についても検討したい。

参考文献

- [1] 金谷健一：理系のための英文練習帳 ーさらなる上達を目指してー、共立出版 (2012)。
- [2] 河合敦夫、杉原厚吉、杉江 昇：英文の誤りを検出するシ

- ステム ASPEC-I, 情報処理論文誌 Vol. 25, No. 6, pp. 1072-1079 (1984).
- [3] 平野孝佳, 平手勇字, 山名早人: 検索エンジンを用いた英文冠詞誤りの検出, 日本データベース学会, Letters, Vol.6, No.3 pp. 1-4 (2007).
- [4] 永田亮, 若菜崇宏, 森広浩一郎, 榊井文人, 河合敦夫, 井須尚紀: 可算/不可算の判定に基づいた英文誤りの検出, 電子情報通信学会論文誌 Vol. J89-D, No. 8, pp.1777-1790 (2006).
- [5] 尾崎弘明, 谷本太郁由, 太田 学: 冠詞の用法規則と検索エンジンを利用した英文冠詞誤りの自動修正, 第5回 Web とデータベースに関するフォーラム (WebDB Forum2012), B4 (2012).
- [6] 綿貫陽, 宮川幸久, 須貝猛敏, 高松尚弘: ロイヤル英文法改訂新版, 旺文社 (2000).
- [7] 尾崎弘明, 太田 学: Web 資源を用いた冠詞の用法に基づく冠詞誤り自動修正, 第4回データ工学と情報マネジメントに関するフォーラム (DEIM2012), E9-2 (2012).
- [8] Error Detection and Correction Workshop 2012
<https://sites.google.com/site/edcw2012/>
- [9] 網嶋祐一, 岡田壮史, 安藤一秋: 検索エンジンを利用した多言語作文支援, 電子情報通信学会, 信学技報, ET, 教育工学, Vol. 107, No. 536, pp. 73-78 (2008).
- [10] 松原, 茂樹, 加藤芳秀, 江川誠二: 英文作成支援ツールとしての用例文検索システム E S C O R T, 独立行政法人科学技術振興機構, 情報管理, Vol. 51, No. 4, pp. 251-159 (2008).
- [11] 有富 隼, 太田 学: 検索エンジンによる英文前置詞誤り修正支援, 日本データベース学会論文誌 9(1), pp. 70-75 (2010).
- [12] 久保田朗, 太田 学: 検索エンジンを用いた英文前置詞誤りの自動検出と修正, 情報処理学会研究報告, データベース・システム研究会報告, 2011-DBS-153 (2), pp. 1-8 (2011).
- [13] 谷本太郁由, 太田 学: 検索エンジンを用いた動詞名詞コロケーションに基づく英文動詞誤りの検出と修正, 情報処理学会研究報告, データベース・システム研究会報告, 2010-DBS-151 (36), pp. 1-7 (2010).
- [14] 谷本太郁由, 太田 学: 検索エンジンを用いた英文動詞誤り検出システム, 情報処理学会研究報告, データベース・システム研究会報告, 2012-DBS-156 (2), pp. 1-8 (2012).
- [15] 乙武北斗, 荒木建治: 単語出現状況の帰納的学習による英文誤りの検出及び自動校正, 電子情報通信学会論文誌, (D), 情報・システム Vol. J90-D, No. 6 pp. 1592-1601 (2007).
- [16] 乙武北斗, 荒木建治: 単語出現状況の特徴を用いた英文冠詞誤りの検出及び自動校正, 情報処理学会研究報告, NL-171, pp.25-30 (2006).
- [17] 永田亮, 井口達也, 脇寺健太, 榊井文人, 河合敦夫, 井須尚紀: 前置詞情報を利用した冠詞誤り検出, 電子情報通信学会論文誌 D-I, Vol. J88-D-I, No. 4, pp. 873-881 (2005).
- [18] 乙武北斗, 吉村賢治, 竹内裕己, 河合敦夫: 事例の自動抽象化に基づくルールを用いた英語冠詞の自動付与手法の提案, 言語処理学会, 第18回年次大会発表論文集, pp. 349-352 (2012).
- [19] 竹内裕己, 河合敦夫, 永田亮, 乙武北斗: 英文自動冠詞付与における前方照応の考慮, 情報処理学会研究報告, Vol. 2011-NL-204 No. 10, pp.1-7 (2011).
- [20] MontyTagger
<http://web.media.mit.edu/~hugo/montytagger/>
- [21] 英和辞典・和英辞典—Weblio 辞書
<http://ejje.weblio.jp/>
- [22] LIBSVM – A Library for Support Vector Machines
<http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [23] 検索 - Yahoo!デベロッパネットワーク I
<http://developer.yahoo.co.jp/webapi/search/>
- [24] 欧羅巴人名録
<http://www.worldsys.org/europe/>
- [25] Genealogy Data: Frequently Occurring Surnames from Census 1990 – Names Files – U.S. Census Bureau
http://www.census.gov/genealogy/www/data/1990surnames/names_files.html
- [26] New York Times
<http://www.nytimes.com/>
- [27] National Center for Biotechnology Information
<http://www.ncbi.nlm.nih.gov/>