

機能表現に基づく日本語事実性解析

成田 和弥[†] 水野 淳太[‡] 上岡 裕大[†] 菅野 美和[†] 乾 健太郎[†]

[†] 東北大学 [‡] 情報通信研究機構 (NICT)

{narita, yudai.k, meihe, inui}@ecei.tohoku.ac.jp junta-m@nict.go.jp

1 はじめに

事実性とは、文中の事象が、実際に起こったことなのかそうではないことなのか、といった情報である¹。

- (1) a. 彼はさきほど部屋を出た。
 b. もう遅いから、彼は先に帰っているだろう。
 c. 問題が発生するのを防いだ。
 d. 彼が怒っていないというのは嘘だ。

例えば、(1a)の事象「出る」は、実際に起こったことだと著者は判断している、と解釈できる。(1b)の事象「帰る」は、「だろう」という機能語により、起こった可能性が高いことだと著者は推量している、と解釈できる。(1c)の事象「発生する」は、「防いだ」という述語により、実際には起こっていないことだと著者は判断している、と解釈できる。(1d)の事象「怒る」は、「ない」および「嘘だ」により、実際に起こったことだと著者は判断している、と解釈できる。事実性を解析する技術は、情報抽出や含意関係認識など、自然言語処理の応用に対して有用であるが、多岐に渡る言語表現を捉えきれず、十分な解析性能は実現されていない。

我々は、主節と従属節における、事実性に影響を与える要素の違いに着目した。(1c)の事象「発生する」や(1d)の事象「怒る」など、従属節に含まれる事象における事実性解析では、「防いだ」や「嘘だ」などの述語、「ない」などの機能表現、さらにはこれらの表現のスコープやフォーカスなど、さまざまな要素を考慮することが必要である。一方で、(1a)の事象「出る」や(1b)の事象「帰る」など、主節に含まれる事象（以下、主事象と呼ぶ）は、「だろう」や「ない」といった機能表現のみで事実性を決定できると考えられる。

これを検証するために、機能表現解析結果を用いたルールベースの事実性解析器を構築し、1,475個の主事象に適用する。機能表現解析結果は、辞書による最長一致、機械学習による解析、正解ラベルの3種類5セットを用い、それぞれに対する事実性解析結果を比較する。その結果、機能表現解析の改善による事実性解析への影響、機能表現のみで事実性が決定可能な主事象の割

¹その事象が真に起こったことかはわからないため、本研究では、著者による事象の成否の判断として事実性を解釈する。また、松吉ら[6]と同様に、事象は行為、出来事、状態の総称であると考えられる。

合を明らかにする。さらに、機能表現の正解ラベルに対する事実性解析の誤り分析によって、事実性解析の今後の方針について検討する。

2 関連研究

事実性およびその周辺情報を付与するための体系に関する研究としては、Sauriらによる FactBank [3] や、松吉らによる拡張モダリティタグ付与コーパス [6] などがある。Sauriらは、事実らしさに対する態度表明者の確信度と、その確信の方向を表す肯否極性の組で事実性を定義し、事象に付与した。松吉らは、<態度>、<真偽判断>等からなる拡張モダリティタグ体系を設計し、それを事象に付与したコーパスを構築した。拡張モダリティタグのうち、<真偽判断>が事実性に対応する。

解析手法に関する研究としては、機械学習に基づく手法や、パターンベースの手法があげられる。江口ら [9] は、拡張モダリティタグの項目間、文内の事象間の依存関係を考慮できる、条件付確率場を用いた拡張モダリティ解析を行った。機能表現は辞書による解析を行い、素性として入れているがその有無による性能比較はしていない。Sauriら [4] は、事実性に影響を与える手がかり表現を利用し、確信度と肯否極性を、依存構造木の根から伝搬させて解析するモデルを提案した。手がかり表現の1種類として *may* や *not* のような機能表現を利用している。de Marneffe ら [1] は、動詞のクラスとともに、機能表現を素性とした解析を行っている。水野ら [8] は機械学習を用いた拡張モダリティ解析を行い、機能表現解析結果を素性に入れることで、性能が向上することを報告している。上岡ら [10] は事実性解析の性能向上を目的として、機能表現の意味ラベル体系を整理し、それを付与したコーパスを構築している。我々は、上岡らが意味ラベル体系を整理したことによる効果を検証する。

3 機能表現意味ラベルを用いた事実性解析

事実性は、Narita et al. [2] と同様に、3種類の確信度 (CT, PR, U) と2種類の肯否極性 (+, -) の組によって表す。従って、各事象の事実性は、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分類される。

本研究で利用するコーパスは、拡張モダリティタグ付与コーパス [6]² と、機能表現意味ラベル付与コーパス [10]³を合わせたものである。これらはいずれも『現代日本語書き言葉均衡コーパス』(BCCWJ)⁴を付与対象としており、そのうちの Yahoo!知恵袋データを利用する。解析対象となる事象は、このうち機能表現が付随する主事象である 1,475 事象である。主事象は、各文において、拡張モダリティが付与されている事象のうち、最も文末に近い事象とした。拡張モダリティから事実性への変換は、Narita et al. と同様にした。事実性の分布は、CT+が 476 事象、PR+が 215 事象、PR- が 51 事象、CT- が 107 事象、U が 626 事象となった。拡張モダリティタグ付与コーパスでは、CT+が全体の 65%を占めるが、本研究では機能表現が付随する事象のみを対象としているため、U の割合が最も多くなった。

3.1 機能表現解析

機能表現解析結果は、以下の 5 種類を用いる。

1. 『つつじ』[7] のラベル体系に基づいた最長一致による解析結果
2. 上岡ら [10] のラベル体系に基づいた最長一致による解析結果
3. 2. で、曖昧性のある機能表現について、コーパス中で最頻出のラベルを選択した結果
4. 上岡らのラベル体系下で、周辺単語を素性とした CRF を用いて曖昧性解消を行った結果
5. 上岡らによってアノテーションされた正解ラベル

1, 2 では、表層が同じであっても意味が異なる、曖昧性のある機能表現に対して、意味ラベルを一意に定められない。そこで、「走る [よう] です。」に対して〈伝聞、推量-不確実、比況〉のように複数の候補をすべて付与する。ただし、機能表現の左側の形態素の品詞接続制約による候補選択は行う。従って、接続制約を満たす意味ラベル候補がない場合および、ラベル体系に含まれていない機能表現については、意味ラベルは付与されない。これらを利用した事実性解析結果を比較することで、機能表現のラベル体系の違いが事実性解析にもたらす影響について検証する。

3 は、2 で複数の候補が出力された場合において、コーパス中で最も頻度の高い意味ラベルを採用する。機能表現の曖昧性解消におけるベースラインとなる。4 は、CRF を用いることで、周辺単語を考慮して機能表現の曖昧性解消を行う。これらの結果を比較することにより、機能表現の曖昧性解消を行うことで、事実性解析にどれだけ寄与できるのかを検証する。

²<http://cl.cs.yamanashi.ac.jp/nldata/modality/>

³<http://tinyurl.com/ja-fe-corpus>

⁴http://www.ninjal.ac.jp/corpus_center/bccwj/

最後に、機能表現の正解ラベルを用いた事実性解析を行うことで、機能表現解析ができたとしても事実性解析が誤る事例がどの程度存在するのかを検証する。この誤り分析によって、事実性解析の今後の方針を検討する。

3.2 事実性解析

事実性解析は、主事象より後ろにある機能表現の意味ラベルを利用することで決定する。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、という事実性更新ルールを適用する。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意味ラベルを表す。

1. 肯否極性：+→-, -→+ (〈否定〉〈否定意志〉〈否定推量〉〈無意味〉〈不明確〉〈不可能〉〈回避〉〈不必要〉〈放置〉〈困難〉)
2. 確信度：CT→PR (〈推量-不確実〉〈推量-高確実性〉〈否定推量〉〈意志〉〈否定意志〉〈伝聞〉〈様態〉〈容易〉〈困難〉)
3. 確信度：CT→U, PR→U (〈疑問〉〈勧誘〉〈勧め〉〈願望〉〈依頼〉)

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があれば該当する更新ルールを適用する。すべての機能表現の更新ルールを適用することで、主事象の事実性を決定する。なお、疑問符も事実性に影響を与える要素として考えられるが、疑問符があっても、肯定の場合も少なくないため、本研究では採用していない。

(2) うまく進めないようですね。

例えば(2)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+のままとなる。次に、「ようです」は〈推量-不確実〉の機能表現であり、更新ルール 2 に該当するため、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、更新ルール 1 に該当するため、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。

また、複数の意味ラベルを採用する機能表現解析結果を用いる場合、いずれかの意味ラベルが更新ルールに対応していれば、対応するすべての更新ルールを適用する。同じ更新ルールをもつ機能表現が複数含まれる場合には、更新は 1 度のみとする。例えば(2)の機能表現「ようです」に〈伝聞、推量-不確実、比況〉というラベルが付与されている場合、〈伝聞〉と〈推量-不確実〉が更新ルールを持っている。そのためそれぞれのルールを適用するが、この例では、同じ更新ルール 2 に該当するため、一度しか適用を行わず、結果として先ほどと同様の出力が得られる。

4 評価実験・考察

表1に、機能表現解析結果ごとの事実性解析器の結果を示す。ここでは、確信度、肯否極性それぞれの軸における精度、再現率、F値、ならびにそれらのマクロ平均を示した。確信度がUの場合には、肯否極性を決定できないため、確信度の正解がUである事例、および、確信度の出力がUである事例は肯否極性の評価対象から除外した。

まず、『つつじ』に基づく最長一致を利用した場合と、上岡らの体系に基づく最長一致を利用した場合とを比較することで、機能表現の意味ラベル体系の違いが事実性解析にもたらす影響を検証する。表1に示すように、上岡らの体系を利用した方が、性能が向上している。機能表現の意味ラベルを再整理し、辞書の拡充をしたことで、特に、有標のラベルであるPR, U, -の再現率を向上させることができ、事実性解析に対して効果的であることを示している。

次に、上岡らの体系に基づく最長一致を利用した場合と、最長一致に加えて、曖昧性のある機能表現についてコーパス中の最頻出ラベルを選択した場合、および、CRFを用いた場合とを比較することで、機能表現の曖昧性解消が事実性解析にどれだけ貢献するのかを確認する。CRFによる解析結果を利用した場合、上岡らの体系に基づく最長一致を利用した場合と比較して性能が向上した。

(3) 5階くらいから落ちて助かったんでしたよね。

(最長一致: U, 最頻出: U, CRF: CT+, 正解: CT+)

(3)では、主事象「助かる」の事実性はCT+である。上岡らの体系に基づく最長一致では、最頻出を用いても「よね」を〈疑問〉と判断するため、誤解析された。CRFでは、「でした」や句点といった周辺情報をもとに、「よね」を〈態度〉と正しく判断することができたため、事実性も正しく解析することができた。このように、機能表現解析をより精緻に行うことが、事実性解析に対して有効であることが確認できた。

最後に、正解ラベルを用いることで、機能表現の曖昧性が完全に解消できたとしても解けない事例はどの程度あるのか、それらを解くためにはどのような情報が必要なのかを分析する。表1を見ると、機能表現が正しく与えられているにも関わらず、事実性解析が誤る事例は少なくない。確信度または肯否極性が誤って解析された562事例のうち、CT+に誤分類された事例が149事例、CT+以外に誤分類された事例が413事例であった。これらの事例のうち、それぞれ50事例に対して、誤り分析を行った結果を表2に示す。

(4) a. ビタミンは、野菜や海草から補給するべきです。

(正解ラベルに基づく解析: CT+, 正解: U)

b. 入札前に確認すべきでしたね。

(正解ラベルに基づく解析: CT+, 正解: CT-)

(4)は、機能表現の更新ルールが不足しているために誤った事例である。(4a)の機能表現「べき」「です」はそれぞれ〈当為〉〈判断〉の意味をもっているが、これは、〈当為〉に関する更新ルールが不足していたことによる誤りである。(4a)は〈当為〉を更新ルール3の適用対象に加えれば解決するが、(4b)の誤解析は防げない。(4b)の機能表現「べき」「でし」「た」「ね」はそれぞれ〈当為〉〈判断〉〈完了〉〈態度〉の意味をもっている。この誤解析に対応するため、〈当為〉のみでは更新ルール3を適用するが、〈当為〉と〈完了〉の組み合わせの場合には、更新ルール1を適用することが考えられる。このように、1つ1つの意味ラベルを各更新ルールに割り当てただけでなく、意味ラベルの組み合わせを考慮すべきである。今後、他にどのような組み合わせによって事実性が変化するかを調査する必要がある。

(5) a. やはりこの御時世、きついのではない
ででしょうか?

(正解ラベルに基づく解析: U, 正解: PR+)

b. どうやって判別してるんでしょうか?

(正解ラベルに基づく解析: U, 正解: CT+)

(5)は、現在の機能表現の意味ラベルの粒度が不十分であり、更新ルールを誤って適用した事例である。いずれの事例も、波線部に対して〈疑問〉が付与されているため、主事象の事実性はUと解析された。しかしながら、(5a)は、問いかけではあるものの、推量の意味合いが強いため、正解はPR+となっている。(5b)は、前提として起こった事象である「判別する」の方法を問う文であるため、CT+が正解である。このような事象の事実性を解析するために、〈疑問〉を〈疑問-推量〉〈疑問-方法〉のように意味ラベルを細分化すべきなのか、あるいは事実性解析の段階で文脈を用いて区別すべきなのかは、容易に決定できない。前者は、意味ラベルを無闇に細分化すると、曖昧性解消が困難になるため、適切な粒度を考えなければならない。後者は、例えば質問タイプ判定[5]の技術を利用して、質問タイプによって適用するルールを変更することが考えられるが、質問タイプと機能表現の組み合わせを考える必要がある。

以上のように、機能表現だけでは解析が難しい事例が見られた。このような事例を解くために、機能表現から事実性を演算するためのルールの拡張、一部の副詞や助詞などを考慮した解析、同じ意味ラベルをもつ機能表現の用法の区別に取り組んでいく必要がある。

5 おわりに

本研究では、主事象における事実性解析に着目し、実験を行った。機能表現解析結果を用いたルールベースの事実性解析器を構築し、1,475個の主事象に適用することで、「だろう」や「ない」といった機能表現のみで、主事象の事実性を決定できるかを検証した。機能表現解

表 1: 機能表現解析結果に基づく事実性解析の評価

機能表現 解析結果	事例数	確信度				肯否極性		
		CT	PR	U	マクロ平均	+	-	マクロ平均
		583	266	626	-	691	158	-
『つつじ』 に基づく 最長一致	精度	46.06 (474/1,029)	34.33 (69/201)	70.20 (172/245)	50.20	86.50 (615/711)	89.23 (58/65)	87.87
	再現率	81.30 (474/583)	25.94 (69/266)	27.48 (172/626)	44.91	98.87 (615/622)	37.66 (58/154)	68.27
	F_1	58.81	29.55	39.49	42.62	92.27	52.97	72.62
上岡らの 体系に 基づく 最長一致	精度	51.25 (389/759)	44.96 (107/238)	62.34 (298/478)	52.85	85.81 (514/599)	82.86 (58/70)	84.34
	再現率	66.72 (389/583)	40.23 (107/266)	47.60 (298/626)	51.52	97.72 (514/526)	40.56 (58/143)	69.14
	F_1	57.97	42.46	53.99	51.52	91.38	54.46	72.92
上岡らの 体系に 基づく 最長一致 (最頻出)	精度	54.05 (374/692)	48.66 (91/187)	64.09 (382/596)	55.60	87.11 (480/551)	82.14 (69/84)	84.63
	再現率	64.15 (374/583)	34.21 (91/266)	61.02 (382/626)	53.13	96.97 (480/495)	49.29 (69/140)	73.13
	F_1	58.67	40.18	62.52	53.79	91.78	61.61	76.70
CRF	精度	60.66 (313/516)	54.59 (107/196)	62.65 (478/763)	59.30	95.29 (405/425)	76.26 (106/139)	85.78
	再現率	53.69 (313/583)	40.23 (107/266)	76.36 (478/626)	56.76	92.47 (405/438)	84.13 (106/126)	88.30
	F_1	56.96	46.32	68.83	57.37	93.86	80.00	86.93
正解ラベル	精度	64.83 (341/526)	56.02 (121/216)	67.26 (493/733)	62.70	96.26 (438/455)	79.22 (122/154)	87.74
	再現率	58.49 (341/583)	45.49 (121/266)	78.75 (493/626)	60.91	93.19 (438/470)	87.77 (122/139)	90.48
	F_1	61.50	50.21	72.55	61.42	94.70	83.28	88.99

表 2: 誤りの種類の分布

		出力	
		CT+	CT+以外
機能表現	意味ラベルの粒度	10	21
	機能表現の アノテーション誤り	6	4
事実性	機能表現の 更新ルールの不足 または誤適用	9	2
	機能語相当表現	9	2
	副詞・助詞など	4	5
	機能表現の省略	5	0
	事実性の アノテーション誤り	3	14
その他 (形態素解析誤りなど)		4	2

析結果は、辞書による最長一致、機械学習による解析、正解ラベルの3種類5セットを用い、それぞれに対する事実性解析結果を比較した結果、機能表現の意味ラベル体系を再構築し、辞書を拡充することで、事実性解析を改善できること、機能表現の曖昧性解消が事実性解析に貢献すること、機能表現だけでは解くことが難しい事例がある程度存在していることを示した。機能表現だけでは解析が容易ではない事例を解くために、機能表現から事実性を演算するためのルールの拡張、一部の副詞や助詞などを考慮した解析、同じ意味ラベルをもつ機能表現の用法の区別に取り組んでいく必要がある。

謝辞 本研究は、文部科学省科研費(23240018)の一環として行われた。

参考文献

- [1] Marie-Catherine de Marneffe, Christopher D. Manning, and Christopher Potts. Did it happen? the pragmatic complexity of veridicality assessment. *Computational Linguistics*, Vol. 38, No. 2, pp. 301–333, 2012.
- [2] Kazuya Narita, Junta Mizuno, and Kentaro Inui. A lexicon-based investigation of research issues in japanese factuality analysis. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pp. 587–595. Asian Federation of Natural Language Processing, 2013.
- [3] Roser Saurí and James Pustejovsky. FactBank: a corpus annotated with event factuality. *Language resources and evaluation*, Vol. 43, No. 3, pp. 227–268, 2009.
- [4] Roser Saurí and James Pustejovsky. Are you sure that this happened? assessing the factuality degree of events in text. *Computational Linguistics*, Vol. 38, No. 2, pp. 261–299, 2012.
- [5] 磯崎秀樹, 東中竜一郎, 永田昌明, 加藤恒昭, 奥村学 (監修). 質問応答システム. コロナ社, 2009.
- [6] 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.
- [7] 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.
- [8] 水野淳太, 成田和弥, 乾健太郎, 大竹清敬, 鳥澤健太郎. 拡張モダリティ解析器の試作と課題分析. ALAGIN & NLP 若手の会 合同シンポジウム, 2013.
- [9] 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回年次大会発表論文集, pp. 852–855, 2010.
- [10] 上岡裕大, 成田和弥, 菅野美和, 水野淳太, 乾健太郎. 日本語文における機能表現意味ラベル付与と事実性解析への効果. 情報処理学会 第77回全国大会予稿集, 2015.