統計的手法に基づく形容詞または形容動詞の修飾先の決定

橋本泰一　白井清昭　徳永健伸　田中穂積
東京工業大学大学院 情報理工学研究科 計算工学専攻
{taiichi, kshirai, take, tanaka}@cl.cs.titech.ac.jp

[名詞$_1$ の 形容詞/形容動詞 名詞$_2$] というパターンの文における形容詞または形容動詞の修飾先は, 名詞$_1$ であるのか 名詞$_2$ であるのかが曖昧である. 本論文では, コーパスから得た共起事例をもとに, [名詞$_1$ の 形容詞/形容動詞 名詞$_2$] というパターンの文の形容詞または形容動詞の修飾先を決定する原則 (名詞関係依存の原則) に従って, 形容詞または形容動詞の決定する手法を提案する. 評価実験の結果, 約 90%の正解率で形容詞の正しい修飾先を決定できた.

Determing adjectival and adjectival noun dependencies using corpus statistics

Taiichi Hashimoto, Kiyoaki Shirai, Takenobu Tokunaga, Hozumi Tanaka
Department of Computer Science, Graduate School of Information Science and Engineering,
Tokyo Institute of Technology
{taiichi, kshirai, take, tanaka}@cl.cs.titech.ac.jp

In Japanese sentences of the form [N_1 *no* $Adj/AdjN N_2$], the adjective or adjectival noun can modify either N_1 or N_2. This paper describes a method to determine adjectival or adjectival noun dependencies in such constructions, via dependency criteria and corpus statistics. Based on evaluation, our method is able to determe in adjective dependencies with an accuracy of about 90%.

1 はじめに

自然言語を解析するためには, 構文的制約や意味的制約, 文脈に関する制約など, 様々な制約が必要である. これらのうち, 前後に現われる文を考慮しないで一つの文を解析する際には, 文法のような構文的制約がよく用いられる. しかしながら, 文を正確に解析するためには, 構文的制約を用いるだけでは不十分であり, 単語の意味的制約を用いる必要がある.

次の二つの例文における形容詞の修飾先について考えてみよう.

a. 六甲のおいしい水

b. 背丈の高い草

例文a, bは, どちらも [名詞$_1$ の 形容詞 名詞$_2$] というパターンに当てはまる. aでは, 形容詞「おいしい」がその後の名詞「水」を修飾している. 一方, bでは, 形容詞「高い」は前の名詞「背丈」を修飾している.

統語的な係り受け関係を考えれば, 日本語では, 単語は常に後ろにある単語に係るとする場合が多く, 2つの形容詞はともに後ろの名詞に係るとみることもできる. しかしながら, 意味的な修飾関係を考えれば, 例文bの「高い」は「背丈」を修飾すると捉える方が自然な解釈である.

例文a, bは, 同じパターンの文であるのにも関わらず, 一方では形容詞が後ろの名詞を修飾し, もう一方では前の名詞を修飾している. したがって, 形容詞の修飾先を決定するには, 文法のような構文的制約だけではなく, 単語の意味的制約を利用することが必要不可欠である. すなわち, [名詞$_1$ の 形容詞 名詞$_2$] という文における形容詞の係り先は文中に現われる単語に強く依存する. これは, 英語における前置詞句付加問題と共通している.

田中らは, [名詞$_1$ の 形容詞/形容動詞 名詞$_2$] というパターンの文の形容詞または形容動詞の修飾先を決定する名詞関係依存の原則と呼ばれる原則を提案している [7]. この原則は, 人間が形容詞または形容動詞の修飾先を決定するには非常に分かり易い原則である. しかしながら, 名詞関係依存の原則を計算機によってどのように処理するべきかについては明らかにされていない. 本論文では, コーパスから得られる共起事例をもとに, 名詞関係依存の原則に従って形容詞また形容動詞の修飾先を決定する手法を提案する.

本論文の構成は以下の通りである. 2節では, 田中らが提案した名詞関係依存の原則について述べる. さらに [名詞 の 名詞] という共起事例のコーパスにおける出現頻度を利用して, 名詞関係依存の原則に基づいて形容詞または形容動詞を処理する手法について提案する. 3節では, コーパスから [名詞 の 名詞] という共起事例を抽出する方法について述べる. 4節では, 提案手法を用いて形容詞の修飾先の曖昧性を解消する実験について報告する. また, その結果を考察し, 本手法の問題点について論じる. 5節では, 先行研究との比較について述べる. 6節では, 本論文のまとめと今後の課題について述べる.

2 名詞関係依存の原則

2.1 名詞関係依存の原則

田中らは, [名詞$_1$ の 形容詞/形容動詞 名詞$_2$] という文の形容詞または形容動詞の修飾先は, 名詞$_1$ と 名詞$_2$ の意味的依存関係と関連があると述べている [7]. このことを図1に示す4つの例文で説明する.

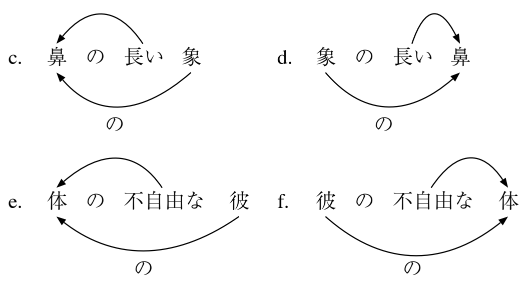


図 1: 意味的依存関係の例

図1中の例文c, dの矢印はそれぞれ「長い」が「鼻」を修飾する,「象」が「鼻」を修飾するという意味的依存関係を表わしている. 二つの例文c, dの違いは,「象」と「鼻」が入れ換わっている点だけである. この例から次の2つの事実が観測できる.

i. [名詞$_1$ の 形容詞 名詞$_2$] と [名詞$_2$ の 形容詞 名詞$_1$] とでは, 形容詞と名詞の間の意味的依存関係は等しい. (どちらの例も「長い 鼻」という意味的依存関係をもつ)

ii. [名詞$_1$ の 形容詞 名詞$_2$] と [名詞$_2$ の 形容

詞 名詞$_1$] とでは, 2 つの名詞の間の意味的依存関係は等しい．(どちらの例も「象の鼻」という意味的依存関係をもつ)

以上から, 形容詞の修飾先は二つの名詞の間の意味的依存関係に深く関係していることがわかる．すなわち, 形容詞は他の名詞の修飾を受ける名詞を修飾する．例文 e, f からも先の 2 つの事実が観測でき, 形容動詞においても同じことが言える．このことから, 田中らは, 形容詞または形容動詞の修飾先を決定する以下の 2 つの原則を定めている．

[名詞$_1$ の 形容詞/形容動詞 名詞$_2$] という文があるとき

原則 1 [名詞$_1$ の 名詞$_2$] が成立する場合, 形容詞または形容動詞は 名詞$_2$ を修飾する．

原則 2 [名詞$_2$ の 名詞$_1$] が成立する場合, 形容詞または形容動詞は 名詞$_1$ を修飾する．

ところが, 上の 2 つの原理だけでは, 形容詞または形容動詞の修飾先を決定できないことがある．例えば, 図 2 の例文に先の原則を当てはめると, 「人気 の チーム」も「チーム の 人気」も成立するため, 形容詞が前後いずれの名詞も修飾できることになる．このような場合, 形容詞または形容動詞の修飾先は, 形容詞と名詞, 形容動詞と名詞の間に存在する意味的依存関係の強さによって決定する．図 2 の場合, 「高い」は「チーム」よりも「人気」を修飾する方が自然であると考えられる．すなわち, 「高い」は「チーム」よりも「人気」との意味的依存関係が強い．以上のことから, 田中らは, さらに次の原則を定めている．

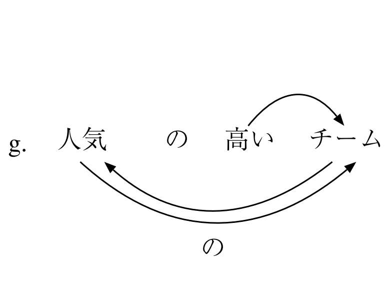


図 2: 原則 1, 2 がともに当てはまる例

原則 3 [名詞$_1$ の 名詞$_2$] と [名詞$_2$ の 名詞$_1$] のどちらも成立する場合には, 形容詞または形容動詞は意味的関係の深い名詞を修飾する．

以上 1, 2, 3 の原則は名詞関係依存の原則と呼ばれている．

2.2 名詞関係依存の原則に基づく形容詞・形容動詞の修飾先の決定

名詞関係依存の原則では, [名詞$_1$ の 名詞$_2$] や [名詞$_2$ の 名詞$_1$] が成立するか否かによって形容詞または形容動詞の修飾先を決定する．しかしながら, 2 つの名詞 a, b について, [a の b] が成立するか否かを計算機によってどのように判定するのかは明らかにされていない．

本研究では, 品詞タグ付きコーパスから [名詞 の 名詞] の共起事例を抽出し, その頻度をもとに [名詞 の 名詞] が成立するか判定する．具体的には, [名詞$_1$ の 名詞$_2$] と [名詞$_2$ の 名詞$_1$] という共起事例の出現頻度を比較し, 頻度の高い方を選択する．したがって, 2.1 項で述べた形容詞または形容動詞の修飾先を決定する名詞関係依存の原則は, 以下の規則に置き換えることができる．

fr(a, b) を [a の b] という共起事例のコーパスにおける出現頻度とする．[名詞$_1$ の 形容詞/形容動詞 名詞$_2$] という文があるとき,

規則 1 $\mathrm{fr}(名詞_1, 名詞_2) > \mathrm{fr}(名詞_2, 名詞_1)$ のとき, 形容詞 (形容動詞) は 名詞$_2$ を修飾する．

規則 2 $\mathrm{fr}(名詞_1, 名詞_2) < \mathrm{fr}(名詞_2, 名詞_1)$ のとき, 形容詞 (形容動詞) は 名詞$_1$ を修飾する．

規則 3 $\mathrm{fr}(名詞_1, 名詞_2) = \mathrm{fr}(名詞_2, 名詞_1)$ のとき, 形容詞 (形容動詞) と意味的依存関係の強い名詞を修飾する．

例えば「対応 の 鈍い 政府」という文があったときに, コーパスには「対応 の 政府」が 10 回, 「政府 の 対応」が 1000 回出現しているとする．このとき, 規則 2 が適用され, 「鈍い」は「対応」を修飾すると判断される．

名詞関係依存の原則において, 原則 3 は形容詞または形容動詞と名詞の間の意味的依存関係に着目し, 形容詞または形容動詞の修飾先を決定する．これに対し, 原則 1, 2 は, 形容詞または形容動詞の前後にある名詞の間の意味的依存関係を手がかりとしている点に特徴がある．言い換えれば, 原則 1, 2 は, 形容詞または形容動詞の修

飾先を決定する際に, 形容詞または形容動詞そのものを考慮しない. このように, 単語の修飾先を決定する際に, その単語そのものを考慮しない名詞関係依存の原則 1, 2 は, 興味深い原則である.

我々は, この原則 1,2 に特に着目し, [名詞 の 名詞] の出現頻度をもとに形容詞の修飾先を決定する規則 1, 2 の有効性を確かめることを主な目的とする. また, これ以後, 規則 3 については特に取り扱わないことにする.

3 [名詞 の 名詞] の抽出

本節では, [名詞 の 名詞] という共起事例を品詞タグ付きコーパスから収集する手法について述べる.

3.1 共起事例の抽出

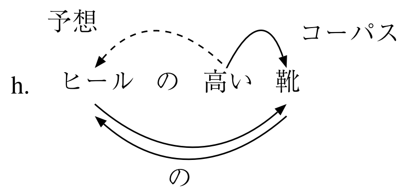
品詞タグ付きコーパスから [名詞$^+$ の 名詞$^+$] というパターンにマッチする単語列を共起事例として抽出する. ただし, [名詞$^+$] は, 1 つ以上の名詞の連続を表す. 名詞の連続はまとめて一つの複合名詞として扱う. ただし, 「こと」「ところ」などの形式名詞は共起事例として収集しない. また, 「名詞$_1$ の 名詞$_2$ の ... 名詞$_n$」にパターンマッチする場合には, 名詞$_1$ ... 名詞$_{n-2}$ の修飾先は一意には決定できないため, 意味的依存関係が曖昧でない [名詞$_{n-1}$ の 名詞$_n$] のみを共起事例として抽出する.

3.2 修飾・被修飾関係の曖昧な共起事例の除去

「ヒール の 高い 靴」という例文を考えてみよう.

この例文の形容詞「高い」の正しい修飾先は, 名詞「ヒール」である. 「ヒール の 靴」は成立しないが, 「靴 の ヒール」は成立すると考えられ, コーパスにおける出現頻度も「ヒール の 靴」よりも「靴 の ヒール」の方が高いと思われる. しかし, 予備実験で調べたところ「靴 の ヒール」よりも「ヒール の 靴」の出現頻度が高かった. コーパスで「ヒール の 靴」が出現する文を調べたところ, 必ず「ヒール の 靴」の前に「高い」「低い」などの形容詞が付いていた.

「高い ヒール の 靴」のように [形容詞 名詞



予想　　：fr(ヒール, 靴) < fr(靴, ヒール)

コーパス：fr(ヒール, 靴) > fr(靴, ヒール)

図 3: 共起事例の頻度が予想に反する例

$_1$ の 名詞$_2$] という並びが存在したとき, 「名詞$_1$ の 名詞$_2$」という意味的依存関係を満たさず, [名詞$_2$ の 名詞$_1$] という意味的依存関係を満たすことがある. しかし, これが常に成り立つわけではない. 例えば, 「新しい 社員 の 制服」という文であれば, 「社員 の 制服」は成立するが, 「制服 の 社員」は成立しない. したがって, [形容詞 名詞$_1$ の 名詞$_2$] という並びのとき, [名詞$_1$ の 名詞$_2$] が成立するのか, [名詞$_2$ の 名詞$_1$] が成立するのかは曖昧である (図 4). このような理由から, 本研究では, [形容詞 名詞 の 名詞] という単語列から [名詞 の 名詞] という共起事例を抽出しないことにする.

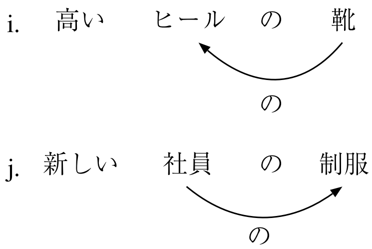


図 4: 修飾・被修飾関係が曖昧な例

3.3 複合名詞の縮退

3.1 項では, 名詞の連続は一つの複合名詞として扱うとした. しかしながら, 複合名詞を個別の事象として扱う場合, 共起事例に出現する名詞が非常に多様になり, データの過疎性の問題が生じやすい. そこで, 複合名詞が現われたとき, その中で最も中心的な意味を持つ名詞 (以下, これを複合名詞の主辞と呼ぶ) に縮退して, 共起事例を抽出する.

例えば, 以下の 2 つの例文の意味はともに「○○法を改正する」という意味であり, どの法律を改正するのかが異なるだけである.

k. 預金保険法 の 改正

l. 廃棄物処理法 の 改正

このような場合, それぞれを別の事例として学習するのではなく, どちらも「法 の 改正」として共起事例を抽出する.

また, 複合名詞を構成する最後の名詞 (「預金/保険/法」であれば「法」) がその複合名詞の主辞であることが多い. このことから複合名詞がコーパス中に出現した場合, その複合名詞を名詞を構成する最後の名詞に縮退して共起事例を抽出する.

3.4 固有名詞と数詞の抽象化

複合名詞と同様に, 固有名詞もまたデータの過疎性の問題を引き起こす要因と成りうる. 例えば, 以下の例文において, m と n の意味に大きな違いはなく,「日本」と「アメリカ」という国名が異なるだけである.

m. 日本 の 経済

n. アメリカ の 経済

この場合 「< 国名 > の 経済」として共起事例を抽出するのが望ましい.

また, 次の例文においても,「15%」と「20%」の違いはそれほど重要ではないため,「< 数 >% の 伸び」として共起事例を抽出するのが望ましい.

o. 15% の 伸び

p. 20% の 伸び

そこで, 固有名詞を「人名」「人名 (姓)」「人名 (名)」「組織名」「地域名」「国名」「その他」の 7 つのクラスに抽象化し, [名詞 の 名詞] という共起事例を抽出する. 4 節に述べるように, 我々は RWC 品詞タグ付きコーパスから共起事例を抽出する [3]. 先に挙げた 7 つのクラスは, RWC コーパスの名詞の品詞タグの細分類に基づいている. 一方, 数詞はすべて「数」というクラスに抽象化して共起事例を抽出する.

4 実験

提案手法の有効性を評価するために実験を行った. 今回の実験では [名詞$_1$ の 形容詞 名詞$_2$] における形容詞の修飾先の決定のみを対象とし, [名詞$_1$ の 形容動詞 名詞$_2$] における形容動詞の修飾先を決定する実験は行っていない.

4.1 共起事例の抽出

[名詞 の 名詞] という共起事例の抽出に RWC 品詞タグ付きコーパスを利用した. RWC コーパスは, 毎日新聞の 1991 年から 1995 年までの 5 年分の新聞記事に対して品詞タグが付与されたコーパスである. 得られた共起事例は, のべで 4,360,031 個, 異なりで 2,656,222 個であった.

テストデータとして, 毎日新聞の 1995 年の 1 月から 4 月までの記事から [名詞$^+$ の 形容詞 名詞$^+$] にパターンマッチする単語列 (1,889 個) を抽出し, 人手により形容詞の正しい修飾先を付与した. [名詞 の 名詞] という共起事例を抽出したときと同様に, [名詞] が形式名詞の場合はテストデータから除去した. また, 形容詞が「ない」「よい」「良い」「いい」の場合もテストデータに含めなかった. これらの形容詞はほとんど一意に修飾先を決定できるためである. すなわち, 形容詞が「ない」のときには, すべて前の名詞を修飾する場合がほとんどである. また, 形容詞が「よい」「良い」「いい」のときには, 後の名詞が「点」または「ところ」であればそれを修飾し, そうでなければ前の名詞を修飾する場合が多い.

4.2 評価実験

以下の 5 つの手法を用いて, テストデータの形容詞の修飾先を決定した. ただし, 手法 B から E については, [名詞$_1$ の 名詞$_2$] と [名詞$_2$ の 名詞$_1$] の出現頻度が同じ場合には誤りとした. また, [名詞$_1$ の 名詞$_2$] と [名詞$_2$ の 名詞$_1$] の頻度がともに 0 である場合は, 修飾先を決定しない. 実験結果を図 5 に示す.

A. すべての形容詞が 名詞$_1$ を修飾する.(ベースライン)

B. 3.1 項で述べた方法で収集した [名詞 の 名詞] の共起頻度をもとに, 2.2 項で述べたアルゴリズムで形容詞の修飾先を決定する.

	A	B	C	D	E
正解率	52.8%	91.2%	92.1%	88.7%	87.7%
	(998/1,889)	(456/500)	(456/495)	(926/1044)	(1,015/1,157)
被覆率	100.0%	26.4%	26.2%	55.3%	61.2%
	(1,889/1,889)	(500/1,889)	(495/1,889)	(1044/1,889)	(1157/1,880)

$$\text{正解率} = \frac{\text{修飾先が正しく決定できた文の数}}{fr(\text{名詞}_1, \text{名詞}_2) \neq 0 \text{ または } fr(\text{名詞}_2, \text{名詞}_1) \neq 0 \text{ となる文の数}} \times 100$$

$$\text{被覆率} = \frac{fr(\text{名詞}_1, \text{名詞}_2) \neq 0 \text{ または } fr(\text{名詞}_2, \text{名詞}_1) \neq 0 \text{ となる文の数}}{\text{全文数}} \times 100$$

図 5: 実験結果

C. B の手法に加えて, [形容詞 名詞 の 名詞] という単語列から抽出された共起事例を用いない. (3.2 項)

D. C の手法に加えて, 複合名詞を主辞に縮退する.(3.3 項)

E. D の手法に加えて, 固有名詞と数詞を抽象化する.(3.4 項)

すべての形容詞が 名詞$_1$ を修飾すると仮定すると正解率は 52.8%となった. それに比べ, コーパスの [名詞 の 名詞] の出現頻度を用いて修飾先を決定することで, 正解率は 91.2%まで向上した. しかし, 被覆率は 26.4%と低い値を示した. この原因は, テストデータ中に現われる複合名詞や固有名詞が [名詞$_1$ の 名詞$_2$] や [名詞$_2$ の 名詞$_1$] という共起事例が現れないためであった.

コーパスで「形容詞 名詞 の 名詞」にパターンマッチする「名詞 の 名詞」を学習データから削除することで, 正解率は 92.1%に向上した. しかし, 修飾・被修飾関係が曖昧な共起事例を除去することで, 不正解となるものを少なくすることができたが, 正解となっていたものがデータ不足のため結果を出力できないとなる場合も存在した.

さらに, 複合名詞を主辞で縮退して共起事例を抽出することと, 固有名詞と数詞を抽象化して共起事例を抽出することで, 被覆率は向上したが正解率は低下した. しかしながら, 正しい修飾先を決定できた形容詞の数 (手法 C, D, E でそれぞれ 456, 926, 1,015) を増やすことができた.

4.3 本手法の改良点

実験結果を分析した結果, 本手法には以下の改良が必要であることがわかった.

- 名詞の意味クラスの利用
  複合名詞を主辞で縮退し, 固有名詞, 数詞を品詞タグで抽象化することで, 被覆率を約 62%まで向上させることができた. さらに被覆率を向上させるためには, 意味クラスを利用して単語を抽象化することが考えられる.

- 「固有名詞 (その他)」の細分類
  「固有名詞 (その他)」で抽象化された単語は, 外来語や片仮名語など非常に多様であったため, 単語が持つ意味情報が [名詞 の 名詞] の出現頻度に反映されず, 形容詞の修飾先を誤っている場合が多かった. 「固有名詞 (その他)」を名詞の意味に基づいて細分する必要がある.

- 複合名詞の縮退の方法の改良
  「ぶり」「など」「程度」「性」「感」などの接尾語を最後に含む複合名詞の場合, 本手法ではこれを複合名詞の主辞とみなして縮退して, 共起事例の抽出を行った. この場合にも誤りを起こすことが非常に多い. これは, 複合名詞の縮退の方法に問題があり, 縮退の方法についてもっと検討する必要がある.

5　先行研究との比較

5.1　形容詞または形容動詞の修飾先決定問題

[名詞$_1$ の 形容詞/形容動詞 名詞$_2$] の形容詞または形容動詞の修飾先に関する先行研究として, 菊地らの研究がある [6]. 菊地らは, 形容詞または形容動詞とその前後の名詞の特性を分析し, 形容詞または形容動詞の修飾先を決定する 7 つのルールを提案している. このルールの特徴は, 形容詞または形容動詞と前後の名詞の関係から修飾先を決定するという点である. このルールを用いることで, 約 97%の正解率で修飾先を決定できることを示している.

この研究では対象となる形容詞が 21 語, 形容動詞が 29 語と非常に少ない. これに対し, 本研究で提案する手法は, 前後の名詞のみに依存した手法であるため形容詞または形容動詞には依存しない. そのため, 実験でテストデータとして用いた形容詞は 203 語であり, 多くの形容詞を対象としている.

本研究では, 形容詞または形容動詞に依存しない手法を提案したが, 形容詞または形容動詞の情報を完全に無視することはできない. 例えば, 形容詞「ない」はほとんどの場合, 前の名詞を修飾するという特徴がある. このような有効な情報は利用すべきである. 形容詞または形容動詞の特徴を捉えた菊地らのルールを本手法と組み合わせることで更なる正解率の向上が望める.

5.2　英語の前置詞句付加問題

[名詞$_1$ の 形容詞/形容動詞 名詞$_2$] という文の形容詞または形容動詞の修飾先を決定する問題は, 単語の意味情報を用いなければ解決できないという点で, 英語の前置詞句付加問題と類似している. この問題をコーパスを用いて解決する手法として, Hindle らの手法 [4], Ratnaparkhi の手法 [5], Brill らの手法 [1],Collins の手法 [2] などがある. Brill らと collins らの手法は, 単語の正しい係り先が付与されている構文木付きコーパスから単語の共起情報の学習を行い, その学習結果をもとに前置詞句の係り先を決定している. 一方, Hindle らと Ratnaparkhi の手法は, 正しい係り先が付与されていないコーパスから単語の共起情報の学習を行う. Hindle らは, 構文解析を行い, その解析結果から共起情報を学習し, その共起情報をもとに前置詞句の係り先を決定する. 一方, Ratnaparkhi は, ヒューリスティックを利用して品詞列から曖昧性のない共起情報のみを学習している.

本手法は, [名詞$_1$ の 形容詞/形容動詞 名詞$_2$] という文の形容詞または形容動詞の正しい修飾先が付与されたコーパスを用いて単語の共起情報を学習するのではなく, 品詞タグ付きコーパスから [名詞 の 名詞] というパターンにマッチする共起事例をもとに修飾先を決定する. したがって, 本手法は Ratnaparkhi の手法と類似している点が多いと言える.

6　おわりに

本論文では, 名詞関係依存の原則と [名詞 の 名詞] という共起事例のコーパスにおける出現頻度に基づき, [名詞$_1$ の 形容詞/形容動詞 名詞$_2$] の形容詞または形容動詞の修飾先を決定する手法を提案した. また, 実験により約 90%の精度で形容詞の正しい修飾先を決定できることを示した. さらに, 意味的依存関係が曖昧な共起事例を削除することで正解率の向上を図った. また訓練データの過疎性の問題を解消するために, 複合名詞の縮退と固有名詞, 数詞の抽象化を行った.

今回の実験では, [名詞$_1$ の 形容詞 名詞$_2$] という単語列のみを実験の対象とした. 今後, [名詞$_1$ の 形容動詞 名詞$_2$] という単語列における形容動詞の修飾先を決める場合にも, 提案手法が有効であるかを検討する必要がある.

本研究で提案した手法では, 形容詞の情報を全く利用していない. 今後は, 形容詞または形容詞の情報を本手法に組み込むことを考える必要がある. また, 本手法で用いた複合名詞の縮退の方法や単語の抽象化についても問題点があり, さらなる改良が必要である.

参考文献

[1] Eric Brill and Phi Resnik. A Rule-Based Approach To Prepositional Phrase Attachment Disambiguation. In *Proceedings of the Fifteenth International Conference on Computational Linguistics*, 1994.

[2] Michael Collins and James Brooks. Prepositional Phrase Attachment through a Backed-off Model. In *Proceedings of 3rd Workshop on Very Large Coropra*, 1998.

[3] Koiti Hasida, Hitoshi Isahara, Takenobu Tokunaga, Minako Hashimoto, Shiho Ogino, Wakako Kashino, Jun Toyoura, and Hironobu Takahashi. The RWC Text Databases. In *Proceedings of the First International Conference on Language Resources and Evaluation*, pp. 457–462, 1998.

[4] Donald Hindle and Mats Rooth. Structual Ambiguity and Lexical Relations. *Computational Linguistics*, Vol. 19, No. 1, pp. 103–120, 1993.

[5] Adwait Ratnaparkhi. Statistical Models for Unsupervised Prepositional Phrase Attachment. In *Proceedings of COLING - ACL '98*, 1998.

[6] 菊地浩三, 伊東幸宏. 連体形イ・ナ形容詞に先行する格助詞句の係りに関するルールの抽出. 言語処理, Vol. 6, No. 3, pp. 75–99, 1999.

[7] 田中穂積, 荻野孝野. 形容詞もしくは形容動詞の修飾先の名詞を決める原則について. 計量国語学, Vol. 12, No. 5, pp. 191–203, 1980.