

述部機能表現に対する意味ラベル付与

上岡 裕大^{1,a)} 成田 和弥^{1,b)} 水野 淳太^{2,c)} 乾 健太郎^{1,d)}

概要：機能表現を正しく認識することは、拡張モダリティ解析や事実性解析をはじめとする応用タスクに有用である。しかし、解析ツールを開発するための基礎となるコーパスが利用可能でないことや、機能表現解析に関わる問題が明らかでないなどの課題がある。そこで、本研究では、応用タスクとして日本語事実性解析を例に挙げ、その精度向上を目的として既存の機能表現辞書の再構築、機能表現意味ラベル付与コーパスの作成および、そのコーパスを用いた機能表現解析を行った。

Semantic Tagging of Functional Expressions in Japanese Predicative Phrases

YUDAI KAMIOKA^{1,a)} KAZUYA NARITA^{1,b)} JUNTA MIZUNO^{2,c)} KENTARO INUI^{1,d)}

Abstract: Semantic tagging of functional expressions is useful for a broad range of NLP applications such as extended modality analysis and factuality analysis. However, automatic semantic tagging of functional expressions has some technical challenges. For example, the corpus of functional expressions is unavailable. In this paper, for improvement of factuality analysis, we present three contributions; (i) a functional expression dictionary based on the existing dictionary, (ii) an available corpus of functional expressions and (iii) an automatic semantic tagger for functional expressions.

1. はじめに

日本語には、語彙的意味をほとんど担わず、主に統語的關係を示す役割を持つ機能語が存在する。機能語には、一つの語からなるものだけでなく、複数の語が組み合わさってはじめて意味をなす複合辞がある。本稿では、これらを総称して機能表現と呼ぶ。機能表現は、文を解釈する上で必須のものであり、事実性解析や機械翻訳などにおいて重要な役割を果たしている。

機能表現の一例を示す。日本語では、主に述語に後続する機能表現によって、種々の主観的情報が表される。例えば、(1) では、事象「壊れる(る)」に「てしまっ」「た」「かも

しれない」という機能表現が後続し、この文を記述するより前の時制において、パソコンが壊れた可能性があることを表している。

(1) パソコンが壊れ てしまったかもしれない。

これまでに、それぞれの機能表現が表す意味を元に、事実性や拡張モダリティ [1] を解析するタスクが取り組まれ、機能表現の解析誤りに起因する誤りが多いことが報告されている [2]。こうした応用タスクにおける精度向上のためには、まず、機能表現を正しく解析する必要がある。

機能表現解析では、内容語と機能表現の区別、機能表現の範囲の同定、機能表現が表す意味の認識が課題となる。例えば、(1) において、まず「てしまったかもしれない」が「壊れる(る)」に後続する機能表現であることを認識する。次に、範囲の同定で、「てしまっ」「た」「かもしれない」が、それぞれ一つの機能表現であることを判別し、最後に、それぞれの機能表現が持つ意味として、無意志、完了、推量-不確定という意味ラベルを認識する。日本語機能表現解析システムを構築するにあたって、次の4つの問題が考えら

¹ 東北大学大学院 情報科学研究科
Tohoku University

² 情報通信研究機構
National Institute of Information and Communications
Technology (NICT)

a) yudai.k@ecei.tohoku.ac.jp

b) narita@ecei.tohoku.ac.jp

c) junta-m@nict.go.jp

d) inui@ecei.tohoku.ac.jp

れる。

A)機能表現には、同一の表現であっても表す意味が異なる表現が存在する。例えば、(2)、(3)に共通して出現する「ではないか」という表現は、それぞれ異なる意味を表す。

(2) 社会人となって結婚までした ではないか。(感嘆)

(3) あれほど確認した ではないか。(疑問)

B)複合辞かどうかの判断が難しい。例えば、「ではないか」という表現は、(2)、(3)ではそれがひとつの表現単位であったが、次の(4)では「うではないか」を複合辞と考え、勧誘の意味を表していると考えられる。また、(5)ではそれぞれ意味の異なる3つの表現の組み合わせであると考えられる。

(4) しっかり見つめ直そう ではないか。(勧誘)

(5) 忙しくてそれどころ ではないか。(判断、否定、感嘆)

C)意味ラベルの粒度は応用タスクに依存する。例えば、以下の(6)、(7)には、いずれも推量を表す表現が現れる。しかし、2つの表現の間には確実性の違いがあり、(7)のほうがより確実性の高い推量を表している。こうした表現に対して、同一の意味ラベルを付与するのか、より細分化した意味ラベルを付与するのかは、解析結果を利用する後段のタスクに依存する。意味ラベルを細分化すると、より精密な解釈が可能となるが、それらを自動分類することは難しくなる可能性が高い。

(6) 読者は彼を愚かだと思 うかもしれない。(推量)

(7) 恐らく掘立小屋かなんか に違いなかった。(推量)

D)解析器を開発するための基礎となるコーパスが一般に利用可能ではない。先行研究において、日本語機能表現を網羅的に収録した辞書の編纂や、機能的・内容的用法の曖昧性解消は行われてきた。しかし、解析器を構築するにあたり、一般に利用可能なコーパスは我々の調べた限り存在しない。

本研究では、事実性解析への応用を主眼として、機能表現への意味ラベル付与を行う。具体的には、機能表現意味ラベル付与コーパスを作成し、接続制約と品詞情報に基づくルールベースモデルと、条件付確率場に基づく系列ラベリングのそれぞれによって機能表現解析を行い、本課題における問題を分析する。CRFによる機能表現解析では、 F_1 近くを達成し、高い精度を実現した。一方で、機能表現の曖昧性に関する問題点も明らかになった。

2. 関連研究

2.1 辞書やリソースに関する研究

日本語の機能表現に関しては、辞書の構築や、範囲の検出、意味の自動認識など、さまざまな研究がされてきた。まず、大規模な数の機能表現を処理する基礎として、松吉らによる日本語機能表現辞書『つつじ』[3]がある。つつじでは、言語学的文献を参考にして得た見出し語341件について、種々の異形を考慮し、約17,000種類の機能表現を収録している。つつじの見出し体系は9つの階層を持つ階層構造となっており、見出し語、意味、派生、表記のゆれなどを考慮している。既存の機能表現リストとの比較により、各々の見出し語に対して、ほぼすべての異形を網羅しているとされる。

また、いくつかの機能表現を含むモダリティ関連表現の曖昧性解消に向けて、川添らによるMCNコーパスの構築も行われている[4]。川添らは、言語学的なテストにより、実テキスト上で、モダリティ表現、否定表現、条件表現などのモダリティ関連表現の曖昧性を解消し、語義を特定したコーパスの構築を目指している。

2.2 機能表現の範囲の同定に関する研究

機能表現の範囲の同定に関する研究としては、土屋ら(2007)[5]や鈴木ら(2012)[6]などがある。土屋らは、機能表現同定を、形態素を単位とするチャンク同定問題として定式化した。そして、SVMを用いて形態素解析結果から機能表現を検出し、F値で約0.92という高精度な検出器を実現した。

鈴木らは、大規模な均衡コーパスである『現代日本語書き言葉均衡コーパス』*1(以下、日本語コーパスと呼ぶ)において、機能的用法、内容的用法の曖昧性を持つ機能表現を対象として、機械学習により用法を判定する手法を適用し、その性能を評価した結果を報告した。当文献における機能表現とは、我々の扱う機能表現のうち複合辞に相当するものである。条件付確率場を利用したチャンキングによる実験の結果、0.97近いF値を達成した。また、鈴木らは、別の文献において、機能的、内容的等の用法判定を行う上で、代表的機能表現の用例や、つつじに基づく接続制約などを利用することで、性能の改善が見込めることを報告している[7]。しかしながら、彼らの研究はいずれも、機能的用法、内容的用法の区別にとどまっており、機能表現の意味を特定するまでには至っていない。

2.3 機能表現が表す意味の自動認識に関する研究

機能表現が表す意味まで自動認識する研究は、我々の調べた限り、小早川ら(2009)[8]と今村ら(2011)[9]の研究

*1 http://www.ninjal.ac.jp/corpus_center/bccwj/

のみである。小早川らは、形態素解析結果から、つつじのエントリーをもとに、機能表現の可能性のある部分に辺を張ることで、ラティスを生成した。そして、マルコフモデルを用いた最適経路選択を行うことで、機能表現の意味を解析した結果、0.88 近い F 値を達成した。また、考察として、評判分析に応用する上で、どのような機能表現が検出できると効果的か、つつじの意味体系をどのように変更すべきか、について議論している。彼らの研究では、NHK 教育テレビで放送された番組に対する反響を集め、その中から、機能表現の意味タグを付与したコーパスを作成しているが、公開されていない。

今村らは、形態素解析結果に対して、述部を同定し、つつじの意味ラベルを機能表現に付与する意味ラベルタグを構築した。彼らの研究では、つつじと識別モデルに基づく最尤選択を組み合わせて機能表現を同定しており、辞書を用いてラティスを作成することにより、異なるドメイン間で学習と評価を行った場合でも、F 値で 0.85 という高い精度で意味ラベルを自動付与することに成功した。しかしながら、新聞記事とブログ記事に対して、機能表現と意味ラベルを手で付与した、学習、評価用のコーパスは公開されておらず、タグ付けにおける課題も明らかにされていない。

2.4 事実性解析への応用

機能表現解析は、事実性解析や機械翻訳などの応用タスクにおいて、重要な役割を果たしている。本稿では、事実性解析への応用を主眼として、機能表現への意味ラベル付与を行う。事実性とは、文中の事象が実際に起こったことなのか、あるいは起こる可能性を述べたものなのか、に関する情報である。

(8) a. それは故障ではありませ ん。

b. 彼は先に帰っている かもしれない。

例えば、(8a) では、「それが故障である」ことが、実際には起こっていないことであることを述べており、(8b) では、「彼が先に帰っている」ことについて、その可能性が高いことを述べている。このとき、(8a) の「ん」が否定を表す機能表現であること、および、(8b) の「かもしれない」が推量を表す機能表現であることを認識することが、事実性解析のための大きな手がかりとなる。

江口ら (2010) [10] は、事実性や価値判断などを統合した拡張モダリティタグ体系を設計し、事象間の依存関係を考慮した、機械学習による解析システムを構築した。その中で、つつじに基づく機能表現の情報が、素性として有効であったことが報告されている。

Narita et al. (2013) [2] は、日本語に対するルールベースの事実性解析器を構築し、それを拡張モダリティ付与コー

パス [1] の一部に適用した。彼らの解析器は、Saurí et al. (2012) [11] の英語における解析器を、日本語に適応させたものであり、機能表現のように、事実性解析の手がかりとなる表現に基づく、事実性の更新ルールを構成的に組み合わせることで、決定的に事実性を解析するモデルである。その誤り分析を行うことで、機能表現の曖昧性解消が、事実性解析において重要な課題の一つであると述べている。

3. 意味ラベル付与コーパスの構築

機能表現解析のために、日本語コーパスの Yahoo!知恵袋中の述部機能表現に対して意味ラベルを付与し、学習および評価に用いる述部機能表現意味ラベル付与コーパスを構築した。本節では、コーパスの構築方法について述べる。構築したコーパスは、日本語コーパスに対する差分として <http://tinyurl.com/ja-fe-corpus> より入手可能である。

品詞体系は IPA 品詞体系に従う。形態素情報は、日本語コーパスに付与されている品詞は UniDic 品詞体系に基づくため、形態素解析器 [12] で解析し直して利用する。従って、形態素解析誤りを含む。

3.1 アノテーション対象

意味ラベル付与の対象とする機能表現は、述部の機能表現である。述部の判定は人手により行う。事態性名詞に続く機能表現も解析対象とし、述部に続く機能表現については、主体を表す「が」や対象を表す「を」など、述語項構造解析の範疇である表現に対してもラベルを付与した。また、「お 勧めする」のような接頭辞は、解析対象とせずラベルを付与しない。

全ての述部機能表現に意味ラベルを付与することは付与コストが大きいため、解析難易度が高いと考えられる事象を中心に付与する。そのような事象として Narita et al. [2] の事実性解析器によって、主節に解析誤りがあると判断された文を選択する。そのような文は 1,100 文存在したが、鈴木ら [6] に倣い、口語調の崩れた日本語や方言、誤字など含む 50 文を除外した 1,050 文に対して、意味ラベルを付与する。

3.2 意味ラベルの表現方法

本研究では、形態素を最小単位として意味ラベルを付与した。機能表現のうち複合辞は、以下に示す IOB2 フォーマットによってその範囲も付与した。

- I 述部機能表現に含まれる先頭以外の形態素
- O 述部機能表現に含まれない形態素
- B 述部機能表現の先頭の形態素

意味ラベルは、機能表現辞書つつじ [3] で定義される 89 種類の意味カテゴリをベースに、解析結果を事実性解析に

利用することを前提として、細分化や追加を施した結果、117種類の意味ラベルを用いる。具体的には、既存の意味カテゴリのうちの5カテゴリを細分化した。例えば、(6)、(7)のような推量表現について、事象の確実性を判別するために、つつじの意味カテゴリ「推量」を「推量-不確実」「推量-高確実性」「推量-反実」の3つの意味ラベルに細分化した。また、対応する意味カテゴリが存在しなかった表現に対して、「結果状態」「受身」「使役」「容易」「困難」など、計21ラベルを追加した。細分化や追加を施した意味ラベルの一覧を表1^{*2}、表2、表3に示す。

意味ラベルの付与例を表4に示す。付与対象となる機能表現は、「てしまっ」「た」「かもしれない」の3種類である。述語「壊れ」や句点「。」には、述部機能表現に含まれないことを表すラベル「O」を付与する。また、完了を表す「た」には、述部機能表現の先頭の形態素であることを表す「B」を用いて、ラベル「B-完了」を付与する。複合辞に対しては、「かもしれない」のように、先頭の形態素「か」にラベル「B-推量-不確実」を付与し、続く形態素には、述部機能表現に含まれる先頭以外の形態素を表す「I」を用いて、ラベル「I-推量-不確実」を付与する。

表4 意味ラベルの付与例

形態素	意味ラベル
壊れ	O
て	B-無意志
しまっ	I-無意志
た	B-完了
か	B-推量-不確実
も	I-推量-不確実
しれ	I-推量-不確実
ない	I-推量-不確実
。	O

4. 意味ラベル付与コーパスの分析

本節では、作成したコーパスについて、その内容の詳細について述べる。

4.1 統計情報

コーパスの統計情報を表5に示す。付与対象の1,050文に対して、アノテーションを施した述部の数は2,295述部であった。コーパス全体に現れる述部機能表現数は3,775であり、その異なり総数は654であった。ここで、機能表現数とは、「かもしれない」のような複合辞も1個と数える。したがって、述部あたりの機能表現数は、平均して1.6個である。また、コーパス中で使用されたラベルの種類数は、117種類のうち89種類であった。

表5 コーパスの統計情報

収録文数	1,050
述部数	2,295
機能表現延べ総数	3,775
機能表現異なり総数	654
ラベル異なり総数	89

表6 意味ラベル別出現回数(上位10件)

ラベル名	出現回数	ラベル名	出現回数
否定	275	態度	163
丁寧	246	順接確定	156
判断	242	結果状態	158
理由	222	依頼	155
疑問	203	過去	153

表7 機能表現別出現回数(上位10件)

機能表現	出現回数	機能表現	出現回数
た(過去)	147	た(完了)	94
ない(否定)	132	ます(丁寧)	93
ので(理由)	104	です(判断)	85
と(内容)	102	て.下さい(依頼)	84
て(順接確定)	99	が(逆接確定)	60

4.2 意味ラベルの分布

コーパス中に含まれる意味ラベルごとの出現数を表6に示す。89種類のラベルのうち、出現数上位10種類のラベルが全体の約半数を占めた。

次に、機能表現ごとの分布を表7に示す。この中で「た」のように曖昧性があり、かついずれのラベルの出現頻度も高い表現の解析が重要である。

最後に、本コーパスに出現する機能表現のうち、2種類以上の意味を持つ、つまり意味ラベルに曖昧性のある機能表現は、654種類のうち、93種類であった。

5. 機能表現解析

解析性能の現状を明らかにし、誤り分析を行うことを目的として、条件付確率場(CRF)を利用した系列ラベリングによる機能表現解析器の評価を行った。本節では、実験の内容及びその結果について述べる。なお、CRF実装はCRFsuite [13]を用いた。

5.1 実験設定

実験では、3節で構築した意味ラベル付与コーパスを正解データとして、10分割交差検定によるクロズドメインテストを行った。機能表現を1単位とし、BIタグの連続をまとめて一致したラベルを正答とし、以下で定義する評価値を算出した。また、実用的な精度として、述部単位、文単位での精度も算出した。これらの評価では、述部または文中に含まれる全ての意味ラベルが正解データと一致した場合にのみ正答とした。さらに、機能表現の範囲の同定

*2 表現の一覧は公開したコーパスに含まれている

表 1 表現を追加した意味カテゴリーの例

意味カテゴリー	追加した表現の例
添加	のですが, のだけど
順接仮定	として
願望	ように
勧め	たほうがいい, ましょう, たらどう
依頼	てくれ, ていただけませんか
伝聞	といえます, といわれている, みたいです
不許可	ちゃだめ
判断	んです, のです
強調	は, も, こそ
意志	ておく, にする
事前	前, 直前
事後	後
目的	ように
比況	がごとく, みたい

表 2 細分化を施した意味カテゴリーの一覧

元カテゴリー	細分化したカテゴリー	表現例
付帯	付帯-続行	まま, ままで, たなりで
	付帯-並行	ながら, つつ
感嘆	感嘆	もの, こと, なんて
	感嘆-確認	ね
疑問	疑問	うが, ではないか, か
	疑問-確認	だって
推量	推量-不確実	かもしれない, かもわからない, らしい
	推量-高確実性	にちがいない, にきまっている, にそういない
	推量-反実	ところ
進行	進行-習慣	ことにしている
	進行-慣習	ことになっている

表 3 追加した意味カテゴリーの一覧

意味カテゴリー	表現例
結果状態	た, ている, である, ておく
受身	れる, られる
自発	れる, られる
使役	せる, させる
丁寧	ます, です
困難	にくい, がたい, つらい
容易	やすい, よい, いい
様態	そう
態度	ね, ねえ, なあ
無意志	てしまう
試行	てみる
慣習	ものだ
内容	と, って, という
手段	て
反復	て, で
条件	早くて, て初めて
比較	方, というより
命令	なさい
過去	た
原因	により
例示	とか, など
逆接必要	でない

が正しく行われているかを検証するため、ラベルの意味部分を考慮せずに BI タグが一致した表現を正答とする評価も同時に行った。

$$\text{適合率} = \frac{\text{正しく解析された機能表現数}}{\text{いずれかのラベルが付与された機能表現数}}$$

$$\text{再現率} = \frac{\text{正しく解析された機能表現数}}{\text{評価データに存在する機能表現数}}$$

$$\text{F 値} = \frac{2 \cdot \text{適合率} \cdot \text{再現率}}{\text{適合率} + \text{再現率}}$$

学習にあたって、本稿では、形態素情報のみを素性とした。具体的には、元コーパスに記述された形態素解析結果を元に、対象となる語の形態素情報(表層・品詞・品詞細分類)、前後 2 単語の形態素情報、前後 1 単語の形態素情報とのパイグラムを用いた。

CRF による機能表現解析の有効性を確認するため、ベースライン手法として、機能表現辞書を用いた人手の規則による解析を行う。本手法では、機能表現辞書に登録された接続情報のうち、左接続情報を制約として、制約を満たす表現の中で最長の表現を選択する。今回、アノテーション対象とした文は、ルールベースの解析手法を用いた場合に誤りが認められた文であるため、機械学習に対して不利な結果になる。

5.2 評価・考察

まず、BI タグのみを考慮した機能表現の範囲同定性能を表 8 に示す。いずれの手法も高い適合率・再現率となった。CRF による解析でも範囲同定の誤った事例を以下に示す。

(9) a. 心配する 事もないです。

b. 心配する 事も ないです。

(9) において、正解データでは (a) のように「事もないです」を 1 つの機能表現としてラベルが付与している。しかし、今回の実験における CRF の出力では、これらを別の機能表現として (b) のように解析した。こうした複合辞を分割して解析してしまう例が複数見られた。

誤りの原因は、2 つ考えられる。まず 1 つ目は、学習事例における機能表現としての出現数が少ないことである。例に挙げた「事もないです」という表現は、学習事例に一度しか出現しない。また、2 つ目は、内容的用法をとりうる機能表現を含んでいることである。「事もない」という表現中に含まれる「ない」は、内容的用法もとり得るため、学習により機能表現でないとしてラベル O を付与してしまったと考えられる。一方で、ベースライン手法での検出では、これらの表現を正しく検出できた。これらに対する

表 8 機能表現の範囲同定の実験結果

	適合率	再現率	F 値
ベースライン	0.8608(3179/3693)	0.8209(3099/3775)	0.8404
CRF	0.9249(3089/3340)	0.8015(3052/3808)	0.8588

表 9 機能表現解析の実験結果

	適合率	再現率	F 値
ベースライン	0.4663(1722/3693)	0.4559(1721/3775)	0.4610
CRF	0.7447(2494/3349)	0.6577(2483/3775)	0.6985

表 10 述部単位及び文単位での機能表現解析精度

	述部単位での精度	文単位での精度
Rule	0.3451(792/2295)	0.2390(251/1050)
CRF	0.5542(1272/2295)	0.3838(403/1050)

解決案として、出現数の少ない表現に関して学習事例を増やすことや、ベースラインで用いた接続制約を学習素性として使うことが考えられる。

次に、表 9 および表 10 に機能表現解析の評価結果を示す。機能表現単位の解析結果について、ベースラインは F 値が 46.10 であるのに対して、CRF による解析性能は F 値が 69.85 となり、解析性能を大きく改善することができた。述部単位、文単位の評価においても CRF がベースラインを上回る結果となった。

意味的曖昧性のある機能表現に対する解析結果の詳細を表 11 に示す。曖昧性のある表現に対して、その表現の出現回数、(A) ベースライン・CRF とともに正答した数、(B) ベースライン・CRF とともに誤った数、(C) ベースラインでは誤ったが CRF では正答した数、(D) ベースラインでは正答したが CRF では誤った数を示す。例えば、「て. いる」という表現の場合、「進行-習慣」「進行-慣習」「継続」「結果状態」の 4 つの意味をとりうる。CRF によって改善されたのは 3 事例と多くはないが、CRF によって悪化した事例はなかった。

全体として 1,987 個の曖昧性のある機能表現について、ベースラインに対して、約 44.57%(1132/2541) の改善が見られた。一方で、約 5.4%(137/2541) の事例はベースラインでは正解していたが、CRF では不正解となった。また、約 28.5%(725/2541) の事例は、いずれの手法を用いても正しく解析することができなかった。

CRF が誤った例として (10) のような文がある。(10) では、「ている」という表現が含まれ、正解データでは「進行-慣習」のラベルが付与されているが、CRF の出力では、「結果状態」のラベルに分類された。「ている」に対して付与され得る意味には、他に、「進行-習慣」「継続」もあるが、正解データ中に登場するほとんどは「結果状態」であるため、その他のラベルの正答率が低くなるという結果になった。曖昧性のある表現であっても、その意味には偏りがあることが判明した。したがって、出現数の少ない意味を持つ表現に関しての訓練事例を選出して追加することで、性

表 11 意味的曖昧性のある機能表現の解析結果

意味候補数	機能表現 (出現数)	A	B	C	D
12	て (176)	10	61	101	4
8	で (25)	0	23	1	1
7	ように (20)	4	10	3	3
6	が (117)	2	45	68	2
5	と (143)	79	23	28	13
	でしょう (62)	35	7	19	1
4	た (249)	0	86	158	5
	ている (66)	49	13	3	1
	か (59)	0	19	31	9
	の (49)	1	15	32	1
	し (38)	0	6	29	3
	てい (38)	29	8	1	0
	な (17)	7	7	1	2
3	たら (55)	0	25	30	0
	ても (37)	23	10	2	2
	てる (24)	17	2	0	5
	ています (13)	5	4	2	2
	てます (8)	4	2	1	1
	計 2541 表現	547	725	1132	137

能が改善する可能性がある。

(10) だから仮売り上げと言っているはずです。

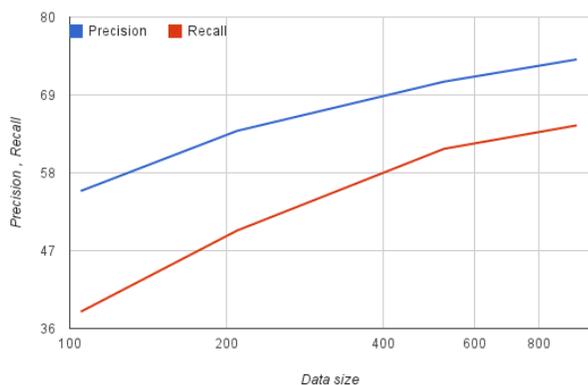


図 1 学習曲線

訓練事例数を変えた場合の学習曲線を図 1 に示す。訓練事例の数を変えて交差検定を行ったところ、訓練事例を増やすことで精度の向上が見られた。また、曲線の傾きからも事例数を増やすことは解析精度の向上に有効な手段であると考えられる。

6. 事実性解析への応用

機能表現解析の応用先の一つである事実性解析において、機能表現解析の影響を調査する。事実性解析器としては、Narita et al. (2013) [2] で定義された事実性の更新ルールのうち、機能表現に関するルールのみを適用することによ

り、著者から見た事実性を決定的に解析する。機能表現に依存する事象に限定するため、解析対象は主節に含まれる事象（主事象と呼ぶ）のみとする。主事象以外的事象は、機能表現以外に「～するのをやめる」のように事実性に影響を与える述語を考慮しなければならない。一方で、主事象における事実性は、「おそらく」のように推量を示唆する副詞、および、否定接頭辞がある場合を除き、機能表現のみによって決定することができる。

事実性ラベルとしては、Narita et al. と同様に、確信度 (CT, PR, U) と肯否極性 (+, -) の組によって表される体系を利用する。即ち、CT+ (実際に起こっている)、PR+ (起こっている可能性が高い)、PR- (起こっている可能性が低い)、CT- (実際には起こっていない)、U (起こっているかどうか不明) の 5 種類のラベルのいずれかに、各主事象の事実性を分類する。

6.1 解析モデル

本解析モデルでは、主事象に付随する機能表現の意味ラベルを利用することで、主事象の事実性を決定する。例えば、否定の機能表現が付随している場合には、肯否極性を反転する、という更新ルールを利用する。その他の更新ルールは、以下の 3 種類を用いる。カッコ内は対応する機能表現の意味ラベルを表す。

1. 肯否極性: $++ \rightarrow -$, $-- \rightarrow +$ (「否定」「無意味」「不可能」など)
 2. 確信度: CT \rightarrow PR (「推量-不確実」「推量-高確実性」「意志」など)
 3. 確信度: CT \rightarrow U, PR \rightarrow U (「疑問」「願望」「依頼」など)
- 我々は、各意味ラベルをもとに、機能表現を 3 種類のルールのいずれかに割り当てた。無標のラベルである CT+ から始めて、文末から順に機能表現を参照し、更新ルールの割り当てられた機能表現を発見したら更新ルールを適用し、最終的な事実性を決定する。事実性に影響を与える機能表現が存在しない場合には、CT+ のままとする。

(11) うまく進めないようですね。

例えば (11) では、「進む」という事象に対して、「ない」「ようです」「ね」の 3 つの機能表現が付随している。このとき、後ろから順に、機能表現の意味ラベルに対応した更新ルールを適用していく。まず、「ね」は態度の機能表現であり、事実性に影響しないため、事実性は CT+ のままとする。次に、「ようです」は推量-不確実の機能表現であり、<確信度: CT \rightarrow PR> の更新ルールを持つため、事実性を PR+ に更新する。最後に、「ない」は否定の機能表現であり、<肯否極性: $++ \rightarrow -$, $-- \rightarrow +$ > の更新ルールを持つため、事実性を PR- に更新する。最終的に「進む」の事実性は PR- となる。

表 12 機能表現解析結果を組み入れた事実性解析器の評価

機能表現解析手法	事例数	CT+	PR+	PR-	CT-	U	マイクロ平均	マクロ平均
		412	189	46	58	345	1,050	-
ベースライン	適合率	0.5736	0.5248	0.8333	0.3175	0.5714	0.5543	0.5641
	再現率	0.7184	0.2804	0.1087	0.3448	0.6029	0.5543	0.4111
	F 値	0.6379	0.3655	0.1923	0.3306	0.6029	0.5243	0.4226
CRF	適合率	0.5923	0.6000	0.8235	0.3977	0.6287	0.5924	0.6084
	再現率	0.7087	0.3492	0.3043	0.6034	0.6232	0.5924	0.5178
	F 値	0.6453	0.4415	0.4444	0.4795	0.6259	0.5924	0.5273
正解ラベル	適合率	0.6201	0.5984	0.7368	0.4048	0.6403	0.6095	0.6001
	再現率	0.6893	0.3862	0.3043	0.5862	0.6812	0.6095	0.5295
	F 値	0.6529	0.4695	0.4308	0.4789	0.6601	0.6095	0.5384

6.2 評価・考察

表 12 に、機能表現を利用した事実性解析器の各ラベルの適合率、再現率、F 値、および、それらのマイクロ平均、マクロ平均を示す。機能表現の意味ラベルとしては、5 節で用いたベースラインモデルによる解析結果、CRF による解析結果、および正解ラベルの 3 種類を用いた。

実験の結果、CRF による解析結果を利用した場合が、正解ラベルを利用した場合の性能にかなり近くなった。次に、ベースラインモデルの結果を利用した場合でも、ある程度の性能が得られた。Narita et al. では解析誤りとなった事例を対象としているが、本研究で行った機能表現の意味ラベルの追加と細分化、および接続制約を用いたことによって改善されたと考えられる。正解ラベルを利用した場合に着目すると、まだ改善の余地があることが分かる。機能表現が正しく解析されているにも関わらず、事実性が正しく解析されない原因を調査するため、誤り分析を行った。

正解ラベルを利用しても、事実性解析が誤った事例としては、機能表現が持つ意味ラベルの適用ルールが適切でなかったものや、副詞が事実性に影響を与えている場合が見られた。

- (12) a. それがチャートの上位に入ってくる原因
じゃないでしょうか? (システム: U, 正解: PR+)
- b. どうやって色を判別してるんでしょうか? (システム: U, 正解: CT+)

(12) の下線部は、いずれも疑問が付与された機能表現を持つため、事実性は U と解析された。(12a) は、著者が推測した結果について同意を求めている問いかけであるため、正解は PR+ である。(12b) は、前提として起こった事象である「判別」の方法の問いかけであり、正解は CT+ である。このような疑問の解釈の違いによる誤りが多く見られた。それ以外の意味ラベルにおいても、文脈によって解釈が変わる事例が存在した。

7. 結論

本研究では、事実性解析への応用を主眼として、機能表
2014 Information Processing Society of Japan

現への意味ラベル付与を行った。まず、既存の機能表現辞書の意味カテゴリを見直し、新たな機能表現辞書を構築した。構築した辞書を元に、Yahoo!知恵袋のテキスト中に含まれる機能表現に対して意味ラベルを付与したコーパスを作成した。作成したコーパスから、どのような機能表現に曖昧性があるのかを分析したところ、全体の 13% の機能表現に曖昧性があることが明らかになった。

また、作成したコーパスを元に、CRF を用いた機能表現解析を行った結果、ベースラインに対して有意な結果が得られた。曖昧性のある機能表現に対する解析性能は十分であるとは言えないが、学習曲線から学習事例を増やすことによって精度の向上が見込めることを明らかにした。

さらに、応用タスクとして、作成したデータを用いて事実性解析を行った結果、CRF による解析結果を用いた場合にも実用的な性能が得られた。

今後の課題として、まず、曖昧性のある機能表現に関して学習事例を追加することが考えられる。また、事実性解析以外の応用タスクにおいても機能表現解析が有効であるかを確認したい。

参考文献

- [1] 松吉俊, 江口萌, 佐尾ちとせ, 村上浩司, 乾健太郎, 松本裕治: テキスト情報分析のための判断情報アノテーション, 電子情報通信学会論文誌 D, Vol. J93-D, No. 6, pp. 705-713 (2010).
- [2] Narita, K., Mizuno, J. and Inui, K.: A Lexicon-based Investigation of Research Issues in Japanese Factuality Analysis, *In Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP 2013)*, pp. 587-595 (2013).
- [3] 松吉俊, 佐藤理史, 宇津呂武仁: 日本語機能表現辞書の編纂, 自然言語処理, Vol. 14, No. 5, pp. 123-146 (2007).
- [4] 川添愛, 田中リベカ, 戸次大介: MCN コーパス: モダリティ関連表現の曖昧性解消のためのアノテーションと言語学的テストの利用, テキストアノテーションワークショップ・コンテスト (2012).
- [5] 土屋雅稔, 注連隆夫, 高木俊宏, 内元清貴, 松吉俊, 宇津呂武仁, 佐藤理史, 中川聖一: 機械学習を用いた日本語機能表現のチャンキング, 自然言語処理, Vol. 14, No. 1, pp. 111-138 (2007).
- [6] 鈴木敬文, 阿部佑亮, 宇津呂武仁, 松吉俊, 土屋雅稔:

- 『現代日本語書き言葉均衡コーパス』における複合辞の検出と評価, 第1回『コーパス日本語学ワークショップ』予稿集, pp. 365-372 (2012).
- [7] 鈴木敬文, 阿部佑亮, 宇津呂武仁, 松吉 俊, 土屋雅稔: 代表・派生関係を利用した日本語機能表現の解析方式の評価, 言語処理学会第18回年次大会予稿集, pp. 598-601 (2012).
- [8] 小早川健, 関場治朗, 木下明德, 熊野 正, 加藤直人, 田中英輝: 単語格子とマルコフモデルによる日本語機能表現の解析: 日本語機能表現辞書「つつじ」を用いて, 電子情報通信学会技術研究報告, pp. 15-20 (2009).
- [9] 今村賢治, 泉朋子, 菊井玄一郎, 佐藤理史: 述部機能表現の意味ラベルタガール, 言語処理学会第17回年次大会論文集, pp. 2-5 (2011).
- [10] 江口萌, 松吉俊, 佐尾ちとせ, 乾健太郎, 松本裕治: モダリティ、真偽情報、価値情報を統合した拡張モダリティ解析, 言語処理学会第16回年次大会, pp. 852-855 (2010).
- [11] Saurí, R. and Pustejovsky, J.: Are you sure that this happened? assessing the factuality degree of events in text, *Computational Linguistics*, Vol. 38, No. 2, pp. 261-299 (2012).
- [12] Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying conditional random fields to Japanese morphological analysis, *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp. 230-237 (2004).
- [13] Okazaki, N.: CRFsuite: a fast implementation of Conditional Random Fields (CRFs) (2007).