

文間関係認識に基づく賛成・反対意見の俯瞰

水野 淳太^{†1} 渡邊 陽太郎^{†2} エリック ニコルズ^{†2}
村上 浩司^{†3} 乾 健太郎^{†2} 松本 裕治^{†1}

情報検索技術の発展により、あるトピックに関連する多様な情報を容易に入手できるようになった。しかしながら、これらの文書に記述されている情報には、不正確な記述、偏りのある意見などが多数混在している。そのため、個々の情報や意見の信憑性を判断するためには、多様な情報源からの意見との整合性を調べる必要がある。しかし、限られた時間で数多くの情報源を調べることは難しいため、ユーザが持っている先入観が正常な判断を妨げてしまう場合がある。我々は、そのような状態を避けるために、言論マップ生成課題に取り組んでいる。これは、検索された文について、まず、トピックに対する賛成意見であるのか、それとも反対意見であるのかを分類し、次に、賛成および反対する根拠を含むかどうかを認識し、それらを俯瞰的に示すというものである。本課題において最も重要な問題は、1組の文対が与えられたときに、その間の意味的关系を分類する文間関係認識である。これは近年さかんに研究されている含意関係認識と重なる部分が多い。しかしながら、ウェブ上の実文に対して既存の含意関係認識を適用しても、その分類性能は限定的であるという報告がある。そこで、我々は、評価用データセットとその分析に基づく文間関係認識モデルを構築した。本論文では、検索された文において、クエリの内容に対応する部分を正しく同定することが、最も重要な技術的課題であること、また、いくつかの制約を変化させることで、関係分類の精度と再現率を制御できることを示す。

Organizing Agreeing and Conflicting Opinions Based on Semantic Relation Recognition

JUNTA MIZUNO,^{†1} YOTARO WATANABE,^{†2}
ERIC NICHOLS,^{†2} KOJI MURAKAMI,^{†3} KENTARO INUI^{†2}
and YUJI MATSUMOTO^{†1}

Internet users are presented with a wealth of information on a variety of topics, but finding reliable opinions is a challenge that demands increasingly more of their time, knowledge and patience. In this paper, we present a system that helps users make sense of the the opinions on the Web through automatic

viewpoint detection. Our system organizes opinions into supporting and opposing viewpoints and detects meta-information such as evidence that is useful for evaluating the reliability of each viewpoint. The most important problem in viewpoint detection is recognizing semantic relations between user queries and Web texts, which is similar in nature to entailment recognition, a task extensively researched in recent years. However, the performance of current approaches on Web texts is limited. We, therefore, develop a semantic relation recognition system targeting Web texts by constructing and conducting detailed analysis of an evaluation dataset. In this paper, we show that the most important technical issue for semantic relation recognition targeting Web texts is correctly detecting the region in retrieved Web texts that correspond to the content of the user query, and we demonstrate that it is possible to control the precision and recall of semantic relation classification by changing various restrictions.

1. はじめに

情報爆発時代を迎えた今日、ウェブ上には大量のテキスト情報が存在し、そこでは様々なトピックに関して多様な意見が述べられている。情報検索技術の発展により、あるトピックに関連する文書集合を容易に入手できるようになった。しかしながら、これらの文書に記述されている情報は、そのすべてが真実というわけではなく、不正確な記述、偏りのある意見などが多数混在している。そのため、個々の情報や意見の信憑性を判断するためには、多様な情報源からの意見との整合性を調べる必要がある。しかし、限られた時間で数多くの情報源を調べることは難しいため、ユーザが持っている先入観が正常な判断を妨げてしまう場合がある。それを確認バイアスが高い状態であるという。

我々は、そのような確認バイアスを軽減するために、言論マップ生成課題¹⁷⁾に取り組んでいる。これは、ユーザが与える言明に対して、それに関連する文をウェブ上から収集し、その言明に賛成する意見や、反対する意見を抽出し、俯瞰的に示すものである。反対意見も同時に見せることが確認バイアスの軽減に有効であることは心理学的にも明らかになっている^{5),11)}。

^{†1} 奈良先端科学技術大学院大学

Nara Institute of Science and Technology

^{†2} 東北大学

Tohoku University

^{†3} 楽天技術研究所

Rakuten Institute of Technology

そのような俯瞰表示を行うためには、1組の文対について、一方が他方に対してどのような意見であるかを認識する技術が必要である。2文間の意味的關係を分類する課題を文間関係認識と呼ぶ。文間関係認識は近年さかんに研究されている含意関係認識⁶⁾と重なる部分が多く、含意関係は賛成意見に、矛盾関係は反対意見に含まれる。したがって、含意関係認識に取り組むことで文間関係認識、ひいては言論マップ生成課題の実現可能性を探ることができると考えられる。しかしながら、日本語での含意関係認識に相当する課題は、NTCIRの評価型ワークショップ RITE^{*1}が初めてであり、依然としてデータセットを構築している段階にある。

そこで、本研究では、まず、人手で収集したいくつかのクエリについて関連文(検索対象文)をウェブ上から獲得、関係ラベルなど必要な情報を付与することで評価データセットを構築した。次に、その間の意味的關係を分類する文間関係認識課題を設計し、それを解くためのモデルを構築した。現時点で、図1に示すように、オープンメインな任意のクエリに対して、言論マップを構築することが可能となっている。この例では、「うがいは風邪予防に効果的だ」というクエリに対して、ウェブ上の賛成意見や反対意見などを分類し、俯瞰的に提示している。それらの根拠となる情報はハイライトし、各関係の下に配置する。また、本論文では言及しないが、条件付き賛成意見と条件付き反対意見をまとめた一部反対意見への分類にも取り組んでいる¹⁸⁾。

情報信憑性判断の支援に対するこうしたアプローチを進めるためには、3つの研究フェーズがあると考える。第1に、図1に示すような賛成意見、反対意見への分類、およびそれらの根拠を抽出する技術を開発する。第2に、分類した意見をさらにクラスタリングすることによる論点の整理、情報発信者の特定などによって、より高次の組織化を行う。第3に、実際の心理学的実験により、開発した技術が情報信憑性判断の支援に貢献することを確かめる。

本論文では、第1のフェーズについて論じる。具体的には、文間関係認識において、2文間で類似・関連する単語の意味的な対応も正しくとることで、多くの問題は簡単な関係分類手法で正しく分類が可能であることを示す。また、アライメントにおけるいくつかの制約を変化させることで、関係分類の精度と再現率を制御できることを示す。

以下、2章では関連研究について述べ、3章で言論マップ生成課題の設計と言論マップ生成システムについて述べる。4章で約1,500文対からなる評価データセットに対して本シス

*1 http://artigas.lti.cs.cmu.edu/rite/Main_Page



図1 言論マップ生成例

Fig. 1 A screenshot of the Statement Map system.

テムを適用することで技術的課題を明らかにし、5章で今後の課題について述べる。

2. 関連研究

ユーザによる情報の信頼性判断についての関連研究として、WISDOM¹⁾と Dispute Finder⁸⁾があげられる。WISDOMでは、情報発信者の同定と評価極性(良いと評価するか、悪いと評価するか)の判定に基づいて意見を分類する。評価極性による意見の分類は、WISDOMに限らず、すでに多数の研究がある¹⁹⁾、言論マップにおける意見の分類は、クエリと検索対象文との間の意味的關係の分類であるという点で異なる。たとえば、「南極の氷は減っている」というクエリに対して、「南極の氷はトータルでは増えます」という対立意見は、評価極性ではとらえられない。

Dispute Finderではまず、多数のユーザが集合知的な仕組みを使い、世の中で論争になっている言明を集め、それらの言明を肯定するページや否定するページの情報とともにデータ

ベースに登録する。次に、ユーザが閲覧中の文章にデータベース中の言明が含まれているかどうかを判定し、含まれていたらユーザに警告し、肯定ページ・否定ページの情報を提示する。ただし、そのようなページは人手で収集することを仮定しており、そのカバレッジには自ずと限界がある。我々は、クエリに関連する文集合をウェブ上から獲得し、自動的に分類することで、オープンドメインな任意のクエリに対応することができる。

複数文書中の文間の関係解析には、CST (Cross-Document Structure Theory)²¹⁾がある。RST (Rhetorical Structure Theory)²⁶⁾に基づく談話構造解析が単一文書内の構造を解析するのに対し、CSTはこれを文書横断構造解析に拡張するものであり、18種類の意味的關係が定義された。CSTでは、Zhangらが比較する2文からのみ素性を抽出して素性空間に表現し、主要な関係を1つの分類器により分類を行った²⁷⁾。また、正解ラベルが付与されたデータが少ないためラベルなしデータも用いたBoostingを利用した手法を提案した²⁸⁾がどちらにおいても精度は高くなく、類義語や反義語などの語彙知識の適用、素性の洗練、各意味的關係認識のための個別処理の検討など、多くの課題を残した。このことから、複数の意味的關係クラスへの分類手法そのものの確立も重要な課題となる。複数文書要約では、反対意見に着目した研究も報告されている¹²⁾。そのような要約を実現するためには、反対意見の分類が重要であり、評価極性に基づく手法が提案されている²⁰⁾。彼らの目的は矛盾のない要約文の生成であるが、我々はウェブ上の様々な意見を俯瞰的に提示することが目的である。また、我々は評価極性では計ることのできない事実的な内容にも着目している。

前述のとおり、文間関係認識は含意関係認識と重なる部分が多い。含意関係は賛成意見に、矛盾関係は反対意見に含まれる。含意関係認識は多様な文を対象とした総合的な課題であり、多数の論文がすでに発表されているが、いまだ十分な分類性能は得られていない。実際に、ウェブ上の実文に適用したところ、特に矛盾関係はほとんど認識できなかったという報告もある⁷⁾。したがって、まずは問題を丁寧に切り分け、その技術的課題を分析することが重要である。

さらに、含意関係認識では、一方の文において他方の文の内容が対応する部分にのみ着目しており、一方の文にのみ含まれる付加的な情報には着目されていない。我々は、付加的な情報の中でも一方の文の内容が成立するための根拠となる情報に着目し、根拠関係というこれまでの含意関係認識にはない新しい関係に拡張している。根拠関係は、典型的な賛成意見や反対意見には含まれていない有用な情報をユーザに提供するため、この関係を認識し、提示することは重要である。

根拠関係認識はWHY型質問応答と重なる部分が多い。Higashinakaら¹⁰⁾は、回答文候

補に対して、(1)理由を示唆する表現を含む、(2)質問文と内容的に類似している、(3)質問文との間に何らかの因果関係がある、という3つの尺度を適用することで回答を選択した。Verberneら²⁴⁾はRSTコーパスに含まれる内容に対するWHY質問を構築し、RSTの談話構造によって58%の質問に答えられることを示した。そのアプローチは、“satellite”が因果関係に相当し、それが“nucleus”を修飾することを利用し、質問と類似する“nucleus”に対する“satellite”を回答として出力する。我々の根拠認識アプローチはこれらの手法と類似しているが、クエリと“nucleus”に相当する部分との対応付けに対して、後述するライメントを用いてより厳密に対応をとる点が異なる。

3. 言論マップ

3.1 言論マップ生成課題

まず、言論マップ生成課題について、いくつかの定義をする。

第1に、意味的關係はクエリと検索対象文から得られる情報だけで決定し、検索対象文の前後の文などは参照しないという制限を設ける。他の文を参照すると、照応解析や共参照解析などが必要となり、問題が複雑化してしまうためである。

次に、分類の対象とする関係の粒度について、含意関係認識はその目的から〈含意〉〈矛盾〉〈不明〉の3つの関係に分類するが、言論マップ生成課題では〈同意〉〈同意根拠〉〈対立〉〈対立根拠〉〈その他〉の5つの関係に分類する。それぞれの関係の定義は「言明間意味的關係コーパス作成のための意味的關係仕様書³¹⁾」に基づく。

以上の条件の下で、入力されたクエリに関連する文を検索し、得られた文集合を前述の5つの関係に分類するのが言論マップ生成課題である。本論文の目的は、この問題に対する我々のアプローチを報告し、そこから明らかになった技術的課題を論じることであり、そのため、以下の3つの方針を前提とする。

- (1) 述語項構造解析やモダリティ解析といった深い意味解析を行う。これらの解析精度はまだ十分とはいえないが、最新の研究成果を用いる。
- (2) 単語間の意味的關係を判断するうえで、大規模な語彙知識は必要不可欠である。したがって、既存の利用可能な語彙知識を積極的に利用する。
- (3) この課題にはどのような問題が含まれているのかを分析することが重要であるため、安易に機械学習に頼るのではなく、制約を段階的に緩和する手法を用いて分析を容易にする。それにより、問題の性質を明らかにする。

以下の節では、どのようにして本課題を実現するか、具体的な手法について述べる。

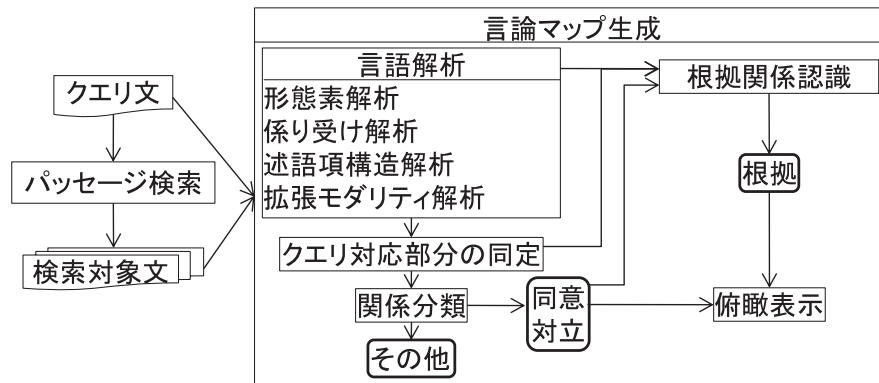


図 2 言論マップ生成システム

Fig. 2 The architecture of the Statement Map system.

3.2 言論マップ生成システム

言論マップの生成過程を図 2 に示す。入力には、自然文クエリを想定している。入力されたクエリに対してパッセージ検索を行い、大量の検索対象文を獲得する。クエリと検索対象文は、それぞれ含意関係認識の仮説 (hypothesis) とテキスト (text) に対応する。最後に、クエリと各検索対象文の間の意味的关系を分類し、俯瞰表示を行うことで、図 1 のような言論マップを得ることができる。パッセージ検索は既存のウェブ検索エンジンを用いることで比較的容易に構築することができる。したがって、本研究において重要なのは、2 文間の意味的关系分類をどのように行うかという点である。

2 文間の意味的关系を同定する課題として代表的な課題は含意関係認識である。文間関係認識は含意関係認識よりも多様な意味的关系を取り扱うが、基本的なアプローチには共通する部分が多い。そのアプローチは、変形に基づく方法とアライメントに基づく方法に大きく分けられる¹⁶⁾。変形に基づく方法は、受動態と能動態の変換や上位語と下位語の入れ替えなどの変形を行い、一方の文が他方の文に変形可能である場合に (含意)、それ以外の場合は (その他) に分類する。この方法は文対が含意関係にあるかどうかを判断することが中心であるため、多様な関係に分類するには不向きである。そこで、本研究ではアライメントに基づく方法を採用する。この方法は、1) 2 文それぞれを言語解析し、2) 文間で類似・関連する単語に対応付け (単語アライメント)、3) 1), 2) の情報を利用して関係分類を行う。

含意関係認識では、仮説、テキストのいずれもある程度の長さの文である²⁾。そのため、

述語項構造解析だけでなく、共参照解析や照応解析といった高度な言語処理技術が要求され、問題が発散する傾向にある。

そこで、本研究では 2 つの条件を設けることで問題を限定し、最も注力すべき課題を明らかにしていく。1 つめの条件として、クエリを単文、すなわち、1 つの述語といくつかの項からなる文に限定する。2 つめは、検索対象文はクエリ中の名詞 (事象を示す名詞を除く) を必ず含み、かつ比較的長い文であるという条件である。たとえば、「キシリトールは虫歯予防に効果がある」というクエリに対して、「キシリトール」、「虫歯」は検索対象文に含まれていなければならない。「効果」は名詞だが、「ある」が機能動詞であることから、「効果が_ある」で事象を示すと考え、「予防」は事象性名詞であるため、検索対象文に必ずしも含まれなくてよい。一方で、クエリ中の述語については条件を設けない。それにより、検索対象文はクエリとの関連性が高く、かつ多様な述語が含まれる。

また、先行研究では、名詞のアライメントと、述語のアライメントは同じレベルの問題として取り扱われている¹⁵⁾。しかし、名詞のアライメントは、その問題の大部分が、該当する 2 単語間の関係が語彙知識中に存在するかどうかである。それに対して、述語は語彙知識だけでなく、一方が複文節からなる対応 (例: 効果が_ある - 効果的だ) や、名詞に述語の内容が含意される (例: 牛乳は_健康に_良い - 牛乳は_健康食品だ - 前者の「良い」は後者の「健康食品」に含意される) など、対応の仕方が多様である点が異なる。したがって、名詞のアライメントと述語のアライメントは分けて考えるべき課題であり、本研究では述語のアライメントに注力する。

以下では、まず各要素技術について図 3 を例に、簡潔に述べる。次に、最も重要なクエリ対応部分の同定と関係分類について詳述する。

パッセージ検索 ウェブ検索エンジンの TSubaki²³⁾ を用いてクエリに関連した文集合を獲得する。獲得された文には、タイトル文や単語の羅列といった文として不自然なものも含まれているため、文字数や文末の品詞といった条件によって、それらを除外する。獲得方法の詳細は (4.1 節) で述べる。図 3 では、 Q に対して関連文である T を抽出することである。

言語解析 クエリと検索対象文に対して、形態素解析¹⁴⁾、係り受け解析¹³⁾、述語項構造解析²⁵⁾、拡張モダリティ解析³⁷⁾、評価極性判定を行う。

評価極性判定は、名詞については日本語評価極性辞書 (名詞編)³⁴⁾ を、述語については日本語評価極性辞書 (体言編)²⁹⁾ を用いて行う。述語については辞書を引くだけでなく、その述語が存在または非存在を示唆する場合は、そのガ格の項の名詞が持つ評価極性に

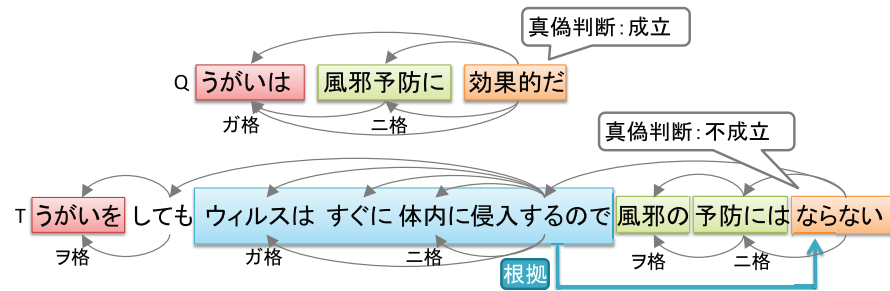


図 3 関係分類の実例
Fig. 3 An example of semantic relation classification.

よって述語の評価極性を決定する。たとえば、「(Positive) がある」や「(Positive) が多い」といった場合、その述語にも名詞と同じ評価極性 (Positive) を付与する。逆に、「(Positive) がない」といったように、述語が評価極性を反転させる作用を持つ場合、述語の評価極性は Negative となる。また、述語が否定を示す場合もその評価極性が反転することから、拡張モダリティ情報の真偽判断もあわせて考慮する。さらに、「～になる」のように状態変化を示唆する述語の場合、その二格の項の名詞の状態に変化すると考え、二格の名詞の評価極性を付与する。

図 3 に言語解析の結果を示す。依存構造は文の上側の矢印で示され、述語項構造は下側の矢印とラベルで示される。拡張モダリティ解析は、すべての述語「予防 (Q)」「効果的だ (Q)」「して (T)」「侵入 (T)」「する (T)」「予防 (T)」「ならない (T)」に対して行う (図では簡単のため、2 述語の真偽判断のみ示した)。評価極性判定では、「風邪 (Q)」に Negative、「予防 (T)」に Positive といった評価極性を付与し、状態変化を示唆する「ならない (T)」には、その二格である「予防 (T)」の評価極性 Positive と、不成立という真偽判断情報から Negative を付与する。

クエリ対応部分の同定 検索対象文において、クエリが内容的に対応する部分を同定する。

詳細は 3.3 節で述べる。図 3 では、Q 中のすべての文節が、T に対応する。

関係分類 関係分類は 2 段階の分類によって行う。詳細は 3.4 節で述べる。

- (1) 〈その他〉と〈関係あり〉に分類する。〈関係あり〉は〈同意〉〈対立〉などを含めて、クエリと検索対象文の間に〈その他〉以外の意味的關係が存在することを示す。この分類は、検索対象文中にクエリが内容的に対応する部分があるかどうかによって判断される。

- (2) 〈関係あり〉に分類された文対を〈同意〉と〈対立〉に分類する。この分類は、クエリ中の単語と、対応付けられた検索対象文側の単語とが反義語や否定の関係にあるかどうかによって判断される。

図 3 では、まず T 中で Q に対応する部分があることから、〈関係あり〉に分類される。次に、対応する述語「効果的だ (Q)」と「ならない (T)」の評価極性が異なることから、〈対立〉に分類される。

根拠関係認識 〈同意〉〈対立〉のいずれかに分類された文対の検索対象文中でクエリの内容に相当する部分に対して、その根拠となる情報が修飾関係にある文対をそれぞれ〈同意根拠〉と〈対立根拠〉に細分類する。詳細は 3.4.1 項で述べる。図 3 では、「ウィルスはすぐに体内に侵入するので」が「風邪の予防にはならない」ことの根拠であることから、Q と T の間の意味的關係は〈対立根拠〉に分類される。

俯瞰表示 最終的に図 1 のように関係ごとに俯瞰的に表示する。それぞれの検索対象文から元ページを参照できるようにハイパーリンクを設定する。

3.3 クエリ対応部分の同定

本節では、クエリの内容が、検索対象文のどこに対応するかを同定する手法について述べる。後段の関係分類では、その対応部分を見ることで、2 文間の意味的關係を分類する。先行研究では、2 文間で類似・関連する単語間に対応付ける単語アライメントによって、クエリ対応部分の同定を行っている^{15),22)}。しかし、文の意味も考えると、類似単語であっても対応付けを行うべきではない場合がある。それに対して、単語アライメントに加えて文構造も素性として入力する関係分類が提案されている⁹⁾。しかし、我々はこの問題を関係分類とは切り分けるべきであると考えている。

まず、対応付ける単位について、英語を対象とした場合は単語単位で対応付ける単語アライメントが一般的だが、日本語では「形態素」か「文節」のいずれかである場合が多い。どちらがより良いアライメント単位であるかは自明ではないが、本研究では文節単位で対応付ける文節アライメントを採用する。文節単位で対応付ける場合、「効果的だ」と「効果がある」のように、同じ意味だが、文節数が異なる場合の対応が必要となる点が問題となる。本研究では、1 対多の文節アライメントを許すことでこの問題に対処している。形態素単位で対応付けた場合でも、「予防」と「予め防ぐ」のように、形態素数が異なるという、本質的に同じ問題があり、どちらがより良いアライメント単位であるかは明らかではない。形態素単位で対応付けた場合との差異については今後の課題とする。

次に、以下の 3 つの手順によってクエリ対応部分を同定する。

- (1) 文節アライメント クエリ-検索対象文間で類似・関連する単語を含む文節を対応付ける．
- (2) 局所構造アライメント クエリ側の2文節間の意味的關係と、その2文節が(1)で対応付けられた検索対象文側の2文節間の意味的關係を対応付ける．
- (3) 文節アライメント選択 (1)で対応付けられた文節アライメントのうち、(2)で対応付けられた文節アライメントのみを選択する．次に、クエリ中の主述語(文末の述語)に対応する検索対象文側の述語を決定し、クエリと対応する部分を同定する．

これらの手順は、大きく2段階に分けられる．まず文節アライメントによって、クエリと対応する部分の候補をあげ、次に局所構造アライメントと文節アライメント選択によって候補の中から対応部分を選択する．文節アライメントは先行研究でも用いられているが、局所構造アライメントは本研究において新たに導入した．次項以降では、これら3つの手順について詳しく述べる．

3.3.1 文節アライメント

まず、クエリと検索対象文の間で類似・関連する文節を対応付ける．類似・関連性の判断には以下の3つの類似度を利用する．これらのうち(1)、(2)は従来の含意関係認識においても用いられてきたが、(3)は本研究において新たに導入した．

(1) 表層的な類似度

クエリ側の文節中の内容語が検索対象文中の1文節にすべて含まれるとき完全一致として対応付け、すべてではなく一部であるとき部分一致として対応付ける．

(2) 語彙知識に基づく類似度

表層的には類似していないが、既存の語彙知識中でエントリが存在する文節を対応付ける．語彙知識は日本語 WordNet³⁾、動詞含意関係データベース³⁸⁾、Wikipedia から自動抽出した上位下位関係シソーラス³⁶⁾、日本語異表記データベース^{30),35)}、事象知識ベース³²⁾を用いた．

語彙知識中で関連のある単語対には類義語や上位下位といった様々な関係が定義されている．前提関係以外のすべての関係を利用して、文節アライメントを行う．その際、語彙知識中で直接関係のある単語対のみを利用する．つまり、3つ以上のエントリを利用して、別の単語を介してエントリを拡張することはしない．たとえば、2つの同義関係のエントリがあり、かつそれらのエントリに含まれる4単語のうち2単語が同一の場合、つまりA-BとB-Cというエントリがある場合は、A-Cも同義関係である可能性が高いが、上位下位関係などについてはこのような単純な推論は行えないためである．

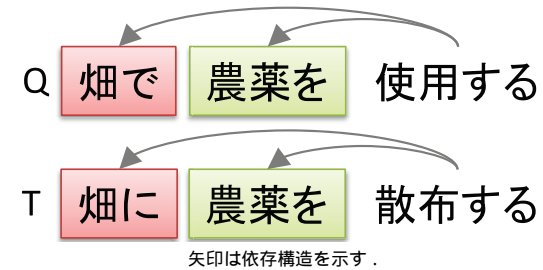


図4 文節アライメントにおける文構造に基づく動的推定例
Fig. 4 An example of structural similarity based alignment estimation.

(3) 文構造に基づく動的推定

実世界には大規模な語彙知識を用いてもカバーされていない単語対が存在する．特にドメイン固有の単語において顕著である．たとえば、「畑で農薬を使用する」と「畑に農薬を散布する」において、「使用する」は「散布する」を含意するが、前節の語彙知識には該当するエントリは存在しない．ここで、分類の対象となる2文は、少なくともトピックを共有しており、内容的な関連性が高い．したがって、2つの述語が持つ項が2項以上類似している場合、それらの述語は類義語である可能性が高いと予測される．本予測に基づき、文構造の類似度に基づいて文節アライメントを付与する．具体的には、2つの述語に対して、係り元の文節や項が、2文間において2文節または2項以上文節アライメントされている場合、その2つの述語に対して文節アライメントを付与する．

図4に本手法の適用例を示す．まず、QとTの間で「畑で(Q)」と「畑に(T)」および「農薬を(Q)」と「農薬を(T)」が、(1)で定義した表層的な類似度に基づいて対応付けられる．次に、「使用する(Q)」と「散布する(T)」に着目すると、これらの係り元の文節はそれぞれ「畑で(Q)」と「農薬を(Q)」、「畑に(T)」と「農薬を(T)」である．これら2組の文節対は、文節アライメントされているため、「使用する(Q)」と「散布する(T)」は文構造という観点で類似しており、文節アライメントを付与する．

また、2つの述語が以下であげる属性レベルで同じ場合、そうでない場合に比べて文節アライメントされやすいと考え、構造的な類似性が低い場合(1文節または1項)でも対応付ける．

(a) 存在/非存在を示唆

2つの述語がいずれも「ある」や「少ない」など、存在または非存在を示唆する述語である場合は、類似文節または項が1組であっても対応付ける。どのような述語が存在/非存在を示唆するかは、人手でリスト化することによって決定する。

(b) 評価極性を持つ

2つの述語がいずれも評価極性を持つ述語である場合、すなわち、2つの述語が同じ評価極性を持つ、または異なる評価極性を持つ場合は、(a)と同様に、条件を弱める。

この中で、(1)の制約は最も確実だが、これだけでは多様な表現に対応することはできない。そこで、(2)の制約も用いることで、多様な表現へ対応させることができる。しかし、(2)もドメインによっては成立しない語彙知識を含むため、誤った文節アライメントを含むことになる。また、語彙知識の多くは名詞に関するものであり、述語に関してはまだ不足している。そこで、(3)で推定することで述語の文節アライメントのカバー率を向上させることが期待できる。しかし、(3)はあくまで推定手法であり、その手がかりとして重要な述語項構造解析はまだ十分な性能ではない。したがって、誤った文節アライメントを含みやすくなる。

図3では、「うがいは(Q)」は「うがいを(T)」と、「風邪予防に(Q)」は「風邪の_予防には(T)」と(1)によって対応付けられる。「効果的だ(Q)」は(1)および(2)では対応付けられないが、(3)によって対応付けられる。すなわち、「効果的だ(Q)」はPositive、「ならない」はNegativeの評価極性を持つ述語であり、それぞれの係り元文節である「風邪予防に(Q)」と「予防には(T)」は対応付けられているため、(3b)によって対応付けられる。

3.3.2 局所構造アライメント

(1) Q ブラックバス_aは生態系_bを破壊する_c。

T_1 ブラックバス_aは有名な外来魚であり、生態系_bを破壊する_c。

T_2 ブラックバス_aのように獰猛な魚類が生態系_bを破壊する_c。

T_3 ブラックバス_aを駆除することが生態系_bを破壊する₂。

文節アライメントによって、文間で類似・関連する文節間に対応付けを行ったが、文の意味も考えると、類似文節であっても対応付けを行うべきではない場合がある。Changら⁴⁾は、クエリ側で依存関係にある2単語に対して、それが単語対応するテキスト側の2単語も依存関係にある場合のみ単語アライメントとして採用するという制約を含意関係認識に

適用した。しかしながら、 $T_{1a} - T_{1c}$ のように、依存構造を持たない場合には有効ではない。 T_1 では、述語項構造解析によって「ブラックバス-ガ-破壊」という格関係が成り立つことが分かり、 $T_{1a} - T_{1c}$ はQ中の依存構造 $Qa - Qc$ と対応付けられる。次に、 T_2 では、同じように「魚類-ガ-破壊」という格関係が成り立ち、かつ「ブラックバス」が「魚類」の例示関係にあることから、「ブラックバス」と「破壊」の間にもガ格に相当する関係が成り立っていると考えられるため、 $T_{2a} - T_{2c}$ は $Qa - Qc$ と対応付けられる。 T_1, T_2 のように、多くの場合は類似した文節に対応付けを行うことで、意味的な対応もとれている。しかし、 T_3 では、「ブラックバス」と「破壊」の間に格関係は成り立たないため、 $T_{3a} - T_{3c}$ は $Qa - Qc$ と対応付けるべきではない。局所構造アライメントは、 $Qa - Qc$ と $T_{1a} - T_{1c}$ 、および $T_{2a} - T_{2c}$ をペアとして対応付け、 $Qa - Qc$ と $T_{3a} - T_{3c}$ は対応付けられないようなアライメントである。

このようなアライメントを実現するために、後述する開発データに対して、人手で局所構造アライメントの正解を付与し、分析を行った。ここで、クエリは3~5文節程度からなる簡単な文であるため、クエリ側は依存構造上で直接修飾関係にある2文節を対象とするだけで十分であると考えられる。したがって、検索対象文側で、局所構造対応する2文節間にはどのような構造があるのかを分析することが重要である。

その結果、依存構造や述語項構造だけでなく、 T_2 のように例示も考慮しなければならないような構造など、様々な種類があった。この問題を機械学習によって局所構造対応するかどうかの二値分類問題とすることも考えられるが、様々な構造をまとめて扱ってしまうと、正しく機械学習させることは難しい。また、前述のとおり、問題を分析するという観点でも、すぐに機械学習に持ち込むことは得策とはいえない。そこで、以下の4つの制約を実装し、その組合せによって、段階的な制約緩和を実現する。

(1) 検索対象文側で対応する2文節が、依存構造上で直接修飾関係にある。

(2) 2文節間に述語項関係が存在する。このとき、格関係の種類は問わない。これによって、 $Qa - Qc$ と $T_{1a} - T_{1c}$ を対応付けることができる。

(3) 依存構造上で直接修飾関係になくても、いくつかの文節を介して修飾関係にある。しかし、いくつ文節を介してもよいわけではなく、その上限は開発データを分析した結果から、前後4文節までとした。この制約を適用することで、たとえば図5のように、 T_2 では2つの文節を介して修飾関係にある場合に局所構造アライメントを付与することができる。

(4) 2文節のうち、一方が助詞の「は」で終わる文節である。そのような文節は、表層上

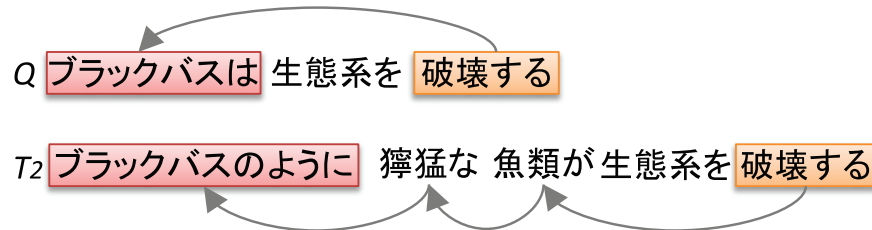


図 5 一方がいくつかの文節を介して修飾関係にある制約
Fig. 5 Alignment restrictions based on dependency path length.

でそれに続く数文節に対して関係があると見なすためである。

3.3.3 文節アライメントの選択

前述の局所構造アライメントでは、クエリ中の各依存構造に対して、検索対象文中のどの 2 文節間の構造と対応付けられるかを判断した。したがって、いずれの文節とも対応付けられなかった文節は、クエリ中の文節とは対応しないと判断できる。次に、局所構造アライメントが付与された文節アライメントのうち、どの文節に着目することが関係分類にとって有用であるかを判断する。

具体的には、まず、クエリ中の述語を中心とした依存構造に対応する構造を最も多く保持している検索対象文中の述語を選択する。次に、その述語と局所構造対応する文節を選択する。

図 6 に例を示す。この例では、クエリの「良い(Q)」に対応する文節は「悪いので(T)」と「良い(T)」の 2 つである。ここで、クエリ中の述語は、「マーガリンは(Q)」と「体に(Q)」の 2 文節を係り元として持つ。この依存構造と同じ構造を局所構造アライメントによってより多く保持している述語は「悪いので(T)」であり、この例では 2 文節とも対応している。「良い(T)」は「体に(T)」の 1 文節のみを保持しているだけであるため、選択されない。最後に、「悪いので(T)」と局所構造対応する文節として、「マーガリンが(T)」と「体に(T)」を選択する。結果として、「1」、「2」、「4」の 3 つの文節アライメントが選択される。

述語の選択が難しい例として、検索対象文中にクエリと同意・対立の両論が記述されている場合がある。この場合、上記の手法では、対応する述語を一意に決定できないため、文末に近い述語が結論である可能性が高いと考え、文末に近い述語を選択する。

	マーガリンは	体に	良い
マーガリンが	①		
体に		②	
悪いので			④
体に		③	
良い			⑤
バターを			
使おう			

丸が文節アライメントを示し、それらをつなぐ線が局所構造アライメントを示す。

図 6 文節アライメントの選択

Fig. 6 Phrasal alignment selection.

3.4 関係分類

以下の 2 つの仮説に基づき、2 文間の関係を分類する。

仮説 1 クエリと検索対象文との間の意味的關係があるとすれば、クエリ中のすべての文節は、文節アライメントと局所構造アライメントによって検索対象文中のいずれかの文節に対応付けることができる。

仮説 2 クエリと検索対象文との間に意味的關係があると判断された場合、具体的な関係は、対応付けられた文節間の意味的關係によって決定できる。

この仮説に基づく関係分類を、以下の 2 段階の分類によって実現する。

(1) 〈関係あり〉と〈その他〉の分類

文節アライメント選択により選択されたアライメント集合を用いて、クエリと検索対象文の間に意味的關係があるかを分類する。仮説 1 より、クエリに含まれるすべての文節が、アライメント集合に含まれている文対を〈関係あり〉、それ以外の文対を〈その他〉に分類する。

(2) 〈同意〉と〈対立〉の分類

1 で〈関係あり〉に分類された場合、仮説 2 に基づいて〈同意〉または〈対立〉に分類する。2 文間の意味的關係が〈対立〉となるのは、述語が否定の關係にある場合(例：

効果的だ(成立)-効果的でない(不成立)), 評価極性が異なる場合(例:効果的だ(Positive)-無駄だ(Negative)), 反義語の関係にある場合(例:減少する-増加する)である。それ以外の場合は〈同意〉に分類する。述語の否定関係は、拡張モダリティの真偽判断によって判断する。評価極性の判定は前述の方法によって行う。反義語であるかは、文節アライメントの判断に用いられた語彙知識の情報によって判断する。

3.4.1 根拠関係の認識

検索対象文がクエリに対して〈同意〉または〈対立〉の意味的關係を持ち、かつその関係を支持する根拠を含んでいる場合の意味的關係を、それぞれ〈同意根拠〉、〈対立根拠〉と定義する。

(2) Q キシリトールは_a虫歯予防に_b効果的だ_c。

T キシリトールは_a虫歯の原因となるミュータンス連鎖球菌の増殖を抑制する効果があるため、虫歯予防に_b効果的だ_c。

(2) では、T において、「あるため(T)」と「効果的だ(T)」は根拠を示す談話標識「ため」によって接続されている。そして、アライメントされている部分のみに着目すると、2 文間の関係は〈同意〉であることから、2 文間の関係は〈同意根拠〉である。

根拠関係の認識手法を図 3 で説明する。まず、根拠を示す談話標識である「から、ので、ため」が検索対象文中に含まれるかどうかを判断する。次に、談話標識を含む文節と、その係り先の文節の情報を用いて、根拠と結論が存在し、かつその間に因果関係があるかどうかを判断する。「侵入する_dので_eならない(T)」がそれを示している。最後に、結論とクエリが対応するかどうかをアライメントを用いて判断する。ここでは、「効果的だ(Q)」が「ならない(T)」と対応している。そして、関係分類によって、この 2 文間の関係が〈対立〉であることが分かっているので、Q と T の意味的關係は〈対立根拠〉に分類される。

4. 評価実験

評価実験では、現状の関係分類性能を確認する。特に、アライメントにおいて、どの程度柔軟に対応付けることで、関係分類性能にどのような影響を与えるのかについて議論する。

4.1 実験設定

実験では、5 種類のクエリに対してウェブから獲得した 1,326 文対を開発データとし、同様に 20 種類のクエリに対してウェブから獲得した 1,467 文対を評価データとした。評価データの一部を表 1 に示す。5 種類のクエリと 20 種類のクエリの間には、トピックレベルで重

表 1 評価データの一例
Table 1 Example evaluation data.

クエリ (クエリ単語)	関係	検索対象文
リサイクルは環境に良い (リサイクル, 環境)	同意 対立	かけがえのない地球の環境をより良くしていく為に、全社員が環境に優しい商品づくり、リサイクル商品づくりとその技術開発に取り組んでいます また、ほぼ全てのリサイクルは環境によいわけではなく、「リサイクル=環境に優しい」とは限らない
温暖化によって海面が上昇する (温暖化, 海面)	同意 同意根拠	地球温暖化が進む海面上昇で、深刻な被害が出始めている 温暖化が原因で南極の水が溶けて海面が上昇するなんてことも聞きますが、海面が上昇するのは、温暖化で温度が上がることで海の温度も上がって、水が膨張するのが主な原因です その当時から、外国に行くとき水道の水は飲んでダメだと言われ、ホテルに泊まるとほとんどの国でミネラルウォーターのペットボトルが飲料水用に置いてあります 東南アジアでは、半透明なボトルで売られている安いミネラルウォーターがありますが、これは薬品を混入させているらしく下痢を誘発する事もあります
ミネラルウォーターは水道水より安全だ (ミネラルウォーター, 水道水)	同意 対立	納豆のダイエット効果については、某番組の影響もあって、ウソだと思ふ人もいますが、 「私はA型人間です」とか「やっぱりB型の性格が出ているわね」などと、血液型によって性格を判断することがよくあります 自分や周りの人間の性格をみて、血液型性格学の記述があたってる、と思うこと時々あるかも知れませんが、これは心理学的トリックによるものです
納豆のダイエット効果はある (納豆, ダイエット)	同意	
血液型で性格が分かる (血液型, 性格)	同意 対立根拠	

なりはない。すべての文対に対して関係ラベルの付与を行った。付与基準は前述の仕様書に基づく。作業者はシステムの開発に関わっていない 1 名である。付与作業の一貫性を計るために、20 クエリのうち 10 クエリについて別の作業者が付与した結果、その一致率は κ 値で 0.72 であった。本実験では、1 人目の作業者が付与した結果を用いる。

評価データの構築方法について説明する。まず、20 種類の自然文クエリを用意する。クエリは、賛成・反対意見およびその根拠情報の俯瞰という言葉マップ生成課題の目的、およびすべての意味的關係の文がある程度の数存在することで正当な評価を行うという目的のために、以下の基準を用いて、上記の作業者とは別の作業者が作成した。

- 単文である。
 - クエリを疑問文と考えたとき、その回答は Yes/No で行える。
 - ウェブ上に、クエリと同意する文および対立する文が存在する。
 - いずれの関係にも科学的根拠が記述された文が存在する、すなわち迷信のように科学的根拠のないクエリは除外される。
 - 同意、対立のいずれが真に正しいかは判明していなくてよい。
- 作成された 20 種類のクエリそれぞれについて、クエリ中のいくつかの単語をクエリ単語

として選択し、検索エンジン^{*1}を用いて文書集合を獲得する。検索対象文が 3.2 節で述べた条件を満たすように、クエリ単語の選択基準と検索対象文の選択基準を以下のように定める。

クエリ単語の選択基準

- 末尾の文節に含まれる形態素はすべて除く。
- 機能語はすべて除く。
- 同一文節中に名詞 + 名詞 (サ変接続) という形態素列がある場合、後者の名詞は除く。
- 最終的に残った形態素の原形をクエリ単語とする。

検索対象文の選択基準

- クエリ単語をすべて含む。
- 20 文字以上, 150 文字以内である。
- 疑問文ではない。
- 文末の形態素の品詞が、動詞, 助動詞, または形容詞である。
- 単語の羅列といった文を除外するために、助詞を 3 つ以上含む。

獲得された文書集合は TSUBAKI のランキングで並べられ、上位から順に文単位で、条件を満たす 70 文程度をそのクエリの評価データセットとする。文分割は TSUBAKI によって行われる。

根拠関係認識について、評価データ 1,467 文中で〈同意〉の関係にある文が 532 文、〈対立〉の関係にある文が 238 文存在する一方で、〈同意根拠〉または〈対立根拠〉の関係にある文対は 45 文対であった。これは、ウェブ上には根拠情報を含む文が少ないことを示している。

これらのデータを言論マップコーパスとして公開している。詳細は <http://www.cl.ecei.tohoku.ac.jp/stmap/> を参照されたい。

実験では、文節アライメントと局所構造アライメントにおいて、用いる制約を変化させることで、関係分類結果がどのように変化するかを観察する。用いる制約の組合せを以下に定義する。

● 文節アライメント

exact 3.3.1 項の (1) で定義した、内容語が完全一致または部分一致する場合にのみ対応付ける、最も強い制約。

表 2 関係分類結果

Table 2 Results of semantic relation classification.

制約の組合せ	同意	対立
exact - dep	76.6% (72/94), 13.5% (72/532)	83.3% (10/12), 4.2% (10/238)
exact+dic - dep	76.5% (75/98), 14.1% (75/532)	73.3% (11/15), 4.6% (11/238)
exact+dic+estimation - dep	67.6% (165/244), 31.0% (165/532)	79.6% (43/54), 18.1% (43/238)
exact+dic+estimation - dep+pat	69.9% (318/455), 59.8% (318/532)	65.3% (62/95), 26.1% (62/238)
exact+dic+estimation - none	68.9% (410/595), 77.1% (410/532)	58.3% (81/139), 34.0% (81/238)
gold standard	84.3% (43/51), 86.0% (43/50)	90.9% (20/22), 62.5% (20/32)

それぞれ左側の値が精度、右側の値が再現率を示す。

exact+dic exact に加えて、3.3.1 項の (2) で定義した語彙知識に基づいて対応付ける。

exact+dic+estimation exact+dic に加えて、3.3.1 項の (3) で定義した文構造の類似度に基づいて対応付ける、最も弱い制約。

● 局所構造アライメント

dep 3.3.2 項の (1) で定義した、依存構造上で直接修飾関係にある場合にのみ対応付ける、最も強い制約。

dep+pat dep に加えて、3.3.2 項の (2) ~ (4) で定義した対応付けも行う。

none 局所構造アライメントを適用しない、最も弱い制約。

さらにそれぞれのアライメントの制約を組み合わせると、“exact - dep” のように記述する。

4.2 実験結果

関係分類の実験結果を表 2 に示す。まず、“exact - dep” は、最も厳密なアライメントを行うものである。この設定では、クエリの各文節とほぼ同じ表現が検索対象文中に存在し、その間の依存構造が一致するという強い制約を満たす局所構造のみ対応付けられるため、クエリとほぼ同じ内容、構造の文が検索対象文に含まれている場合のみ〈関係あり〉に分類される。したがって、精度は高いものの、分類される文の数は少ない。そのため、言論マップ生成課題に適用するのは難しい。

この制約から緩和していく方針は 2 つあり、文節アライメントの制約緩和と局所構造アライメントの制約緩和である。まずは文節アライメントについて 2 段階で制約を緩和していく。

1 段階目で、語彙知識に基づくアライメントも行うことで、文節アライメントの制約を緩和する。その結果を “exact+dic - dep” に示す。大規模な語彙知識を利用しているにもかかわらず再現率の向上がわずかであるのは、検索対象文はクエリ中の名詞 (事象を示す名詞

*1 TSUBAKI (<http://tsubaki.ixnlp.nii.ac.jp/>) を用いた。

を除く)をすべて含むという条件のためである。この条件のため、名詞のアライメントは、ほぼ表層的な類似度によって付与され、語彙知識は必要とならない。一方、述語の対応付けについては語彙知識が有効であるが、この結果から十分にはカバーできていないことが分かる。

2段階目で、文構造の類似度に基づく文節アライメントの動的推定を行うことで、文節アライメントの制約を緩和する。その結果を“exact+dic+estimation - dep”に示す。動的推定を適用しない“exact+dic - dep”と比較すると、再現率は〈同意〉が14.1%から31%、〈対立〉が4.6%から18.1%と大幅に向上していることから、このような推定手法が有効であることが示された。

次に局所構造アライメントの制約を2段階で緩和していく。

1段階目として、局所構造アライメントをより柔軟に対応付けた結果が“exact+dic+estimation - dep+pat”である。局所構造アライメントの制約を弱めることにより、“exact+dic+estimation - dep”と比較して再現率が大幅に向上し、特に〈同意〉の再現率は約2倍に向上している。この結果から、依存構造の一致は制約としては強すぎであることが分かる。ここで、柔軟な対応付けを行うことで、誤った対応付けが増えることによる関係認識精度の低下が懸念される。しかし、〈対立〉については精度低下がみられるものの、〈同意〉に関しては若干ではあるが精度が向上している。

2段階目は、最も緩和された制約として、局所構造アライメントを適用せず、単純に文節アライメントのみを利用して関係分類を行うというものである。その結果を“exact+dic+estimation - none”に示す。本制約では3.3.3項で述べた文節アライメント選択手法は適用できないため、検索対象文中で最も文末に近い述語の文節アライメントを選択する。“exact+dic+estimation - dep+pat”と比較すると、構造的な対応をとらないためにすべての関係について精度が低下した。特に〈対立〉の精度低下が大きい。このことから、高精度に関係を分類するためには、局所構造の対応付けが重要であることが分かった。

しかしながら、自動的にアライメントを行った結果では分類性能はまだ不十分である。そこで、局所構造アライメントおよび、対応する文節間の意味の関係については人手で正解を付与することで、関係分類における問題について考察する。しかし、すべての文対にアライメントの正解を付与するのはコストがかかるため、全体の1割程度に対して正解を付与した。1割の事例は、関係ラベルの割合は維持されるよう、ランダムに選択した。正解のアライメント情報を与えたときの関係分類結果を“gold standard”に示す。結果から、正解のアライメントを与えることで、〈同意〉、〈対立〉の双方の関係の大部分が正しく分類できている

表3 根拠認識結果

Table 3 Results of evidence detection.

制約の組合せ	根拠
exact+dic+estimation - dep	37.5% (6/16) 13.3% (6/45)
exact+dic+estimation - dep+pat	41.7% (20/48) 44.4% (20/45)
exact+dic+estimation - none	38.3% (23/60) 52.1% (23/45)

それぞれ左側の値が精度、右側の値が再現率を示す。

ことが分かる。このことから、関係分類において中心的な役割を果たすのはアライメントであり、これが十分な精度で実現できれば、関係分類は少数のルールの記述で十分であるといえる。ここでの誤分類事例が、関係分類における課題である。それについては次節で考察する。

〈根拠〉の分類性能について、〈同意根拠〉と〈対立根拠〉を区別せずに評価する。表3において、“exact+dic+estimation - dep”に比べて弱い制約で局所構造アライメントを付与する“exact+dic+estimation - dep+pat”では、再現率が大きく向上しており、精度も向上している。これは、主に〈同意〉または〈対立〉の分類性能の向上が寄与している。

(3) Q マーガリンは体に良い

T マーガリンは生体には存在しない油なので、体に悪い影響がある

根拠情報を含む文は、(3)に示すように、主語と述語の間に根拠情報が挿入され、主語と述語の間に修飾関係が存在しない場合が多い。このような文対に対して、弱い制約の局所構造アライメントが有効である。

4.3 考察

まず、表2の“exact+dic+estimation - dep”と“exact+dic+estimation - dep+pat”で、分類結果の異なる事例を分析したところ、“exact+dic+estimation - dep+pat”では正しく分類された事例が176事例であった。このうち、175事例は、〈その他〉と誤分類されていたものが、〈同意〉または〈対立〉に正しく分類された。このことから、柔軟な局所構造対応は、〈関係あり〉と〈その他〉の分類において大変重要な役割を果たしていることが分かる。“exact+dic+estimation - dep+pat”において正しく関係を分類できた事例の例を以下に示す。

(4) Q ヨーグルトは_a体に_b良い_c

T ヨーグルトを_a毎日欠かさず食べる習慣ができ、体調も_b良くなりました_c。

「ヨーグルトを(T)」と「良くなりました(T)」の間には直接の依存構造が存在しないため、“exact+dic+estimation - dep”では局所構造アライメントされない。その結果、「ヨー

グルトは(Q)」と「ヨーグルトを(T)」の間の文節アライメントは、文節アライメントの選択において取り除かれ、意味的關係は〈その他〉と分類されていた。“exact+dic+estimation-dep+pat”では、3.3.2 項の(3)によって局所構造アライメントが行われることで〈同意〉に分類された。

一方で、“exact+dic+estimation-dep”では正しく認識できたが、“exact+dic+estimation-dep+pat”では誤分類された事例は31事例であった。これらは、〈その他〉の文対が、〈同意〉または〈対立〉に誤って分類されたものが26事例であり、文節アライメントの動的推定誤りが主な原因であった。以下にその一例を示す。

(5) Q 酢を_a飲むと_b身体が_c柔らかくなる_d

T スポーツの後や_c身体が_c疲れたときに_a酢と_a糖を一_c緒に_b飲んだり_b、食べたりすると、回復が早くなります

(5)において、「柔らかくなる(Q)」と「疲れた(T)」の文節アライメントは誤りである。これは、評価極性が異なり(Positive - Negative)、かつ「身体が(Q)」と「身体の(T)」が、それぞれ依存関係にあるため対応付けられた。文節アライメントの動的推定の付与基準を弱める条件について、改善していく必要がある。

次に、表2の“gold standard”について、同じ文対の“exact+dic+estimation-dep+pat”との間で分類結果の異なる事例を分析した結果について述べる。このような事例は57事例であり、そのうち45事例は“gold standard”では正しく分類されていた。この理由は大きく2つに分類できる。

- (1) 過剰にアライメントしてしまった結果、〈その他〉の事例を〈関係あり〉に誤分類
- (2) アライメントすることができず、〈関係あり〉の事例を〈その他〉に誤分類

局所構造アライメントの false positive によってアライメントの選択を誤った事例や、〈関係あり〉と正しく判定できたが、正解の文節に対してアライメントが付与できなかった事例を含めると、45事例中35事例がアライメントに起因するものであった。一方で、双方で選択している文節アライメントは同一であり、対応する文節の意味的關係を正しく認識できなかったことで、関係の分類を誤った事例も存在したが、2事例とわずかであった。以上のことから、関係分類性能への影響はアライメントの性能が支配的であることが分かった。

57事例中7事例は“gold standard”では誤って分類されたが、“exact+dic+estimation-dep+pat”では正しく分類された。この理由も上記の2つと同様だが、問題の性質は異なる。まず、いずれの事例も、アライメントに基づく手法では解くことが困難な事例である。1つめの理由で正しく分類された事例は、アライメントの正解を明示的に付与することが困難

な事例である。“exact+dic+estimation-dep+pat”は、誤った場所にアライメントを付与し、それによって〈同意〉または〈対立〉と判断され、意図せず正しく分類された。2つめの理由で正しく分類された事例は、アライメントの正解を付与することは可能だが、それとは別の部分の表現によって〈その他〉である事例である。“exact+dic+estimation-dep+pat”は、アライメントすることができなかったために、意図せず正しく分類された。

残りの5事例は、いずれの制約でも誤って分類された。

最後に、クエリごとの分類性能の違いについて考察する。本実験では、できるだけ多様なクエリに対して評価実験を行い、かつ正解ラベル付与のコストを抑えるために、20種類のクエリそれぞれに対して、70文程度の検索対象文を評価データとした。そのため、〈同意〉や〈対立〉の文が極端に少ない場合もあり、特に〈対立〉の文は存在しない場合もあった。したがって、定量的に比較することは難しいため、定性的な考察を行う。〈対立〉について、その判断が述語の評価極性によって可能な「納豆ダイエットは効果がある」「にがりには健康に良い」といったクエリは精度・再現率ともに高い傾向にあった。逆に、「温暖化によって海面が上昇する」といった、述語が反義語の関係にあることによって関係が決定できる場合が多いクエリについては、分類性能は限定的であった。〈同意〉についても〈対立〉と同じ傾向は見られたが、顕著ではなかったことから、語彙知識における同義・類義関係よりも、反義語の関係が不足していることが分かった。

4.4 関係分類における課題

関係認識の課題において、次にどのような問題に取り組むべきかを調査するため、正解のアライメント情報を利用した関係分類の誤り分析を行った。この分析の結果、誤分類の多くは事実性や情報の発信源に関連するものであった。

(6) Q お酢を飲むと体が柔らかくなる。

T お酢を飲むと体が柔らかくなるというのは誤解である。

この例では、QはTに完全に対応するため、システムはこの文対を〈同意〉に誤分類した。しかし、Tの従属節内の言明のモダリティは、Tの「誤解」という非叙実述語によって否定されている。このようなモダリティは、我々の拡張モダリティ解析³⁷⁾によって解析されるが、その性能は限定的である。今後の課題としては、モダリティに関連する関係分類の問題に取り組んでゆく必要がある。

5. おわりに

オープンドメインな任意のクエリに対して、それに関連するウェブ上の様々な意見を同

意, 対立, 根拠の関係に基づいて整理し, 俯瞰する言論マップ生成システムを開発した. このようなシステムは我々の知る限り, 世界的に見ても前例がない.

本論文では, 言論マップ生成課題において, 最も重要な要素技術である文間関係認識の技術的課題に焦点を当て, 制約を段階的に緩和するモデルの開発と問題の分析について報告した. 実験では, 構築したモデルを 1,467 文対からなる評価データに適用した. その結果, 段階的に制約を緩めていくことで関係分類の精度と再現率をコントロールできることが分かった. しかし, 現時点ではアライメントそのものの性能が十分ではない. そこで, アライメントまでの段階については正解データを与えたところ, シンプルな関係分類であっても大多数の問題は正しく分類できることが分かった. したがって, 文間関係認識において最も重要な課題がアライメントであることが改めて明らかになった.

今後の課題として, まず第 1 に, 現時点ではアライメントの性能が不十分であるため, 制約の一部に機械学習を用いることで関係分類性能の改善を行う. 第 2 に, 言論マップ生成における意味的關係の粒度について, 多くの賛成, 反対意見を分類できるようになってきたため, それぞれの中でさらに情報発信者を特定したりして, より高次の組織化を行う. 第 3 に, 構築した文間関係認識モデルを用いた言論マップ生成システムに対して, 大規模なユーザ評価を行うことがあげられる. 予備的なユーザ評価では, 既存のウェブ検索エンジンより良い評価が得られたが³³⁾, より多くの設問を設定し, 詳しく分析をしていく必要がある.

謝辞 本研究は, (独) 情報通信研究機構の委託研究「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」の一環として実施した. また, 本研究の一部は次の研究費の支援を受けている: 科研費基盤研究 A「ネット上の医療・健康情報の健全な利活用を可能にするヘルスリテラシー支援環境」(課題番号: 23240018, 代表: 乾健太郎), 科研費若手研究 B「自然論理と確立論理の融合による文間関係認識」(課題番号: 23700157, 代表: 渡邊陽太郎).

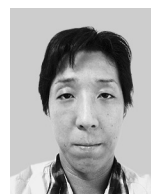
参 考 文 献

- 1) Akamine, S., Kawahara, D., Kato, Y., Nakagawa, T., Inui, K., Kurohashi, S. and Kidawara, Y.: WISDOM: A Web Information Credibility Analysis Systematic, *Proc. ACL-IJCNLP 2009 Software Demonstrations*, Association for Computational Linguistics, pp.1-4 (2009).
- 2) Bentivogli, L., Clark, P., Dagan, I., Dang, H.T. and Giampiccolo, D.: The sixth PASCAL recognizing textual entailment challenge, *Proc. Text Analysis Conference (TAC)* (2010).
- 3) Bond, F., Isahara, H., Fujita, S., Uchimoto, K., Kuribayashi, T. and Kanzaki, K.: Enhancing the Japanese WordNet, *Proc. 7th Workshop on Asian Language Resources*, pp.1-8 (2009).
- 4) Chang, M.-W., Goldwasser, D., Roth, D. and Srikumar, V.: Discriminative Learning over Constrained Latent Representations, *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp.429-437 (2010).
- 5) Cox, L.A.T. and Popken, D.A.: Overcoming confirmation bias in causal attribution: a case study of antibiotic resistance risks, *Risk Analysis an Official Publication of the Society for Risk Analysis*, Vol.28, No.5, pp.1155-1172 (2008).
- 6) Dagan, I., Glickman, O. and Magnini, B.: The PASCAL Recognising Textual Entailment Challenge, *1st PASCAL Machine Learning Challenges Workshop*, Vol.3944, pp.177-190 (2005).
- 7) de Marneffe, M.-C., Rafferty, A.N. and Manning, C.D.: Finding Contradictions in Text, *Proc. ACL-08: HLT*, Association for Computational Linguistics, pp.1039-1047 (2008).
- 8) Ennals, R., Trushkowsky, B. and Agosta, J.M.: Highlighting Disputed Claims on the Web, *Proc. WWW 2010*, pp.341-350 (2010).
- 9) Harabagiu, S., Hickl, A. and Lacatusu, F.: Negation, contrast and contradiction in text processing, *Proc. 21st National Conference on Artificial Intelligence*, pp.755-762 (2006).
- 10) Higashinaka, R. and Isozaki, H.: Corpus-based Question Answering for why-Questions, *Proc. International Joint Conference on Natural Language Processing*, pp.418-425 (2008).
- 11) Keselman, A., Browne, A.C. and Kaufman, D.R.: Consumer health information seeking as hypothesis testing, *Journal of the American Medical Informatics Association: JAMIA*, Vol.15, No.4, pp.484-495 (2008).
- 12) Kim, H.D. and Zhai, C.: Generating comparative summaries of contradictory opinions in text, *Proc. 18th ACM Conference on Information and Knowledge Management*, pp.385-394 (2009).
- 13) Kudo, T. and Matsumoto, Y.: Japanese Dependency Analysis using Cascaded Chunking, *CoNLL 2002: Proc. 6th Conference on Natural Language Learning 2002 (COLING 2002 Post-Conference Workshops)*, pp.63-69 (2002).
- 14) Kudo, T., Yamamoto, K. and Matsumoto, Y.: Applying conditional random fields to Japanese morphological analysis, *Proc. EMNLP*, pp.230-237 (2004).
- 15) MacCartney, B., Galley, M. and Manning, C.D.: A Phrase-Based Alignment Model for Natural Language Inference, *Proc. 2008 Conference on Empirical Methods in Natural Language Processing (EMNLP-08)*, pp.802-811 (2008).

- 16) Mirkin, S., Dagan, I. and Padó, S.: Assessing the Role of Discourse References in Entailment Inference, *Proc. ACL 2010*, pp.1209–1219 (2010).
- 17) Murakami, K., Nichols, E., Matsuyoshi, S., Sumida, A., Masuda, S., Inui, K. and Matsumoto, Y.: Statement Map: Assisting Information Credibility Analysis by Visualizing Arguments, *Proc. 3rd ACM Workshop on Information Credibility on the Web (WICOW 2009)*, pp.43–50 (2009).
- 18) Ohki, M., Nichols, E., Matsuyoshi, S., Murakami, K., Mizuno, J., Shouko, M., Inui, K. and Matsumoto, Y.: Recognizing Confinement in Web Texts, *Proc. IWCS 2011*, pp.215–224 (2011).
- 19) Pang, B. and Lee, L.: Opinion Mining and Sentiment Analysis, *Foundations and Trends in Information Retrieval*, Vol.2, No.1-2, Now Publishers Inc. (2008).
- 20) Paul, M., Zhai, C. and Girju, R.: Summarizing Contrastive Viewpoints in Opinionated Text, *Proc. 2010 Conference on Empirical Methods in Natural Language Processing*, pp.66–76 (2010).
- 21) Radev, D.R.: Common Theory of Information Fusion from Multiple Text Sources Step One: Cross-Document Structure, *Proc. 1st SIGdial Workshop on Discourse and Dialogue*, pp.74–83 (2000).
- 22) Sammons, M., Vydiswaran, V.G.V., Vieira, T., Johri, N., Chang, M.-W., Goldwasser, D., Srikumar, V., Kundu, G., Tu, Y., Small, K., Rule, J., Do, Q. and Roth, D.: Relation Alignment for Textual Entailment Recognition, *Proc. Recognizing Textual Entailment 2009* (2009).
- 23) Shinzato, K., Shibata, T., Kawahara, D., Hashimoto, C. and Kurohashi, S.: TSUBAKI: An open search engine infrastructure for developing new information access methodology, *Proc. 3rd International Joint Conference on Natural Language Processing (IJCNLP2008)*, pp.189–196 (2008).
- 24) Verberne, S., Boves, L., Oostdijk, N. and Coppen, P.-A.: Evaluating discourse-based answer extraction for why-question answering, *Proc. 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '07*, pp.735–736 (2007).
- 25) Watanabe, Y., Asahara, M. and Matsumoto, Y.: A Structured Model for Joint Learning of Argument Roles and Predicate Senses, *Proc. ACL2010*, pp.98–102 (2010).
- 26) William, M. and Thompson, S.: Rhetorical Structure Theory: towards a functional theory of text organization, *Text*, Vol.8, No.3, pp.243–281 (1988).
- 27) Zhang, Z., Otterbacher, J. and Radev, D.: Learning Cross-Document Structural Relationships Using Boosting, *CIKM '03*, pp.124–130 (2003).
- 28) Zhang, Z. and Radev, D.: Combining Labeled and Unlabeled Data for Learning Cross-Document Structural Relationships, *IJCNLP '05*, pp.32–41 (2005).
- 29) 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: 意見抽出のための評価表現の収集, 自然言語処理, Vol.12, No.3, pp.203–222 (2005).
- 30) 黒田 航, 風間淳一, 村田真樹, 鳥澤健太郎: Web 文書にも対応できる日本語異表記の認定基準, 言語処理学会第 16 回年次大会, pp.990–993 (2010).
- 31) 奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座: 言明間意味の関係コーパス作成のための意味的關係仕様書 Version 1.0 (2009), 入手先<http://www.cl.ecei.tohoku.ac.jp/stmap/sem_corpus.html>.
- 32) 松吉 俊, 村上浩司, 増田祥子, 松本裕治, 乾健太郎: 事象間関係知識の整備と類似・対立認識への応用, 情報処理学会研究報告, 自然言語処理研究会報告, Vol.2008, No.90, pp.15–22 (2008).
- 33) 水野淳太, Nichols, E., 渡邊陽太郎, 村上浩司, 松吉 俊, 大木環美, 乾健太郎, 松本裕治: 言論マップ生成技術の現状と課題, 言語処理学会第 17 回年次大会, pp.49–52 (2011).
- 34) 東山昌彦, 乾健太郎, 松本裕治: 述語の選択選好性に着目した名詞評価極性の獲得, 言語処理学会第 14 回年次大会, pp.584–587 (2008).
- 35) 小島正裕, 村田真樹, 風間淳一, 黒田 航, 藤田 篤, 荒牧英治, 土田正明, 渡辺靖彦, 鳥澤健太郎: 機械学習と種々の素性を用いた編集距離の小さい日本語異表記対の抽出, 言語処理学会第 16 回年次大会, pp.928–931 (2010).
- 36) 隅田飛鳥, 吉永直樹, 鳥澤健太郎: Wikipedia の記事構造からの上位下位関係抽出, 自然言語処理, Vol.16, No.3, pp.3–24 (2009).
- 37) 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).
- 38) 橋本 力, 鳥澤健太郎, 黒田 航, デサーガステイン, 村田真樹, 風間淳一: WWW からの大規模動詞含意知識の獲得, 情報処理学会論文誌, Vol.52, No.1, pp.293–307 (2011).

(平成 23 年 4 月 12 日受付)

(平成 23 年 9 月 12 日採録)



水野 淳太 (学生会員)

2006 年豊橋技術科学大学工学部情報工学課程卒業。2009 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在, 同大学院大学情報科学研究科博士後期課程在学中。2010 年より東北大学大学院情報科学研究科システム情報科学専攻特別研究学生。自然言語処理の研究に従事。



渡邊陽太郎 (正会員)

2004年信州大学工学部情報工学科卒業。2007年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。2010年同大学院大学情報科学研究科博士後期課程修了。同年より、東北大学大学院情報科学研究科助教、現在に至る。博士(工学)。自然言語処理の研究に従事。ACL, 人工知能学会, 言語処理学会各会員。



エリック ニコルズ (正会員)

東北大学大学院情報科学研究科研究員。2002年メリーランド州立大学カレッジパーク校情報科学科・日本語文学科卒業。2005年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。2008年同大学院大学情報科学研究科博士後期課程指導認定退学。同大学院大学研究員を経て、2010年より現職。博士(工学)。ACL, 言語処理学会各会員。機械翻訳, 知識獲得, 情報信憑性検証技術の研究に従事。



村上 浩司

1996年室蘭工業大学情報工学科卒業。2004年北海道大学大学院工学研究科博士課程単位取得退学。同年ニューヨーク大学計算機科学科助手, 2006年東京工業大学統合研究院特任助教, 2008年奈良先端科学技術大学院大学特任助教を経て, 2010年より楽天株式会社楽天技術研究所 NY に所属。博士(工学)。自然言語処理の研究に従事。電子情報通信学会, 音響学会, 言語処理学会各会員。



乾 健太郎 (正会員)

東北大学大学院情報科学研究科教授。専門は計算言語学。自然言語処理, 知識処理に関する研究に従事。1995年東京工業大学大学院博士課程修了。同大学院助手, 九州工業大学助教授, 奈良先端科学技術大学院大学准教授を経て, 2010年より現職。言語処理学会理事, 雑誌, 自然言語処理, Computational Linguistics 各編集委員。



松本 裕治 (正会員)

1977年京都大学工学部情報工学科卒業。1979年同大学大学院工学研究科修士課程情報工学専攻修了。同年電子技術総合研究所入所。1984~1985年英国インペリアルカレッジ客員研究員。1985~1987年(財)新世代コンピュータ技術開発機構に出向。京都大学助教授を経て, 1993年より奈良先端科学技術大学院大学教授, 現在に至る。工学博士。専門は自然言語処理。人工知能学会, 日本ソフトウェア科学会, 言語処理学会, 認知科学会, AAAI, ACL, ACM 各会員。