

言語情報編集のための 広義モダリティ解析に向けて

奈良先端科学技術大学院大学
情報科学研究科准教授

乾 健太郎

PROFILE

奈良先端科学技術大学院大学 情報科学研究科 准教授。専門は計算言語学、言語情報や知識の自動編集に関する研究に従事。1995年東京工業大学博士課程修了。同大学助手、九州工業大学助教授を経て、現職。情報処理学会自然言語処理研究会幹事、Computational Linguistics 編集委員。人工知能学会 1997年度論文賞、COLING/ACL-2006 Best Asian NLP Paper Award、言語処理学会年次大会優秀発表賞等、受賞。博士（工学）。

✉ inui@is.naist.jp

☎ 0743-72-5241

奈良先端科学技術大学院大学
情報科学研究科特任助教

松吉 俊

PROFILE

2003年京都大学理学部卒。2008年同大学院情報科学研究科博士後期課程修了。現在、奈良先端科学技術大学院大学情報科学研究科特任助教。京都大学博士（情報学）。自然言語処理の研究に従事。

✉ matuyosi@is.naist.jp

☎ 0743-72-5243

1 はじめに

情報抽出や質問応答、複数文書要約や情報分析など、大量の文書集合から特定の種類の情報を抽出し、抽出した情報を照合、集約するといった言語情報編集を計算機で実現するためには、一般に意味解析、談話解析と呼ばれる、言葉の意味に踏み込んだ処理が必要である。本稿では、言語学でモダリティと呼ばれている意味情報に焦点を当て、モダリティ解析の課題設計とモデル構築に関する最近の取り組みを報告する。

自然言語の文章では、出来事や状態などの事象に対する言及に加え、それらの事象について書き手や第三者がどんな心的態度をとっているかが言語的に表現され、伝達される。例えば(1)では、「(供給事業が)進む」という事象は現在成立しているが、「業者が(事業に)協力する」という事象や「(燃料を)販売する」という事象は成立しておらず、未来においても成立の可能性が高くないことが述べられている。

(1) 環境省主導で進んでいるバイオエタノールとガソリンの混合燃料の供給事業に、石油元売り業者が協力を拒否、今夏に予定している販売開始のめどが立たない状況に陥っている。

このように、書き手や第三者が個々の事象を事実と認識しているか否か、また望ましいことと認識しているか否か、といった情報がモダリティである。こうした情報

を自動的に識別することは、情報抽出を含む広い範囲の言語情報編集に必要な処理であり、その課題設計やコーパス作成にもにわかに関心が高まっている[1, 8, 9, 10, 12]。ただし、モダリティと一口に言っても、どのような種類の情報をカバーし、どのような粒度で識別するかは研究グループによって大きく異なり、課題設定間相互の関係も不明確である。研究者間で一定の合意が得られるまでには、さらに理論的な分析とデータの蓄積を重ねる必要がある。

2 モダリティと否定とその周辺

モダリティは概ね次のように分類される[15, 6]。

- 真偽判断のモダリティ** 断定か、推量かを表す。
- 価値判断のモダリティ** 必要か、許可できるか。
- 表現類型のモダリティ** 叙述、意志、行為要求、勧誘、疑問、感嘆のいずれかの態度を表す。
- 丁寧さのモダリティ** 普通体か、丁寧体かを表す。
- 伝達態度のモダリティ** 聞き手の存在に対する話し手の意識のありようを表す。
- 説明のモダリティ** 文と先行文脈との関係づけ。

「表現類型のモダリティ」は、書き手の基本的な態度を、「真偽判断のモダリティ」は、事象の真偽に対する書き手の**確信度**を、「価値判断のモダリティ」は、事象成立を書き手が望んでいるか否かを表すと見ることがで

きる。一方、丁寧さ、伝達態度、説明のモダリティについては、情報抽出、分析、編集等の当面の応用を想定する限り緊急性はない。

上記モダリティに加えて重要なのは**肯否の極性**である。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[7]。文章に否定表現が存在する場合、どの部分が否定されているのか、すなわち、どこが否定の**焦点**であるのかを知ることは、文章の意味を正確に捉える上で重要である。また、周辺項目として、次のような項目も様々な応用で有用であると期待できる。

●**態度表明者** モダリティの態度を表明している人物や団体を指す。デフォルトは文章の書き手であるが、伝聞や引用の場合にその情報の発信者を特定することが重要になる。

●**仮想性** 仮想世界の話であるかどうか。

●**時制** 真偽が確定している過去または現在の事象か、本質的に真偽が定まらない未来の事象か。

●**真偽アスペクト** 真偽が一方から他方へ変化するアスペクトを持っているかどうか。

上記項目のうち、先行研究がどれをカバーしているかをまとめると表 1 のようになり、それぞれ一部しかカバーできていない。例えば、SauriらのFactBank[12]では、事象とその時制、肯否、モダリティをマークアップするTimeML[10]の体系の上に、事象を対象として、態度表明者ごとに、事実らしさに対する態度表明

者の確信度と独自の肯否極性をマークアップする。ただし、TimeMLにおけるモダリティは、英語の助動詞(must, may, should など)をそのままラベルとするに留まっており、他の言語への適用が困難である。これに対し我々は、表 1 の末尾に示すように、上記項目のうち丁寧さ、伝達態度、説明のモダリティを除くすべての項目をカバーする課題設計を考える。

3 広義モダリティ情報のタグ付け

上記の背景に基づいて新規に設計したタグ体系の概要を述べる。タグ付与の例は表 2、詳細は次の URL で公開している作業基準マニュアルを参照されたい。

<http://cl.naist.jp/nltools/modality/>

3.1 タグ付与の対象

FactBank などと同様に我々も、入力文章が言及するすべての事象について、個別にモダリティ情報を付与することを考える。ここで事象とは、行為、出来事、状態の総称であり、例えば次の文では「雑誌を購入するコト」、「来週から購入を中止するコト」、「(略)と思うコト」を事象と解釈し、すべてにタグを付与する。

(2) 来週からこの雑誌の購入を中止しようと思う。

なお、文献[15]に従い、述語に受動と使役のヴォイス

	対象	態度表明者	時制	仮想性	肯否極性	否定の焦点	確信度	推量の焦点	真偽アスペクト	表現類型	価値判断	事象の範囲
Rubinら[9]	語や句	○	○	×	×	×	○	×	×	○	×	×
TimeML[10]	事象	×	○	○	○	×	○	×	×	○	×	○
Prasadら[8]	談話関係と事象	○	×	×	○	○	○	×	×	○	×	○
Sauriら[11]	事象	○	×	×	○	×	○	×	×	×	×	×
FactBank[12]	事象	○	○	○	○	×	○	×	×	○	×	○
川添ら[1]	語や句	○	×	○	○	×	○	×	×	×	×	×
Lightら[4]	文	×	×	×	×	×	○	×	×	×	×	×
Medlockら[5]	語や句	×	×	×	×	×	○	×	×	×	×	×
BioScope[14]	語や句	×	×	×	○	×	○	×	×	×	×	×
本研究	事象	○	○	○	○	○	○	○	○	○	○	×

表 1 広義モダリティ情報の種類と各研究の対象範囲



ID	文（下線が事象の対象の核となる述語）	態度表明者	時制	仮想	態度	真偽判断	価値判断	焦点
1	京都からも近いので京都観光したいけど宿は温泉がいい。	wr	非未来	0	叙述	成立	0	0
2	日本でも、いよいよ裁判員制度が始まります。	wr	未来	0	叙述	成立	0	0
3	投票はどうやら三半規管の機能を安定させるために <u>続ける</u> のだそうだ。	wr_arb	非未来	0	叙述	成立	0	0
4	吸入ステロイド剤でも急性副腎不全が起きることを知らなければ、 <u>医師も診断</u> できないことがある・・・	wr	非未来	帰結	叙述	低確率	0	0
5	副腎皮質ホルモンから作られるステロイド剤の劇的な効果に医師達が驚き、 <u>競って使</u> い始めたのです。	wr	非未来	0	叙述	不成立から成立	0	0
6	ステロイド剤を使用している場合は、 <u>急激に使用</u> を中止するとリバウンド現象が起こり、症状が・・・	wr	非未来	条件	叙述	成立から不成立	0	0
7	・・・現在、クローン技術を応用して臓器を供給する試みを否定することは困難であろう。	wr	未来	0	意志	高確率	ポジティブ	0
8	化膿を防ぐために大切なのは、入浴した時に患部を <u>濡ら</u> さないように気をつけることです。	wr	未来	0	意志	低確率	ネガティブ	0
9	新宿行って、 <u>なんか食べ</u> に行きたい。	wr	未来	0	欲求	0	ポジティブ	0
10	ぜひ一度、「還元水」の持つ効果を実感して <u>みて</u> ください！	wr	未来	0	働きかけ-直接	0	ポジティブ	0
11	あなたは <u>その時に彼女に</u> 真実を伝えるべきだった。	wr	非未来	0	叙述	不成立	ポジティブ	0
12	彼は東京に仕事で行ったわけではない。	wr	非未来	0	叙述	成立	0	否定（仕事で）

表2 モダリティタグの付与例

を含めたものを事象と見なす。

3.2 モダリティ情報タグ

表1に挙げた項目は、以下に述べる7つの項目にさらに整理することができる。文章中の各事象に対し、そのモダリティ情報をこの7つ組で表現する。

態度表明者 発信された情報の信憑性を判断する手がかりとして、態度を表明する人物を特定することは重要である。本研究では、Wiebeら[16]が導入した「態度表明者の入れ子構造」を用いて、事象に対する態度表明者（の列）を記述する。ここで、“wr”は態度表明者が書き手であることを、“wr_arb”は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ことを表す。

時制 事象に対する真偽判断が推量の場合、その事象の時制により、次のように解釈することができる。

- ・時制が未来：その推量は、いまだ真偽が定まっていないことによる断定保留を表す。
- ・時制が未来でない：その推量は、事象の真偽を確認していないことによる断定保留を表す。

そこで、本研究では、事象の時制が未来であるかどうかを取り扱う。上記の2つの区別に必要である

のは、態度表明時に対する相対的な時制であるので、「態度表明時を基準とする時制」を記述する。

仮想 情報抽出においては、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区別することが求められるため、仮想性を取り扱う。

態度 「表現タイプのモダリティ」は、態度表明者の中心的な態度を表す重要な項目であるので、これを〈態度〉に記述する。叙述、意志、欲求、働きかけ-直接/間接/勧誘など、8種類のラベルを用意した。

真偽判断 「真偽判断のモダリティ」は、事象の真偽に対する態度表明者の確信度を表すため、応用に有用であるので、これを〈真偽判断〉に記述する。“成立/不成立”は肯定/否定の断定、“高/低確率”は肯定/否定の推量を表し、これらの間の遷移を“成立から不成立”のように表現する。

価値判断 〈価値判断〉では、言語学における「価値判断のモダリティ」の根幹にある、「事象成立の望ましさ」という概念を、極性情報¹として取り扱う。言語学の先行研究[6]において、上記のカテゴリーの表現は、基本的意味の面から、「必要」、「許可・許容」、「不必要」、「不許可・非許容」の4つ

¹ 語句に対して設定される評価極性（例えば、「給料が上がる」は評価極性ポジティブ、「事故に遭う」は評価極性ネガティブ）とは異なる。例えば、「あのとき事故に遭ってよかった。」における事象「あのとき事故に遭うコト」に対する〈価値判断〉は、“ポジティブ”であると判断する。

に分類される。本研究では、これらの意味を独自の体系により〈態度〉に記述し、態度表明者が事象成立を望ましいと判断しているかを〈価値判断〉に記述する。この記法により、〈態度〉、〈真偽判断〉、〈価値判断〉に関して、表現力があり、かつ、見やすいタグ体系の構築を目指している。

焦点 事象が“不成立”であるとき、実際に否定されているのは、事象を構成する一部の要素であり、暗に、同類の別の事象の成立が含意されることがある。例えば、次の文では、態度表明者は、「太郎が仕事で東京に行くコト」は不成立であると判断しているが、この文からは、「太郎が東京に行くコト」の成立が読み取れる。

(3) 太郎は東京に仕事で行ったのではない。

このような含意の存在を表現するため、否定、推量、問いかけを対象として、その焦点を記述する。

4 コーパスへのモダリティタグ付与

上述のタグ体系に基づき、ブログ記事を含む Web 文書からサンプルした約 1 万 3 千文（約 3 万 9 千事象）に実際にタグ付けを行った。ただし、本コーパスでは、事象の範囲を明確にマークアップすることはせず、述語に対してタグを付与することで、その述語を核として持つ事象にそのタグを付与したと見なす。また、補助動詞や、機能語相当表現 [2] の一部など、事象と見なせない表現はタグ付与の対象としない。

紙面の制約で詳細な統計は示せないが、各項目における値の分布には大きな偏りが見られた。例えば、〈態度〉と〈真偽判断〉のそれぞれ約 90% が“叙述”、“成立”で占められる。しかしながら、実際の応用においては、残りの 10%（例えば“不成立”）を正確に識別することが重要と考えられる。

タグ付与作業の信頼性を評価するために、ランダムにサンプルした 300 事象について 2 名の作業者が独立に付与したタグの一致率を調べたところ、7 つの項目に対する κ 統計量 [3] の平均で 0.71 という高い値を得た。この結果から、少なくとも人間はこの課題を十分

2 依存グラフの最適化は、今後の課題である。

な精度で行えることが分かる。なお、タグ付与にかかる時間は 1,000 事象あたりおよそ 5 時間であった。

5 解析モデルの構築

モダリティ解析の自動化については、他の言語処理タスクと同様、人手による語彙的・統語的知識を用いる手法 [11, etc.] や機械学習による手法 [5, etc.] が提案されている。例えば、Sauriá [11] は、項と述語間の知識、否定辞の有無、モダリティ表現を利用して、事象の入れ子の外側から、順に事象の事実性を解析するアルゴリズムを提案している。こうした既存研究は、表 1 で示したように、いずれもモダリティ情報の一部の項目だけを解析対象とするに留まっており、我々の課題設定のような包括的なモダリティ解析についてはほとんど報告がない。

そこで以下では、前述のタグ付きコーパスで解析モデルを訓練・評価する実験の現状を報告し、残る課題を考察する。

5.1 実験方法

文の構文解析結果と事象の核となる述語の位置を入力として、その事象のモダリティ情報タグを出力する解析システムを構築した。学習モデルには条件付確率場を利用した。具体的な学習・実行ツールとしては、タグの項目間の依存関係を学習できる Graphical Models in Mallet (GRMM) [13] を用いた。本実験では、モダリティ情報の中心的な項目である〈態度〉をハブにして残り 6 つの項目を連結させる依存グラフ²を構成し、〈態度〉と他の各項目の依存関係を学習した。なお、モデルのハイパーパラメタは初期設定を利用した。

素性集合としては、次の 2 種類を用いた。

基本素性 核となる述語が存在する文節（中心文節）、その前後の文節、および、中心文節の係り先 2 つまでの文節における形態素情報（基本形、品詞、活用形）の 2-gram と 3-gram。

知識素性 人手により構築した語彙統語パターン（モダリティ関連 182 個、肯否関連 32 個、条件関連



27 個)の有無。

比較のため、次の4つのシステムを構築した。

ベースライン 1 コーパスにおいて一番頻度が高いタグを決定的に出力する。

ベースライン 2 上記の語彙統語パターンに基づく人手規則により、決定的にタグを出力する。

GRMM 1 基本素性のみで GRMM を利用する。

GRMM 2 GRMM 1 に知識素性を加える。

訓練、評価には、前述のコーパスの一部 10,545 文 (24,858 事象) を利用し、3 分割交差検定を行なった。なお、自由な文字列を含むタグを持つ〈態度表明者〉と〈焦点〉に関しては、それぞれ、“wr” とそれ以外の 2 択、「焦点あり」と「焦点なし」の 2 択とした。また、〈真偽判断〉タグのうち、真偽の変化を表すタグの出現はコーパスにおいてあまりにも低頻度であったため、それらを含む文は訓練・評価から除外した。

5.2 結果と考察

各項目における「正解率」、および各タグに対する精度と再現率の調和平均である F 値の、各項目での平均をとった「F 値の平均」を表 3 に示す。この表から、すべての項目において、GRMM 2 が他のシステムより高い値を示していることが分かる。正解率に関して検定を行なった結果、〈時制〉、〈仮想〉、〈態度〉、〈真偽判断〉、〈価値判断〉において、GRMM 2 と他の 3 つのシステムの間有意差 (有意水準 1%) が認められた。GRMM 2 は、GRMM 1 と同等、もしくは、有意に高い性能を示しているため、知識素性は有効であったと言える。表 3 の結果は十分なものとは言えないが、この実験により、試作したシステムの性能向上のために

解くべき課題が明らかになった。以下、GRMM 2 の出力結果を元にこれを考察する。

まず、係り先の形態素情報のエンコード方法に起因する問題が見られた。例えば、表 2 の ID=1 の「近い」の例では、係り先文節内の「たい」に引っ張られ、〈態度〉が“欲求”に誤って判定される。また、文内を比較的自由に浮遊する副詞などの扱いも同様に課題が残った。例えば、表 2 の ID=2 の「始まる」では、副詞「いよいよ」をうまく手がかりとできず、〈時制〉を“非未来”と誤ってしまう。このように今後は、モダリティ表現や副詞のスコープをうまく学習できるような素性空間を構成する必要がある。また、「いよいよ」や「中止」(例えば表 2 の ID=6) のような、モダリティ解析に有用な手がかり表現の知識を整備する必要がある。

さらに、紙面の制約で具体的には示せなかったが、タグの分布に大きな偏りがあり、出現頻度の低いクラスの学習が困難になっているという問題もある。能動学習を採用するなど、分類にとって有益な事例、とくに相対的頻度の小さいクラスの事例を選択的にサンプルし学習する枠組みを検討したい。

6 むすび

本稿では、自然言語文章中のモダリティ情報の解析という問題を取り上げ、包括的な課題設計、コーパス作成、モデル構築を進める我々の取り組みを報告した。

今後は、まずタグ付きコーパスの作成をさらに拡大する必要がある。これについては、科研費特定領域研究「日本語コーパス」プロジェクト³で構築が進んでいる

システム	態度表明者		時制		仮想		態度		真偽判断		価値判断		焦点	
	正解率 (%)	F 値	正解率 (%)	F 値	正解率 (%)	F 値	正解率 (%)	F 値	正解率 (%)	F 値	正解率 (%)	F 値	正解率 (%)	F 値
ベースライン 1	99.7	0.50	92.2	0.48	96.7	0.33	94.4	0.11	87.9	0.19	94.9	0.32	99.6	0.50
ベースライン 2	98.4	0.62	—	—	92.8	0.56	95.1	0.50	89.3	0.54	95.9	0.59	—	—
GRMM 1	99.6	0.60	92.9	0.69	96.5	0.42	95.0	0.40	89.2	0.48	95.4	0.51	99.8	0.66
GRMM 2	99.7	0.68	95.6	0.82	98.1	0.57	96.7	0.60	94.9	0.68	97.0	0.64	99.8	0.66

表 3 実験結果 (左:正解率 (%), 右:F 値の平均)

“現代日本語書き言葉均衡コーパス” にタグ付けし、データを公開していく予定である⁴。データの蓄積と並行して、タグの仕様に関する検討も進める必要があるだろう。また、解析モデルについては、前節で考察した内容を踏まえ、モダリティ情報の解析に有用な手がかり語に関する知識を整備する予定である。

謝辞

本研究は、(独)情報通信研究機構委託研究「電気通信サービスにおける情報信憑性検証技術に関する研究開発」、および科研費特定領域研究「情報爆発時代に向けた新しいIT 基盤技術の研究」の公募研究「経験マイニング技術の高度化と実用化」(課題番号: 21013036、代表: 乾健太郎)、科研費若手研究(スタートアップ)「類義述語句同定のための語彙的知識の体系化と集積」(課題番号: 20800029、代表: 松吉俊)の支援を受けた。

参考文献

- [1] 川添愛、齊藤学、片岡喜代子、戸次大介。確実性判断に関わる意味的文脈アノテーション。情報処理学会研究報告書、2009-NL-189, pp. 77-84, 2009.
- [2] 国際交流基金、財団法人日本国際教育協会(編)。日本語能力試験出題基準【改訂版】。凡人社、2002.
- [3] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, Vol. 33, No. 1, pp. 159-174, 1977.
- [4] Marc Light, Xin Ying Qiu, and Padmini Srinivasan. The language of bioscience: Facts, speculations, and statements in between. In *Proceedings of BioLink 2004 workshop on linking biological literature, ontologies and databases*, pp. 17-24, 2004.
- [5] B. Medlock and T. Briscoe. Weakly supervised learning for hedge classification in scientific literature. In *the 45th ACL*, pp. 992-999, 2007.
- [6] 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.
- [7] 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、2007.
- [8] R. Prasad, N. Dinesh, A. Lee, A. Joshi, and B. Webber. Annotating attribution in the Penn discourse treebank. In *the COLING/ACL Workshop on Sentiment and Subjectivity in Text*, pp. 31-38, 2006.
- [9] V. Rubin, E. Liddy, and N. Kando. Certainty identification in texts: Categorization model and manual tagging result. In *Computing Attitude and Affect in Text: Theories and Applications*, pp. 61-74. Springer-Verlag, 2005.
- [10] R. Saurí, J. Littman, B. Knippen, R. Gaizauskas, A. Setzer, and J. Pustejovsky. *TimeML Annotation Guidelines Ver. 1.2.1*. <http://www.timeml.org/>, 2006.
- [11] R. Saurí and J. Pustejovsky. Determining modality and factuality for text entailment. In *the International Conference on Semantic Computing*, pp. 509-516, 2007.
- [12] R. Saurí and J. Pustejovsky. FactBank: A corpus annotated with event factuality. In *Language Resources and Evaluation*, 2009.
- [13] C. Sutton. Grmm: Graphical models in mallet., 2006. <http://mallet.cs.umass.edu/grmm/>.
- [14] G. Szarvas, V. Vincze, R. Farkas, and J. Csirik. The BioScope corpus: Annotation for negation, uncertainty and their scope in biomedical texts. In *the Workshop on Current Trends in Biomedical Natural Language Processing*, pp. 38-45, 2008.
- [15] 益岡隆志。日本語モダリティ探究。くろしお出版、2007。
- [16] J. Wiebe, T. Wilson, and C. Cardie. Annotating expressions of opinions and emotions in language. In *Language Resources and Evaluation*, 39 (2-3), pp. 165-210, 2005.

3 <http://www.tokuteicorpus.jp/>

4 タグ付与の作業基準やモダリティ情報解析に関する最新情報を、次のサイトで公開中である。構築したコーパスも同サイトから公開する予定である。
<http://cl.naist.jp/nltools/modalilty/>