

委託事業報告書

モダリティタグ付きコーパスの特徴分析

成田 和弥 福原 裕一 乾 健太郎

東北大学 大学院情報科学研究科

2012年3月8日

目次

1	はじめに	1
1.1	本研究の背景	1
1.2	本研究の目的	3
1.3	本論文の構成	4
2	事実性解析課題	5
2.1	事実性	5
2.2	本研究における事実性解析の問題設定	9
2.2.1	態度表明者	9
2.2.2	時制	10
3	関連研究	12
3.1	タグ体系およびコーパス構築に関する研究	12
3.2	解析手法に関する研究	17
3.3	生物医学分野における研究	19
3.4	事実性に影響を与える要素に関する研究	20
3.4.1	機能語	20
3.4.2	内容語	21
3.4.3	スコープ	21
4	日本語事実性解析器の構築	23
4.1	事実性の定義	24
4.2	使用する語彙知識	25
4.3	解析モデル	28
5	実験・誤り分析	31
5.1	実験	31
5.2	評価	32
5.3	肯否極性における誤りの分析	35

5.3.1	語彙知識のカバレッジの問題	35
5.3.2	語彙的手がかりの多義性の問題	36
5.3.3	手がかりとなる表現のスコープの問題	37
5.3.4	その他の誤り	38
5.4	確信度における誤りの分析	38
6	誤りの情報をもたせたデータの試作	40
6.1	分析結果に基づく今後の指針	40
6.2	誤り情報のラベル付データの作成	43
7	おわりに	45
7.1	まとめ	45
7.2	今後の課題	45
	参考文献	47

目 次

1	Sauriら (2009)[1, 2] の 2 軸で表す事実性 (縦軸: Modality, 横軸: Polarity)	13
2	江口ら (2010) [3] が用いた条件付き確率場のグラフ構造	18
3	日本語事実性解析器の概要	23
4	(??) に対する事実性解析アルゴリズムの動き	29
5	今後の日本語事実性解析器	41

表 目 次

1	確信度と肯否極性の組み合わせによる事実性の値	25
2	機能表現辞書の意味クラス例と事実性への影響	26
3	手がかり表現辞書の見出し語例と対応する事実性	27
4	<真偽判断> ラベルと事実性との対応	32
5	それぞれのラベルに対する評価	33
6	Confusion Matrix	33
7	確信度に関する評価	34
8	確信度に関する Confusion Matrix	34
9	肯否極性に関する評価	34
10	肯否極性に関する Confusion Matrix	34
11	事象の位置に着目した正解率	38

1 はじめに

1.1 本研究の背景

自然言語処理の分野において，文章を解析するための技術は古くから研究されており，これまでに様々なツールが開発されてきた．例えば，形態素解析器および構文解析器は，その最も基本的なツールであり，誰もが自由に利用可能な解析器が存在する．形態素解析器としては，Chasen¹や MeCab²，JUMAN³，構文解析器としては，CaboCha⁴や KNP⁵といったものが利用可能である．そのような状況ではあるものの，文に記述されている事象が，実際に起こったことなのかそうでないことなのか，といった情報を，高精度で解析できるツールは，現在のところ，利用可能ではない．

しかしながら，文に記述されている事象が，実際に起こったことなのかそうでないことなのかは，重要な情報であり，このような情報を解析する技術は，情報抽出や含意関係認識などの応用に対して有用である．例えば，与えられた文章から情報を抽出するにあたり，その事象が事実であるのか否かを知ることは，必須の情報である．

- (1) a. 最近 商品 A を使い始めた。
- b. 商品 A を使ってみたい。
- c. 商品 A を使うことに抵抗がある。
- d. 商品 A を使ってみたけど難しい。

(1) の例文は，いずれも「商品 A を使う」という文脈が現れている文である．この中で，実際に商品 A を使っているユーザーの情報のみを抽出したい（今回は，(1a)，(1d) が実際に使っている，という解釈ができる）場合に，それぞれの文に

¹<http://chasen-legacy.sourceforge.jp/>

²<http://mecab.sourceforge.net/>

³<http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN>

⁴<http://code.google.com/p/cabocha/>

⁵<http://nlp.ist.i.kyoto-u.ac.jp/index.php?KNP>

において「商品 A を使う」という事象が実際に起こっていることなのか、といった情報を解析することは、非常に重要であるといえる。経験マイニング [4] や、言論マップ [5, 6]、インフルエンザ流行予測 [7] 等の先行研究においても、実際に起こったことなのかそうでないことなのか、といった情報は、重要な情報として扱われている。

本研究において扱う、事実性とは、文中のある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だけなのか、に関する情報である。本研究では、文献 [8] と同様に、事象は行為、出来事、状態の総称であると考える。

(2) 来週からこの雑誌の購入を中止しようと思う。

(2) では、「(雑誌を) 購入するコト」、「(来週から雑誌の購入を) 中止するコト」、「(来週から雑誌の購入を中止しよう) 思うコト」がすべて事象であり（下線部は、それぞれの事象の中心となる述語である）、このような事象を対象として、事実性を付与する。

事実性は、ある事象が実際に起こったことなのか、といった情報であるが、その事象が真に起こったことかはわからないため、著者あるいは文章中の登場人物による事象の成否の判断として事実性を解釈する。

(3) a. 彼はさきほど部屋を出た。

b. もう遅いから、彼は先に帰ったんだらう。

c. 問題が発生するのを防いだ。

例えば、(3a) 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。(3b) の事象「帰る」は、「だらう」という表現により、起こった可能性が高いことだと著者は判断している、と解釈できる。(3c) の事象「発生する」は、「防いだ」という表現により、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。

前述の通り、実際に起こったことなのかそうでないことなのか、といった事実性の解析は、情報抽出や含意関係認識などの自然言語処理の応用に対して有用であるが、実際は日本語事実性解析を行う上で、多岐に渡る言語表現を捉えきれて

おらず、いまだ十分な解析性能は実現されていない。また、内在する問題の分析・整理が不十分であることも原因の一つとなっている。そこで、高精度な日本語事実性解析器の実現をするために、事実性解析を行う上で、何が問題となっているかを分析し、明らかにする必要がある。

1.2 本研究の目的

我々は、事実性に影響を与える、多様な言語表現の存在に着目した。(3b)の事象「帰る」の事実性は「だろう」という表現に影響され、(3c)の事象「発生する」の事実性は「防いだ」という表現に影響されるように、事実性に影響を与える手がかりとなる表現が多数存在している。

(4) 彼女はその商品を買わなかつたことを否定した。

また、(4)では、事実性に影響を与える表現である「ない」および「否定する」の両方が作用することにより、事象「買う」は、実際に起こったことであると著者は判断している、と解釈できる。このように、事実性に影響を与える表現が構成性を持ち、その組み合わせによって事実性を決定できる場合が、少なからず存在している。

そこで我々は、事実性に影響を与える語彙知識に焦点を当て、その語彙知識の組み合わせによって解決できる問題、できない問題を明らかにする。本研究では、その細分類として、現在利用可能な語彙知識によってどれだけの問題がカバーできるのか、カバーできなくても語彙知識の拡充および拡張によって解決できるのか、語彙知識の曖昧性はどの程度問題となるのか、局所的な語彙知識だけで解決できない問題はあるのか、といった論点に着目する。語彙知識と構成性に基づく日本語事実性解析器を構築することで、これら論点に基づいた問題の分析を行った。

具体的には、Sauríら (2007) [9] の英語を対象とした事実性解析モデルをもとに、既存の語彙知識を利用した日本語事実性解析器を構築し、その誤り分析を行った。Sauríらの解析モデルは、事実性に影響を与える表現を手がかりとし、それらの組み合わせによって事実性解析を行う。このモデルは、語彙知識の組み合わせによって構成的に事実性を解析でき、後述する確信度と肯否極性に事実性を分割し

て、個別に分析を行えるといった特長がある。そのため、問題の分析という本研究の目的に適している。

1.3 本論文の構成

本論文の構成は以下の通りである。

まず、2節において、事実性とは何かを、関連概念であるモダリティおよび否定と共に述べ、本研究でどのように事実性解析問題を扱うか、を説明する。3節では、事実性およびその周辺における先行研究について述べる。4節において、既存の語彙知識と Sauriら (2007) [9] の構成性に基づくモデルをベースに、日本語事実性解析器を構築する。5節では、日本語事実性解析における問題を分析・整理するため、4節で構築した日本語事実性解析器の評価実験を行い、その誤り分析を行う。そして、6節において、誤り分析の結果に基づいた今後の指針と、そのための誤り情報を付与したデータを試作する。最後に、7節でまとめを行う。

2 事実性解析課題

前節にて、事実性とは何かを例文を用いて簡潔に説明し、その解析の有用性を述べた。本節では、まず 2.1 節で、本研究で扱う事実性がどのような定義なのかを、関連概念と共に述べる。そして、2.2 節で、本研究において、どのように事実性を扱い、どのような問題を解くかを述べる。

2.1 事実性

事実性とは、文中のある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だけなのか、に関する情報である。その事象が真に起こったことかはわからないため、著者あるいは文章中の登場人物による事象の成否の判断として事実性を解釈する。

また、非未来における事実性と、未来における事実性とでは、意味合いが異なる。もしその事象が未来のことでないならば、態度表明者から見て、事象が実際に成立した（あるいは、実際には成立していない）ことであるという断定、あるいは、事象の真偽を確認していないことによる推量（断定保留）を表す。しかしながら、もしその事象が未来のことであるならば、事象に対する断定は出来ず、真偽が定まっていないことに対する可能性を述べた推量（断定保留）を表すことになる。

事実性に大きく関連する概念として、情報発信者の主観的な態度（モダリティ）、および、肯定または否定があげられる。本研究における事実性は、言語学における「真偽判断のモダリティ」と、肯定または否定の組み合わせに相当しているといえる。

言語学において、文全体の意味は、事態を表す意味領域、および、話し手の態度を表す意味領域に分けられ、前者は「命題」、後者は「モダリティ」と呼ばれる⁶[10, 11, 12]。

⁶Fillmore (1968) [10] は、文が proposition (命題) と modality (モダリティ) の 2 つで構成される、という見方をした代表的な研究であり、「Sentence = Proposition + Modality」という書き換え規則を導入している。

(5) a. ねえ、どうやら昨夜激しく雪が降ったようだよ。

b. [ねえ、どうやら [昨夜激しく雪が降った] ようだよ]

c. [ねえ [どうやら—ようだ] よ]

例えば、(5a)の文は、(5b)のような意味領域からなり、「昨夜激しく雪が降った」が事態を表す命題の領域、「ねえ、どうやら—ようだよ」が態度を表すモダリティの領域である。さらに、このモダリティの領域は、文の意味的階層構造という観点から、事態に対する判断を表す領域、および、表現・伝達を表す領域に分けられ、前者を「判断のモダリティ」、後者を「発話のモダリティ」と呼ぶ⁷。(5a)におけるモダリティの領域は、(5c)のような2つの領域に分けられ、「どうやら—ようだ」の部分が判断のモダリティの領域を、「ねえ—よ」の部分が発話のモダリティの領域を表している。

益岡 (2007) [11] は、モダリティを以下の7種類のカテゴリーに分類している。

真偽判断のモダリティ 事態が成り立つかどうかの真偽性(断定か、非断定か)を表す。非断定の判断は、断定こそできないものの何らかの判断は下すという「定判断」と、疑問文など、真偽の判断がまったく下せない「不定判断」に大別される。さらに、定判断には、「だろう」「～(よ)う」などの表現によって、真であるとの確信が持てなかったり、聞き手との関係で断定を差し控えたりする場合を表す「断定保留」、「かもしれない」(可能性)、「にちがいない」(必然性)といった、確からしさの程度を表す「蓋然性判断」、「ようだ」「らしい」など、ある証拠に基づいて推定を行うことを表す「証拠性判断」、「はずだ」といった、推論から得られる当然の帰結を表す「当然性判断」という4つの下位類に分類される

⁷命題の領域についても、さらに細分類がされ、事態のタイプを表す「一般事態」の領域、および、その事態が位置する時空間を表す「個別事態」の領域に分けられる。

(6) [昨夜 [激しく雪が降る] た]

(5b)における命題の領域は、(6)のような2つの領域に分けられ、「激しく雪が降る」の部分が「激しい降雪」という事態のタイプを表し、「昨夜—た」の部分がその事態が位置する時空間(今回は時間)を表す。

価値判断のモダリティ 事態が是認されるかどうかの妥当性（現実像か，理想像か）を表す．理想像を表す形式は「べきだ」「ほうがよい」などの「適当」，「なければいけない」「しかない」などの「必要」，「てもよい」「てはいけない」などの「容認・非容認」といった意味に分類される

発話類型のモダリティ 文法的な観点から見た，表現・伝達の類型的な機能を表す．話し手の認識を表す「演述型」，話し手の内面にある感情や意志を表す「情意型」，聞き手に判断を求める，あるいは，話し手の判断が定まらない状況を表す「疑問型」，聞き手に何らかの対応を求める「要求型」，話し手の感情の発露を表す「感嘆型」のいずれかの態度を表す

丁寧さのモダリティ 対話文において，聞き手に対する丁寧さの有無を表す「です」「ます」といった表現がこれに該当する

対話態度のモダリティ 聞き手に対する情報提示の調整を表す「ね」「よ」「よね」などの終助詞がこれに該当する

説明のモダリティ 種々の説明を表す「のだ」「わけだ」「ものだ」などの表現があり，各表現によって用法が変化する．例えば「のだ」には「叙述様式説明」「事情説明」「帰結説明」「実情説明」「当為内容説明」といった用法が存在する

評価のモダリティ 述部以外の付加部に出現するモダリティ要素であり，事態に対する評価を表す「あいにく」「さいわい」「親切にも」のように，通常文頭に現れる

「判断のモダリティ」の下位カテゴリとして「真偽判断のモダリティ」および「価値判断のモダリティ」，「発話のモダリティ」の下位カテゴリとして「発話類型のモダリティ」および「丁寧さのモダリティ」，「対話態度のモダリティ」を認定し，さらに，特殊なモダリティとして，判断のモダリティの領域と発話のモダリティの領域の両方に関係する「説明のモダリティ」，および，述語以外の部分に現れる要素である「評価のモダリティ」を認定している．

文献 [12] では，モダリティを以下の4つのタイプに大別している．

文の伝達的な表し分けを表すモダリティ 聞き手に情報を伝達する「叙述のモダリティ」, 聞き手から情報を聞き出そうとする機能をもつ「疑問のモダリティ」, 話し手の行為の実行を表す「意志のモダリティ」, 話し手の行為の実行を前提として, 聞き手に行為の実行を求める「勧誘のモダリティ」, 命令など, 聞き手に行為の実行を求める「行為要求のモダリティ」, 物事に触れて引き起こされる話し手の感動の気持ちを表す「感嘆のモダリティ」といったタイプに分類される。このモダリティ全体で, 益岡 (2007) [11] の「表現類型のモダリティ」に相当する

命題が表す事態のとらえ方を表すモダリティ 命題によって表される事態に対する, 必要, 不必要, あるいは許容できる, できないといった, 話し手の評価的なとらえ方を表す「評価のモダリティ」, 断定や推量, 可能性や必然性, 推定や伝聞など, 事態に対する話し手の認識的なとらえ方を表す「認識のモダリティ」といったタイプに分類される。「評価のモダリティ」は益岡 (2007) [11] の「価値判断のモダリティ」⁸, 「認識のモダリティ」は「真偽判断のモダリティ」に相当する

先行文脈と文との関係付けを表すモダリティ その文を先行文脈と関係があるものとして示すことによって, 先行文脈の内容が聞き手に理解させやすくする働きをもつ。

(7) 遅れてすみません。渋滞していたんです。

(7) の文では, 先行文脈で表されている, 話し手が遅れたという事実に対して, 道路が渋滞していたという後続文で表されている事実が, 「のだ」によって関係付けられて示されている。このモダリティは, 益岡 (2007) [11] の「説明のモダリティ」に相当する

聞き手に対する伝え方を表すモダリティ 聞き手に対してその文を通常スタイルで伝えるか, 丁寧なスタイルで伝えるかという, スタイルの選択に関わ

⁸益岡 (2007) [11] にも「評価のモダリティ」という分類が存在するが, 文献 [12] における「評価のモダリティ」とは別のものである。

る「丁寧さのモダリティ」、話し手の認識状態を示したり、聞き手に伝えるにあたっての微調整をしたりする「伝達態度のモダリティ」といったタイプに分類される。それぞれ、益岡 (2007) [11] の「丁寧さのモダリティ」、「対話態度のモダリティ」に相当する

しかしながら、この分類では、益岡 (2007) [11] の「評価のモダリティ」に相当するものが考慮されていない。以上のように、我々の調べた限りでは、用語も含めて、完全に統一された見解は存在しないようであるが、概念としてはおおむね対応がみられる。

この中で、事象の真偽に対する書き手の確信度を表した「真偽判断（評価）のモダリティ」は、本研究における事実性と非常に近い概念である。また、文献 [13] によると、事態の成立を表すことを肯定といい、事態の不成立を表すことを否定という。以上のことから、本研究における事実性は、言語学における「真偽判断のモダリティ」と、肯定または否定の組み合わせに相当する。

2.2 本研究における事実性解析の問題設定

前節では、事実性の定義を関連概念と共に述べた。しかしながら、本研究においては、問題を簡略化するため、その事実性をある程度限定して扱うこととした。以下、本節では、態度表明者、時制といった側面から、事実性解析問題を限定していく。

2.2.1 態度表明者

2.1 節で、「その事象が真に起こったことかはわからないため、著者あるいは文章中の登場人物による事象の成否の判断として事実性を解釈する」と述べた。この事象の成否の判断を行っている「著者あるいは文章中の登場人物」が態度表明者である。

- (8) a. 太郎は花子が帰ったことを知っている。
- b. 太郎は花子が帰ったことを知らない。

例えば、(8a)、(8b)では、ともに「帰る」という事象が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。しかしながら、これはあくまで著者からみた事実性であり、文中の登場人物（例えば太郎）からみた事実性も存在する⁹。(8a)では、太郎が知っていることなので、太郎も「帰る」という事象が実際に起こったことだと判断している、と解釈できる。しかし、(8b)では、太郎が知らないことなので、太郎は「帰る」という事象が実際に起こったことかどうかわかっていない、と解釈できる。

このように、厳密に事実性を扱う上では、著者あるいは文章中の登場人物といった、態度表明者を考慮する必要がある。しかしながら、本研究では、タスクを簡略化し、分析をしやすくするため、文中の登場人物からみた事実性は考えず、著者の事実性判断を認識する問題に限定する。

2.2.2 時制

2.1節で、未来、非未来における事実性の意味合いの違いを述べた。しかしながら、その事象が未来に起こることか、非未来に起こっていることか、といった時制を判別するだけでも、大きな問題となると考えられる。

そこで、本研究においては、時制を判断する問題を切り離し、時制を考慮した解析を行わない。すなわち、システム内において、未来、非未来に関する事実性の意味合いの区別を考慮せず、同等に扱う。

また、文中で事実性の変化が含意される場合が存在する。そのような場合、明確な変化が示されていれば、変化後の事実性を主要な要素と捉え、その事象の事実性とする¹⁰。

(9) a. 彼は商品 A を使用するのを中止した。

b. 彼は商品 A を使用していた。

⁹真には文中の登場人物の判断はわからないため、厳密には「著者からみた登場人物からみた」事実性である。

¹⁰松吉ら (2010) [8] の拡張モダリティタグ体系では、このような、変化が含意される場合、「成立から不成立」のように、変化を含む真偽判断ラベルを定義している。

例えば，(9a)では，過去に「使用する」という事象が，著者から見て実際に起こっているものの，それを中止しているので，「使用する」という事象の事実性は更新され，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．それに対して，(9b)では，過去に「使用する」という事象が，著者から見て実際に起こったことだとわかるが，その後の変化が明示されていないため，事実性の更新を行わず，その時点で事象が実際に起こったと著者が判断している，という解釈に留める．

3 関連研究

モダリティおよび否定をはじめとした，さまざまな概念が事実性に関連している [14]．本節では，事実性に関連する情報を付与するためのタグ体系・コーパス構築，および，その解析手法，そして機能語やスコープなどの事実性に影響を与える要素に関してまとめる．近年このような研究は，純粋な自然言語処理の分野ばかりでなく [15, 16, 9, 2, 1, 17, 18, 4, 8, 3, 19, 20, 21, 22, 23]，生物医学分野においても盛んに行われている [24, 25, 26]．

本節では，3.1 節で，タグ体系およびコーパス構築に関する研究，3.2 節で，その解析手法に関する研究，3.3 節で，生物医学分野における研究，3.4 節で，事実性に影響を与える要素に関する研究について述べる．

3.1 タグ体系およびコーパス構築に関する研究

事実性およびその周辺情報をマークアップするためのタグ体系およびその付与基準の関連研究として，Prasad らによる Penn Discourse TreeBank (2008) [17] における attribution タグや，Sauri らによる FactBank (2009) [1, 2]，松吉ら (2010) [8] や川添ら (2011) [20, 21] などがある．

Prasad ら (2008) [17] は，Penn Discourse TreeBank (PDTB) の談話関係とその項に対して，情報の発信源とその情報が事実であるかどうかを判断するために，attribution という属性タグを付与している．attribution タグは以下の 4 種類の要素からなる．

Source 情報の発信源や判断している主体を表し，テキストの著者 ("Wr")，テキストにより導入される特定の動作主 ("Ot")，任意の個人や組織 ("Arb")，前の関係を継承 ("Inh") のいずれかの値をとる

Type 対象が事実である度合いを推論するための意味クラスであり，主張 ("Comm")，信念 ("PAtt")，事実 ("Ftv")，成り行き ("Ctrl") のいずれかに分類される

Scopal Polarity 作用の及ぶ範囲(スコープ)を考慮した否定の有無 ("Neg" ,"Null")

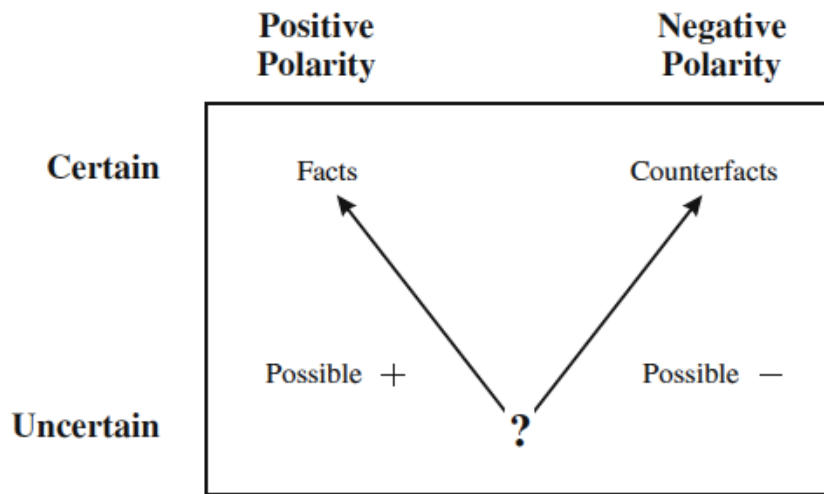


図 1: Sauríら (2009)[1, 2] の 2 軸で表す事実性 (縦軸: Modality, 横軸: Polarity)

Determinacy 上記 3 つの要素のいずれかが, より広い文脈で非決定的になりうるかどうか (“Indet”, “Null”)

彼らは, Type において, 動詞の意味クラスのみを用いて, 事実性を表している。しかし, 助詞の “may” や, 副詞の “probably” のように, 動詞以外にも事実性に影響する要素があるため, 事実性を動詞の意味クラスのみで表すのは不十分である。

Sauríら (2009)[1, 2] は, 事象を対象とし, 以下の 2 つ組のタグによって事実性を定義した。

Modality 事実らしさに対する態度表明者の確信度。Certain, Probable, Possible, Underspecified の 4 種類で表す

Polarity 事象に対する確信の方向。Positive, Negative, Underspecified の 3 種類で表す

そして, 事象とその時制, 肯否, モダリティが付与された TimeML[16] に対して, Modality と Polarity を態度表明者 (source) ごとに付与する枠組みを提案し, FactBank と呼ばれるコーパスを構築した。この 2 軸で表す事実性に基づく概念図を図 1 に示す。本研究では, この Sauríらの枠組みに基づいて事実性を定義し (4.1 節で後述する), Modality を確信度, Polarity を肯否極性と呼ぶ。

松吉ら (2010) [8] は、以下の 7 項目からなる拡張モダリティタグ体系を設計し、それを付与したコーパスを構築した。

態度表明者 態度を表明している人物。「wr:筆者」(態度表明者が書き手である)や「wr:筆者_arb:不特定」(態度表明者が不特定の個人や集団である、と書き手が述べている)、「wr:筆者_1:太郎」(態度表明者が太郎である、と書き手が述べている)のように記述する

時制 態度表明時に対する相対的な時制。「未来」または「非未来」(過去、現在、脱時間的のいずれか)の 2 種類で表現する

仮想 文章に記述される情報が事実であるのか、それとも、単なる仮想的な話であるのかを示す。「条件」(事象が条件として仮想的に述べられている)、「帰結」(事象が仮想的な条件の帰結として述べられている)、「0」(2 つのいずれでもない)の 3 種類で表現する

態度 命題に対する態度表明者の主観的な態度であり、益岡 (2007) [11] における「発話類型のモダリティ」に相当する。「叙述」(事象が表す内容や態度表明者の判断などを情報の受け手に伝える)、「意志」(態度表明者が、自分自身の行為の実行、もしくは、非実行を決定)、「欲求」(態度表明者が、自分自身の行為の実行、もしくは、非実行を望んでいる)、「働きかけ-直接」(態度表明者が、直接、相手に対して、行為の実行、もしくは、非実行を求める)、「働きかけ-間接」(態度表明者が、間接的な表現手段を用いて、行為の実行、もしくは、非実行を相手に求める)、「働きかけ-勧誘」(態度表明者が、態度表明者の行為を前提として、行為の実行、もしくは、非実行を相手に誘いかける)、「許可」(態度表明者が、事象の成立、もしくは、不成立を許容する)、「問いかけ」(態度表明者にとって不明なことがあるために、その事象に対して態度表明者の判断が成り立たない)の 8 種類で表現する

真偽判断 事実の真偽に対する態度表明者の確信度であり、Sauriら (2009)[1, 2] の事実性に相当するといえる。「成立」(肯定の断定)、高確率(肯定の推量)、低確率(否定の推量)、「不成立」(否定の断定)、0(詳細不明)の 5 種類に

加え、これらの間を表す真偽の変化を含意するラベルとして、「成立から不成立」、「高確率から低確率」、「低確率から高確率」、「不成立から成立」の4種類、計9種類で表現する

価値判断 事象成立の望ましさを表す極性情報「ポジティブ」(態度表明者が、事象成立が望ましいと判断している)、「ネガティブ」(態度表明者が、事象成立は望ましくないと判断している)、「0」(事象成立の望ましさについての態度表明者の判断が、文章中に記述されていない)の3種類で表現する

焦点 否定、推量、問いかけが実際に作用を受ける要素「否定(仕事で)」のように、対象とする形態素(列)を記述することで表現する。

彼らはその後、拡張モダリティタグ体系をさらに整理し、松吉ら(2011)[19]では、<態度表明者>、<相対時>(時制に対応)、<仮想>、<態度>、<真偽判断>、<価値判断>の6項目で拡張モダリティを設計している。項目の1つであった「焦点」は、モダリティの周辺情報であるが、ラベル選択ではなく、対象文内の形態素(列)を直接記述する必要があるため、拡張モダリティの補足情報として記述されている。また、拡張モダリティの補足情報として、「ほとんど~ない」や「めったに~ない」など、否定を伴う程度・頻度の表現を「程度」として記述している。この拡張モダリティタグ体系は、事象のモダリティ情報を整理して表すことができている。しかしながら、我々はこの拡張モダリティタグ体系を用いなかった。この理由としては、この中で実際に重要であるのは、<真偽判断>と呼ばれる事実性であると、我々は考えているからである。この拡張モダリティタグ体系を用いた解析(3.2節で後述する)は、項目間の依存関係を考慮することが有用であるが、そのため処理が複雑化してしまう。そこで我々は、この拡張モダリティタグ体系の中で最も重要であると考えた、事実性のみを焦点を当てた解析、分析を行う。

川添ら(2011)[20, 21]は、テキストに現れる事実とそれ以外の情報との区別、また推量や仮定などの間に見られる確実性の差を自動的に識別することを目指している。そのために、以下のように「確実性」に影響を与える言語表現を分析・分類し、それに従ってそれらの言語表現およびその影響を与える範囲(スコープ)

をアノテーションしたコーパスを構築している。

様相表現 「ようだ」や「だろう」のような、文の内容に対する書き手/語り手の認識・判断を表す表現であり、益岡(2007) [11] で「真偽判断のモダリティ」と呼ばれているものに対応する。叙実表現，証拠推量表現，認識的推量表現，他人の認識を表す表現，不定判断・疑問表現，比況表現，反叙実表現に分類される

否定表現 「わけではない」のような通常の否定と、「～(というのは)正しくない」のようなメタ否定 [27] を区別している。通常否定の解釈は命題を否定する解釈だが，メタ否定の解釈は先行発話そのものが不適切であると異を唱える解釈である

条件表現 「たら」や「なら」、「としても」のような条件を表す表現であり，事実的条件表現，予測的条件表現，認識的条件表現，一般的条件表現，半事実条件表現に分類される

この「確実性」は、「テキストの書き手が判断する，命題の内容が真である確率」という意味で定義されており，我々の事実性と対応している。

小橋ら(2011) [23] は，階層意味論という側面から，事柄が真なのか偽なのか，事柄が望ましいか望ましくないか，といった心的態度(モダリティ)を分類し，アノテーションを行っている。彼らは「べき(だ)」「だろう」「か」といった助動詞や終助詞などの，モダリティに属する文法形式をアノテーションするだけでは心的態度を網羅することはできず「ことを確信している」「と非常に良さそうだ」等，文法形式以外の心的態度を表す表現も対象としたアノテーションを行うことで，心的態度を網羅することを目的とした。その際，階層意味論で規定されるモダリティの概念を用いることで，一貫したアノテーションを行えると考え，アノテーションの不一致を引き起こす要因とその対策について論じた。彼らは，以下のように心的態度を細分類している。

真偽判断のモダリティ 事柄の真理値について肯定的あるいは否定的に断定，推定する心的態度

判断保留のモダリティ 疑問，質問態度など，事柄の真理値について判断を保留し中立的な立場を表明，含意する心的態度

是非判断のモダリティ 真理値を評価している点では真偽判断と一緒だが，事柄が既定的（その情報が既に談話の世界に提示済）である点で異なる

価値判断のモダリティ 事柄に対する情緒的な反応や評価を指す心的態度

拘束判断のモダリティ 事柄が指し示す未来の行為を拘束することに関する書き手，話し手の立場を表す心的態度

本研究で扱う事実性は、「真偽判断のモダリティ」「判断保留のモダリティ」「是非判断のモダリティ」に対応している（彼らはこの3つをまとめて「真偽判断系」の心的態度と呼んでいる）。彼らの調査では，この真偽判断系において，アノテーションの不一致を引き起こす場合が多く，その要因としては，

- 述語が書き手または他者の心的態度を表すもの
- 後続する表現が他者の心的態度や条件，伝聞を表すもの

があげられており，これらに関してはアノテーションする対象から外すべきだとされている。

3.2 解析手法に関する研究

解析手法に関する研究としては，機械学習に基づく手法 [24, 25, 4, 18, 3, 19] や，人手で構築した語彙的・統語的な知識を利用したパターンベースの手法 [24, 9] があげられる。

原ら (2008) [18] は，事象の事実性情報を，<時間情報（極性を含む），モダリティ，モダリティの時間情報（極性を含む）> の3種類で表現し，SVMを学習器に用いた解析手法を提案している。素性は，予測対象の文節，その前後の文節，文全体を区別した上で，品詞と原型を組み合わせたものを用いた。

Inuiら (2008) [4] は，原らの提案するタグ体系を整理統合し，条件付き確率場を学習器として用いた解析手法を提案した。実験の結果，SVMを用いるよりも，

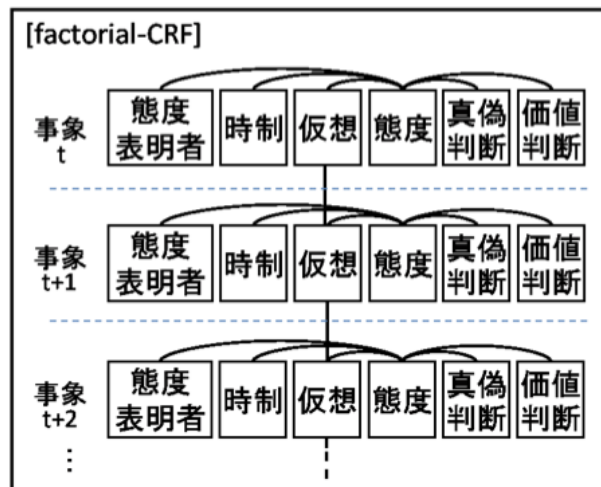


図 2: 江口ら (2010) [3] が用いた条件付き確率場のグラフ構造

タグ間の依存関係を考慮できる条件付き確率場を用いたほうが、精度が高いことが示されている。

江口ら (2010) [3] は、拡張モダリティタグ体系における項目間には強い依存関係があり、また、同じ文に存在する複数の事象間にも依存関係があることに着目した。そこで彼らは、項目間、および事象間の依存関係を考慮できる条件付き確率場を用いた拡張モダリティ解析システムを構築した。図 2 に、彼らが用いた条件付き確率場のグラフ構造を示す。核となる述語が存在する文節、その前後の文節などの形態素情報の n-gram や、意志動詞の有無、拡張モダリティに影響を与える述語を収録したモダリティ解析手がかり表現辞書などを素性として用い、項目間、事象間の依存関係を考慮したモデルの有効性を示している。

松吉ら (2011) [19] は、モダリティ解析の精度向上に必要な技術や言語資源について理解を深めるため、最大エントロピーモデルを用いた拡張モダリティ解析システムを試作した。その中の 1 つのタグである「態度」に着目して誤り分析を行い、語義曖昧性解消や連体節内の述語に及ぼす影響の解明、節間の意味的關係の認識などが、「態度」に関するモダリティ解析の精度向上に向けた課題であることを述べている。

Sauri ら (2007) [9] は、事象の成立に影響を与える手がかり表現を利用し、態度

表明者ごとに，確信度と肯否極性で表される事実性を，依存構造木の根から伝搬させて解析する，パターンベースの決定的アルゴリズムを提案した．このモデルは，手がかり表現の情報を持つ語彙知識が与える事実性を組み合わせて，事象の事実性解析を行うという構成性をもつため，どのような解析過程でその結果が出力されたのかが明確である．また，事実性を確信度と肯否極性という2軸に分割して解析を行うため，それぞれを個別に分析することができる．

機械学習に基づく手法は，出力結果がどのような要素に基づいて出力されたかがわかりづらく，本研究の目的とする，日本語事実性解析における問題の分析に対して適当であるとはいえない．そこで本研究では，構成性に基づいた Sauriら (2007) の手法を日本語に適用することで，特に語彙知識に関する問題の分析を行った．

3.3 生物医学分野における研究

生物医学分野においても，近年このような研究は，盛んに行われている [24, 25, 26]．生物医学分野では，例えば，感染症情報など緊急の判断が必要とされる情報の取得の際に，不要な情報を取り除いて効率的に情報を取り出すことや，情報の確実性・信憑性判断に関わる人的コストを減らせる，などの応用が考えられる [20]．

Light ら (2004) [24] は，生物医学分野の論文のアブストラクト内の，推測を表す文を人手によってアノテーションしたコーパスを構築した．そして，*suggest*, *potential*, *likely*, *may* などの14のキーワードの有無により，推測を表す文かどうかを判断した結果が，SVMによる学習結果と同等のパフォーマンスを示した，と報告している．しかしながら，日本語においては，言語表現が多様であり，このようなキーワードによる単純な判断は難しいものと思われる．

Medlock ら (2007) [25] も，生物医学分野のテキストを対象に，推測を表す文を人手によってアノテーションしたコーパスを構築した．彼らは，Light らと異なり，推測を表す文に出現する，特徴的な表現を認識することに重きを置き，スコープはその表現を含む文全体とみなしている．そして，特徴的な表現を自動的に認識する教師あり学習の確率モデルを提案した．

生物医学分野のテキストを対象としたコーパスも構築されている。BioScope (2008) [26] は、否定表現、様相表現、そして、それらのスコープをマークアップしたコーパスであり、3.4.3 節で後述するような、スコープを特定する研究等に利用されている。

3.4 事実性に影響を与える要素に関する研究

事実性に影響を与える要素として、機能語、内容語、否定や推量の作用する範囲（スコープ）といった要素が存在する。

(10) a. もう遅いから、彼は先に帰っただろう。

b. 問題が発生するのを防いだ。

例えば、(10a) の事象「帰る」の事実性は、「だろう」という機能語に影響を受け、(10b) の事象「発生する」の事実性は、「防いだ」という内容語に影響を受けている。また、このような事実性に影響を与える表現において、その作用する範囲（スコープ）を特定する研究も進められている。

本節では、3.4.1 節で機能語、3.4.2 節で内容語、3.4.3 節でスコープについて、それぞれの関連研究を述べる。

3.4.1 機能語

事実性に影響を与える表現として、「～ない」「～だろう」などの機能表現があり、このような日本語機能表現の意味に関連した研究も多く進められている [28, 29, 30, 31]。

例えば、機能表現を網羅的に集めた辞書として、日本語機能表現辞書「つつじ」 [28] が利用できる。この辞書は、日本語の機能表現の表層形約 17,000 種に対して、その ID、意味、文法的機能、音韻的变化などを網羅的に収録した辞書であり、機能表現の意味として、「対象」や「目的」、「名詞化」など、89 種類のラベルが定義されている。その中には「推量」や「否定」、「疑問」など、事実性に影響を与え

る意味の分類も多数収録されている。また、これらの機能表現は表層による曖昧性を持つため、その意味を同定する研究も行われている [29, 30]。

田辺ら (2011)[31] は、日本語機能表現の中で、特に文末表現に焦点をあてた推論を行っている。彼らは、日本語文末には、必要性を表す「なければならない」、様態を表す「ようだ」のように、発話者の、バラエティに富んだ主観を表す表現が含まれていることから、機能語性 Multiword Expression を意識した適切な文末表現の扱いが必要であると考えた。そして、含意関係認識における機能語の重要性に着目し、日本語文末表現に対応する意味構造間で定義できる含意関係（推論ルール）について論じている。

3.4.2 内容語

事実性に影響を与える内容語に関する研究としては、江口ら (2010) [3] が構築した、モダリティ解析手がかり表現辞書がある。彼らは、コーパス作成時に、拡張モダリティに影響する動詞、形容詞が存在することを明らかにした。そして、こうした動詞、形容詞が直前の事象に与える影響を記述した、モダリティ解析のための手がかり表現辞書を作成した。

しかしながら、内容語が事実性にどのような影響を与えるかを整理した研究は、我々の調べた限りでは他に見当たらず、研究の余地が多分に残されている。

3.4.3 スコープ

事実性を決定する上で、否定や推量などの作用が及ぶ範囲（スコープ）を決定することは重要だと考えられる。

否定表現および推量表現のスコープを同定する研究は、近年盛んに行われており [32, 33]、例えば、3.3 節で述べた BioScope (2008) [26] は、否定表現、様相表現、そして、それらのスコープをマークアップしたコーパスであり、スコープを特定する研究に広く利用されている。また、CoNLL-2010 Shared Task において *Learning to detect hedges and their scope in natural language text* というテーマが扱われた [34]。そこでは、機械学習を用いた手法 [35, 36] や人手によるルール

ベースの手法 [37, 38] , 機械学習と人手のルールの両方を合わせた手法 [39] などが提案されている .

また , 本年も *SEM Shared Task 2012 において *Resolving the Scope and Focus of Negation* というテーマが扱われている¹¹ .

¹¹<http://www.clips.ua.ac.be/sem2012-st-neg/>

4 日本語事実性解析器の構築

日本語事実性解析における問題を分析・整理するため，Sauriら (2007)[9] の英語事実性解析モデルをもとに，日本語に対する事実性解析器を構築した．図 3 に，日本語事実性解析の流れを示す．今回構築した日本語事実性解析器は，構文解析結果を入力とし，各事象に対する事実性を出力する．Sauriらの構成性に基づく解析モデルをベースに，松吉ら (2007) の日本語機能表現辞書「つつじ」[28] および江口ら (2010) のモダリティ解析手がかり表現辞書 [3] といった既存の語彙知識を利用した解析モデルを構築した．

本節では，4.1 節で事実性を定義し，4.2 節で使用する語彙知識，4.3 節で解析モデルについて述べる．

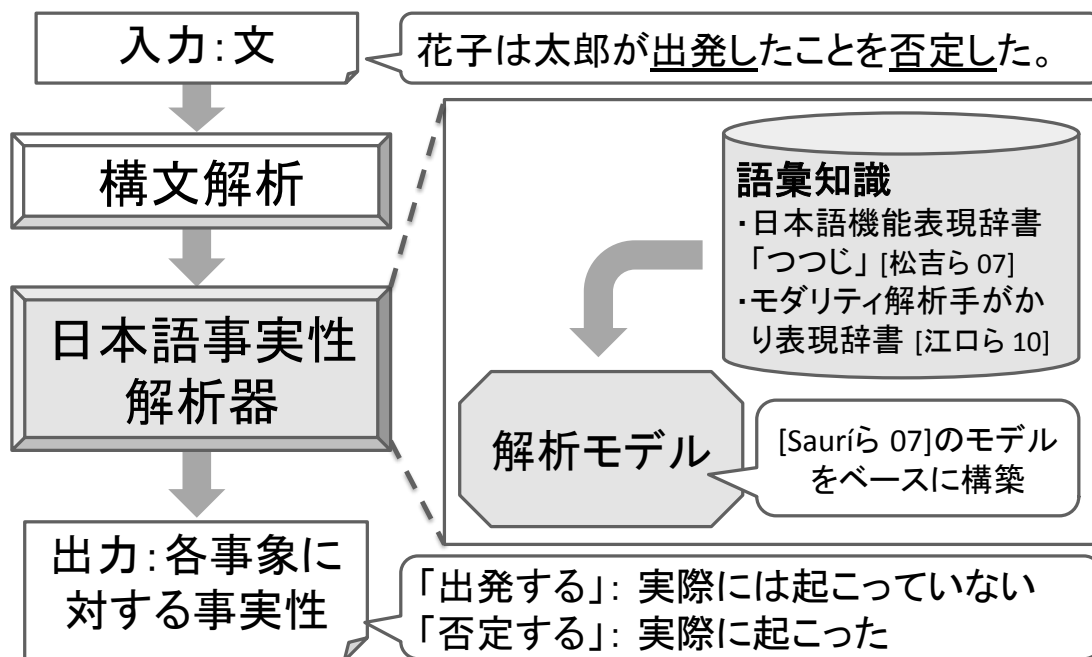


図 3: 日本語事実性解析器の概要

4.1 事実性の定義

2.1 節で、言語学的な側面からを定義した。本節では、システムに適用する際、どのように事実性を定義・形式化するか述べる。

Sauríらは、事実らしさに対する確信度を Certain (CT)・Probable (PR)・Possible (PS)・Underspecified (U) の4段階、その確信の方向を表す肯否極性を positive (+)・negative (-)・underspecified (u) の3値として扱い、これらの組み合わせによって事実性の値を定義している。例えば(3a)の事象「出る」は、著者から見て実際に起こったことだと解釈でき、その事実性はCT+と表される。同様に(3b)における事象「帰る」は、著者から見て起こった可能性が高いことだと解釈でき、その事実性はPR+、(3c)における事象「発生する」は、著者から見て実際には起こっていないことだと解釈でき、その事実性はCT-と表される。

英語では、PRは *probable*、PSは *possible* といった表現によって解釈しているが、日本語では表現が多彩であるため、PRとPSの境界が曖昧で、その区別は容易でないことが予想される。

(11) 彼が無事に着いたと私は信じている。

例えば(11)において、事象「着く」に対する著者の確信度は明らかにCTでもUでもないが、PRであるのかPSであるのかは解釈できない。そこで我々は、PRとPSを1つの値PRとしてまとめて扱った。

またSauríらは、事実性が明らかでないときに、部分的に不明である場合にはCT_u、完全に不明である場合にはU_uという事実性を定義している(PR_u、PS_u、U₊、U₋は利用不可値と考えられている)。ここにおける「部分的に」とは、起きているか起こっていないかがその文からは明確ではないが、確実にどちらかではある場合をさす。

(12) a. 彼は花子が来たかどうかを知っている。

b. 彼は花子が来たかどうかを知らない。

例えば(12a)において、実際に花子が来ているか来ていないかは、この文だけではわからない。しかし、どちらかであるのは確実であるので、Sauríはこのよう

表 1: 確信度と肯否極性の組み合わせによる事実性の値

確信度 \ 肯否極性	positive (+)	negative (-)
Certain (CT)	実際に起こっている (CT+)	実際には起こっていない (CT-)
Probable (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 (PR-)
Underspecified (U)	不明 (U)	

な事象「来る」の事実性を CT_u として扱っている．それに対して，(12b) の事象「来る」のような，完全に不明確である事実性を U_u と表現している．我々は，問題の簡単化のためこれらの値を区別せず，事実性が明確ではない場合を U としてまとめて扱うこととした．

以上 2 点を変更し，それ以外は Sauríらの定義を利用した．即ち，確信度を Certain (CT)・Probable (PR)・Underspecified (U) の 3 段階，肯否極性を positive (+)・negative (-) の 2 値として扱い，これらの組み合わせによって事実性を表す．ただし，前述のとおり確信度が U の場合には，確信度 U のみで事実性を表現する．これら組み合わせをまとめたものを図 1 に示す．

Sauríらは著者以外の態度表明者 (source) から見た事実性も考慮しているが，2.2.1 節において述べたとおり，タスクを簡略化し分析をしやすくするため，本稿では著者の事実性判断を認識する問題に限定し，これ以降は態度表明者を省略する．

4.2 使用する語彙知識

Sauríらのモデルは，確信度と肯否極性の組で表される事実性を，factuality marker と呼ばれる事実性に影響を与える表現を利用して解析する．例えば *not* は肯否極性を反転させる factuality marker，*may* は確信度を下げる factuality marker である．機能語だけでなく述語についても factually marker が考えられ，例えば *know*

表 2: 機能表現辞書の意味クラス例と事実性への影響

機能表現辞書中の意味クラス	該当する機能表現	事実性への影響
否定	ない ぬ わけでない	肯否極性: $+ \rightarrow -$, $- \rightarrow +$
推量	だろう かもしれない らしい	確信度: $CT \rightarrow PR$
疑問	か かな ではないか	確信度: $CT \rightarrow U$, $PR \rightarrow U$

that という表現は *that* 節の内容が成立していることを前提としているので, *know* は *that* 節内の事象を $CT+$ とする factuality marker と考えることができる. これらのような語彙的な factuality marker だけではなく, 目的節や関係詞節等の統語的な factuality marker も考慮している.

日本語においても「～ない」は肯否極性の反転, 「～だろう」は確信度の減少, というように, 同等の表現が存在する. このような表現を集めた語彙知識として, 日本語機能表現辞書「つつじ」(機能表現辞書) [28] およびモダリティ解析手がかり表現辞書(手がかり表現辞書) [3] を利用した.

機能表現辞書は, 文の構成に関わる要素である機能表現を, 意味, 文法的機能, 音韻的变化などに応じて網羅的に収録した辞書である. 機能表現辞書の意味クラスおよびその表現の例, それらの事実性に対する影響を表 2 に示す. この辞書は意味クラスによって表現を分類しており, それぞれの表現に対して活用や表記の異なりなどを考慮して構成されている. この中で「否定」と分類される表現は肯否極性を反転, 「推量」と分類される表現は確信度を減少, 「疑問」と分類される表現は確信度を U に割当, といった役割を持つ factuality marker として利用できる. このような事実性に影響を与える手がかりとなる機能表現として, 機能表現辞書の一部である 5,345 表現を利用した. ここでは, 「否定」「推量」「疑問」「意

表 3: 手がかり表現辞書の見出し語例と対応する事実性

見出し語	事象の形式	肯否環境	真偽判断	事実性
防ぐ	スル形	肯定	不成立	CT-
		否定	成立	CT+
	タ形	肯定	-	-
		否定	-	-
忘れる	スル形	肯定	不成立	CT-
		否定	成立	CT+
	タ形	肯定	成立	CT+
		否定	成立	CT+

志」「願望」などの事実性に影響を与える意味クラスに基づいて表現を選出した。

手がかり表現辞書は、拡張モダリティに影響を与える 3,692 述語を収録しているが、これは事実性とも大きな関わりをもっている。手がかり表現辞書の項目の例を表 3 に示す（実際の辞書には態度・真偽判断・価値判断の 3 つ組が記述されているが、表では事実性と関わりの深い真偽判断に関する項目のみを示した）。この辞書は、各述語が直前の事象に与える影響を、直前の事象の時制（未来：スル形「～する」、非未来：タ形「～した/していた/している」）および肯否環境ごとに収録している。例えば (3c) の「防いだ」という述語は、肯定環境下では不成立 (CT-)、否定環境下では成立 (CT+) というように、事象「発生する」の肯否極性を反転させる。(13) の「忘れる」は、述語の直前の事象の時制を考慮した例である。

(13) a. 彼は発言したのを忘れている。

b. 彼は発言するのを忘れている。

「発言する」に対して、(13a) では「発言する」の事実性は CT+ となるが、(13b) では「発言する」ことが実際には起こっておらず、事実性は CT- となる。

Algorithm 1 日本語事実性解析システム

```
1: chunk = root node in DEPENDENCY TREE
2: set initial value (CT+) of contextual-factuality
3: function traverseDependencyTree(chunk)
4: # Step1: 機能表現による更新
5: if chunk has functional expressions then
6:   update contextual-factuality with functional expressions dictionary
7: end if
8: # Step2: 事象をもつかの確認
9: if chunk has an event then
10:  get event-factuality from contextual-factuality
11: end if
12: # Step3: 手がかり表現による更新
13: if chunk has key expressions then
14:   update contextual-factuality with key expressions dictionary
15: end if
16: for all chunks c adjacent to chunk not visited do
17:   traverseDependencyTree(c)
18: end for
19: end
```

4.3 解析モデル

前節であげた、事実性に影響を与える手がかりとなる表現を集めた語彙知識を利用し、確信度と肯否極性を依存構造木の根から伝搬させて事実性を解析する。事象間には依存関係があり、事実性を伝搬させていくことで、それを反映したモデルとなっている。Sauríらの英語を対象とした解析モデルでは、単語をノードとした依存構造木を走査するが、日本語では文節を意味の区切りと考えるのが一般的であるため、文節をノードとした依存構造木を走査する。

このアルゴリズムを Algorithm 1 に示す。ここでは、文全体で伝搬させる事実性を *contextual factuality*、各事象に割り当てる事実性を *event factuality* と呼ぶ。入力は一文の構文解析結果、出力は各 *event factuality* であり、*contextual factuality* は *event factuality* を伝搬させるための変数である。

まず、多くの事象の事実性は CT+ であるため、初期値として *contextual factuality* に CT+ を割り当て (line 2)、文末の文節から依存構造木の走査を開始する。

Algorithm lines	依存構造木	手がかり表現	contextual factuality	event factuality
2			CT+	
(4-7) 8-11 12-15	知らない	polarity: - (ない) modality: CT (知る) polarity: + (知る)	CT- CT+	知る: CT-
(4-7) (8-11) (12-15)	相手は ことを		CT+	
(4-7) 8-11 12-15	断念した	polarity: - (断念する)	CT+ CT-	断念する: CT+
(4-7) 8-11 (12-15)	彼が 出場を		CT-	出場: CT-

図 4: (14) に対する事実性解析アルゴリズムの動き

各文節で行うことは、機能表現辞書を用いた更新 (Step1: line 4-7)・事象がその文節内に存在するかの確認 (Step2: line 8-11)・手がかり表現辞書による更新 (Step3: line 12-15) の3ステップである。以下では、それぞれのステップについて述べる。

Step1 (line 4-7) では、機能表現辞書と着目している文節とを照らし合わせ、機能表現が見つかった場合はそれに応じて contextual factuality を更新する。例えば「～ない」のような否定表現があれば肯否極性を反転させ、「～ではないか」のような疑問表現があれば確信度を U に更新する。Step2 (line 8-11) では、その文節内に事実性を割り当てるべき事象が存在するかどうかを確認し、存在するならば現在の contextual factuality をその事象の event factuality として割り当てる。Step3 (line 12-15) では、手がかり表現辞書と着目している文節内の述語とを照らし合わせ、手がかり表現が見つかった場合は、その文節よりも前の文節に事実性を伝搬させるために contextual factuality を更新する。この順番で処理を行うのは、文節内の事象は機能表現に影響を受け、その事象自体がそれ自身に影響を与えることはなく、現在見ている文節より前の文節に影響を及ぼすためである。

例として、(14) に対するアルゴリズムの動きを図 4 に示す。

(14) 彼が出場を断念したことを相手は知らない。

まず contextual factuality の初期値 CT+ を割り当て、依存構造木を文末から走査していく。最初の文節「知らない」に対して、Step1 では極性を反転させる否定の機能表現「ない」が存在するので、contextual factuality を CT- に更新する。Step2 において「知る」は事実性を割り当てるべき事象であるため、そのときの contextual factuality (CT-) を「知る」の event factuality として出力する。Step3 では「知る」という述語が手がかり表現辞書内に存在するので、それに基づいて contextual factuality を CT+ に更新する。各文節についてこれを繰り返し、「知る」の事実性が CT-、「断念する」の事実性が CT+、「出場」の事実性が CT- という出力が最終的に得られる。

5 実験・誤り分析

5.1 実験

日本語事実性解析における問題を整理・分析するために、構築した解析器をウェブ上の数千文に対して適用し、その評価を行った。

実験では、拡張モダリティタグ付与コーパス [8, 19] の一部である、Yahoo!知恵袋 (OC) に含まれる 6,404 文 (形態素数: 110,649) に対してアルゴリズムを適用した。このコーパスは現代日本語書き言葉均衡コーパス (BCCWJ)¹² に対して、< 態度表明者 > , < 相対時 > , < 仮想 > , < 態度 > , < 真偽判断 > , < 価値判断 > の 6 項目からなる拡張モダリティを付与したものである。

この中の < 真偽判断 > は、肯定か否定か、断定か推量かの組み合わせを表しているものであり、我々の事実性に相当する。態度や相対時といったラベルも事実性に影響を与える場合があるが、今回の実験では考慮しない。表 4 に < 真偽判断 > ラベルと我々の事実性との対応を示す。また、我々の事実性には、「成立から不成立」や「高確率から低確率」などの変化を伴うラベルに相当するものは存在しない。これら変化を含意するラベルは、2.2.2 節で述べたように、変化後が主要素であると考え、変化後のラベルに対応づけた（「成立から不成立」→ CT-）。

(15) 支給が停止されました。

例えば (15) の事象「支給」は、「成立から不成立」と真偽判断ラベルがつけられている。この場合には、現在は「支給」されていないという解釈ができるので、CT- に対応づけた。

本実験で対象とする事象は、態度表明者が著者となっているものの中で、限定修飾・機能表現のように「対象外」が付与されていない 14,917 事象（約 2.33 事象/文）である。正解の形態素情報を入力し、構文解析を行った結果を、構築した事実性解析器の入力とする。尚、事実性を付与すべき事象の同定は、あらかじめ正解を与えた。

¹²<http://www.tokuteicorpus.jp/>

表 4: <真偽判断> ラベルと事実性との対応

確信度 \ 肯否極性	+	-
CT	成立 不成立から成立 (CT+)	不成立 成立から不成立 (CT-)
PR	高確率 低確率から高確率 (PR+)	低確率 高確率から低確率 (PR-)
U	0 (U)	

5.2 評価

それぞれのラベル (CT+・PR+・PR-・CT-・U) に対する二値分類として、解析器の評価を行った。例えば CT+ に対しては、CT+ を正しく出力できているものを TP (True Positive)、CT+ を出力しているが正解は異なるものを FP (False Positive)、正解には CT+ がついていないが正しく CT+ を出力できていないものを FN (False Negative) とする。他のラベルも同様にし、TP、FP、FN の数と、各ラベルに対する以下の 3 値を表 5 に示す。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$F_1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

各正解ラベルに対してどのような出力が与えられたかを表す Confusion Matrix を表 6 に示す。表 5 には、多値分類と考えたときの精度 (表 6 の太字部分の和を Total で割った値=Micro-Average) も併記する。この結果によると、PR- は確信度と肯否極性の両方に対して更新を行わなければならないため、難易度が高いことがわかる。

また、確信度および肯否極性の 2 軸を個別に分析するため、確信度および肯否

表 5: それぞれのラベルに対する評価

	CT+	PR+	PR-	CT-	U	Micro-Average
TP	7,140	141	7	631	1,893	(9,812)
FP	1,467	286	71	583	2,698	(5,105)
FN	2,791	794	168	408	944	(5,105)
Precision	0.83	0.33	0.09	0.52	0.41	0.66
Recall	0.72	0.15	0.04	0.61	0.67	0.66
F_1	0.77	0.21	0.06	0.56	0.51	0.66

表 6: Confusion Matrix

正解 \ 出力	CT+	PR+	PR-	CT-	U	Total
CT+	7,140	237	25	408	2,121	9,931
PR+	528	141	8	23	235	935
PR-	47	6	7	67	48	175
CT-	97	1	16	631	294	1,039
U	795	42	22	85	1,893	2,837
Total	8,607	427	78	1,214	4,591	14,917

極性を個別に評価し、その Confusion Matrix を作成した。これを表 7~10 に示す。事実性が U の場合には、肯否極性が割り当てられないため、事例数から除いている。これら結果を見ると、CT が PR、U よりも多く正解できており、+ が - よりも正解できている傾向がわかる。それぞれを比較すると肯否極性の方が結果が良くなり、確信度の判定は人間でも難しい場合がある、という直感にあっていることがわかる。

表 7: 確信度に関する評価

	CT	PR	U	Micro-Average
Precision	0.84	0.32	0.41	0.69
Recall	0.75	0.15	0.67	0.69
F_1	0.80	0.20	0.51	0.69

表 8: 確信度に関する Confusion Matrix

正解 \ 出力	CT	PR	U	Total
CT	8,276	279	2,415	10,970
PR	665	162	283	1,110
U	880	64	1,893	2,837
Total	9,821	505	4,591	14,917

表 9: 肯否極性に関する評価

	+	-	Micro-Average
Precision	0.98	0.61	0.93
Recall	0.95	0.83	0.93
F_1	0.96	0.70	0.93

表 10: 肯否極性に関する Confusion Matrix

正解 \ 出力	+	-	Total
+	8,046	464	8,510
-	151	721	872
Total	8,197	1,185	9,382

5.3 肯否極性における誤りの分析

4節で述べた解析器は、確信度と肯否極性を独立して認識し、その組み合わせによって事実性を表している。2つのうち肯否極性の解析誤りは、成立と不成立のような真逆の誤りにつながるため、事実性を利用するアプリケーションにおいて、確信度の誤りよりも深刻である場合が多い。そこで、肯否極性に関する誤り事例の中から、ランダムに200事例を選出し、分析を行った。ここで議論したい目的としては、以下の3点があげられる。

- I. 現在利用可能な語彙知識（機能表現辞書・手がかり表現辞書）によってどれだけの問題がカバーできるのか。カバーできなくても語彙知識の拡充および拡張によって解決できるのか
- II. 語彙的手がかりが曖昧性を持つ場合は多いのか
- III. 局所的な語彙知識の組み合わせだけで解決できない問題は存在するのか

以上の論点に基づいて誤り分析を行った結果、以下のように誤りを分類することができた。

- i. 語彙知識のカバレッジの問題（内容語 17%，機能語 9%）
- ii. 語彙的手がかりの多義性の問題（9%）
- iii. 手がかりとなる表現のスコープの問題（37%）
- iv. その他（28%）

本節では、これらの誤りについて例文とともに詳述する。

5.3.1 語彙知識のカバレッジの問題

語彙知識のカバレッジの問題としては、(16)のような例がある。これは内容語のカバレッジに関する誤りの例である。

(16) サーバーは接続を解除しました。（正解：CT-，出力：CT+）

解析器は「接続」の肯否極性を+と出力したが、正解は-である。これは、事象の肯否極性を反転させる述語である「解除する」が手がかかり表現辞書に存在しないことが原因である。

また、例えば表記揺れのように、語彙知識として存在はしているが、うまく引くことができない問題も存在し、(17)はその例である。

(17) 関東だから違うかもしんない。(正解：PR+，出力：PR-)

「違う」の肯否極性の正解は+であるが、解析器は-と出力してしまった。辞書には、推量の機能表現「かもしれない」は載っているが、口語的な表現「かもしんない」としては引けないためである。この種類の誤りは少なく、機能表現辞書では表記の異なりも考慮して収録できているため、多くの表現をカバーできていることがわかった。

以上の問題は、対応する語彙知識を充実させ、その表現を正しく認識できれば解決できる誤りである。しかし、語彙知識を充実させれば全ての問題が解けるわけではない。語彙知識の単純な適用だけでは難しい問題として、語彙的手がかかりの多義性の問題や、手がかかりとなる表現のスコープの問題が存在する。これらの問題は、語彙知識のカバレッジの問題よりも、割合の大きな重要な問題となっている。

5.3.2 語彙的手がかかりの多義性の問題

語彙的手がかかりの多義性に起因する誤りの例を(18)に示す。

(18) 歩けば交通費はかからないか、いい運動にもなるしね!(正解：CT-，出力：U)

この例では、「ないか」が勧誘の「ないか」として認識されてしまい、Uを出力するが、実際にはこれは勧誘の意味ではなく、否定の「ない」と自己確認の「か」との組み合わせである。「ないか」が勧誘の意味ではないことは、「交通費がかかる」が意思的な行為ではないことからわかる。このような局所的な文脈により、意味を同定するようなアプローチがまず考えられる [29, 30]。しかし、「ないか」が勧

誘の意味ではないことと、「ない」が否定の意味であることは判別できても、「か」が並立の意味なのか自己確認の意味なのかを判別することは、文外の情報に依存するため、容易ではない。局所的な文脈だけを用いて意味を同定するアプローチによって、どこまでの問題が解決でき、そこから先の問題はどの程度残るのかを明らかにする必要がある。

5.3.3 手がかりとなる表現のスコープの問題

(19) のような、手がかり表現が影響を与える範囲（スコープ）の問題は多く存在し、全体の誤りの 4 割近くを占めている。

(19) 車が見えたがぶつかるのを避け切れなかった。(正解：CT+，出力：CT-)

(19) では、「避け切れなかった」のスコープが問題となる。「避け切れなかった」のは「ぶつかる」であり「見える」ではないので、「見える」の事実性は「避け切れなかった」に依存しないはずである。しかし、現在のアルゴリズムでは、係り元すべてをスコープとしているため、余計に伝搬がなされてしまい、間違った出力がされてしまう。スコープ外には伝搬させないようにする必要があるが、スコープを完全に特定するのは容易ではない。まず、「避け切れなかった」の項に関する知識が必要となる。そして、この項の範囲を特定する必要があり、このためにも世界知識が必要となる。単純に項に関するマーカーを特定すればよいのではなく、その項がどこからどこまでの範囲なのかを特定する必要がある、難しい問題である。

このスコープの誤りの多さを裏付けるものとして、表 11 に、事象の位置に着目した正解率を示す。例えば、(19) においては、「避け切る」が文末の事象、「見える」「ぶつかる」が文末以外の事象にあたる。文末の事象は、その文節内の機能語にしか影響されないが、文末以外の事象には、それ以外の要素も影響される。このそれぞれの要素の影響を確認するため、文末の事象の正解率と文末以外の事象の正解率とを比較する。

文末の事象は、「たぶん」などの副詞を考慮できていないため、十分ではないが、正解率は高いといえる。これは、機能語の影響をうまく反映できている、という

表 11: 事象の位置に着目した正解率

	正解率 (正解数 / 事象数)
文末の事象	0.80 (4691 / 5885)
文末以外の事象	0.57 (5121 / 9032)

ことを示している。それに対して、文末以外の事象は、あまり正解率が高くない。ここに、内容語の問題とともに、スコープの問題が大きな要素であることが示されている。

5.3.4 その他の誤り

その他の誤りとして、構文解析誤りやアノテーション誤り、連体修飾部の問題などが存在する。依存構造木に大きく従ったアルゴリズムであるので、構文解析の時点で誤ってしまうと、結果として事実性解析がうまくいかず、謝った出力を行ってしまう。

また、(20) は、連体修飾部が問題となる誤りである。

(20) 信用できない人には貸しません。(正解: CT-, 出力: CT+)

「貸しません」の肯否極性 - が連体修飾部内の事象に対しても伝搬されてしまうため、間違った出力がされてしまう。このような限定修飾の場合の事実性などについて、判断可能であるかはそもそも問題である。

5.4 確信度における誤りの分析

確信度においても、いくつかの誤り事例の分析を行ったところ、肯否極性と同様の誤り分類が多く見られた。語彙的手がかりの多義性に起因する誤りの例を(21)、語彙知識のカバレッジ、スコープ、アルゴリズムの拡張と、多くの問題が含まれる例を(22)に示す。

(21) 何分かおきに音で知らせてくれるような機械です。(正解: CT+, 出力: PR+)

(22) おそらく、親戚は自慢したいだけの見栄っ張りです。(正解：PR+，出力：CT+)

(21) では、「知る」の確信度を CT ではなく，PR と判定してしまっている．本来は比況の意味である機能表現「よう」を誤って推量と認識してしまったことが原因である．

(22) では「見栄っ張り」に対して「おそらく」がその確信度を下げるが，今回利用した語彙知識では副詞はカバーできていない．このような確実性に影響する副詞は，川添ら (2011)[20, 21] によって体系化されており，ある程度カバーできる．また，カバレッジの問題を解決できても「おそらく」は「見栄っ張り」だけでなく「親戚」や「自慢したい」に係る解釈も可能であるので，副詞が持つスコープを特定する必要がある．さらに，現在のアルゴリズムでは，依存構造木の根から事実性を伝搬させていくので，副詞のように，後続の文節に影響を与える手がかり表現にも対応できるよう，アルゴリズムの拡張も必要となる．アルゴリズムの拡張として，例えば前からも事実性を伝搬させていくことが考えられる．しかしながら，単純に前からも事実性を伝搬させていくことが，可能かどうかを検討する必要がある．可能である場合にも，後ろからの事実性伝搬との相互作用をどのようにするか，など考慮すべき点は多く，この拡張は今後の課題として考える．

6 誤りの情報をもたせたデータの試作

前節まで、既存の語彙知識を利用して、Sauriら [9] の構築した構成性に基づく解析モデルを日本語に適用し、日本語事実性解析における問題を探ってきた。本節では、その結果をもとに、6.1 節で、今後どのように日本語事実性解析器の精度向上を目指していくかを議論する。そして、6.2 節で、今後の精度向上に向け、誤り情報をラベル付けしたデータを試作したことを報告する。

6.1 分析結果に基づく今後の指針

前節の分析結果により、語彙知識のカバレッジの問題や、手がかりとなる表現のスコープの問題などが主問題となっていることがわかった。この結果に基づき、現在の日本語事実性解析器を改善し、精度向上を目指す。図 5 に、今後目指す日本語事実性解析器を示す。

まず、現在の解析モデルは、図 5 の左上に示すように、係り受け構造をもとに事実性解析を行っている。例えばこの図では、「出発していないことを否定した」というような文があり、「出発する」という事象の事実性 (event factuality) を解析する場合を考える。このとき、現在のアルゴリズムでは、以下の 4 つの要素に基づいて、事実性解析が行われる。

親の辞書情報 親の内容語により、手がかり表現辞書をもとに contextual factuality を更新する。例では、「否定する」が親の内容語にあたり、肯否極性を反転させる

機能語の辞書情報 解析したい事象に付随する機能語により、機能表現辞書をもとに contextual factuality を更新する。例では、「～ていない」が解析したい事象「出発する」の機能語にあたり、そのうちの「ない」という表現により、肯否極性が反転する

親のラベル 親まで伝搬されてきた contextual factuality をもとに、解析したい事象まで伝搬、更新を行う

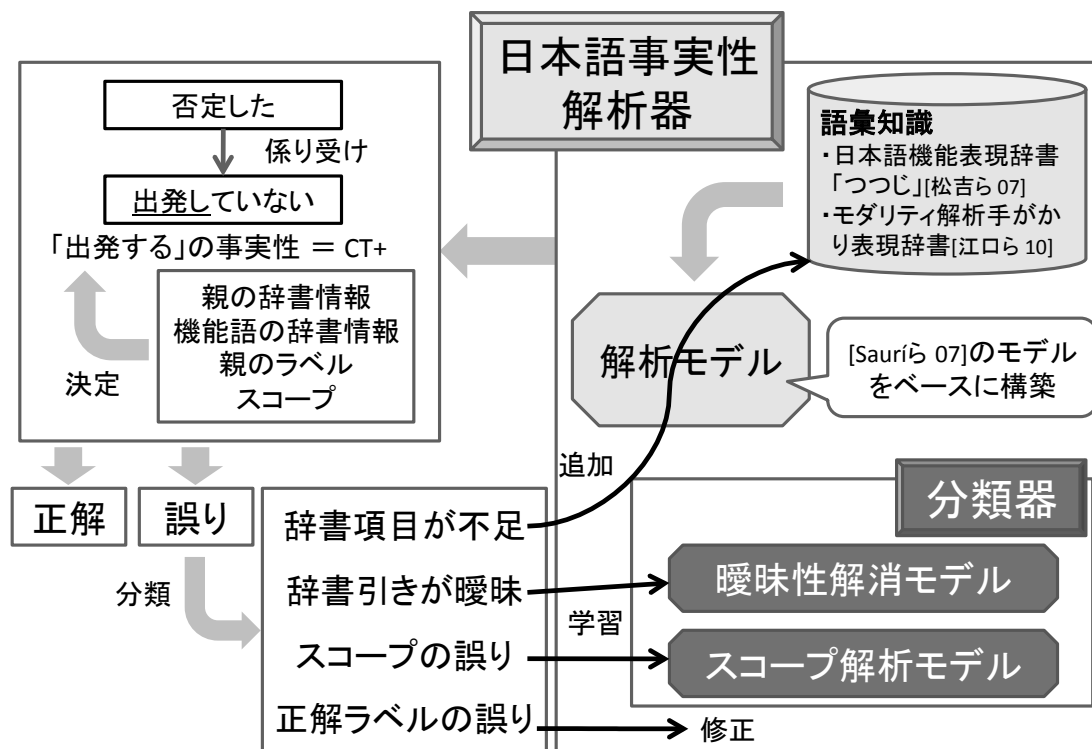


図 5: 今後の日本語事実性解析器

スコープ そもそも contextual factuality を，この係り受け構造に沿って伝搬させるべきなのか否か．現在のモデルでは，基本的にすべての場合において伝搬させている

これらの要素のいずれかにより，誤った事実性が出力された場合，誤りの原因としては，5.3 節であげたように，以下の 4 つのタイプに大別される（図 5 左下）．

- i. 語彙知識のカバレッジの問題：辞書項目が不足している
- ii. 語彙的手がかりの多義性の問題：辞書を引く際に，その曖昧性が問題となっている
- iii. 手がかりとなる表現のスコープの問題：手がかり表現が影響を与える範囲（スコープ）が問題となっている

iv. その他の問題：正解ラベルがそもそも誤っている¹³

これらの誤りを修正するために、それぞれの誤りごとに問題を切り分けると、例えば以下のような方法が考えられる。

- i. 辞書項目が不足：不足している辞書項目を、既存の語彙知識に追加する
- ii. 辞書引きが曖昧：曖昧性が問題となっている誤り事例を学習し、分類器を構築、日本語事実性解析器に組み込む
- iii. スコープの誤り：スコープが問題となっている誤り事例を学習し、分類器を構築、日本語事実性解析器に組み込む。この際、ここからここまでがスコープの範囲だということを、正確に把握するのではなく、どの場合に事実性が伝搬されるべきで、どの場合に事実性が伝搬されるべきでないか、という二値分類として学習を行う
- iv. 正解ラベルの誤り：正解ラベルをアノテーションしなおす

以上のように、誤りの種別に問題を切り分け、それぞれに対応していくことを考えると、各誤り事例がどの誤りに属するか分類する必要がある。また、曖昧性、および、スコープの問題を解決するモデルを構築するために、その学習データも必要となる。

そこで我々は、各誤り事例に対し、どの誤り原因による間違いだったのか、という誤り情報をラベル付けし、誤り原因に応じて、事実性解析器全体の改善を行うこととした。更に、事実性解析 誤り情報のラベル付データの作成 各誤り原因に応じた改善、というサイクルを繰り返すことで、日本語事実性解析器の精度向上を目指す。次節で、試作した誤り情報のラベル付データについての分析結果を示す。

¹³5.3 節において、その他に分類されたうち、構文解析誤り、および、連体修飾部など、そもそも事実性の判断が可能かどうかを、検討する必要がある場合に関しては、日本語事実性解析器内において、すぐに対応できる問題ではないため、今回の分類から外し、正解ラベルの誤りのみ取り上げている

6.2 誤り情報のラベル付データの作成

前節で述べたように，今後の日本語事実性解析器の精度向上に向け，各誤り事例に対し，どの誤り原因による間違いだったのか，という誤り情報をアノテーションしたデータを試作した．現在，著者以外のアノテータが，アノテーション作業をはじめており，ランダムにサンプリングしたおよそ 500 の誤り事例に対し，アノテーションを試験的に行った．

4.3 節のアルゴリズムでは，依存構造木に沿って contextual factuality を最後まで伝搬させる．そのため，一度誤ってしまうと，そのまま誤った contextual factuality が伝搬され，さらにそこに誤りが重なることで複雑化し，どの原因により誤ったかが分析しづらい．このような，誤った contextual factuality が伝搬されることによる，誤りの複雑化を防ぐため，文内の部分的な依存構造木に着目し，「親のラベル」として，伝搬されてきた contextual factuality の代わりに，親の事象の正解ラベルを利用した¹⁴．

この設定に基づき，アノテーションを行った結果，スコープに関する誤りが 57%，語彙知識の不足，多義性に起因する誤りが 16%，そもそも正解ラベルが誤っている，と考えられるものが 27%であった．これは，おおよそ 5.3 節で行った分析と同様の傾向を示している．

その中でも，正解ラベルが誤っている，という誤り情報が付与された事例が，5.3 節で行った分析と比較して，多く存在してしまっていることがわかる．この理由としては，拡張モダリティタグ付与コーパス [8, 19] では，対象の文章から事象の核となる述語を抽出するプログラムによって，事象の認識をおこなっている，という理由が考えられる．この文章から述語を抽出するプログラムは完全ではないため，事象の核となる述語以外の表現が誤って抽出され，その結果，そもそも事象とは言いがたいものに拡張モダリティタグが付与されている場合が多く存在していることがわかった．今後，このような場合には正解ラベルを修正し，対応する．

¹⁴この実験設定での正解率は，67.5%であり，今までの実験設定での正解率は 65.8%であったことから，誤ったまま contextual factuality を伝搬させたことによる誤りも，若干ではあるが存在することがわかった．

次に、半数以上を占めている、スコープの誤りについて分析を行った。

(23) a. 電源が切れて、再起動しません。

b. ど忘れして思い出せません。

c. パソコンを買ったばかりで、お気に入りに何も入っていません。

スコープの誤りとしては、例えば、(23) のような、「A して B しません」といった、それぞれの事象間に順序はあるものの、独立した事象となっている構造の文がほとんどだった。いずれも、正解は CT+ であるが、出力は CT- となり、「～ません」という表現がもつ否定が、誤って伝搬されたことが原因となっている。また、わずかだが、複数の誤りがつけられた事例も存在した。

今回試作した結果をもとに、判断の揺れが少なくなるようなアノテーション基準を設定し、今後の日本語事実性解析器の精度向上に向け、アノテーションを進めていきたいと考えている。

7 おわりに

7.1 まとめ

本論文では、情報抽出や含意関係認識などの、自然言語処理の応用に対して、有用な情報である、事実性の解析に関して議論した。

日本語事実性解析において、事実性に影響を与える語彙知識に焦点を当て、語彙知識の組み合わせに関する種々の問題を分析・整理することを目的として、Sauríら (2007) [9] の構成性に基づくモデルをもとに、既存の語彙知識を利用した日本語事実性解析システムを構築し、実験・誤り分析を行った。その結果、

- i. 語彙知識のカバレッジの問題
- ii. 語彙的手がかりの多義性の問題
- iii. 手がかりとなる表現のスコープの問題

などの問題が存在し、特に、手がかり表現の影響を及ぼす範囲（スコープ）が大きな問題となっていることを明らかにした。

7.2 今後の課題

6.1 節で述べたように、前節であげた分析結果に基づき、現在の日本語事実性解析器を改善し、精度向上を目指す。現在試作している、誤り情報のラベル付データの作成を進め、今後は、実際にそのラベル付データに従い、辞書の拡充、および、ラベル付データから学習した、曖昧性解消モデル・スコープ解析モデルを、実際に日本語事実性解析器に取り入れ、実験、誤り分析をおこなう。

Sauríら (2007) [9] のモデルに基づいて構築した、事実性解析モデルの拡張も必要である。現在の、依存構造木の後ろから事実性を伝搬させていくモデルでは、「たぶん」といった副詞などの、前から後ろに影響を与える表現に、適応する必要があるからである。また、事象が「行為」を表すか、「状態」を表すか、といったタイプの違いに基づいた事実性の異なりについても、検討する必要がある。

連体修飾部や，条件節における事実性など，そもそも何を事実性として捉えるかは，まだまだ議論の余地がある．また，例えば「後悔する」のような表現がもつ，前提という概念をうまく取り入れることができれば，事実性の表現できる幅が広がるのではないかと考えている．

今後は更なる分析とともにこれらの課題を進め，語彙知識および言語現象を加味した，高度な事実性解析器の構築を行う．さらに，事実性から発展させ，高度なモダリティ解析器の構築につなげていきたいと考えている．

参考文献

- [1] Roser Saurí and James Pustejovsky. FactBank: a corpus annotated with event factuality. *Language resources and evaluation*, Vol. 43, No. 3, pp. 227–268, 2009.
- [2] Roser Saurí. FactBank 1.0 Annotation Guidelines, 2008.
- [3] 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.
- [4] Kentaro Inui, Shuya Abe, Kazuo Hara, Hiraku Morita, Chitose Sao, Megumi Eguchi, Asuka Sumida, Koji Murakami, and Suguru Matsuyoshi. Experience Mining: Building a Large-Scale Database of Personal Experiences and Opinions from Web Documents. In *the 2008 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 314–321, 2008.
- [5] Koji Murakami, Eric Nichols, Suguru Matsuyoshi, Asuka Sumida, Shouko Masuda, Kentaro Inui, and Yuji Matsumoto. Statement map: Assisting information credibility analysis by visualizing arguments. In *Proceedings of the 3rd Workshop on Information Credibility on the Web (WICOW2009)*, pp. 43–50, 2009.
- [6] 水野淳太, 渡邊陽太郎, Eric Nichols, 村上浩司, 乾健太郎, 松本裕治. 文間関係認識に基づく賛成・反対意見の俯瞰. 情報処理学会論文誌, Vol. 52, No. 12, 2011.
- [7] 荒牧英治, 増川佐知子, 森田瑞樹. Twitter Catches the Flu : 事実性判定を用いたインフルエンザ流行予測. 情報処理学会研究報告 2011-NL-201 No.1, pp. 1–8, 2011.

- [8] 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.
- [9] Roser Saurí and James Pustejovsky. Determining Modality and Factuality for Text Entailment. *International Conference on Semantic Computing (ICSC 2007)*, pp. 509–516, 2007.
- [10] Charles Fillmore. The Case for Case. In Emmon Bach and Robert T Harms, editors, *Universals in Linguistic Theory*, Vol. 25, pp. 1–88. Holt, Rinehart and Winston, 1968.
- [11] 益岡隆志. 日本語モダリティ探求. くろしお出版, 2007.
- [12] 日本語記述文法研究会 (編). 現代日本語文法 4. くろしお出版, 2003.
- [13] 日本語記述文法研究会 (編). 現代日本語文法 3. くろしお出版, 2003.
- [14] Roser Morante. Modality and Negation in Natural Language Processing. In *Tutorial of IJCNLP*, 2011.
- [15] V Rubin, E Liddy, and N Kando. Certainty Identification in Texts: Categorization Model and Manual Tagging Results. *Computing Attitude and Affect in Text*, pp. 1–16, 2005.
- [16] Roser Saurí, Jessica Littman, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky. TimeML Annotation Guidelines, Version 1.2.1, 2006.
- [17] Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. The penn discourse treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, pp. 2961–2968. Citeseer, 2008.
- [18] 原一夫, 乾健太郎. 事態抽出のための事実性解析. 情報処理学会研究報告, 2008-FI-89, 2008-NL-183, pp. 75–80, 2008.

- [19] 松吉俊, 佐尾ちとせ, 乾健太郎, 松本裕治. 拡張モダリティタグ付与コーパスの設計と構築. 言語処理学会第 17 回年次大会発表論文集, pp. 147–150, 2011.
- [20] 川添愛, 齊藤学, 片岡喜代子, 崔栄殊, 戸次大介. 確実性判断のためのアノテーション済みコーパスの構築. 言語処理学会第 17 回年次大会発表論文集, pp. 143–146, 2011.
- [21] 川添愛, 齊藤学, 片岡喜代子, 崔栄殊, 戸次大介. 言語情報の確実性に影響する表現およびそのスコープのためのアノテーションガイドライン Ver.2.4. Technical report, Department of Information Science, Ochanomizu University, OCHA-IS 10-4, 2011.
- [22] Roser Morante and Walter Daelemans. Annotating Modality and Negation for a Machine Reading Evaluation. In *Proceedings of CLEF*, 2011.
- [23] 小橋洋平, 坂野達郎. 階層意味論に基づいた心的態度のアノテーション. 自然言語処理, Vol. 18, No. 4, pp. 323–350, 2011.
- [24] Marc Light, Xin Ying Qiu, and Padmini Srinivasan. The language of bio-science: Facts, speculations, and statements in between. In *Proceedings of BioLink 2004 workshop on linking biological literature, ontologies and databases*, pp. 17–24, 2004.
- [25] Ben Medlock and Ted Briscoe. Weakly supervised learning for hedge classification in scientific literature. In *Annual Meeting of Association of Computational Linguistics*, Vol. 45, pp. 992–999, 2007.
- [26] György Szarvas, Veronika Vincze, Richárd Farkas, and J. Csirik. The BioScope corpus: annotation for negation, uncertainty and their scope in biomedical texts. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, pp. 38–45, 2008.
- [27] L. R Horn. Metalinguistic negation and pragmatic ambiguity. *Language*, Vol. 61, No. 1, pp. 128–174, 1985.

- [28] 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, pp. 123–146, 2007.
- [29] 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. 言語処理学会第 17 回年次大会発表論文集, pp. 308–311, 2011.
- [30] 鈴木敬文, 阿部佑亮, 宇津呂武仁, 松吉俊, 土屋雅稔. 大規模階層辞書と用例を用いた日本語機能表現の解析. 『現代日本語書き言葉均衡コーパス』完成記念講演会予稿集, pp. 105–110, 2011.
- [31] 田辺利文, 首藤公昭. 日本語文末表現による推論について. 言語処理学会第 17 回年次大会発表論文集, pp. 159–162, 2011.
- [32] Roser Morante, Sarah Schrauwen, and Walter Daelemans. Corpus-based approaches to processing the scope of negation cues: an evaluation of the state of the art. In *Proc. of IWCS 2011*, pp. 350–354, 2011.
- [33] Emilia Apostolova, N. Tomuro, and D. Demner-Fushman. Automatic extraction of lexico-syntactic patterns for detection of negation and speculation scopes. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pp. 283–287, 2011.
- [34] R. Farkas, Veronika Vincze, G. Móra, J. Csirik, and G. Szarvas. The CoNLL-2010 shared task: learning to detect hedges and their scope in natural language text. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 1–12, 2010.
- [35] Roser Morante, V. Van Asch, and W. Daelemans. Memory-based resolution of in-sentence scopes of hedge cues. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 40–47. Association for Computational Linguistics, 2010.

- [36] Xinxin Li, Jianping Shen, Xiang Gao, and Xuan Wang. Exploiting rich features for detecting hedges and their scope. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 78–83. Association for Computational Linguistics, 2010.
- [37] E Velldal, L. Øvrelid, and S. Oepen. Resolving speculation: MaxEnt cue classification and dependency-based scope rules. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 48–55. Association for Computational Linguistics, 2010.
- [38] Halil Kilicoglu and Sabine Bergler. A High-Precision Approach to Detecting Hedges and Their Scopes. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 70–77. Association for Computational Linguistics, 2010.
- [39] Marek Rei and Ted Briscoe. Combining manual rules and supervised learning for hedge cue and scope detection. In *Proceedings of the Fourteenth Conference on Computational Natural Language Learning — Shared Task*, pp. 56–63. Association for Computational Linguistics, 2010.