

意味素に基づく隠喩の名詞句“ A の B ”の意味解析

栗原 健^{†1} 松本 和幸^{†2}
土屋 誠司^{†3} 任 福 継^{†2}

新聞記事や日常会話など、自然言語を扱う様々な場面で比喩表現が使用されている。我々は比喩表現を使うことで微妙なニュアンスの意志を伝えることが出来る。しかし、比喩、特にその中の隠喩の意味解析は難しく、従来では有効な実現例は示されていない。そこで、本研究では隠喩の名詞句“ A の B ”の意味解析手法を提案する。本論文では、隠喩の名詞句“ A の B ”を対象とした辞書である隠喩辞書を独自で構築し、その辞書を検索することで意味解析を行う。提案手法の特徴として、意味素と単語の組をクエリとして与えることで、柔軟な検索が出来ることが挙げられる。評価実験を行い、提案手法によって 70%程度の精度で隠喩の意味を出力できることを確認した。

Semantic Analysis of Metaphor “ A-no-B ”based on Sememe

TAKERU KURIHARA,^{†1} KAZUYUKI MATSUMOTO,^{†2}
SEIJI TSUCHIYA ^{†3} and FUJI REN^{†2}

The metaphor expression is used in various scenes that treat the natural language like the newspaper articles and the daily conversation, etc. We can express our intention by using metaphor expressions. However, it is very difficult to analyze metaphor expression by a computer. This paper constructs two kind of metaphor dictionaries. We also analyze meaning of metaphor “ A-no-B ” by searching metaphor dictionaries. Using the method described in this paper, we showed that this method was able to analyze meaning of metaphor A-no-B ” with accuracy about 70%.

1. はじめに

新聞記事や日常会話、Weblog など、自然言語を扱う様々な場面で比喩表現が使われている。言語表現において比喩表現は多くの曖昧性を持った表現であり、その比喩表現を使うことで我々は微妙なニュアンスの意志を伝えることが出来る。そのため、新聞記事や日常会話などを自然言語処理の処理対象とした場合、比喩解釈は必要不可欠な処理である。

比喩解釈が可能となると様々な有益な影響を及ぼすと考えられる。例えば、機械翻訳の分野において、「先生の卵」を「teacher's egg」と翻訳するような比喩表現を字句通りの意味で直訳することで生じる誤訳を回避することが出来る。また、比喩表現の意味を解析することで、比喩表現を非比喩の表現へと置き換えて翻訳することによる翻訳精度の向上が見込まれる。

従来、比喩に関する研究は直喩が主な対象であった。榊井らは顕現性落差と意外性という尺度を導入することで、テキスト中に存在する比喩表現の認識を行った¹⁾。また、比喩解釈には概念間の共通属性値集合のランキング精度が重要であると考え、WWW から得られた属性値知識によって再ランキングすることで概念対から得られるクローズアップ属性値の精錬を行った。その結果、人間の判断を約 80%シミュレートできたと報告している²⁾。それに対して、直喩のように喩詞が存在しない隠喩は、解析が難しく有効な実現例はない。そこで、本研究では隠喩を対象とした。特に、隠喩の中でも名詞句“ A の B ”で明示される表現に限定した。隠喩を含む名詞句“ A の B ”の解析は難しく、池原らは、比喩的な表現の名詞句“ A の B ”の翻訳を行なう場合、個別に特化した枠組みが必要であると報告している³⁾。

本研究では、隠喩の名詞句“ A の B ”の意味解析を行なうためには、隠喩の名詞句“ A の B ”に特化した辞書が必要であると考えた。そこで、隠喩の名詞句“ A の B ”を対象とした隠喩辞書を独自で構築し、その辞書を検索することによる意味解析手法を提案する。そして、比喩の名詞句“ A の B ”と非比喩の名詞句“ A の B ”を入力し、その出力結果から本手法の性能評価を行なう。隠喩辞書の特徴として意味素と単語の組をクエリとして与えることで、より柔軟な検索が出来ることが挙げられる。

†1 徳島大学大学院先端技術科学教育部
Graduate School of Advanced Technology and Science, The University of Tokushima

†2 徳島大学大学院ソシオテクノサイエンス研究部
Institute of Technology and Science the University of Tokushima

†3 同志社大学理工学部インテリジェント情報工学科
Department of Intelligent Information Engineering and Sciences

以下、第2節では、今回構築した隠喩辞書の登録内容及び構築方法について、第3節では、隠喩辞書を用いた意味解析手法について説明する。第4節で実験結果について説明し、第5節でまとめる。

2. 隠喩辞書の構築

隠喩の名詞句“ A の B ”の意味解析を行なうため、隠喩の名詞句“ A の B ”を対象とした辞書を独自で構築した。これを隠喩辞書とする。隠喩の名詞句“ A の B ”には2種類の表現が存在する。例を表1に示す。1つは、「二の舞」や「コロンプスの卵」などの慣用的な意味をもった表現 (TypeI) である。もう一方は、それ以外の表現 (TypeII) であり、「先生の卵」や「政策の目玉」などがある。

2.1 隠喩の名詞句“ A の B ”について

TypeI は、名詞句“ A の B ”という纏まりで1つの意味を持つ。そのため、TypeI の名詞句“ A の B ”の意味を検索する場合には、登録された名詞句“ A の B ”に対して1対1で対応する意味を登録した辞書が必要である。それに対して、TypeII は、単語 A と単語 B の間に修飾関係が存在する。そして、修飾語である単語から想像される意味が被修飾語を修飾することによってその意味が理解される。例えば、「先生の卵」という表現は「卵」から想像される「～に対して未熟である」、「～になるために勉強中の人物」などの意味が「先生」を修飾することで「先生になるために勉強中の人物」という意味だと理解される。このとき、被修飾語である「先生」をシソーラス上において共通の上位概念を持つ単語である「教授」に置換した場合でも、その意味は「教授になるために勉強中の人物」と理解される。つまり、被修飾語は共通の上位概念を持つ単語ならば置換可能であると考えられる。このことから、TypeII の名詞句“ A の B ”に対しては、被修飾語である単語の上位概念と修飾語である単語の組、それに対する修飾語から想像される意味が登録された辞書が必要である。以上のことから、隠喩辞書には TypeI の名詞句“ A の B ”と TypeII の名詞句“ A の B ”を対象とした2種類の枠組みが必要であると考えられる。そこで、隠喩辞書は2種類構築した。それらを慣用隠喩辞書、意味素対応隠喩辞書とする。以下に隠喩辞書を構築するために利用した資源と2種類の隠喩辞書の詳細を述べる。

2.2 利用資源

比喩表現は人間の感性によって日々創造される新しい表現と、日常的に使用される慣用的な表現がある。前者はその表現自体に新規性がある。それに対し、後者は新規性はないがその表現は多種多様である。そのため、隠喩辞書を構築するためには、新規性と日本語の網羅

性に着目した資源が必要である。

本研究では、日本語の網羅性に着目した資源として国語辞書を使用した。国語辞書は約23万8000語が収録された大辞林第三版⁴⁾を採用した。国語辞書には日本人が日常的に使用する単語の殆どが登録されているため、国語辞書を参照することにより日本語表現の殆どを網羅できると考えられる。国語辞書から名詞句“ A の B ”を抽出した場合には、掲載されている説明文を対応する意味として隠喩辞書に登録した。また、新規性に着目した資源として WWW データを利用した。Web 上の Weblog や最新のニュース記事などでは様々な新しい表現が創造される。そのため、WWW データを利用すれば新規性のある名詞句“ A の B ”が抽出できると考えられる。

2.3 慣用隠喩辞書

慣用隠喩辞書は TypeI の名詞句“ A の B ”を対象とした辞書である。この隠喩辞書には、「二の舞」や「背水の陣」などの慣用的な意味を持った隠喩の名詞句“ A の B ”と、それらと1対1で対応する意味が列挙されている。慣用隠喩辞書の一部を表2に示す。

慣用隠喩辞書は資源から隠喩の名詞句“ A の B ”を抽出し、それらに対して意味を付与することによって作成される。慣用隠喩辞書には「単語 A の単語 B」という纏まりで意味を持つ名詞句“ A の B ”が登録されているため、検索クエリに名詞句“ A の B ”を用いる。完全一致検索を行い、慣用隠喩辞書に登録されている名詞句“ A の B ”が入力されると、それに対応する意味を出力する。

2.4 意味素対応隠喩辞書

意味素対応隠喩辞書は TypeII の名詞句“ A の B ”の処理を目的とした辞書である。意味素対応隠喩辞書には、隠喩の名詞句“ A の B ”において被修飾語である単語の上位概念と修飾語である単語の組、その組に対応する意味が列挙されている。そして、被修飾語である単語の上位概念には意味素を用いた。シソーラスから単語を検索し、単語が属するノード

表1 名詞句“ A の B ”の例

TypeI	TypeII
二の舞	先生の卵
堪忍袋の緒	日本の癌
腹の虫	トラブルの種
背水の陣	黄金の味
非の打ち所	空白の時間
コロンプスの卵	一握りの砂
本の虫	三ツ星のレストラン

表2 慣用隠喩辞書の一部

名詞句“ A の B ”	意味
鼻の差	ほんの少しの僅かの差
束の間	ほんの少しの間
腹の虫	苛立ちや不満などの感情
背水の陣	全く後が無い状態
目の色	目つきや表情
雀の涙	ほんの少しの量
腐った魚の目	全くやる気の感じられない目

表 3 意味素対応隠喩辞書:(意味素, 単語 B)

意味素	単語 B (修飾語)	意味
企業	横綱	(単語 A) において代表的な企業
案	舵取り	(単語 A) が良い状態になるように誘導する役
企業	害虫	(単語 A) において邪魔な存在
創作物	筋立て	(単語 A) における大まかな内容
感情	爆発	(単語 A) が激しく表出されること
精神	色	(単語 A) を感じさせる表情

表 4 意味素対応隠喩辞書:(単語 A, 意味素)

単語 A (修飾語)	意味素	意味
一夜	人間活動	一晩で終わってしまうような僅かな(単語 B)
不毛	地域	見るべき成果のない(単語 B)
鉄壁	行為	とても守りが強固である(単語 B)
三流	創作物	程度が非常に劣っている(単語 B)
長丁場	事	完了するまでに長い時間のかかる(単語 B)
裸	精神	全く隠すことの無い(単語 B)

を意味素として定義する。シソーラスには、約 30 万単語が 2715 カテゴリからなる最大 12 段の木構造のどこに属するかが定義された日本語語彙大系⁵⁾を用いた。

意味素対応隠喩辞書は資源から隠喩の名詞句“ A の B ”を抽出し、それらから意味素と単語の組を作成し登録することで構築した。検索クエリは意味素と単語の組となる。意味素対応隠喩辞書の一部を表 3, 表 4 に示す。

2.4.1 意味素と単語の組の登録

意味素対応隠喩辞書に登録される意味素には、資源から抽出された名詞句“ A の B ”の被修飾語の意味素が割り当てられる。しかし、隠喩の名詞句“ A の B ”の意味を考えた場合、意味素と単語の組における意味素の対応範囲が広い場合、登録すべき意味素は被修飾語が属するノードのみに限定されない⁶⁾。「～の卵」という表現の名詞句“ A の B ”を例に挙げる。「先生の卵」は「先生になるために勉強中の人物」という意味になる。このとき「先生」が「教師」という意味素を持つ被修飾語となる。そして、被修飾語の「先生」を異なる意味素を持つ単語である「医者(医師)」に置換して「医師の卵」とした場合でもその意味は「医者になるために勉強中の人物」となり、「先生の卵」と同様の意味となる。このことから、意味素対応隠喩辞書に登録する意味素は 1 つの単語に割り当てられる意味素ではなく、さらに上位の意味素を登録すべきである。そこで、意味素対応隠喩辞書に登録する意味素は、複数の単語が共通して属する上位の意味素とした。

図 1 に意味素と単語の組生成の流れを示す。まず、資源より抽出した名詞句“ A の B ”から修飾語にあたる単語を固定し収集する表現を決定する。決定した表現と同じ表現の名詞句

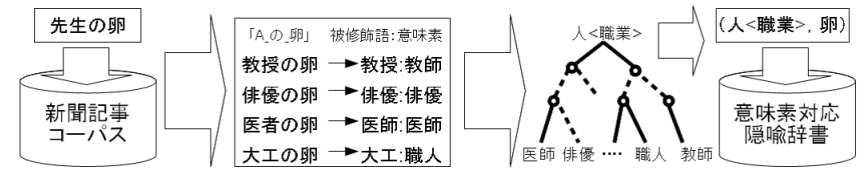


図 1 意味素と単語の組決定の流れ

“ A の B ”を新聞記事コーパス 1 年分⁷⁾から抽出し、その複数の名詞句“ A の B ”の被修飾語である単語に対して意味素を付与する。そして、その複数の単語が共通に属する意味素と修飾語である単語を意味素と単語の組として登録する。図 1 の場合、資源から取得した名詞句“ A の B ”が「先生の卵」となり、修飾語が「卵」となる。そして、「A_の_卵」という表現の名詞句“ A の B ”を新聞記事コーパスから抽出し、「医者」、「大工」などの被修飾語である単語に対して意味素を付与する。その結果、被修飾語に対して付与された意味素が共通に属する意味素である「人<職業>」が登録すべき意味素となり、意味素対応隠喩辞書には(人<職業>, 卵)という意味素と単語の組が登録される。

2.4.2 意味素対応隠喩辞書検索方法

意味素対応隠喩辞書は(意味素:単語 A, 単語 B), (単語 A, 意味素:単語 B) という意味素と単語の組が検索クエリとなる。入力された組の意味素が意味素対応隠喩辞書に登録されている意味素の下位ノードに属しており、なおかつ単語が一致した場合に検索成功となる。検索が成功した場合には、登録されている意味に意味素が割り当てられた単語を代入することで意味を生成する。

意味素対応隠喩辞書は名詞句“ A の B ”を検索クエリとするのではなく、単語 A と単語 B において被修飾語である単語を意味素として抽象化し、意味素と単語の組を直接検索クエリとすることで柔軟なマッチングが可能なのが特徴である。

3. 意味解析手法

本手法では、図 2 で示されるように、名詞句“ A の B ”が入力されると慣用隠喩辞書の検索、検索クエリの生成、意味素対応隠喩辞書の検索の処理過程を経ることで隠喩の名詞句“ A の B ”の意味解析を行う。また、本手法は隠喩の名詞句“ A の B ”を対象としているので、入力は名詞句“ A の B ”に限定する。

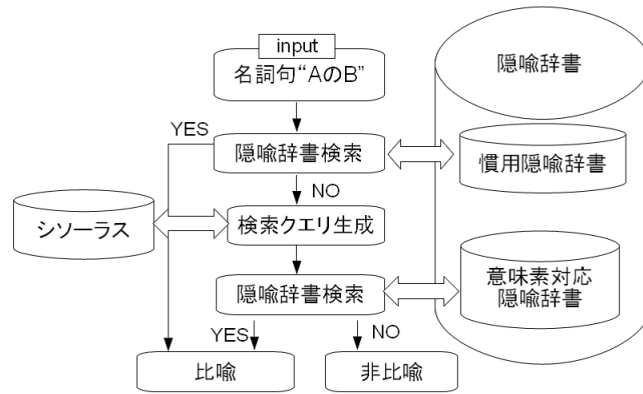


図 2 提案手法の概要

3.1 慣用隠喩辞書検索

入力として名詞句“ A の B ”が得られると、まず、慣用隠喩辞書の検索が行われる。慣用隠喩辞書を検索することで入力名詞句が慣用的表現の隠喩の名詞句“ A の B ”であるかを判断する。ここで、検索が成功した場合、慣用隠喩辞書に登録された対応する意味が入力名詞句の意味として出力される。検索が失敗した場合、入力名詞句の単語 A と単語 B に対して意味素を決定し、検索クエリを生成する。

3.2 検索クエリの生成

慣用隠喩辞書の検索が失敗した名詞句“ A の B ”の単語 A と単語 B をシソーラスから検索し意味素を決定する。このとき、単語が属するノードを意味素とする。そして、決定した意味素を元に入力名詞句から意味素と単語の組を生成する。意味素と単語の組は、(単語 A: 意味素, 単語 B), (単語 A, 単語 B: 意味素) の 2 種類が生成できる。

「先生の卵」が入力された場合には、「先生」には「教師」という意味素、「卵」には「肉・卵」という意味素が決定される。その結果、「先生の卵」からは(教師: 意味素, 卵)と(先生: 肉・卵: 意味素)という意味素と単語の組が生成される。このように生成した意味素と単語の組を意味素対応隠喩辞書の検索クエリとする。

3.3 慣用隠喩辞書検索

入力名詞句“ A の B ”から生成された意味素と単語の組を検索クエリとし、意味素対応

隠喩辞書から検索する。意味素対応隠喩辞書の検索が成功した場合は、意味素対応隠喩辞書に登録された意味に被修飾語である単語を代入して出力結果とする。検索が失敗した場合には、入力された名詞句“ A の B ”は非比喻の名詞句“ A の B ”であるとする。このように、慣用隠喩辞書と意味素対応隠喩辞書のどちらか検索が成功した場合には、入力された名詞句“ A の B ”は比喻の名詞句“ A の B ”であると判断し、隠喩辞書に登録された意味が解析結果として出力される。そして、どちらの隠喩辞書でも検索が失敗した場合には、入力された名詞句“ A の B ”は非比喻の名詞句であると判断する。

4. 評価実験

4.1 実験方法

名詞句“ A の B ”に本手法を適用することで本手法の有効性を評価する。具体的には、隠喩の名詞句“ A の B ”を入力した場合に正しく意味を出力できるか、非比喻の名詞句“ A の B ”を入力した場合に非比喻の表現であると判断するかどうかを確認する。

実験データには Web から人手により収集した隠喩の名詞句“ A の B ”を 1000 句、非比喻の名詞句“ A の B ”を 1000 句用意した。実験に使用する隠喩辞書の規模は慣用隠喩辞書が 1267 句、意味素対応隠喩辞書が 3318 組である。

4.2 実験結果

隠喩の名詞句“ A の B ”と非比喻の名詞句“ A の B ”を本手法に適用した結果を表 5 に示す。本手法に隠喩の名詞句“ A の B ”1000 句を入力した場合、773 句が隠喩辞書での検索に成功し比喻の名詞句“ A の B ”であると判断された。その 773 句の意味が、本来の出力されるべき意味であるか人手で判断した。その結果、出力結果まで正しい名詞句“ A の B ”は 715 句であった。本手法を適用し、正しく意味が出力された比喻の名詞句“ A の B ”を表 6 に示す。表 6 を見ると、「長蛇の列」のような TypeI の名詞句“ A の B ”も「心配の種」や「書類の山」のような TypeII の名詞句“ A の B ”も正しく意味が出力されていることが確認できる。

一方、非比喻の名詞句“ A の B ”1000 句に本手法を適用した場合、1000 句中 882 句を非比喻の名詞句“ A の B ”であると判断した。

4.3 考察

隠喩の名詞句“ A の B ”1000 句を本手法に適用した結果、773 句を比喻であると判断し、その内 715 句の名詞句の意味を正しく出力できた。ことから全体の約 72% の意味を出力できたことが確認できたため本手法の有効性を示すことが出来た。隠喩の名詞句“ A の B ”

表 5 実験結果

	比喻と判断	非比喻と判断	意味解析成功
比喻	773 句	227 句	715 句
非比喻	124 句	876 句	-

表 8 同一の表現の出力結果

	入力名詞句	被修飾語：意味素	意味解析結果
比喻の名詞句	先生の卵 医者卵	先生：教師 医者：医師	先生になるために勉強中の人物 医者になるために勉強中の人物
非比喻の名詞句	恐竜の卵 鶏の卵	恐竜：爬虫類 鳥：鳥	恐竜になるために勉強中の人物 鶏になるために勉強中の人物

表 7 極端に上位の意味素と単語の組が設定された例

意味素	単語 B	意味
具体	卵	(単語 A) になるために勉強中の人物
抽象	種	(単語 A) の原因となる物事
名詞	山	(単語 A) が最も盛り上がる場面
名詞	宝庫	(単語 A) が大量多種ある場所
名詞	山	(単語 A) が大量に積み上がっている状態

表 6 意味解析成功名詞句の一部

入力名詞句	意味解析結果
チームの要	チームの中で重要とされる人物
心配の種	心配の原因となる物事
争いの種	争いの原因となる物事
情報の重み	情報の重大さ
法律の抜け穴	法律を上手く逃れる手段
書類の山	書類が大量に積みあがっている状態
貧乏のどん底	貧乏が一番最悪な状態
長蛇の列	長々と続く行列

1000 句中 715 句を正しく意味解析できた要因として、意味素対応隠喩辞書の検索が意味素の変化に頑健であることが挙げられる。意味素と単語の組に複数の単語が属する共通の意味素を登録することで、表 6 に示されるように「心配の種」や「争いの種」といった被修飾語の意味素が異なる同一の表現であっても正しく意味を出力できた。しかし、複数の単語が共通して属する意味素を登録する場合、意味素対応隠喩辞書登録に極端に上位の意味素と単語の組が登録される問題がある。また、慣用隠喩辞書と意味素対応隠喩辞書のどちらにも登録されていない表現が入力された場合や、未知語を含む名詞句「A の B」が入力された場合には意味解析が失敗してしまう問題がある。

4.3.1 極端に上位の意味素と単語の組の登録

意味素対応隠喩辞書に極端に上位の意味素と単語の組が設定されることの原因として、ソーラス上における単語の分布が広範囲になった場合、被修飾語である単語が共通して属する意味素が極端に上位となってしまうことが挙げられる。意味素対応隠喩辞書に登録された極端に上位の意味素と単語の組の一部を表 7 に示す。意味素対応隠喩辞書に登録すべき意味素は 1 つの被修飾語が属する意味素のみに限定されないため、さらに上位の意味素を登録する必要がある。しかし「具体」や「抽象」等のような極端に上位の意味素と単語の組が

登録されている場合、「先生の卵」と「恐竜の卵」を意味素対応隠喩辞書から検索するとどちらも比喻の名詞句と判断してしまう。比喻の名詞句「A の B」だと誤判断された 124 句の内 111 句について極端に上位の意味素と単語の組が登録されていることが原因であった。

同じ表現で複数の意味を持つ名詞句「A の B」が入力された場合に意味の誤判断は発生する。「A の山」という表現は、「(単語 A) が大量に積みあがった状態」と「(単語 A) が最も盛り上がる場面」という 2 つの意味を持っているため、意味素対応隠喩辞書には「A の山」という表現だけで 2 種類の登録がある。この場合、どちらが出力する意味として妥当であるか判断できないため、本来出力すべき意味とは異なる意味を出力してしまう。比喻と判断した名詞句「A の B」773 句のうち 58 句が本来出力すべき意味とは違う出力結果であった。

意味素対応隠喩辞書に極端に上位の意味素と単語の組が設定されることにより、上記のような意味の誤判断が発生することが確認できた。これらの問題は、比喻と判断した出力結果の信頼性を低下させる要因となる。例えば、設定された極端に上位の意味素と単語の組と同一の表現の名詞句「A の B」を入力した場合、比喻の名詞句「A の B」と非比喻の名詞句「A の B」の区別なくすべて比喻の名詞句であるとしてその意味を出力してしまう。実際に、登録されている極端に上位の意味素と単語の組(具体, 卵)と同一の表現を入力した結果を表 8 に示す。表 8 に示されるように、比喻の名詞句である「先生の卵」と非比喻の名詞句である「恐竜の卵」をどちらも比喻の名詞句であると判断し意味を出力している。このことから、本手法に名詞句「A の B」を適用して比喻と判断された出力結果は、その意味が正しいということに対して信頼性が低いことが確認できる。そのため、結果を出力する前に、比喻と判断された名詞句「A の B」とそれに対応する意味に対して、その意味が妥当かどうかの検証が必要であると考えられる。

4.3.2 隠喩辞書に登録されていない表現について

慣用隠喩辞書と意味素対応隠喩辞書のどちらにも登録されていない表現が入力されることによる検索失敗が、比喻の名詞句「A の B」を入力し非比喻の名詞句「A の B」であると判断された 227 句中 176 句存在した。その一部を表 9 に示す。本手法は慣用隠喩辞書と意

表 9 隠喩辞書に登録されていない名詞句

入力名詞句	被修飾語：意味素
成功の翼	成功：成功
時間の逆流	時間：時間
疑惑の濁流	疑惑：負信用
歴史の歪曲	歴史：学問分野・学科

表 10 未知語を含む名詞句

入力名詞句	被修飾語（未知語）
ブーイングの嵐	ブーイング
DVD の山	DVD
ライトノベルの第一歩	ライトノベル
値千金の満塁弾	満塁弾

意味素対応隠喩辞書を検索することによる意味解析を行う。その性質上「疑惑の濁流」や「歴史の歪曲」等の慣用隠喩辞書と意味素対応隠喩辞書に登録されていない表現の名詞句“ A の B ”が入力された場合は非比喩の名詞句“ A の B ”であると判断してしまう。しかし、比喩表現は、日々人間によって新たな表現が創作されるものである。そのため、本手法の実用化を考えた場合、未知の表現の名詞句“ A の B ”が入力されることは必然であると考えられる。そのため、未知の比喩表現に対する新たな枠組みが必要である。

4.3.3 未知語を含む名詞句“ A の B ”

未知語を含む名詞句“ A の B ”が入力された場合、慣用隠喩辞書の検索が成功しなければ意味素対応隠喩辞書の検索が出来ず意味解析が失敗する。ここで未知語はシソーラスに定義されていない単語である。未知語を含むために意味解析が失敗した名詞句“ A の B ”の一部を表 10 に示す。意味素対応隠喩辞書は検索クエリとして意味素と単語の組を使用する。そのため、意味素対応隠喩辞書を検索する場合には、前処理として入力された名詞句“ A の B ”から意味素と単語の組を生成する。しかし、表 10 に示されるような「ブーイングの嵐」や「DVD の山」などの被修飾語が未知語である名詞句“ A の B ”が入力された場合、被修飾語である「ブーイング」や「DVD」といった未知語に対して意味素が決定できないため意味素と単語の組を生成することが出来ない。その結果、意味素対応隠喩辞書の検索が失敗してしまう。未知語を含むことが原因で意味解析が失敗した名詞句“ A の B ”は、比喩の名詞句“ A の B ”を入力し非比喩の名詞句であると判断された 227 句中 21 句存在した。

隠喩辞書に登録されていない表現の名詞句“ A の B ”が入力された場合と同様、本手法の実用化を考えた場合未知語を含む名詞句“ A の B ”が入力されることは必然である。そのため、単語間の関連の強さなどから未知語に対して意味素を決定する方法についての検討が必要である。

5. ま と め

本稿では、自然言語を扱う様々な場面で頻出する比喩表現、特に、隠喩の名詞句“ A の

B ”を対象とし、その意味解析手法を提案した。隠喩の名詞句“ A の B ”の意味解析には隠喩の名詞句“ A の B ”に特化した辞書が必要と考え、独自で隠喩辞書を構築し、その辞書を用い意味解析を行った。また、隠喩の名詞句“ A の B ”には慣用的な表現とそれ以外の表現が存在し、それぞれに別々の枠組みが必要なことから隠喩辞書として慣用隠喩辞書と意味素対応隠喩辞書の 2 種類を構築した。

評価実験による検証の結果、比喩の名詞句“ A の B ”1000 句に本手法を適用した場合には 773 句を比喩と判断した。その内、出力された意味が正しかったものは 715 句であり、全体の約 72% の意味を解析できたことが確認できた。また、非比喩の名詞句“ A の B ”1000 句に本手法を適用した場合には 887 句を正しく非比喩であると判断した。以上のことから、本手法の有効性を確認できた。しかし、意味素対応隠喩辞書において極端に上位の意味素と単語の組が登録されることによる誤判断があった。これは、比喩と判断された名詞句“ A の B ”の出力結果の信頼性を低下させるため、結果を出力する前に妥当性・比喩性の検証を行う必要がある。また、慣用隠喩辞書と意味素対応隠喩辞書に登録されていない表現が入力された場合でも検索が失敗してしまうことから、隠喩辞書からの類似表現の推定などの新たな枠組みが必要である。

今後は、比喩と判断した名詞句“ A の B ”に対して、妥当性や比喩性の検証を行い、出力結果の信頼性を向上させる予定である。

参 考 文 献

- 1) 榊井文人, 福本淳一, 椎野努, 河合敦夫: 確率的判定尺度を用いた比喩性検出手法, 自然言語処理, Vol.J89-D, No.4, pp.860-870, (2002)
- 2) 榊井文人, 福本淳一, 荒木健治: 比喩解釈を目的とする World Wide Web を利用した属性値の適合性判定手法とそのフィードバック, 電子情報通信学会論文誌, Vol.J89-D, No.4, pp.860-870, (2006)
- 3) 池原悟, 村上仁一, 宮元健司: 「A の B」型名詞句の日英翻訳規則について, 情報処理学会論文誌, Volume43, Number7, pp.2300-2308, (2002)
- 4) 松村明: 大辞林第三版, 株式会社三省堂, (2007)
- 5) 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 小倉健太郎, 大山芳史, 林良彦: 日本語語彙大系 CD-ROM 版, 岩波書店, (1999)
- 6) 栗原健, 土屋誠司, 任福継: 意味素の関連を考慮した隠喩の名詞句“ A の B ”の意味推定手法, 平成 21 年電気学会電子・情報・システム部門大会講演論文集, pp.1408-1411, (2009)
- 7) 日本経済新聞出版社: 日経全文記事データベース 2004 日経 4 誌 DVD-ROM2004 版, (2004)