コーパスからの形容詞概念階層の構築と評価 —実データによる形容詞オントロジーの構築にむけて—

神崎 享子[†]・馬 青^{††}・山本 英子^{†††}・白土 保[†]・井佐原 均^{†,††}

本研究は、実データに基づいた形容詞の観点からみた概念体系の自動構築をめざし、その一環として、形容詞概念の階層関係構築に焦点を当てたものである。包含関係の尺度によって上位下位関係を求め、その単語間の上位下位関係に基づき概念階層を自動構築した。結果については、カバー率などの表層的な面と、階層の作られ方についての質的な面での評価を行った。階層の質的な面での評価にあたって、概念の継承関係と事例（形容詞）の各概念の成員としての連続性という観点から心理実験を行い、既存の人手によって作られた EDR 辞書の階層と比較した。実験手法は Scheffe の一対比較法を用いた。その結果、自動構築がよい、あるいは既存の辞書と有意差がないと判断された階層は、全体の 43% となった。抽出した概念数の不足や階層構築の際の問題点など課題も抱えているが、自動生成の階層が既存辞書の階層に対して、その結果の半分弱の階層で問題を提起するという意味で、ベースラインとなる数値と考える。

キーワード：オントロジー，概念，形容詞，自己組織化マップ，階層関係，シェッフエ法

Acquiring Concept Hierarchies of Adjectives from Corpora: Towards Construction of Ontology of Adjectives from a real data

KYOKO KANZAKI[†], QING MA^{††,†}, EIKO YAMAMOTO^{†††,†},
TAMOTSU SHIRADO[†] and HITOSHI ISAHARA^{†,††}

The method of organizing word meanings is a crucial issue with lexical databases. We are aiming to extract the semantic structure of concepts of adjectives from corpora automatically. The first step to achieving this is to obtain the concepts of adjectives from corpora, for which we used abstract nouns. We constructed linguistic data by extracting semantic relations between abstract nouns and adjectives from corpus data. This paper describes how to hierarchically organize abstract concepts of adjectives mainly using the Complementary Similarity Measure (CSM) which calculates inclusion relations (hypernym/hyponym relations) between words. To estimate hypernym/hyponym relations between words, we compared three hierarchical structures of abstract concepts of adjectives: according to CSM, CSM with frequency (Freq),

[†] 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター, National Institute of Information and Communications Technology

^{††} 龍谷大学理工学部, Faculty of Science and Technology, Ryukoku University

^{†††} 神戸大学大学院工学研究科, Graduate School of Engineering, Kobe University

and an alternative similarity measure based on coefficient overlap. We evaluated automatically generated concept hierarchies of adjectives with those in EDR, and found that 43% of those automatically generated were better than EDR.

Key Words: *Ontology, Concept, Adjectives, Self-Organizing Map, Hierarchy, Scheffe's paired comparison*

1 はじめに

計算機科学でいう「オントロジー」とは、ある行為者や行為者のコミュニティに対して存在しうる概念と関係の記述であり「概念」というのは、何らかの目的のために表現したいと思う抽象的で単純化した世界観である (Gruber 1992). 認知科学では、「概念」について外延的意味 (事例集合で定義された意味) と内包的意味 (属性の集合から定義された意味) の見方があるとする (安西 1989). 我々の認知活動の中で、概念化は、語、文、文脈、動作の仕方、事柄、場面など、様々なレベルで行われている。では、なぜ対象の概念化が必要かという点、河原では、Medin and Goldstone を引用して「概念」の機能を次のように述べている (Medin and Goldstone 1990; 河原 2001)「現在の経験を、あるカテゴリの成員とみなす (分類) ことで、その経験を意味のあるまとまりとして解釈し (理解と説明)、そこから将来に何がおきるか (予測) や関連する別の知識 (推論) を引き出すことが可能になる (コミュニケーション)。その他、複数の概念を表す語を組み合わせることで新たな概念を生成したり、新たな概念の記述を生成してから、その記述にあう事例を検索することもできる」。つまり、人間や計算機が効率的に柔軟な活動をするために、概念と、(言語化する・しないにかかわらず) 概念の具体化された表現 (あるいは事例) の総体である「オントロジー」は重要な役割を担っているといえる。

我々が対象とする言語的オントロジー、特に、語彙の概念を体系化したオントロジーは、概念体系や意味体系と呼ばれ 10 年以上前から人手で構築されてきた (EDR 電子化辞書 (日本電子化辞書研究所 1995) や分類語彙表 (国立国語研究所 1964) など)。その目的は、ある特定のアプリケーションでの利用ではなく、我々の言語知識を体系化することであり、その知識体系を利用して計算機に予測・推論・事例の検索・新たな概念の理解など、深い意味処理をさせることを目的としている。本研究がめざす「形容詞のオントロジー」の目的も、従来の語彙的なオントロジーの目的と同様に、計算機や人間が、形容詞を使って表現する知識の体系化をはかるものである。ここで本研究の「形容詞」とは、形容詞と形容動詞を含むものとする。

従来のものと異なる点は、実データからの獲得を図るため、運用の実態を反映したオントロジーを得ようとする点である。人間の内省による分析の場合、概念記述を行う個々人の言語的経験から、概念体系の粒度や概念記述に差異がでてくる。心理実験のように、複数の人が同じタスクをすれば共通の傾向もとれるが、通常のプロジェクトでは同じ個所に多くの人を投入

することは不可能である。自動獲得の目的は、できるだけ実際の言語データから言語事実を反映した結果を得ることである。一つ一つのテキストは個々人の記述だが、それを量的に集めれば、複数の人のバリエーションを拾うことができ、結果的に多くの人の言語運用の実態をとることができる。言語データから意味関係を反映した概念体系を捉えられれば、人間の内省によって作られたオントロジーや言語学的知見、意味分類などと比較することは意義があるのではないかと考える。

ところで、コーパスからの語彙のクラスタリングや上位下位関係の自動構築などについては、Webの自動アノテーションやインデックス、情報検索など、その目的は様々であるが、そのほとんどが、名詞や動詞を対象にした分類や関係抽出である。形容詞や副詞に関する研究はまだ少ない。しかし、形容詞や副詞が語彙のオントロジーにとって重要でないわけではなく、たとえば、WordNetで形容詞の意味情報が手薄であることを指摘し、イタリア語形容詞の意味情報を導入することで、ヨーロッパの複数言語で共同開発しているEuroWordNetの抽象レベルの高い概念体系(Euro WordNet Top Ontology)に変更を加えることを試みている研究がある(Alonge, Bertagna, Calzolari, Roventini, and Zampolli 2000)。

オントロジーの主要な関係の一つに、類義関係と階層関係がある。形容詞概念を表すような抽象的な名詞の類義関係については、馬らなどの研究がある(馬, 神崎, 村田, 内元, 井佐原 2001)。しかし、形容詞概念の階層関係については、まだ研究が進んでいない。本研究では、形容詞概念の階層関係に着目し、コーパスから取得した概念から階層を構築する方法と、妥当そうな階層を得るための評価について述べる。

本研究で扱う概念数は約365概念であり、それに対しEDR電子化辞書の形容詞の概念数が約2000概念ほどと考えると取り扱うべき概念はさらに増える可能性があるが、本研究は、現時点よりも多くの概念数を扱うために、まず、現段階での概念数で、階層構築とその評価方法について実験および考察を行ったものである。

我々は、第2節でオントロジーのタイプの中で本研究がめざすオントロジーについて述べ、第3節で先行研究の言語学的考察から、形容詞の概念を語彙化したような表現があることを述べ、形容詞の概念をコーパスから抽出する。第4節では第3節で抽出したデータをもとに複数の尺度での階層構築と、得られた階層のうち、妥当そうな階層を判別するための条件を述べ、第5節で心理実験によってEDRの形容詞概念階層と比較評価を行う。第6節でオントロジー構築に向けての今後の展望をのべ、第7節でまとめを行う。

2 オントロジーのタイプ

「オントロジー」という用語が喚起する定義や種類、目的などは多様化しているため、本研究がどのようなオントロジーで、何を目的として作るのかを明らかにする必要がある。

Sowa は、主なオントロジーのタイプとして、terminological ontology, prototype-based ontology, formal ontology などあげている (Sowa 2003). Terminological ontology は、概念の上位下位関係や部分全体関係などの関係が特定され、他の概念との相対的な関係が決められているタイプのものである。このタイプのオントロジーは概念を完全に定義するものではない。

意味公理や論理的な定義よりも、プロトタイプや事例によって弁別されるオントロジーが、prototype-based ontology である。つまりプロトタイプ集合や事例集合が相対的な関係をもって分類されているタイプのものである。

また、論理形式で書かれる意味公理や定義によって概念が弁別されるオントロジーが formal ontology である。論理の複雑さの制限はない。一般的な terminological ontology と formal ontology の違いは、種類というより、概念間の関係の深さが異なり、formal ontology は規模が概して小さいが、深い記述がされるため推論をサポートするのに用いられる。

Biemann はそれぞれのタイプを図示し (図 1)、長所と短所を述べている (Biemann 2005)。

このうち、formal ontology は直接推論に使えるものの、コード化に労力がかかり、大規模になると不整合もおきる危険性がある。一方、terminological ontology や prototype-based ontology は自動化がしやすく作りやすい。しかし、prototype-based ontology は、概念ラベルがないので、QA システムなどには使いにくい。このオントロジーは語のクラスタリングによってすぐに導出されるので terminological ontology より構築しやすいが、利用しにくい。

その他の種類として、シソーラスがあげられる。シソーラスは関連語のまとまりをもち、prototype-based ontology に似ている。しかし、関連語などの中には互いに異なった関係も含まれている場合がある。

テキストからオントロジーを学習するのに利用できる手法として、分布特性によるクラスタリングをはじめ、表層的な統語情報や特定の言語表現パターンなどを利用して、階層関係や部分全体関係などの単語間の関係を獲得する研究などがある (Hindle 1990; Hearst 1992; Caraballo 1999; Berland and Charniak 2000)。これらの研究と関係が深い、Semantic lexicon の構築としてのオントロジーの学習という観点もある。Semantic lexicon は、カテゴリと事例のセットという形であるが、オントロジーのように内部的に構造化されていない。そのアルゴリズムのほとんどが bootstrapping のアプローチである (Riloff and Shepherd 1997; Thelen and Riloff 2002; Lin and Pantel 2002; Pantel and Ravichandran 2004)。

以上、Sowa, Biemann に従ってオントロジーのタイプを概観したが、我々が目的とする形容詞のオントロジーは、形容詞が事例となり、その事例が共通してもつ概念がラベルとなり、その概念ラベルが構造化されたオントロジーという形である。上記でいえば、Semantic lexicon のように形容詞の事例とそれらが共通にもつ概念を一つのユニットとして、それが、ラベルつき prototype-based ontology のように構造化されたものである。これは、EDR 電子化辞書や分類語彙表などの語彙の概念体系／意味体系と同様の形である。ただし、概念名については、EDR

Prototype-based ontology

Formal ontology

Axioms:

food(brie), food(camembert), food(turkey), food(meatballs), food(chili con carne), meat(turkey), meat(minced meat), part_of(minced meat, chili con carne), part_of(minced meat, meatballs)

veg_food(x) = { x | food(x) ∧ (¬part_of(y,x) ∧ meat(y)) ∧ ¬meat(x) }

non_veg_food(x) = { x | food(x) ∧ ((part_of(y,x) ∧ meat(y)) ∨ meat(x)) }

Possible to derive: "turkey" and "chili con carne" are non-vegetarian foods

図 1 菜食主義用食物と非菜食主義用食物を区別する食物オントロジー
(Biemann 2005 p.79 からの引用)

category: sets of its instances	
A:	N1, N2, N3, ...
B:	N21, N22, N23, ...
C:	N31, N32, N33, ...

図 2 Semantic lexicon

電子化辞書は説明的記述を使い、『分類語彙表』の項目では抽象的な語句あるいは代表的な形容詞を使っているが、本研究では、概念ラベルを抽象的な名詞で表現している。オントロジーは柔軟に概念と表現を結びつける必要があるので、概念間の関係は一種類ではなく複数想定され、木構造よりネットワーク的な形態の可能性もある。しかしまずは、概念の類義関係と階層関係をとらえていきたいと考える。本稿では、その一環として、特に、形容詞の概念の上位・下位関係について焦点をあてる。

3 言語表現に現れる「概念とその具体事例」という関係

3.1 言語現象

日本語では、高橋が、以下のような例を言語学的に観察している (高橋 1975).

- (1) やぎは性質がおとなしい
- (2) ぞうは鼻が長い

(1) の「性質」は、主体「やぎ」からみると主体のある一面を表すので「側面語」であり、(2) は、主体「ぞう」からみると身体の一部であるので「部分語」と呼び分けて、統語的には同じ構造でも、語の文中での役割の違いを指摘した。そして、側面語については、主語が示すものの側面を表すとともに、述語が示す属性の類概念 (上位概念) を表す単語であると考察している。また根本は、「色が白い」「速さがはやい」「年が若い」「背が高い」などは、同義反復的な性格が強いと述べている (根本 1969)。このように、我々の言語活動の中でも、形容詞の上位概念を語彙化した表現がみられる。

本研究では、(1) の「性質—おとなしい」や「色—白い」「速さ—はやい」などの関係を、「概念とその具体事例」という関係と捉え、このような抽象的な名詞と形容詞の共起をコーパスから抽出する。本稿では、このような抽象的な名詞を「概念」とよび、共起する形容詞を「事例」と考える。

3.2 コーパスからの抽出方法

次に抽出方法であるが、上記の例のような「N は X が Adj」というパターンは今回は利用していない。この構文の N と X, Adj の意味関係は様々である。それは、3.1 節の (1) と (2) が構文上同じ形をしているにもかかわらず、N や Adj に対する X の意味関係が違っていることからわかる。従って、コーパスから「N は X が Adj」の文型を集めると、量が多いが、「概念とその具体事例」という関係を雑多な関係の中から取捨選択することが難しい。これにはなんらかの基準が必要である。(なるべくある程度のノイズがあっても自動化したり、あるいは人間が簡単に判別できる基準を求める必要があるが、それは今後の課題とする。)

- 1 抽出方法は、毎日新聞 94 年、95 年の 2 年分のデータから、「X トイウ Y」の句を抽出することからスタートする。「トイウ」の直前に現れる表現は、内容節 (あるいは内容語) である。寺村、益岡は連体修飾表現を分析しているが、特に益岡では「トイウ」内容節をとる場合は、Y が X の属する範疇だとしている (寺村 1991; 益岡 1994)。たとえば、「質実剛健という気風」の例では、「気風」(Y に相当) は「質実剛健」(X に相当) の属する範疇ということである。

そこで、何かを範疇化する可能性をもつ「Y」の収集を目的として、「X トイウ Y」を使って「Y」をコーパスから自動的に抽出した。このプロセスは、具体事例によって説明され

る被修飾名詞の収集を意図している。具体的には、コーパスを形態素解析 (JUMAN を使用) したあと、「トイウ」の前後の単語を抽出した。最終的に Y にくる被修飾名詞は 15,391 語となった。

- 2 次に、「トイウ」を介在して内容語を取る Y を集めた後、Y と共起する形容詞をコーパスから抽出する。たとえば、「温和な性格」とは表現しても、「トイウ」を介在させた「温和という性格」のような表現はあまり見られない。そこで、単に Y と共起する形容詞を取り出すことにした。使用したデータは、毎日新聞 11 年分、日本経済新聞 10 年分、日経産業金融流通新聞 7 年分、読売新聞 14 年分である。また、新聞以外でも形容詞と Y (名詞) の共起をとるため、新潮文庫 100 選、新書 100 冊についても使用例を人手によって調べ追加した。このように取り出した共起関係には、雑多な共起関係が含まれているので、本研究で対象とした、X が具体事例で Y がそれを範疇化した語という共起対 (「赤い」と「色」のような共起対) を、最終的に人手で整理した。

人手によるデータ整理は、一人の作業者が行った。判定基準は、神崎、神崎・井佐原で記述されている、連体用法にみられる形容詞と名詞の統語的・意味的關係の中から、以下の關係を採用した (神崎 1997; 神崎, 井佐原 1999)。

I Adj (形容詞) + X (名詞) ⇒ X が Adj

X を限定する表現 (その X, 「N は X が～」など) をとらなくても「X が Adj」に変換可能なもの。

被修飾名詞が属性を表現するタイプ。(例) ゆるやかな傾斜

II 主語述語關係に変換ができない

II-1 形容詞は、被修飾名詞の意味を構成する一部の意味だけにかかり、類似した意味を重ねて強調する働き (例) 古い昔

II-2 被修飾名詞の指示対象の内容を表す (例) 悲しい思い

I) は直接的な属性—属性値の關係であり、II-1) は被修飾名詞の意味の一部を形容詞で顕在化し重ねることで強調する關係、II-2) は被修飾名詞の指示対象の内容を具体化する關係となっている。どれも被修飾名詞の意味を含意しつつ形容詞表現で具体化する關係であるため、これら 3 つのパターンを採用した。

以上のように、上記プロセスの 2 のステップでは、最終的に人間の内省で取捨選択しており、やはり労力がかかるが、1 のステップで、被修飾名詞を「何かを範疇化する可能性のある語」に限定することで、抽出対象とする形容詞と概念の共起対を多く含んだデータを得ようとした。

人間の取捨選択で得られた形容詞と Y (名詞) 共起対の総数は 36,023 共起対、異なりが 10,524 共起対である。概念数は 365 で、最大共起形容詞数は、「こと」の 1,594 語である。出現する共起形容詞数に対する各概念の例は以下のようになる。

表 1 形容詞の共起数ごとの概念例

共起形容詞数 (個)	概 念 例
1-10	曲線, 縁, 緊張関係, 時刻, 質, 騒音, 大胆さ, 弾力性, 読み方, 家柄, 輝き, 勾配, 時間, 背丈, 評価
11-30	価値, かかわり, 外形, 才能, 境遇, 割合, 活気
31-50	体格, 味, におい, 能力, 影響, 音, 心境, 資質
51-70	気品, 形式, 見方, 位置, 規模, 形状, 情勢, そぶり
71-90	物腰, 美, 面持ち, 感触, 状況, 次元, 分野
91-110	風情, 気配, 気風, 身
110-200	作り, 見地, 外見, 性質, 局面, 程度, 関係, 体質
200-500	観点, 段階, 感情, 意味合い, 傾向, 態度, 方向
500-1000	様子, 雰囲気, 気分, 感じ, 感覚, 状態, 側面, 意味
1001-1594	ところ, 印象, 面, 性格, こと

そして, 抽出した概念と形容詞グループを最終的に以下のようなリストにまとめた.

表 2 抽出した概念と形容詞グループ

概念	形容詞
感触	柔らかい, 湿っぽい, なめらかな, ...
感性	ナイーブな, 繊細な, しなやかな, ...
形	四角い, 丸い, 平たい, 曲線的な, ...

概念名となる抽象名詞と形容詞集合のリストを作る際, 同じ抽象名詞であっても, 明らかに指示対象が異なる場合は, 番号をつけて区別した. たとえば, 物理的な「形」(たとえば「丸い形」)と, 形式的な意味の「形」(たとえば「おだやかな形で〜」)では, 「形」の指すものが異なるので, 「形1」「形2」のように区別し, 形容詞集合をわけた.

一方, 形容詞についてはどのような名詞と共起しているか, だけを見ており, ここで形容詞の多義は区別していない.

4 概念の階層関係の構築—階層構築の手法と閾値の選定基準—

本節では, 包含関係を求める尺度を使い, 第3節で抽出した形容詞の概念の階層構築を行う.

概念間の包含関係を求める尺度として, Hagita and Sawaki が開発し山本・梅村が言語データへ応用した補完類似度 (Complementary Similarity Measure, CSM) と, 頻度を考慮した CSM

(Freq), そして, CSM 以外の尺度としてオーバーラップ相関係数 (Overlap Coefficient, Ovlp) を使う (Hagita and Sawaki 1995; 山本, 梅村 2002; Manning and Shütze 1999).

2つの概念間の包含関係を計算したのち階層構築を行うが, 予備実験を行った結果, 全ての包含関係の概念ペアを使うと明らかにランダムな長い階層ができた. そこで, 包含関係が希薄な概念ペアが階層関係の精度を低くすることを阻止するため, 包含関係の値に閾値を設定することにした.

しかし, 閾値設定の問題点として, ゴールドスタンダードがないこのタスクで, 複数の尺度や閾値の組み合わせから構築された階層の明らかな差異を, 一見して判定できない場合, どのように妥当そうな階層を特定したらよいのだろうか. 第5節では心理実験によって階層評価を行うが, 事前に明らかに不適当な尺度と閾値の組み合わせを除外するために, 緩やかな階層の評価方法が必要になる. 本節では, 包含関係の尺度を用いて階層構築を述べた後, 包含関係の尺度と閾値の組み合わせの妥当性をいくつかの観点から眺め, 妥当そうな類似尺度と閾値をある程度特定する方法を述べる.

4.1 包含関係の尺度

包含関係の尺度として, 3種類の方法を用いた.

まず, 補完類似度 (CSM) について述べる. 補完類似度は一対多関係を推定する尺度として提案されたが, 事象間の一対多関係は, 包含関係 (あるいは上位下位関係) を表すので, 包含関係を推定する尺度とも考えられる. 山本, 梅村では, {沖縄県, 那覇市} のような一対多関係を推定するタスクを行っており, そのタスクでは, コーパスからの関係抽出に用いられた他の類似尺度や連想規則の抽出に用いられる類似尺度など (たとえば相互情報量, コサイン関数, ダイス相関係数など) よりもよい結果を示したことが報告されている (山本, 梅村 2002).

この補完類似度を用いて, 対象としている概念の包含関係, つまり上位下位関係を推定する. 補完類似度は以下ようになる.

今, 共起形容詞のセットで定義した抽象名詞 F と T があるとする. 我々のデータでは, F と T の特徴ベクトルは, 双方の共起形容詞の出現状況を 0 または 1 で表現したものに相当する. それを以下のように表す.

$$\vec{F} = (f_1, f_2, \dots, f_i, \dots, f_n) \quad (f_i = 0 \text{ または } 1)$$

$$\vec{T} = (t_1, t_2, \dots, t_i, \dots, t_n) \quad (t_i = 0 \text{ または } 1)$$

そして、補完類似度の式は以下ようになる。

$$CSM(F, T) = \frac{ad - bc}{\sqrt{(a+c)(b+d)}} \quad (1)$$

$$a = \sum_{i=1}^n f_i \cdot t_i, \quad b = \sum_{i=1}^n f_i \cdot (1 - t_i),$$

$$c = \sum_{i=1}^n (1 - f_i) \cdot t_i, \quad d = \sum_{i=1}^n (1 - f_i) \cdot (1 - t_i),$$

$$n = a + b + c + d$$

“ a ” は F と T で共通する共起形容詞の数である。

“ b ” は F とは共起するが T とは共起しない形容詞の数である。

“ c ” は F とは共起しないが T とは共起する形容詞の数である。

“ d ” は F と T とも共起しない形容詞の数である。

“ n ” は、ベクトルの次元数となる。

F が T を完全に包含する場合、 $c = 0$ となり、 T が F を包含する場合、 $b = 0$ となるため、 $bc = 0$ となる。補完類似度では、一致情報 (ad) と不一致情報 (bc) の差分をとるので、包含関係にある二語間の類似度は高くなる。

さらに、補完類似度は F から T の類似度と T から F の類似度が非対称であることも特徴の一つである。 F から T をみた補完類似度では、 b は F だけに出現する形容詞の数、 c は T だけに出現する形容詞の数である。逆に、 T から F をみた補完類似度では、 b は T だけに出現する形容詞の数となり、 c は F だけに出現する形容詞の数となる。計算式の分母をみると、 F と T がどちらの方向の類似度を計算するかで、 b と c に代入される数値の大小が逆転し、それに伴って、類似度も非対称になる。

次に CSM と比較する手法として、オーバーラップ相関係数 ($Ovlp$) と頻度つき $CSM(Freq)$ の階層を用いる。

オーバーラップ相関係数について、Manning and Shütze は包含関係を求める尺度として述べている (Manning and Shütze 1999)。これは、二値ベクトル間の類似尺度で、計算式は以下のようになる。

$$\vec{F} = (f_1, f_2, \dots, f_i, \dots, f_n) \quad (f_i = 0 \text{ または } 1)$$

$$\vec{T} = (t_1, t_2, \dots, t_i, \dots, t_n) \quad (t_i = 0 \text{ または } 1)$$

$$Ovlp(F, T) = \frac{|F \cap T|}{\min(|F|, |T|)}$$

$$= \frac{a}{\min(a+b, a+c)} \quad (2)$$

$$a = \sum_{i=1}^n f_i \cdot t_i, \quad b = \sum_{i=1}^n f_i \cdot (1 - t_i),$$

$$c = \sum_{i=1}^n (1 - f_i) \cdot t_i$$

a , b , c などのパラメータは, CSM での定義と同様である.

次に頻度を考慮した補完類似度について計算式を示す. これは, Sawaki, Hagiga, and Ishii が二値画像のための補完類似度を, 多値画像解析のために拡張したものである (Sawaki, Hagiga, and Ishii 1997). Yamamoto, Kanzaki, and Isahara ではこれを言語データに応用している (Yamamoto, Kanzaki, and Isahara 2005). これは, 多値ベクトル間の類似尺度で, 計算式は以下ようになる.

$$\vec{F}_g = (f_{g1}, f_{g2}, \dots, f_{gi}, \dots, f_{gn}) \quad (0 \leq f_{gi} < 1)$$

$$\vec{T}_g = (t_{g1}, t_{g2}, \dots, t_{gi}, \dots, t_{gn}) \quad (0 \leq t_{gi} < 1)$$

$$CSM_g(F_g, T_g) = \frac{a_g d_g - b_g c_g}{\sqrt{n T_{g2} - T_g^2}} \quad (3)$$

$$a_g = \sum_{i=1}^n f_{gi} \cdot t_{gi}, \quad b_g = \sum_{i=1}^n f_{gi} \cdot (1 - t_{gi}),$$

$$c_g = \sum_{i=1}^n (1 - f_{gi}) \cdot t_{gi}, \quad d_g = \sum_{i=1}^n (1 - f_{gi}) \cdot (1 - t_{gi}),$$

$$T_g = \sum_{i=1}^n T_{gi}, \quad T_{g2} = \sum_{i=1}^n t_{gi}^2$$

この定義式において, 各要素は, i には, 抽象名詞 (概念に相当) が i 番目の形容詞と頻繁に共起するかどうかの状況を表す, 共起頻度に基づく重みを用いる. 重みは以下のように求める.

$$Weight(noun, adj) = \frac{Freq(noun, adj)}{Freq(noun, adj) + 1} \quad (4)$$

重みの値域は $0 \leq Weight(noun, adj) < 1$ である.

ベクトル $\vec{F}_g$ を持つ抽象名詞を $noun_F$ と $\vec{T}_g$ を持つ抽象名詞を $noun_T$ とし, 上式の重みで f_{gi} と t_{gi} を表すと, ベクトルは以下のように表される.

$$\vec{F}_g = (f_{g1}, f_{g2}, \dots, f_{gn})$$

$$= (Weight(noun_F, adj_1), Weight(noun_F, adj_2), \dots, Weight(noun_F, adj_n))$$

$$\vec{T}_g = (t_{g1}, t_{g2}, \dots, t_{gn})$$

$$= (Weight(noun_T, adj_1), Weight(noun_T, adj_2), \dots, Weight(noun_T, adj_n)) \quad (5)$$

4.2 概念階層の構築方法

CSM などによって包含関係を計算し、値を正規化して得られたリストの一部を示すと以下のようになる。

表 3 CSM によって推定された包含関係

概念	概念	CSM 値
イメージ	印象	1.0
こと	かたち 1	0.955
印象	イメージ	0.951
印象	感じ	0.936
...	...	...
感じ	印象	0.778

単語 A から単語 B を見たときの CSM 値が、単語 B から単語 A を見たときの CSM 値より大きければ、単語 A が上位語、単語 B が下位語となる。たとえば、本稿では概念とは第 3 節で抽出した抽象的な名詞で定義しており、上記の概念の並びは、左の概念からみた右の概念の包含関係を表している。たとえば、左の概念が「印象」で右の概念が「感じ」の場合は、「印象」からみた「感じ」を示し、上記では CSM 値は 0.936 となる。逆に、方向が逆転し、左の概念が「感じ」で右の概念が「印象」の場合は、「感じ」からみた「印象」の包含関係を表し、0.778 となる。この場合「印象」から「感じ」を見るほうが、CSM 値が高いため、「印象」は「感じ」の上位概念となる。ただし、この場合は両方向からの CSM 値が高いため、かなり事例に重なりがあると考えられる。

上記のような二単語間の包含関係を求めた後、これを利用して階層を構築する。階層構築方法は次のようになる。

- (0) 初期階層として、CSM 値の高い順に二単語をつなげる。

ここでは、仮に単語 A が上位語、単語 B が下位語という関係とする。

階層 (0): A-B (“-” は上位下位関係を示す記号とする)

- (1) まず、階層 (0): A-B の下位語を探索する。

二単語間の CSM 値のリストから、単語 B を上位語として、B の下位語として最大値をとる単語 X を探して、単語 B の後ろに連結し、次に、その単語 X を上位語として、X の下位語として最大値をとる単語 Y を探して、単語 X の後ろに連結する。この操作を下位語がなくなるまで繰り返す。これによって以下のような階層ができる。

階層 (1): A-B-X-Y

- (2) 次に、階層 (0): A-B の上位語を探索する。

二単語間の CSM 値のリストから, 単語 A を下位語として, A の上位語として最大値をとる単語 W を探して, 単語 A の前に連結し, 次に, その単語 W を下位語として, W の上位語として最大値をとる単語 V を探す. この操作を上位語がなくなるまで繰り返す. 階層 (1) と連結することで以下のような階層ができる.

階層 (2): V-W-A-B-X-Y

ただし, 上位下位関係は必ず保存する. もし上位下位関係が逆転した場合はその関係は連結しない.

- (3) 長い階層に完全に含まれる短い階層はマージし, 二つの階層のうち一単語ずつ異なる場合は, 各階層の差異となる二単語の補完類似度の値を測り, 上位下位関係があれば結合した.

例 1) A-B-C-D-E と A-B-D という階層があるとき

A-B-D は, 順序を保存した状態で長い階層に完全に含まれるのでマージし, 短い階層は削除する.

A-B-C-D-E

例 2) A-B-C-D-E と A-B-X-D-E があるとき

C と X の補完類似度の値を求め, C と X に上位下位関係があれば結合した.

A-B-C-X-D-E

- (4) 最後に各階層のトップに「こと」を結合する. 「こと」は全ての形容詞と共に起するとして, 計算時間の便宜上, 「こと」は最後に各階層のトップに結合させることとした.

最終階層: こと-V-W-A-B-X-Y

4.3 妥当そうな手法と閾値の選定

4.2 節で求めた概念間の包含関係を全て使って階層を構築すると, 冗長な意味のない概念階層 (つまり単語の羅列) になり, 閾値があまりに高いと, 概念階層は非常に短くなる (つまり, 連結される名詞があまりにも少なくなる). そこで, 包含関係が希薄な概念ペアが階層関係構築時に悪影響を及ぼすことを阻止するため, 包含関係の値に閾値を設定することにした. CSM については 0.3 と 0.2, Ovlp については 0.3 と 0.2, Freq については 0.2 と 0.1 の閾値を設定し, その閾値以上の概念間の包含関係を用いて階層を構築した. これらの閾値による階層は, 概念を連結したある程度の長さの階層であり, かつ, 明らかにおかしい概念の羅列ではない. この閾値以上でも以下でも, 前述の弊害が出る. 逆にいえば, 前述の閾値からできた階層は, 一見して妥当なのか, 妥当ではないのか, すぐにはわからないともいえる. 手法ごとに多くの階層が生成され, 閾値をいくつか設定すれば, その分だけ, また階層が増えるので, 心理実験などで既存辞書との比較評価を行おうと思えば, 妥当そうな階層を事前に選定した方が効率的である.

そこで, 次のような観点から, 階層を分析した.

(1) 形容詞の階層としてできた階層の割合

第2節で述べた prototype-based ontology で構造化された事例集合（図1）をみるとわかるように、最下位レベルで出現している「チーズ」は、最上位レベルの事例集合にも出現している。

通常、上位概念の特徴は下位概念の特徴に「継承」される。概念の特徴を定義するのが事例集合である場合、下位レベルで出現する事例は、上位レベルの概念の事例にもなる（「すずめ」は、鳥の事例でもあり、動物の事例でもあり、生物の事例である）。

この認知科学的ルールから、自動構築した階層の、最下位概念の事例集合（形容詞の集合）が、最上位概念までの各概念の事例集合に含まれているかを調べ、連続して出現していれば、その階層は、当該形容詞の階層と考える。本稿のデータで考えると、「形容詞の階層としてできた階層」とは、ある形容詞が、最下位から最上位に位置するすべての概念の形容詞集合に出現している場合、その階層を「形容詞の階層としてできた階層」と呼ぶ。もし、ある形容詞が、階層のどこかの概念の形容詞集合の成員でなければ、形容詞の階層とはよばず、手法によって得られた「階層」とよぶ。

この考えに則って、手法ごとに、得られた階層の中で、形容詞の階層として得られた階層が何割あるか、計算した。分母は、ある手法に基づいて構築された全階層であり、分子は「形容詞の階層としてできた階層」である。

$$\text{形容詞の階層として得られた階層の割合} = \frac{\text{形容詞の階層としてできた階層数}}{\text{ある手法に基づいて構築された階層の総数}} \quad (6)$$

(2) 事例としてコーパスから抽出された全形容詞のうち、何語の形容詞に「形容詞の階層」が得られたか。

(3) 階層を構成する概念の割合

概念としてコーパスから抽出した抽象名詞は全部で 365 語あるが、そのうち何割が階層を構成しているか。階層を構成する概念の割合は次のように計算した。

$$\text{階層を構成する概念の割合} = \frac{\text{階層を構成する概念の異なり数}}{\text{抽出した全概念数 (365 語)}} \quad (7)$$

上記3つの観点から各手法の結果を求めると、表4のようになる。

表中で、高い数値の第一位から第三位に「○」、そのうち極端に数値が高い場合は「◎」を数字の前に付与した。また、極端に数値が低いものには「*」を付与した。

表1から、総合的にみると CSM0.2 と Ovp0.3 が、形容詞の階層としてできた階層の割合も、階層ができた形容詞数や 365 の概念のうちで階層を構成する概念の割合もよいとわかる。CSM0.3 は、6 種類の階層の中で形容詞の階層を最も多く作っているが、階層を構成している概念の数は、

表 4 階層の作られ方からみた手法別の階層の特徴

	階層の深さ (最多ノード)	形容詞の階層として できた階層の割合	階層ができた 形容詞数(語)	階層を構成する概念 (抽象名詞)の割合
CSM0.3	13 ^注	◎ 90%	◎ 175 語	* 24%
CSM0.2	13	○ 58%	○ 101 語	○ 90%
Ovlp0.3	9	○ 63%	○ 99 語	○ 86%
Ovlp0.2	9	52%	53 語	○ 92%
Freq0.2	8	53%	70 語	* 28%
Freq0.1	12	* 36%	67 語	82%

注) CSM0.3 の 13 については深さ 10 でほとんどの階層が作られており、深さ 11 や 12 の階層数は 0 なのだが、深さ 13 の階層が 1 つだけ作られた。

が最も低い。これは同じ階層をもつ形容詞のグループが、未分化である可能性がある。また、Ovlp0.2 の階層は、形容詞の階層はあまり作られていないが、対象にしている 365 の概念をほとんど使って、階層を作っている。これは、冗長に階層を作っている可能性がある。程度の差こそあれ同様の傾向がみられるのは Freq0.1 である。形容詞をカバーする階層は少ないが、階層を構成している概念が多いことがわかる。Freq0.2 は、形容詞をカバーする階層は多めであるが、それより顕著な特徴は、階層を構成する概念の種類が少ないことである。程度の差こそあれ、その点では、CSM0.3 に似た傾向がみられる。

上記の結果より、CSM0.2 と Ovlp0.3 は、外見的に妥当そうな階層となっているので、EDR と比較する階層としてこの二者を選択する。また、頻度を考慮した CSM (Freq) については、両者の閾値とも外見上それほど適当ではないが、異なる種類を比較するという点で、形容詞の階層が比較的得られている Freq0.2 を、EDR との比較実験に加えることとする。

5 自動構築の階層と EDR 辞書の概念階層との比較評価

本節では、本研究の提案手法によって自動構築した階層と、EDR 階層の優劣を心理実験によって定量的に比較する。

EDR 電子化辞書は、10 年ほどの年月をかけ、言語学者や辞書編纂者などが携わった大規模な計算機用辞書である。この電子化辞書は概念体系をもち、各語彙は概念 ID によって概念体系とリンクしており、概念 ID から上位概念や下位概念などを辿ることができる。EDR の概念体系には局所的な不整合性や冗長性などの問題はあっても、概念分類と概念間の関係を細かく記述した全品詞にわたる大規模シソーラスである。本研究が対象とする形容詞概念についても、EDR は広範囲にカバーしているため、本実験では EDR を比較対象とした。EDR の概念記述は単語で定義されている場合もあれば、文で説明していることもある。たとえば、EDR の「肯定

的な」という形容詞の概念階層の例を示すと次のようになる。

EDR 階層：「肯定的な」

概念 (3aa966) → 事象 (30f7e4) → 移動 (30f801) → 情報の移動 (30f832) → 情報の受信 (3f96e7)
→ 知る (30f876) → 認知主体と認知対象との認知的距離減少 (3f972c) → (意見などに) 同意しているさま (0f0ae2)

一方、自動生成の階層は、各ノードの概念が抽象名詞で表現されている。たとえば以下のような階層である。階層中、「面 1」のような名詞の横の番号は、「面」の意味を区別した際に付与した番号である。(ここでは、「やさしい面がある」のような形式的な意味の「面」と「丸い面」のような物理的な面とを区別し、前者の形式的な意味の「面」を「面 1」とした。)

自動生成の階層：「肯定的な」

こと → 面 1 → 傾向 → 見方 → 評価

二つの異なる形式の階層比較をした研究としては、概念記述間の単語の一致度によって表層的に評価した研究がある(山本, 神崎, 井佐原 2006)。しかし上記の通り、本研究で自動抽出された階層と EDR 階層の構造は異なり、また概念の比較評価という抽象的な対象を扱うタスクであるため、我々は、表層的な比較ではなく、心理実験による優劣の比較を行った。

4.3 節 (1) でも触れたが、階層関係が成立する重要な条件として、

- 1) 上位概念の下位概念への継承性
- 2) 事例(本稿では形容詞)が、事例集合の成員として最下位概念から最上位概念まで出現する連続性

という二つの条件がある。本実験の目的は、上記二つの観点から、自動構築の階層と EDR 階層の優劣の程度に関する人間の判断を数値化することにより、これら二つの階層間の優劣を統計的に比較することである。数値化手法としては Scheffe の一対比較法を用いた(Scheffe 1952)。Scheffe の一対比較法で得られる値は相対的な値であるため、各々の階層の絶対的な優劣の程度は数値化できないが、異なる階層間の優劣の相対的な程度を数値化することができる。そしてこのような数値化によって、階層間の優劣の判断に関する被験者間のばらつきを考慮した上で、階層間の優劣を統計的に検定することができる。なお、言語データの評価に Scheffe の一対比較法を用いた他の研究として、丸元らの研究があるが、この研究は敬語表現の丁寧さの程度の数値化に Scheffe 法を用いたものである(丸元, 白土, 井佐原 2005)。

5.1 実験データについて

実験には、下記に示す三種類のデータを用いた。

- 1) 30 語の形容詞に対して、CSM0.2, Ovp0.3, Freq0.2 の三手法の全て、あるいは、三手法のうち少なくともいずれか二手法で共通して生成された階層(以下、COMMON と呼ぶ)と、EDR の階層のペア。(COMMON-EDR)

表 5 COMMON-EDR で対象となる形容詞 (30 語)

ID	形容詞	ID	形容詞	ID	形容詞	ID	形容詞	ID	形容詞
1	ユニークな	7	早い	13	ソフトな	19	危うい	25	自由な
2	温かい	8	速い	14	デリケートな	20	危険な	26	若い
3	急な	9	遅い	15	ナイーブな	21	気楽な	27	十分な
4	古い	10	面倒な	16	鋭い	22	肯定的な	28	柔らかい
5	親しい	11	優しい	17	甘美な	23	高い	29	重要な
6	絶望的な	12	きつい	18	頑固な	24	高貴な	30	独創的な

表 6 CSM-EDR で対象となる形容詞 (10 語)

ID	形容詞	ID	形容詞
1	おびただしい	6	孤独な
2	ノスタルジックな	7	誠実な
3	パワフルな	8	理論的な
4	開放的な	9	有害な
5	狭い	10	不幸な

表 7 Ovp-EDR で対象となる形容詞 (10 語)

ID	形容詞	ID	形容詞
1	いぶかしげな	6	粘り強い
2	スリリングな	7	悲壮な
3	一途な	8	不安な
4	果敢な	9	不公平な
5	つらい	10	不敵な

2) 10 語の形容詞に対して, CSM0.2 だけで生成された階層と EDR の階層のペア. (CSM-EDR)

3) 10 語の形容詞に対して, Ovp0.3 だけで生成された階層と EDR の階層のペア. (Ovp-EDR)
 Freq0.2 の階層は CSM0.2 の階層と殆ど同じであったので, 本実験では, Freq0.2 の手法だけで生成された階層は対象外にし, CSM0.2 の階層のみをとりあげた.

1)(COMMON-EDR) で対象となる 30 語の形容詞のリスト, 2)(CSM0.2-EDR) で対象となる 10 語の形容詞のリスト, 3)(Ovp0.3-EDR) で対象となる 10 語の形容詞のリストをそれぞれ表 5, 表 6, 表 7 に示す.

5.2 被験者

言語学者, 辞書編集者, 自然言語処理に関係する人, 合計 20 人が被験者として参加した.

5.3 実験方法

各々の被験者には, 各々の形容詞に対する自動構築の階層と EDR の階層のペアが各行に記された回答用紙が示された. この例を図 3 に示す. 各行の左側には自動構築による階層 (図中, A-B-C), 右に EDR 階層 (図中, P-Q-R) が記される. ただし左が自動構築の階層で, 右が EDR の階層であることは, 被験者には知らせていない. 被験者は, 各行に記された 2 つの階層の妥当性を比較し, “左が妥当” / “どちらかといえば左が妥当” / “どちらともいえない” / “どちらかといえば右が妥当” / “右が妥当” のいずれかを回答するよう求められた.

図 3 被験者に提示した刺激と回答フォーム

階層の判定基準として、被験者には概念の階層関係のルールを示した。これは第 4 節で述べた概念と事例の関係と同様の考え方である。

[被験者に示したルール]

- 1) 事例の成員としての連続性：もし概念が階層関係であるならば、最下位概念の成員となる事例（本実験では形容詞）は最上位概念まで一貫して成員として出現している。
- 2) 概念の継承性：もし概念が階層関係であるならば、下位概念は上位概念から特徴を「継承」しており、下位概念は一方向的に上位概念の特徴を含意している。

ルール 2 は、以下の考察に基づいて設定した。即ち Cruse では次の様に述べている (Cruse 1986)。もし、「A は f(X)」であれば、「A は f(Y)」が成り立つが、その逆の、「A は f(Y)」であっても、「A は f(X)」が必ずしも成り立たない場合には、X が Y の下位語で、Y は X の上位語であるといえる。たとえば、「太郎は犬です」が成り立てば必然的に「太郎は動物です」が成り立つが、「太郎は動物です」が成り立つ場合に、必ずしも「太郎は犬です」が成立するとは限らない。上位から下位へ概念が並んでいるためには、上位から下位への一方向的な概念継承がみられる必要がある。

ルール 1 および 2 によって、下位から上位への事例（共起形容詞）の出現状況と、上位から下位への概念の継承性を考慮して、階層の妥当性を判定してもらったのである。階層の妥当性に関し、Scheffe 法によって数値化された値に基づき、自動生成で得られた階層と EDR の階層の間の平均値の有意差を、下記に示す検定量 T を用いて検定した。

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N} + \frac{s_2^2}{N}}} \quad (8)$$

N は対象とする形容詞の数である、とは Method 1（自動構築）で得られた階層に対して各被験者について得られた数値の全被験者（20 人）にわたる平均と不偏分散である。とは Method 2 (EDR) で得られた数値の全被験者にわたる平均と不偏分散である。

5.4 実験結果

図4に, $\bar{x}_1$ (“◆”で示す), $\bar{x}_2$ (“■”で示す) を標準誤差とともに示す. 図中の横軸は形容詞 ID, 縦軸は計量化した値を表す. 縦軸の値が大きいほど, “より妥当である”ことを意味する. エラーバーはそれぞれの標準誤差を示す. エラーバーが長いほど被験者間のばらつきが大きいことを意味する.

5.4.1 30 形容詞に対する COMMON-EDR 階層ペア

[検定結果]

有意に **COMMON > EDR**

形容詞 ID: 22, 26 (30 個中 2 個)

有意に **COMMON < EDR**

形容詞 ID: 1, 2, 4, 7, 8, 9, 11, 16, 19, 20, 21, 23, 24, 25, 27, 29, 30 (30 個中 17 個)

両 Method 間に有意差なし (ただし **COMMON > EDR**)

形容詞 ID: 3, 13, 14, 18, 28 (30 個中 5 個)

両 Method 間に有意差なし (ただし **COMMON < EDR**)

形容詞 ID: 5, 6, 10, 12, 15, 17 (30 個中 6 個)

[COMMON-EDR 階層比較の考察]

本提案手法で自動生成 (COMMON) された階層が EDR 階層より有意に良かった形容詞は 2 個であった. 一方, 自動生成の階層よりも EDR 階層の方が有意に良かった形容詞は 17 個であっ

図4 COMMON と EDR の階層の全被験者にわたる平均値と標準誤差

た。また、両者の間に有意差がなかった形容詞は 11 個であった。例として形容詞 ID を挙げるが、その形容詞 ID に対する階層については付録に載せる。

本提案手法で得られた階層の方が EDR 階層より良かった形容詞は、例えば、ID-22 の「肯定的な」に対する階層である。

本提案手法で得られた階層より EDR 階層の方が優位に良かった形容詞は、例えば、ID-7 の「早い」に対する階層である。

両方に有意差がなかった形容詞は たとえば、ID-17 の「甘美な」に対する階層である。

以上から、COMMON に関しては、本研究で提案した手法は、43% (30 個中 13 個) の形容詞に関し EDR と同程度あるいは EDR より人間の直感にあう階層が生成できたことが示唆された。

5.4.2 10 形容詞に対する CSM-EDR 階層比較

[検定結果]

有意に $CSM > EDR$

形容詞 ID: 無し (10 個中 0 個)

有意に $CSM < EDR$

形容詞 ID: 1, 5, 6, 9, 10 (10 個中 5 個)

両 Method 間に有意差なし (ただし $CSM > EDR$)

形容詞 ID: 3, 4 (10 個中 2 個)

両 Method 間に有意差なし (ただし $CSM < EDR$)

形容詞 ID: 2, 7, 8 (10 個中 3 個)

図 5 CSM と EDR の階層の全被験者にわたる平均値と標準誤差

[CSM-EDR 階層比較の考察]

CSM でのみ作られた階層では, EDR 階層の方が有意によいと判定された形容詞は 5 個 (10 個中 50%) であり, CSM の階層の方が有意によいと判定された形容詞は 0 個であった. また, CSM 階層と EDR 階層で有意差がないと判定された形容詞は 5 個 (10 個中 50%) であった.

EDR 階層の方が有意によいと判定された階層は, ID-1, 5, 6, 9, 10 の形容詞であったが, 特に, ID-5, 10 は, 明確な有意差がみられる. たとえば, ID-5 の「狭い」の階層を付録に示す.

また, 両手法間で有意差がなかった形容詞のうち, ID-4「開放的な」を付録に示す.

以上から, CSM に関しては, 本研究で提案した手法は, 50% (10 個中 5 個) の形容詞に関し EDR と有意差がない階層を生成したことが示唆された.

5.4.3 10 形容詞に対する Ovlp-EDR 階層比較

[検定結果]

有意に $Ovlp > EDR$

形容詞 ID: 無し (10 個中 0 個)

有意に $Ovlp < EDR$

形容詞 ID: 1, 2, 3, 4, 5, 6, 7, 8, 10 (10 個中 9 個)

両 Method 間に有意差なし (ただし $Ovlp > EDR$)

形容詞 ID: 無し (10 個中 0 個)

両 Method 間に有意差なし (ただし $Ovlp < EDR$)

図 6 Ovlp と EDR の階層の全被験者にわたる平均値と標準誤差

形容詞 ID: 9 (10 個中 1 個)

[Ovlp-EDR 階層比較の考察]

Ovlp-EDR の階層比較では、EDR 階層の方が有意によりと判定された形容詞が 10 個中 9 個あった。それに対し、Ovlp の階層と EDR の階層とで有意差がないと判定された形容詞は、ID-9 のわずか 1 個だけであった。

Ovlp の階層より EDR の階層の方があきらかに有意により階層としては、たとえば、ID-4「果敢な」に対する階層がある。階層例は付録に付す。

両手法間で有意差のない階層（一つのみ）は ID-9「不公平な」に対する階層である。階層例は付録に付す。

以上から、Ovlp に関しては、本研究で提案した手法は、10% (10 個中 1 個) の形容詞に関し EDR と同程度の階層を生成したことが示唆された。

5.5 考察

本実験では、概念階層関係が成立するための重要な条件である「事例の成員としての連続性」と「概念の継承性」の点から、提案手法の階層と EDR 階層の優劣の程度に関して、人間の判断を数値化することにより、統計的に検定した。その結果、EDR の階層の方が、「事例の成員としての連続性」と「概念の継承性」の観点で、COMMON, CSM, Ovlp, のどの手法による階層よりも、概ね良いことがわかったが、COMMON と CSM の階層においては、半分ほどは EDR の階層と同程度くらいであった。特に Ovlp の評価は、全体の 90% ほど EDR の方が有意に良いという判定となり、かなり悪かった。

しかし、表 1 (第 4 節) で表層的な特徴からみると、CSM 手法による階層と比較して悪い結果ではなかった。ここで実験した自動構築手法の階層に共通して言えることであるが、特に Ovlp で顕著にみられた原因があった。

では、その原因は何か。それを探るため、被験者のコメントを参考に分析した。被験者は概して最下位概念と最上位概念との間に中間的な概念がないと、妥当性が低いと判断した。つまり、ルールで提示した継承性が失われていると判定した。表 1 の「階層の長さ」をみるとわかるように、Ovlp はノードの深さが 9 と短い。逆に CSM は、ノードの深さが 13 あり、これはほぼ EDR のノードの深さと同じくらいである。短い階層を作る Ovlp は、中間的なノードが作られていないと判断されたのである。CSM 階層を EDR 階層と比較する場合にも、多かれ少なかれ、この判断が反映されていた。

上記の概念の継承性の問題で考えられるのは、概念数の不足があげられる。第 4 節の表 1 で示したように、CSM0.2 は階層を構成している抽象名詞 (本稿では概念に相当) は、365 語中 325 語で約 90%、Ovlp0.3 は、365 語中 317 語で 86% となっており、閾値設定によって、抽象名詞が大幅に削られているわけでない。では、そもそも 365 語の名詞が不足なのかどうかについてで

あるが、365 語の名詞は、新聞 2 年分から第 3 節で述べたパターン（「X という Y」など）を使い特定の意味関係（「具体事例とそれが属する概念」という意味関係）を抽出したものであり、データスパースネスの問題も考えられる。また、一つの目安として EDR の形容詞の概念数を調べると約 2000 強ほどある。365 概念ではまだ不足していると考えられる。今後さらに抽出方法も含めデータを精査していきたい。

また、我々の手法で取り出した 365 語の抽象名詞は、質的に、形容詞の概念のラベルとして適当かどうかという問題もある。この問題については、試験的に、我々が抽出した抽象的な名詞（本稿で概念と呼んでいるもの）とその形容詞集合の組をランダムサンプリングして 18 組用意し、5 人の被験者に、同じ形容詞集合に対する EDR の概念と比較評価してもらった。この小規模な実験では、提示した形容詞集合に対して EDR の概念より我々が抽出した概念の方が適当という回答が多くみられ（全 90 問（18 組×5 人）中 57 問、63%）、18 組での小規模な実験の範囲ではよい感触を得ているが、さらに実験方法とその結果も含めさらに精査していきたいと考えている。その他、人手でデータを取捨選択した際に、「具体事例とそれが属する概念」という関係以外の関係も含まれている可能性もあり、データを見直し対象外のものを除く必要があろう。

次に、継承性が低いと評価された階層について考察すると、形容詞の多義性と事例数の関係が問題として考えられる。EDR に比べ明確に有意に悪い形容詞は、COMMON の ID-1「ユニークな」ID-4「古い」、ID-7「早い」、ID-9「遅い」、ID-16「鋭い」ID-23「高い」である。どれも基本的な抽象度の高い形容詞である。提案手法での自動階層では、多義的な語についてうまく階層を構築できていない可能性がある。

原因として考えられるのは、一つはそもそも多義の場合の概念をすべてコーパスから抽出しきれていない可能性である。もう一つは、階層構築で考えられる問題点である。ある概念の事例集合に多義的な形容詞が 1 語しかない、つまり他に手がかりとなる事例がない場合、概念間の包含関係の質的な判断を必要とする。たとえば、[背丈] にも [金額] にも「高い」「低い」だけが事例として出現している場合、[背丈] と [金額] の「高い」を区別する必要がある。しかし、階層をみると、現状では、最上位レベルの概念と直接結合したり、多義の別の概念と結びついたり（たとえば、「高い」に対して [背丈] と [金額] が結びつく）、概念の継承性の評価が低いようである。今後の課題として、各概念の事例としての形容詞の網羅性と、事例が少ない概念間の関係（特に事例が多義的な場合の概念間の関係）についての対応について考える必要がある。

自動抽出した語彙の評価方法には、他にも語義のあいまい性解消、情報検索などへの応用によって評価を行うことも考えられる。それらの多角的な観点も含め、概念階層の評価方法についても今後、さらに検討していきたいと考える。

6 今後の展望—類義関係と階層関係をとらえるために—

第3節で取り出した言語データから形容詞概念を体系的にとらえるためには、階層関係と同時に類義関係も考慮する必要がある。これまでに、形容詞と抽象的な名詞を類義関係によって自動分類する研究として、Kohonen の自己組織化神経回路網モデル (Self-Organizing map, SOM) を用いた研究がある (Ritter and Kohonen 1989; 馬他 2001)。Kohonen の自己組織化神経回路網モデルは、高次元入力をもつ2次元配列のノード (ニューロン) で構成され、自己組織化によって高次元データを2次元空間に、その特徴を反映するように、非線形的に射影する。特徴は、得られる分布が可視的でまた連続的であるということである。特に、連続的な分布という性質は、言葉の意味の連続性を反映できると考えられ、馬らの研究に用いられている。出力される2次元平面は、意味的に類似性の高い名詞どうしが近くに配置され、意味的に類似性の低い名詞どうしが遠くに配置されるような、意味的類似性を距離とする2次元表現である。馬らは出力される2次元平面をマップと呼んでいる。馬らの研究では、SOM の分類能力を多変量解析や階層型クラスタリング手法等と比較実験し、その結果、他手法より劣らない、あるいは彼らのタスクに関してはむしろよい結果を得たと報告している。

しかし、提案された SOM による分類では、類義関係以外の意味関係は求めることができなかった。Kanzaki et al. では、上位下位関係 (包含関係) を SOM に導入する実験を行っている (Kanzaki, Ma, Yamamoto, and Isahara 2004)。この実験で SOM の入力データとなる単語の特徴ベクトルは、当該単語に対する全対象単語との包含関係値を多次元ベクトルにしたものである。つまり、この特徴ベクトルを SOM の入力とすることにより、類似した包含関係をもつ単語どうしが類義関係をもつことになる。SOM は通常、類義関係を可視化するので、SOM に包含関係を導入した場合、出力されるマップは、上位下位関係が反映された方向性ある分布になるか、あるいは方向性の全くみられない分布になるかの可能性が考えられる。包含関係を導入した SOM の分布は、評価方法が提案されていないためさらに検討する必要があるものの、抽象度が高い名詞から抽象度の低い名詞へと方向性ある分布を示唆したものとなった。

SOM によるマップに、概念の階層関係が反映されれば、同じような上位下位関係をもつ概念が類義関係となって近くに分布することになり、階層関係と類義関係を同時に反映した分類をえられる可能性がある。結果は、既存のシソーラスなどと比較して差異を考察すると共に、形容詞側から、得られた概念体系に従って形容詞表現の様相をとらえていきたい。

7 まとめ

本研究では、将来的に実データに基づいた形容詞の観点からみた上位下位関係と類義関係をとらえるため、その一環として形容詞の概念の階層関係に着目し、コーパスから取得した概念

から階層を構築する方法と, 妥当そうな階層を得るための評価について述べたものである.

階層の評価は, 事例(形容詞)の各概念の成員としての連続性と概念の継承関係という観点から評価し, まず, 収集した形容詞の出現状況や構築された概念の表層的な面からと, 心理実験で EDR との比較を行う質的な面から評価した.

表層的な観点からの評価で, CSM, Ovlp, Freq の三つの手法と複数の閾値を組み合わせた階層の中で, CSM の手法で閾値 0.2 以上を使った階層と, Ovlp の手法で閾値 0.3 以上を使った階層を外見的に妥当そうな階層として絞った. そして次のステップの心理実験で, EDR 概念階層との比較評価の対象として採用した.

心理実験による EDR 概念階層との比較評価では, 「概念の継承性」と「事例の成員としての連続性」という観点で, 自動構築と EDR とで概念階層の優劣を判定してもらい統計的に数値化した. その結果上記二つの観点からの評価では, EDR の階層の方が, COMMON (Ovlp, CSM, Freq 共通), CSM, Ovlp のどの手法による階層よりも, 概ね, 有意に良いことがわかったが, COMMON と CSM の階層においては, 半分ほどは EDR の階層と同程度くらいであった. 自動生成の階層が既存辞書の階層に対して, その結果の半分弱の階層で問題を提起するという意味で, ベースラインとなる数値と考える.

自動階層の評価が低かった主な要因は, 最下位概念と最上位概念との間に中間的な概念がないと, 妥当性が低いと判断されたことであった. 原因としては, 概念数の不足, 形容詞の概念ラベルの妥当性, 形容詞の多義性などが考えられる. これらの点をさらに精査し, 今後, SOM による概念分布に対応させ, 類義関係とあわせタクソノミー的な構造として, 形容詞の概念体系をとらえていきたい.

オントロジーは必ずしも木構造ではないかもしれないが, 階層関係と類義関係については, それを計算する尺度が利用できることもあり, 最初にこれらの関係を明らかにしていきたいと考える.

参考文献

- Alonge, A., Bertagna, F., Calzolari, N., Roventini, A., and Zampolli, A. (2000). "Encoding information on adjectives in a lexical-semantic net for computational applications." In *the 1st Conference of the North American Chapter of the Association for Computational Linguistics (ACL)*.
- 安西祐一郎 (1989). 岩波講座ソフトウェア科学 認識と学習, 第 16 巻. 岩波書店.
- Berland, M. and Charniak, E. (2000). "Finding Parts in Very Large Corpora." In *38th Annual Meeting of the Association for Computational Linguistics (ACL)*.

- Biemann, C. (2005). "Ontology Learning from Text: A Survey of Methods." *LDV-Forum*, **20** (2), pp. 75-93.
- Caraballo, A. S. (1999). "Automatic Acquisition of a Hypernym-Labeled Noun Hierarchy from Text." In *37th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge University Press.
- Hagita, N. and Sawaki, M. (1995). "Robust Recognition of Degraded Machine-Printed Characters using Complimentary Similarity Measure and Error-Correction Learning." In *the SPIE—The International Society for Optical Engineering*, 2442.
- Hearst, A. M. (1992). "Automatic Acquisition of Hyponyms from Large Text Corpora." In *the 14th International Conference on Computational Linguistics*.
- Hindle, D. (1990). "Noun Classification From Predicate-Argument Structures." In *the 28th Annual Meeting of the Association for Computational Linguistics (ACL)*.
- Kanzaki, K., Ma, Q., Yamamoto, E., and Isahara, H. (2004). "Construction of an Objective Hierarchy of Abstract Concepts via Directional Similarity." In *21st International Conference on Computational Linguistics (COLING)*.
- 神崎享子 (1997). "連体修飾関係を結ぶ形容詞類と名詞." 計量国語学, **21** (2), pp. 53-68.
- 神崎享子, 井佐原均 (1999). "形容詞類の連体用法にみられる連用的な意味." 計量国語学, **22** (2), pp. 51-65.
- 国立国語研究所 (1964). 分類語彙表. 秀英出版.
- Lin, D. and Pantel, P. (2002). "Concept Discovery from Text." In *the 19th International Conference on Computational Linguistics (COLING)*.
- 馬青, 神崎享子, 村田真樹, 内元清貴, 井佐原均 (2001). "日本語名詞の意味マップの自己組織化." 情報処理学会論文誌, **42** (10), pp. 2379-2391.
- Manning, C. D. and Shütze, H. (1999). *Foundations of Statistical Natural language Processing*. The MIT Press. pp. 298-303.
- 丸元聡子, 白土保, 井佐原均 (2005). "動詞待遇表現に対する丁寧さの印象に関する定量的分析—接頭辞オを用いた表現と接頭辞ゴを用いた表現との比較—." 自然言語処理, **12** (5), pp. 71-90.
- 益岡隆志 (1994). 名詞修飾節の接続形式—内容節を中心に—. 田窪行則編『日本語の名詞修飾表現』. くろしお出版. pp. 5-27.
- Medin, D. L. and Goldstone, R. L. (1990). *Concept in Eysenck, M. W. Ed., The Blackwell Dictionary of Cognitive psychology*, Cambridge, MA: Basic Blackwell Inc.
- 根本今朝男 (1969). "「が格」の名詞と形容詞とのくみあわせ." 国立国語研究所報告書, 電子計算機のための国語研究 II.

- Pantel, P. and Ravichandran, D. (2004). “Automatically Labeling Semantic Classes.” In *Human Language Technology/North American Association for Computational Linguistics (HLT/NAACL)*.
- Riloff, E. and Shepherd, J. (1997). “A corpus based approach for building semantic lexicons.” In *the Second Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Ritter, H. and Kohonen, T. (1989). “Self-Organizing Semantic Maps.” *Biological Cybernetics*, **61**, pp. 241–254.
- Sawaki, M., Hagiga, N., and Ishii, K. (1997). “Robust Character Recognition of Gray-Scaled Images with Graphical Designs and Noise.” In *the International Conference on Document Analysis and Recognition. IEEE. Computer Society*.
- Scheffe, H. (1952). “An analysis of variance for paired comparison.” *Journal of the American Statistical Association*, **47**, pp. 381–400.
- 高橋太郎 (1975). “文中にあらわれる所属関係の種々相.” 国語学, **103**, pp. 1–16.
- 寺村秀夫 (1991). 日本語のシンタクスと意味 III. くろしお出版.
- Thelen, M. and Riloff, E. (2002). “A Bootstrapping Method for Learning Semantic Lexicons using Extraction Pattern Contexts.” In *the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Yamamoto, E., Kanzaki, K., and Isahara, H. (2005). “Extraction of Hierarchies Based on Inclusion of Co-occurring Words with Frequency Information.” In *19th International Joint Conference on Artificial Intelligence (IJCAI)*.
- 山本英子, 神崎享子, 井佐原均 (2006). “出現状況の包含関係による語彙の階層構造の構築.” 情報処理学会論文誌, **47** (6), pp. 1872–1883.
- 山本英子, 梅村恭司 (2002). “コーパス中の一対多関係を推定する問題における類似尺度.” 自然言語処理, **9** (2), pp. 46–75.

付録

- 5.4.1. 30 形容詞に対する COMMON-EDR 階層ペア
 - 5.4.2. 10 形容詞に対する CSM-EDR 階層比較
 - 5.4.3. 10 形容詞に対する Ovp-EDR 階層比較
- 上記 3 節に挙げた形容詞の階層を以下に示す.

A 5.4.1. 30 形容詞に対する COMMON-EDR 階層ペア

[自動階層の方が EDR 階層より良いと判定された形容詞の階層]

ID-22 肯定的な

自動階層：

こと→面 1→傾向→見方→評価

EDR 階層：

概念 (3aa966) → 事象 (30f7e4) → 行為 (30f83e) → 対象行為 (444dd) → ものを対象とする行為 (444dd9) → 情報の移動 (30f832) → 情報の受信 (3f96e7) → 知る (30f876) → 認知主体と認知対象との認知的距離減少 (3f972c)

[EDR 階層の方が有意に良いと判定された形容詞の階層]

ID-7 早い

自動階層：

こと→時→速度→時刻

EDR 階層：

概念 (3aa966) → 時 (30f776) → 時間点 (3f9882) → 基準時点の前後の時 (444da3) → ある時間点よりも前の時点 (444cc3) → 基準になる時刻や時期より前であること (3cf6d9)

[両方に有意差がなかった形容詞の階層]

ID-17 甘美な

自動階層：

こと→面 1→イメージ→感覚→雰囲気→空気→気配→魅惑

EDR 階層：

概念 (3aa966) → 事象 (30f7e4) → 状態 (3aa963) → 性状・性向 (3f9871) → 性状 (444e60) → 人のありさま (444db3) → 人の気持ち (30f7b7) → 人の気持ちの様子 (3f98bd) → 自分自身の感情や心境 (444ded) → 愉快・不愉快 (30f961) → 良い気持 (0e89db) → 心地よい (3cf426)

B 5.4.2. 10 形容詞に対する CSM-EDR 階層比較

[EDR 階層の方が有意によいと判定された形容詞の階層]

ID-5 狭い

自動階層：

こと→方→空間→面積

EDR 階層：

概念 (3aa966) → 事象 (30f7e4) → 状態 (3aa963) → 性状・性向 (3f9871) → 事物の属性 (444db1)

→具体物の属性 (3aa962) →具体物の空間的属性 (444ce7) →形の値 (444ce8) →形の広狭 (3f9876) →空間的に小さいさま (1e88a8)

[両手法間で有意差がなかった形容詞の階層]

ID-4 開放的な

自動階層：

こと→面 1 →イメージ→印象→態度→人柄→気質→気風→気性

EDR 階層：

概念 (3aa966) →*事象 (30f7e4) →状態 (3aa963) →性状・性向 (3f9871) →事物の属性 (444db1) →人間の属性 (3aa961) →人の性格や態度 (3f98c9) →態度や性格の値 (444ce3) →善良 (30f96d) →かくしだてをせず，あけっぴろげであること (3cf4b7) →かくしだてをすることなく，あけっぴろげであるさま (3ce57e)

C 5.4.3. 10 形容詞に対する Ovlp-EDR 階層比較

[EDR 階層の方が有意によいと判定された形容詞の階層]

ID-4 果敢な

自動階層：

こと→性格→勇気

EDR 階層：

概念 (3aa966) →事象 (30f7e4) →状態 (3aa963) →性状・性向 (3f9871) →事物の属性 (444db1) →人間の属性 (3aa961) →人の性格や態度 (3f98c9) →態度や性格の値 (444ce3) →豪胆・臆病 (30f96f) →勇敢 (30f96b) →思い切ったことをすること (3cefc5)

[両手法間で有意差がなかった形容詞の階層]

ID-9 不公平な

自動階層：

こと→関係→差別

EDR 階層：

概念 (3aa966) →事象 (30f7e4) →状態 (3aa963) →性状・性向 (3f9871) →さまざまな属性 (444d17) →いろいろな抽象物の属性 (444d27) →かたよりがあって平等でないようす (3cfcd6)

略歴

神崎 享子：1994 年早稲田大学大学院文学研究科日本語日本文化専攻修士課程修了。1998 年同大学院文学研究科日本語日本文化専攻博士課程単位取得後満

期退学。2001年 神戸大学大学院自然科学研究科修了。博士（学術）。1998年～2005年郵政省通信総合研究所。2005年～独立行政法人情報通信研究機構。現在、独立行政法人情報通信研究機構主任研究員。自然言語処理、言語学の研究に従事。言語処理学会、計量国語学会、日本言語学会、日本語学会、会員。

馬 青：1983年北京航空航天大学自動制御学科卒業。1987年筑波大学大学院修士課程理工学研究科修了。1990年同大学院博士課程工学研究科修了工学博士。1990-1993年株式会社小野測器勤務。1993年郵政省通信総合研究所入所。1994年同所主任研究官。2003年龍谷大学理工学部教授。自然言語処理の研究に従事。言語処理学会、情報処理学会、電子情報通信学会、日本神経回路学会、各会員

山本 英子：1998年豊橋技術科学大学大学院工学研究科情報工学専攻修士課程修了。2002年同大学大学院工学研究科電子・情報工学専攻博士後期課程修了。博士（工学）。2002年-2007年10月独立行政法人情報通信研究機構専攻研究員。現在、神戸大学工学研究科プロジェクト奨励研究員。自然言語処理の研究に従事。言語処理学会、情報処理学会各会員。

白土 保：1999年電気通信大学博士（工学）。情報通信研究機構知識創成コミュニケーション研究センター推進室研究マネージャ。専門分野は、言語心理、音楽音響、感性情報処理。言語処理学会、電子情報通信学会、日本音響学会各会員。

井佐原 均：1980年京都大学大学院工学研究科電気工学専攻修士課程修了。博士（工学）。1980年通商産業省工業技術院電子技術総合研究所入所。1995年郵政省通信総合研究所知的機能研究室長。現在、独立行政法人情報通信研究機構上席研究員。自然言語処理の研究に従事。言語処理学会、情報処理学会、日本認知科学会、人工知能学会各会員。

（2007年10月10日 受付）

（2008年1月25日 再受付）

（2008年3月18日 採録）