

自然言語および図形理解のための形容動詞の概念の分析 —2字漢語

岡田直之

(大分大学工学部)

濱島真人⁺

1. まえがき

計算機による自然言語および図形の理解を目的として、形容詞や形容動詞で表される属性概念の分類、分析を進めている^{(1)~(4)}。本稿は、2字漢語の形容動詞の概念について述べている。

知識ベース等の作成において語の概念を捉えるとき、二つの観点がある。一つは語の指し示す外界あるいは内界の対象を参照しながら、個々の概念の成分的特徴ないしは内部構造に立ち入り、分析的に表現するものである。もう一つは概念の用いられ方を参照しながら、ある概念か他の概念とのような関係で結びつくかを調べ、関係的に表現するものである。具体的には前者の場合性質リスト等で、また後者の場合ネットワーク等で表現される。

分析的表現は図形等非言語データを参考する意味処理や異なった言語間の意味処理に好都合であるか、知識ベース作成の際概念の成分的特徴の抽出や内部構造の把握には、鋭い洞察力と豊富な経験を要する。それに対し関係的表現はテキストに現れる語と語の関係を調査する等により、知識ベースの作成作業は分析的表現より機械的に行えるか、得られた概念データが特定の言語に依存してしまう可能性がある。実際の機械処理では場合に応じていずれの観点のデータも必要とされようか、本稿では、図形をも背景としていることから、分析的表現を意図している。

形容動詞のクラスは多くの漢語を含み、ヒリカケ2文字から成る漢語が多い。一般に漢語は和語と比べると概念の把握が難しい。国語学では從来から

⁺ 理在鹿児島ソフトウェア(株)

種々漢語の調査を行つてゐる。森岡は現代漢語の成立を調べ、日本語としての定着の度合いから漢語形態素を分類した⁽⁵⁾。この分類は2字漢語の概念を分析的に表現する上で貴重である。それに対し野村と田中らは3字漢語と4字漢語について、それらから7字漢語あるいは2字漢語からどのように構成されるかに注目し、文法的なうらひに統計的性質を明らかにした^{(6),(7)}。そこでは2字漢語の概念を要素として扱つており、関係的表現を求めるのに役立つ。

本稿では、まず、2字漢語で表される形容動詞の概念を森岡の分類を参考にして三つのクラスに分ける。次に漢語形態素との結びつきの強い二つのクラスについて個々の概念を分析するためのアルゴリズムを提案し、最後にそのアルゴリズムに従つて“複合概念A”と呼ばれるクラスを実際に分析する。このアルゴリズムは形容動詞だけではなく他の品詞の2字漢語に対しても応用が可能である。また得られた概念データは複合概念Aのクラスの定性的かつ定量的性質を明らかにするもので、知識ベースを作成するための基礎資料として有効である。

なお、本稿では直接图形とのつなわりには触れないが、それについては文献⁽³⁾を参照されたい。

2. 2字漢語の形容動詞 とその概念

2.1 2字漢語の形容動詞 まず初めに、例を示そう。

G =

{不快、過大、有限、動的、駿然、
軟弱、豊富、平靜、優秀、重要、

強力，小心，異様，稀代，最適，
巨大，單純，安直，敏感，開闊，
瀟洒，伶俐，慾慾，吝嗇，飄輕，
頑丈，器用，立派，無茶，億却}

注：簡単のため形容動詞の語幹のみ示している。

これららの表す概念を分析する準備として、森岡によつて提案された漢字母即ち造語成分としての漢語形態素の分類を参照しよう。森岡は漢字母を7種類に分類したが、特に関係の深い二つのクラスは以下のように解説できる。

1. 接辞的に用いられる單純結合形式

例：不自然，一般化，科学的

接辞的に接觸して派生語を作る。特に明治以後登場し、造語機能が著しく、現代漢語の一特色をなしています。

2. 音訓の転換のきく單純結合形式

例：山川，花鳥

漢字か意音素の役割を荷ない音と訓がアロモルフ (allomorph 実形態) の関係にあると認められる。ここでアロモルフとは、同一の形態素が文脈によって違った形態をとる現象で、アメ(雨)→アミ(雨), イ(位)→サンシ, イソ(礁)→アリ(糸)のような形態音素的な事実を指す。やまと(山)やかわ(川)などの訓、つまり和語で考え、それを音に転するニレによつて、さんせん(山川)をひの漢語を創造したり理解せしEJであります。現代漢語の上台となっている形態素が圧倒的に多い。以上を参考にして、2字漢語の形容動詞の概念を三つのクラスに分けよう。

2.2 概念のクラス

初めに記号を定義する。

W_i : 2字漢語の形容動詞

w_{ij} : W_i の第j漢字母 ($j = 1, 2$)。音読み。

w_{ij} : w_{ij} に対応する和語・訓読み。

C_i : W_i の概念

C_{ij} : w_{ij} の概念

C'_{ij} : w'_{ij} の概念

R_a : w_{ij} と w'_{ij} の間のアロモルフ的関係+

R_e : C_{i1} と C_{i2} の間の論理的関係

R_s : C_{i1} と C_{i2} の間の構文的関係

1) 派生概念のクラス G_d

W_i が w_{i1} / w_{i2} から派生した語で、 C_{i2} / C_{i1} が C_{i1} / C_{i2} に様相的な意

+ 森岡の示した音・訓の間の“アロモルフ的関係”を、本稿では、一步後退して“アロモルフ的関係”と呼ぶ。

味を添えている。

例：{不快(快くない), 過大, 有限, 動的(動いているような), 驚然} $\subset G_d$

2) 複合概念Aのクラス G_a

三つのサブクラスがある。

G_{a1} : W_i において $w_{i1} Ra w_{i1}$ かつ $w_{i2} Ra w_{i2}$ (即ち $C_{i1} = C_{i1}$ かつ $C_{i2} = C_{i2}$) で, C_i において $C_{i1} Re C_{i2}$ である。

例：{軟弱(軟らかくかつ弱い), 豊富, 平靜, 優秀, 重要} $\subset G_{a1}$

G_{a2} : W_i において $w_{i1} Ra w_{i1}$ かつ $w_{i2} Ra w_{i2}$ で, C_i において $C_{i1} Re C_{i2}$ である。

例：{強力(力が強い), 小心, 異様, 稀代, 最適(最も適する)} $\subset G_{a2}$

G_{a3} : W_i において w_{i1} ($j = 1$ または 2) に対し w_{ij} が存在しないか, w_{ij} が存在してもまじみか薄いか, または $C_{ij} \neq C'_{ij}$ か, のいずれかである。しかし C_{ij} は基礎的でわかり易く, $C_i = C_{i1} Re C_{i2}$ または $C_{i1} Rs C_{i2}$ である。

例：{巨大(巨大で大きい), 単純, 安直(直^{まっ}か安^{いた}い), 敏感, 開闊(開^{ひらく}ひままである⁺⁺)} $\subset G_{a3}$

3) その他のクラス G_r

二つのサブクラスがある。

G_{r1} : G_d または G_a のような分析が可能かもしれないか, w_{ij} ($j = 1$ または 2) が難しい漢字母であるため, C_{ij} が理解しにくい。

例：{瀟洒, 伶俐(さとくかっこいい), 慾慾, 吝嗇, 飄輕} $\subset G_{r1}$

G_{r2} : C_{ij} ($j = 1$ または 2) が故事に因るような場合を除いて, C_i と無関係である。

例：{頑丈, 器用, 立派, 無茶, 億却} $\subset G_{r2}$

以上三つのクラスの中で, G_r は概

+ 「巨^キ」は通常漢字で「値」または「価」を用いる。

+ 「散^ス」は通常“ひまである”的意味で用いない。

念 C_i を漢字字母 w_{ij} と関連づけて捉えるのが困難を、もしくは捉えるべきでないクラスと考える。

3. 2字漢語の概念の分析方法

ある W_i が与えられたとき、 $C_i \in G_d$ または $C_i \in G_a$ かどうかを識別しつゝ、 C_i の特徴を抽出する方法を示そう。

3.1 分析例

初めに分析上の問題点を述べておこう。特に注意しなければならないのは次の2点である。

1) C_{ij} または w_{ij} の難易度

G_{a3} に關し C_{ij} が基礎的でわかり易いか、また G_{r1} に關し w_{ij} が難しいか、には個人差がある。どこで線を引けばよいのか？

2) w_{ij} には一般に多義がある。ある $W_i = w_{i1} w_{i2}$ において、 w_{i1} 、 w_{i2} のそれぞれいずれの多義に注目すればよいのか？

1) に關し岩波国語辞典は、現代生活に必要な約2300の漢字字母を項目として設け、詳しい概念記述を行っていいる⁽⁸⁾。そこでそれらの漢字字母は、他と比べ、難しくなく、かつ概念も基礎的でわかり易いものと見なす。また2)についてもそれらの項目に示されている用例を参照し、当該の概念を見出す手がかりとする。

次に「傲慢だ」を例に取り、同辞典の関連する項目を示す。

あなどる【侮る】相手を軽く見下すには及に及ぶ。みくびる。

おこたる【怠る】①おこける。いいかげんにする。②

ゆだんする。③病気かくசくなる。

おごる【驕る・傲る・奢る】①自分の才能・実力・地位・財産等に得意になり、たかがる。才覚・實力等を主のみとして勝手をふるむをする。②身分以上にせいたくな生活をする。③自分の金で人に飲食させてやる。

ごう【傲】(コウ るごる たかがる) 人を人とも思わない。おごりたかがる。「傲慢・傲岸・傲然・傲諧・傲慢・驕傲」

ごうまん(傲慢) 高ぶって人をあざけり見くだす態度であること。

たかがる(高ぶる・昂る) ①高まる。すむ。亢進する。

②自分がそらうといふ様子を見せる。自慢する。

まん(慢)(マン おこたる わなどる) ①心にしまりがない。なまける。ちこたる。「怠慢・慢然

②物事を大事にしない。からんする。おろそかにする。

わなどる。「慢黒」③他を軽んじて自らをよしとする。たかがる。「慢心・傲慢・物慢・白慢」④

道りかのうい。たらたらと長びく。「慢性・緩慢

慢慢」

注。各項目も複雑、深い部分の候補している。

次に「傲慢だ」を実際に分析しよう。

1) まず「傲慢」(W_i) の項目を参照し、その概念(C_i) が“高ぶって人をあざけり見くだす態度であること”を把握する。

2) 次に「傲」(w_{i1}) の項目を参照し、訓が「おごる・たかがる」であることを知る。このとき項目中に「傲慢」という用例があることに注意しながら、“人を人とも思わない。おごりたかがる。”を C_{i1} における C_{i1} の候補とする。

3) 次に「おごる」を w_{i1} と見なし w_{i1} の項目を参照し、①の①の記述(C_{i1}) がほぼ C_{i1} の候補と一致することを確認する。よって w_{i1} Ra w_{i1} であることがわかり、ここで C_{i1} の候補を正式に C_{i1} と認める。

4) 次に「慢」(w_{i2}) の項目を参照し、訓が「おこたる。わなどる。」であることを知る。用例「傲慢」を通じて③の“他を軽んじて自らをよしとする。たかがる。”を C_{i2} の候補とする。

5) 次に「おこたる」を w_{i2} と見なし w_{i2} の項目を参照するが、和語としての漢字字母が異なり、 C_{i2} も C_{i2} の候補と一致しない。「わなどる」を w_{i2} と見なししても漢字が異なる（しかしこの場合 C_{i2} がほぼ C_{i2} の候補と一致するので）、 w_{i2} の訓としては「あなどる」を採用しておく）。よって「慢」と「あなどる(慢)る」はアロモルフ的関係にないものと見なす。これによ

り $C_i \in G_{a1}$ または $C_i \in G_{a2}$ の可能性はなくなつたが、一応 C_{i2} の候補を正式に C_{i2} と認めておく。

6) C_{i1} と C_{i2} を論理積で結合するところほぼ C_i になる。よって $C_i \in G_{a3}$ と認める。

7) 最後に、 w_{i1} と w_{i2} と共に動詞であること、 $C_{i1} \wedge C_{i2}$ であること、等の構造的性質を記述する。さらに 1), 2), 3), 4) で把握した C_i , C_{i1} および C_{i2} の内容的性質も記述する。

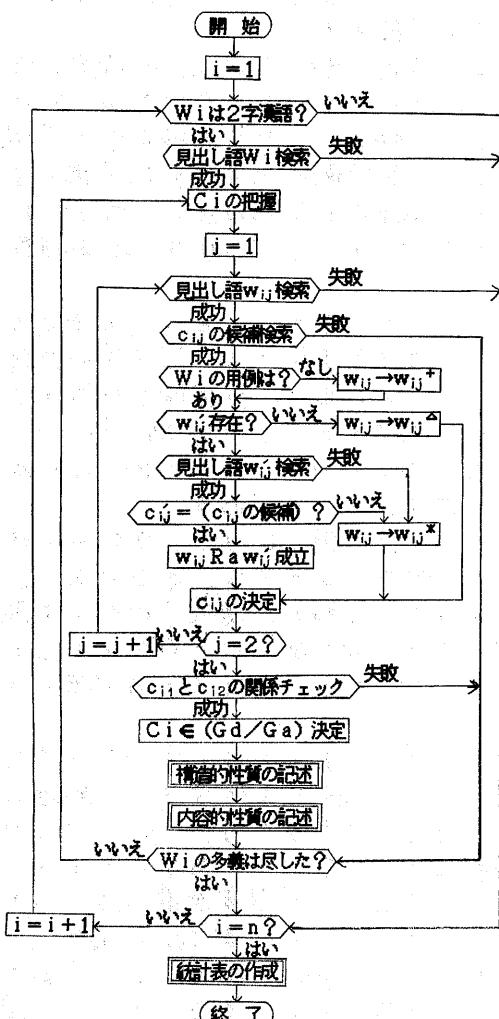


図 1 分析アルゴリズム

3.2 分析アルゴリズム

初めに、分析の対象について述べよう。従来から概念調査は、“分類語彙表”に登録されている語の概念を対象として進めてきた⁽⁹⁾。分類語彙表の相の類には相異なる形容動詞が、それに準ずるものも含めて(4.2節参照)，約1,480登録されている。それらの語の集合を $W = \{W_i\}$ とし、 W の概念(多義があるのが約1,650)の中から G_d または G_a を識別し、個々の概念を分析することを考える。

次に、詳しい分析アルゴリズムを図1に示す。分析の際 R_a の成立しない w_{ij} には右肩に“△”あるいは“*”印を付けて区別している。“□”の処理結果については、次章以降で詳しく述べる。

4. 構造的性質

前章の方法に従って実際に2字漢語の概念を分析した。 G_d の分析結果に関する點は稿を改めて述べることとし、4～6章では G_a の分析結果を示そう。本章では、 C_i における C_{i1} と C_{i2} の結合上の性質について述べる。

4.1 w_{ij} の品詞と C_{ij} の格

C_{i1} と C_{i2} が論理的に結合している場合は、 w_{ij} の品詞が重要である。初めに、分析の際見出されたすべての品詞を、記号と共に列挙する。

V: 動詞, Adj: 形容詞, AV: 形容動詞, Adv: 副詞, N: 名詞, K: 漢字字母

形容詞または形容動詞を組み合せて形容動詞を作ることは、属性概念同志が結合して属性概念になるので自然である。しかし、例えば動詞と形容詞/形容動詞を結合して形容動詞になるのはどう考へればよいのであるか? 事象概念はある状態 S_0 から他の状態 S_1 への変化を捉える概念であり、属性概念は対象 O_1 と対象 O_0 の間の差を捉

える概念である、という考え方を文献(1)で示した。 C_{0j} が事象概念の場合は捉え方を変え、変化によって生じた S_i における属性(尺度)に注目しているものと考える。例えば“尖る”は、本来、尖っていない状態 S_0 から尖っている状態 S_i への変化を捉える概念であるか、“尖鋭だ”にむりでは S_i の属性、即ちある物の先端部が細くて角度が小さい、という尺度に注目しているものと考える。この尺度により、例えば「君のナイフ(O₁)は私のナイフ(O₀)より尖っている」のように、差を捉えることが可能になる。 C_{0j} が名詞などの場合も、属性化しておるものと考える。

なお、論理的関係としては論理積のみが見出された。

次に構文的結合においては、 C_{01}/C_{02} が述語概念 C_{02}/C_{01} にかかるときの格が重要である。分析によって見出されたすべての格を、記号と英に列挙する。

s : 主体, Θ : 客体, O_f : 起点または源, O_t : 目標, O_s : 基所, L : 場所, t : 時間, d : 程度
格要素と述語概念の語順に注目すると、例えば“最適だ”的ように格要素が前にくるものと、“小心だ”的のように後にくるものとかある。いずれの場合も C_{02} が C_{01} の中心になつておるものと考える。

4.2 C_{01} と C_{02} の結合関係

すべての複合概念 A を、 C_{01} と C_{02} の結合関係を明らかにして、表1.1～2.2に示す。用いておる記号の意味は次の通りである。

下線なし: 形容動詞で、ナ型またはナ/1型

—: 名ナ1型

—: トタル型

—o: 和語を含む形容動詞。G_aには属さないが、それに準ずる取り扱い

ができる。

+, *, △: 図1参照。

[]: 多義を区別する。

文献(4)で口語形容動詞の判定基準

表1.1 論理的結合 第一漢字母形容詞

V+V 謙遜, 下劣, 優秀, (対話) 適当 [1], 亂暴, 別々, 狂騒*, 沈痛*, (話声) 絶え絶え [2], 謙讓, 濡潤, 駄々, 駄々, 駄々, 粉々々*, 燐々々*, 黙々々*

(行政) 因縛* [1], 動越*, 徒然*, 乱雜*, 遊刃*, 驚異*, 僦寧*, 高慢*, 謙直*, 慢*慢*, 慢*慢*, 燐*燐* 残虚*, 周到, 放逸, 老練, 駄々要, 駄々當*

奔放*, 流暢, 沈着, 駄々断, 謙暗*, 駄々忍, 率直*, 駄々急*, 廉腐*, 活*発*, 駄々切* [1], 徒容* 燐々々, (態度) 駄々々* [3], 駄々々, 燐々々, 駄々々*

V+Adj 尖锐, (注意) 適宜 [1], 透明, 富強, 浮薄, 急速, 適正*, 謙*嚴, 嚴*激, 困難, 絶無, 優良, 美麗*

慢重*, 謙謹*, 荒涼*, 謙易*, 謙潔*, 駄々多, 廉強, 謙大*, 廉大*, 駄々酷*, 流*麗*, 駄々細*, 容易*

簡*捷*

V+AV 散漫, 秘密, 明朗

富裕 老巧, 簡*約, 謙明*, 駄々華

繁密*

V+Adv 果*敢*

V+N 過尺, 過*縮 謙便, 駄々常

簡*素*, 賢*素*, 賢*朴*, 賢*実*

V+K 沈黙*, 謙懶*

駄々駄*, 謙單*

表1.2 論理的結合 第二漢字母形容詞

Adj+V 単陋, 単劣, 批劣, 畏劣, 快適, 好適, 濡透, 普通, 正當*, 早急*, 悲痛*, 奇異, 早熟*

奇抜*, 鮮新*, 繼続*, 相繼*, 相推*, 繼延*, (保持) 嶄新* [2], 高尚*, 正直*, 畏直*, 奇*奇*, 繼延* 相*異, 畏*點, 可*能*

溫*順, 寛*容, 畏*直*, 簡*活*, 単*屈*, 簡*切*, 輕*潔*, 畏*劣, 駄*直*

Adj+Adj 珍奇, (事件) 奇怪 [1], 危険, 陰惡, 激烈, 热烈, 精细, 詳細, 優美, 美麗, (果物) 甘美 [1], 满足, 単近, 過大, 貧乏, 貧弱, 軟弱, 柔弱, 柔軟, 強硬, 堅固, (手段) 安易 [1], 平易, 賢明, 濡暢, 温暖, 貴重, 嶄重, 聰明, 多大*, 強大*, 強烈*, 满新*, 良好*, 温厚*, 正直*, 均等*, 濡潤*, 過飽*, 鈍重*, 多*忙, 怪*奇, 強*固, 淡*薄*, 善良, 高貴, 高潔, 清涼, 単陋, 畏小, 短小, (身体) 弱小 [1], 畏少, 長大, 畏遠, 批遠*, 広大*, 鮮新*, 憎厭*, 新*銳, 深*遠, 深*遠, 烈々, 淡々, 駄々, (山奥) 深*々* [2]

(つづく)

陥*，陥屈，清淨，醜惡，潔白，貌美，體味，圓潤，
低聲，勇猛，幼稚，溫厚，偉大，重大，高*速，快，
速，迅*速，溫*良，寬*大，崇，高，猛烈，快，好，變，充
全，輕*薄，堅*嚴，堅*烈，蒼*白，精*銳，廉*潔，重*厚，
明*較，敏*捷，貞*淑，優*美，審*々，(胸內的)
審*々〔2〕

Adj+AV 平明，(性格が) 明朗〔1〕，勇壯，悲壯，精巧，
溫和〔2〕，円滑，優雅，悲慘，高雅，正*確，輕*微，古*雅，
醜*惡，潔*淨，細密，精密，嚴密，強*壯，卑賤，輕*妙，
奇*妙，珍*妙，新*鮮，剛*健，強*健，宏*壯，
豪*華

Adj+Adv 勇敢

Adj+N 軽便，(品物が) 粗末〔1〕，堅實，無實，正*常，
重要

粗野

寬*仁

Adj+K 卑俗，猥穢，煩惱，堅牢，低俗，清楚，
廣*廣

(原音が) 奇*特〔1〕

表1・3 構詞の結合 第一漢字字母分析動詞

AV+V 豐滿，穩當，(式典が) 鮎端〔1〕，豐富*

豐沃，柔順，精耕，慈厚，淳厚*

密*接*〔1〕

AV+Adj 壯大，壯麗，壯烈，爽快，明快，冷酷，冷淡，(作
業が) 細緻〔1〕，頑固，懶惰，安全，華麗，明細，豔麗，華美，
靚麗，希薄，(気分が) 輕快〔2〕，邪惡，盛大，微弱，
(身体が) 健*全*〔1〕，微小，微少，確固*

明白，頑冥，安泰，清廉，幸福，顯着，平*等，冗*長，
壯*快，悠*長，微*賤

幽*玄，(身分が) 微*細〔2〕，裕*福*

AV+AV 平靜，平安，平靜，冷靜，杜健，聰明，柔和，溫馨，
安樂，明確，靜謐，穩健，安*穩，平*滑，微久，(態度が)
悠久〔2〕

明瞭，靜寂，健壯，巧妙，淫*蕪*，隆*盛

冗*漫，隆*々〔1〕，悠*々〔1〕，瀟*々

AV+N 同*一，確*實，主要

AV+K 清*純*，巧*麗*，安*開*

邪*魔*

Adv+N 凡*庸*

表1・4 構詞の結合 第一漢字字母分析名詞他

N+V 共通，公財

火*急，輕*易

神*秘

N+Adj 中正，公正，真正，些細，空虛，陰*慘，些少，下*
階*

(フ フ" ハ)

險*險，雄*大，端*正*，野*罕*，與*難*，端*重*

N+AV 公平，空隙，內密

與*難，森*嚴*

縹*密，靈*妙*

N+Adv 平凡*

N+N 誠實，うちうち

肝*要，極*要，獨*獨

神*聖，穩*々，炎*々，縹*々

N+K 便利

險*險

野*蛮*，肝*腎*，肝*肺*

K+V 豪*奢，妥*當，特*別，應*病*，免*暴*，堅*張*

開*散*，卓*拔*，凸*凹*，靈*放*

K+Adj 奇*酷，奇*烈，贊*寂，巨*大，縹*細，凶*惡，
俗*惡*，壯*重*，純*麗*，迂*遠*，純*良*，牢*固*

該*博*，(底が) 賦*蒼*〔1〕

K+AV 豪*壯，莊*嚴，開*靜*，靈*華*

縹*密*

K+N 璞*未，(部下が) 忠*實〔1〕，純*一，純*真，準*一
純*粹*，純*朴*，特*殊*，忠*誠*

K+K 單*純*，殷*々*，靈*勃*，莊*漢*，區*々*，楚*々*，
堂*々*，(大海が) 洋*々*〔1〕，莊*洋*，靈*々*〔1〕

表2・1 構文の結合 格要素が主体

s+V 金*入り

s+Adj 気長，氣短，氣早，氣輕，閒適，(動作が)
身軽〔1〕，色白，(労働力が) 手薄〔1〕，憲世惡，性惡，
皆無，間近，欲深，足強，中高

对*等

s+AV 的確，氣氛，(女性が) 氣丈夫〔2〕，口下手，筆
まめ，筆無情

V+s 异様，乘*氣，移*氣，浮*氣

残念〔2〕，乱*脈，適度*

反*對*，違*筆*

Adj+s 大胆，小心，潔癖，強力，細心，多樣，小*膽*，
(色調が) 多*彩*〔1〕，美味，頭脳，(性分が) 堅氣〔1〕，強
体*，薄*情，大*力，少*食，著*名，多*溼，多*溫，快*調*
(如聲が) 安寧〔1〕，古風*，貧相*，好調*，強情*，正式*，
庄範*，惡*質*，短*氣*，新式*，狹隘*，強欲*，旧*式*，
多*端*，執迷*，多*禁*，多*介*，敏*脆，廉*價，博*學*
寡*欲*，敏*感*，堪*能*

AV+s 同様

幸運*，平氣*

N+s 一樣

本*式*，陽*氣*，本*氣*，內*氣*，陰*氣*

K+s 豪*氣，靈*胆，靈*勢

單*調*，純*情*

表2・2 構文的結合 格要素が主体以外

<u>o+V</u>	動詞
<u>o+Ad j</u>	副詞
<u>V+o</u>	入念、能く。
<u>o f + V</u>	(結果が) 風外 ⁽¹⁾ 、心 ⁺ 外、(先天的に) 固有 ⁽¹⁾ 。
<u>論⁺外⁺</u>	
<u>意⁺外⁺</u>	存 ⁺ 外 ⁺ 、案 ⁺ 外 ⁺
<u>o f + AV</u>	自明
<u>o t + V</u>	目障り ⁺ 、耳障り ⁺
<u>V+o t</u>	道 ⁺ 格 ⁺ 、道 ⁺ 主 ⁺
<u>V+o s</u>	隨 ⁺ 意 ⁺
<u>Ad j+o s</u>	正 ⁺ 則 ⁺
<u>嚴格⁺</u>	
<u>V+I</u>	通俗
<u>AV+I</u>	稀代
<u>t+V</u>	(柄が) 未熟 ⁽¹⁾
<u>t+Ad j</u>	機 ⁺ 敏
<u>d+V</u>	(才覚が) 独創 ⁽²⁾ 、稀有、克明 ⁺ 、必要 ⁺ 道 ⁺
<u>d+Ad j</u>	精快、頻繁、甚大、極悪 ⁺
<u>d+N</u>	極端
<u>V+d</u>	盛りだくさん。
<u>Ad j+d</u>	深 ⁺ 甚 ⁺
<u>痛⁺切⁺</u>	

について述べた。ナ型とナノ型は形容動詞と認められる語を述体修飾形に注目して下位区分したもの、名ナノ型とトタル型は基準を満たさないか形容動詞に準ずる扱いを受けたものである。

5. 内容的性質

ここでは C_i と C_{ij} の、概念内容についての性質を述べる。

5.1 C_i と C_{ij} の内容の記述

概念の内容を一番わかり易く表現する方法は、自然言語で記述することである。3.1節で国語辞典の記述例を示した。分析過程では各 C_i や C_{ij} ごとに辞典から記述文を抜き出し、カードの上に書き写した。

自然言語で語句を記述することは、広い意味でその語句を言い換えていることになる。その際しばしば問題になるのは、記述が單なる言い換えで終ってしまうことである。例えば「徹る」の項目に「高ぶる」、「高ぶる」の項目に「徹る」とだけ説明されていたのでは、それそれの内容を把握するのに不十分である。それに対し「徹る」を、例えば“才能・努力を頼みとして勝手をひるまいとする”と分析的に言い換えることは、より基本的な概念を解釈

表3・1 概念内容の記述—構造的結合

異常：慣れ秀でているさま。
優れる (V) : 力・価値などの程度が他より上になる。
秀でる (V) : 他を抜いて上に出る。
偉くて立派なさま。
偉い (Ad j) : 行いなどが他を抜いて上に出ているさま。
大 ⁺ きい (Ad j) : どこにも悪い所の見当らない程の良い状態
肝要：最も必要なさま。
肝 ⁺ (N) : 一番大切な部分
要 (N) : 一番大切な部分
懇切：行き届いて親切なさま。
懇ろ (AV) : 読がこもっていて、おろそかにしないさま。
切 ⁺ る (V) : しきりに、またはひたすらする。
粗末：品物などの作りが金入りでなく、つまらないさま。
粗い (Ad j) : 心を込めて作られているさま。
末 (N) : 下位の、つまらないもの。
尖銳：先が尖って鋭いさま。
尖る (V) : 先の方が次第に細くなり、先端部の角度がきわめて小さくなる。
鋭い (Ad j) : 刃物などの先が細く、先端部の角度がきわめて小さい。
簡便：簡単で便利なさま。
簡 ⁺ ぶ (V) : 無駄を省いて縮める。
便り (N) : 都合が良いこと。
雜談：知識・思想が雑然と入り混っているさま。
雜 ⁺ じる (V) : 多くのものがまとまることなく集まる。
駆 ⁺ (K) : 入り混っていること。
平明：平たくて明らかなさま。
平たい (Ad j) : わかりよい。
明らか (AV) : はっきりしていて、誰にもそうだと知れるさま。
煩煩：細かすぎて煩らわしいさま。
煩わしい (Ad j) : 気にかかるてうるさい。
瑣 ⁺ (K) : 細かくて小さいこと。
健康：身体が健やかで康らかなさま。
健やか (AV) : 身体がしっかりとして力強い。
康 ⁺ らか (AV) : 身体に故障がない。
確實：確かめることあるさま。
確か (AV) : 問題がない、はっきりしているさま。
実 (N) : 事実の通りであること。
清純：清らかで混じりけのないさま。
清らか (AV) : けがれがなくて、美しい。
純 ⁺ (K) : 混じりけのこと。

表3・2 概念内容の記述—構文的結合

異様 (V+s) : 様子が普通でないさま。
多彩 (Ad j+s) : 彩が多いさま。
幸運 (AV+s) : めぐり合せがいいさま。
陽気 (N+s) : 気が明るいさま。
心外 (o f + V) : 若えから外れるさま。
適格 (o t + V) : きまりに適っているさま。
殊勝 (d+V) : 特に勝っているさま。

でき、本筋の意図する所でもある。

本研究では、概念の抽象化過程を背景にしつゝ、概念を要素的なものと連結合成されたものとに分けている⁽²⁾。連結合成概念に関しては要素的概念を表す語を用いて分析的に記述することができるよう。また要素的概念に関しても、適当な部分集合(それに属する概念を仮に“原子概念”と呼ぼう)を定めることにより、原子概念を表す語を用いて分析的に記述できる可能性がある。

表4・1 概念内容のカテゴリ—要素的事象

カテゴリ	漢字母の概念
感情の変化	憂える(憂), (憂) 痛む
感覚の変化	—
要 位	別れる(离), (离) 接する
向 きの変化	反く(反)
要 形	尖る(锐), (锐) 利く
要 質	肥える(肥), (肥) 湿る 簡ぶ(简), (简) 满たす
要 量	燃く(燃), (燃) 燃れる (多) 彩る
光の変化	—
色の変化	—
熱の変化	—
力・勢い	屈む(屈), (屈) 弯める
音の変化	轟く(轟), (轟) 黙る
出現・消滅	秘める(密), (神) 秘める
開始・終了	絶える(無)
時 間	急ぐ(速), (速) 渋る
状 態	連なる(続)
有 様	透く(明)
その 他	乱れる(亂), (亂) 雜じる

表4・2 概念内容のカテゴリ—複合事象

カテゴリ	漢字母の概念
精神的行為	慮る(外), (高) 尚ぶ
学術・芸術的行為	老いる(朽), (博) 学ぶ
宗教的行为	—
言語的行為	論う(外)
社会的行為	残う(忍), (里) 潰すう
所作・振舞	暴れる(慢), (慢) 暴れる
労働・生産	勤む(勤), (勤) 到る (兼) 讀る
所 有	—
調査・計量	従う(順), (因) 循う
支配・人事	優れる(秀), (殊) 勝る
功 所・勝敗	—
待遇・逃亡	富む(富), (殷) 賑わう
興亡・盛衰	—
そ の 他	(徳) 用いる

原子概念に関してはもはや自然言語で分析的に記述することは不可能で、それで記述しようとすると、成分的特徴を示す記号を定義し、それらを列挙するなどしなければならない。

以上を念頭において表3・1では、国語辞典を参照しつつ C_{ij} と C_{kj} の内容を分析的に記述した。品詞をかご()とく()は、それぞれの概念が要素的か連結合成かを示している。特にく()の場合は、現在原子概念が明らかでないので、他の要素的概念を表す記号を用いて記述することにより、成分的特徴を明らかにすることに努めた。またいすれの場合も、格要素か名詞の場合には国語辞典の記述をそのまま用いた。

表3・2は C_{ij} と C_{kj} の格関係に注意して C_{ij} を記述した例である。

5.2 C_{kj} の内容のカテゴリ

C_{kj} の内容にどのようなカテゴリがある

表5・1 概念内容のカテゴリ—要素的属性

カテゴリ	漢字母の概念
感 情	快く(高), 安らか(柔), 暢く(氣) (愉) 快く(温) 和やか(強) 情
感 覚	敏く(感), 静か(寂) (聰) 敏く(開) 静か(美) 味
場 所	密やか(接), 中(正) (近) 遠く(広) 範
向 き	—
形	長い(大), 平ら(滑), 巨大(大) (尖) 锐く(平) 滑らか(狭) 险
質	軟らかい(柔), 韋(柔) (強) 固く(硬) 密やか(柔) やか(柔) 素
量・程度	多い(種), 薄か(弱), 些少(少) (微) 少ない(微) 密やか(直) 度
光	炎(火) (透) 明るい(透) 淡
色	蒼い(白), 濃やか(黒), 色(白) (青) 白く(淡) 淡
熱	温かい(暖), 淡く(火) (冷) 凉しい(冷) 淡
力・勢い	激しい(烈), 壮健(健), 爽快(氣) (富) 強く(勇) 壮(健) 力
音	静か(廉), 森(闇) (静) 寂か(森) 開
出現・消滅	(内) 密か
開始・終了	—
時 間	早い(急), 未だ(與), 機(動) (出) 迅い(新) 鮮か
状 態	疎(外) (縁) 縮
有 様	寒い(寒), 明らか(白), 苛(酷) (苛) 酷く(平) 明らか
油 像	正しい(当), 同じ(一), 略(簡) (便) 良い(平) 安らか(安) 特
そ の 他	自ら(明), 程(次) (健) 全し, (反) 対

表5・2 概念内容のカテゴリ—複合属性

カテゴリ	漢字母の概念
精神的属性	賢い(明), 頑くな(固), 腊(病) (通) 銛(微) 漫る, (小) 心
学術・芸術的属性	博く(学), 談(博) (浅) 薄く(多) 芸
宗教的属性	嚴か(嚴), 韋妙(妙) (往) 嚫か(神) 聖
言語的属性	口(下手) (能) 弁
社会的属性	幼く(稚), 懇ろ(切), 俗(愚) (通) 厳しい(理) 猥ら(野) 豪
態度・性格	勇ましい(威), 冷やか(冷), 慵懶(長) (廣) 深く(老) 柔み(軟) 腺
・能力	—
労働・生産	賄(貪) (多) 忙しい
所有・取引	貴い(重), 元(長), 壱(末) (低) 廉(廉) 価
調査・計量	偉(大), 微(微) か(既), 下(既)
地位・身分	(高) 貴(貴), (微) 細い
勝敗・優劣	(妙) 妙
待遇・逃亡	—
貧富・盛衰	貧しい(乏), 豊か(満), 殷(賑) (貧) 乏しい(隆) 盛ん(著) 名
そ の 他	—

あるかを知ることは、知識ベース作成上重要である。動詞で表される事象概念と形容詞で表される属性概念については、それぞれ文献(10), (11)と文献

(1), (2) にカテゴリの種類を示した。それらのカテゴリに基づいて各 C_{ij} の内容かどの程度把握できるかを調べた。表 4.1 ~ 5.2 に結果の一部を示す。各表において w_{ij} (w_{i2}) または (w_{i1}) w_{i2} なる表現は、 $W_i = w_{i1} w_{i2}$ について、それぞれ C_{i1} または C_{i2} の内容のカテゴリに注目していることを表す。

表 4.1 と 4.2 は w_{ij} が動詞の場合の結果を示している。動詞で表される各 C_{ij} はほぼ問題なくそれらのカテゴリで把握できた。また表 5.1 と 5.2 は w_{ij} が形容詞、形容動詞、名詞および漢字母の場合の結果を示している。各 C_{ij} は、形容詞に対して提案されたそれらのカテゴリでほぼ把握できた。ごく少數表 4.1 ~ 5.2 で把握できなかった C_{ij} もあつたが、それらは物の概念で、“身(軽)”, “足(弱)”など構文的結合の格要素になるものである。

6. 統計的性質

表 6.1 ~ 6.4 に統計データを示す。表 6.1 は C_{i1} と C_{i2} の結合関係に注目している。形容動詞の概念全体に占める Ga の割合を示すと、約 35 % である。この割合は形容詞の場合の 5 % に比べるとはるかに大きい。結合関係を比較すると、全体としては論理的結合が多いが、和語については構文的結合の多いのが注目される。

表 6.2 は w_{ij} の品詞に注目している。Adj が特に多い。歴史的に見て形容動詞は数の少い形容詞を補って発達したから⁽⁸⁾、二つの形容詞を組み合せて形容動詞を作ることが最も必要でかつ自然であるたものと考える。Adj について V と AV が多いが、これら 3 品詞で全体の 82 % を占める。

表 6.3 は格関係に注目しているが、S 以外は散見される程度と言える。

表 6.4 はアロモルフ的関係に注目している。アロモルフ的関係の成立しま

表 6.1 結合関係による分布 Ga

関係	漢語			和語	計
	ナ	ナノ	名ナノ		
論理的	345	67	45	2	459
構文的	59	35	0	30	124
計	404	102	45	32	583

表 6.2 品詞による分布 論理的結合

w_{i2}	V	Adj	AV	Adv	N	K	計
V	52	27	9	1	8	4	101
Adj	38	103	28	1	8	8	186
AV	10	39	26		3	4	82
Adv					1		1
N	5	14	7	1	9	5	41
K	10	14	5		9	10	48
計	115	197	75	3	38	31	459

表 6.3 格関係による分布 構文的結合

格関係	述語概念	V	Adj	AV	Adv	N	K	計
s		10	58	9		6	5	88
o		3	1					4
o_t		7			1			8
o_s		4						4
o_b		1	2					3
l		1			1			2
t		1	1					2
d		6	6		1			13
計		33	68	11	0	7	5	124

表 6.4 アロモルフ的関係による分布

成立状況	語数
w_{i1}, w_{i2} 共に *も△なし	235
w_{i1}, w_{i2} いずれかに *または△	219
w_{i1}, w_{i2} 共に *または△	97
計	551

注：和語は除いてある。

いなし w_{ij} については次章で検討する。

7. 検討

分析アルゴリズムに関して、その改良と形容動詞以外の 2 字漢語への応用について検討する。

1) 分析アルゴリズムの改良

アロモルフ的関係の成立していない場合について述べる。例えば「傲慢*」における「慢」は「あなどる」と読み、「侮る」の意味で用いられている。*

印か付いたのは辞典の中に「あなど（慢）る」という項目がないためである。しかしある程度国語刀のあら人は「慢る」を“侮る”的意味で理解できる。それに対し例えは「閉△散*」の「散る」を“ひまである”の意味で理解できることは少ない。「慢」のような例は、「容*す→許す」、「愉*しい→楽しい」などかなりの数に上り、きめ細かな分析を望む場合は「散る」などと区別した方がよい。

もう一つ＊印について述べる。例えば漢字母「質」には“問いたずねて理判をただす”という事象概念と，“地のまゝ。歸り気がないこと。”といふ属性概念とかかる。前者には“ただす”という訓があり、後者には対応する訓が見当らない。「質△朴*」における「質」は後者の意味で用いられており、「ただ（質）す」という項目を参照して $C_{ij} \in C_{ij}$ となり、 $w_{ij} \in W_i$ に＊印が付せられると共に動詞に分類された。このような場合は訓にせず、漢字母のまゝ（△印）で扱つた方が自然である。

2) 分析アルゴリズムの応用

本アルゴリズムは形容動詞に関する提案したものであるが、アルゴリズムから自明のように、他の品詞の二字漢語に対しても応用できる。また本稿では岩波国語辞典を用いたが、それは漢字母の達ひ方ならびに概念記述が適切と判断したためである。これについても通常の漢和辞典と国語辞典で処理ができる。その場合は、まず目的に応じて漢字母の集合 K を定める。ある $W_i = w_{i1} w_{i2}$ が与えられると、 $w_{ij} \in K$ の場合に漢和辞典を検索する。漢和辞典の用例を手かかりに C_{ij} の候補を調べ、最後に国語辞典で w_{ij} を検索し、 C_{ij} を詳しく調べる。ただし w_{i2} の用例を調べる際漢和辞典では、通常 w_{i1} の項目を参照しなければならない。

8. まとめ

形容動詞を対象として、二字漢語の概念を分析する方法を提案した。またその方法に従って、約 580 の複合概念 A を詳細に分析した。その結果複合概念 A のクラスの、構造的、内容的な統計的性質が明らかになった。

提案した方法は形容動詞以外の二字漢語に対しても応用可能である。また得られた概念データは自然言語や图形理解に必要な知識ベースの作成に、有効な基礎資料を提供するものである。

本研究の一部は、文部省科学研究費 60580029 による。

文献

- (1) 岡田：自然言語および图形理解のための属性概念の分類—形容詞における要素的概念、情報処理学会論文誌、Vol.26, No.1, pp. 25-32 (1985).
- (2) 岡田：自然言語および图形理解のための属性概念の分析—形容詞の概念の抽象化過程、情報処理学会論文誌、Vol.26, No.3, pp. 497-504 (1985).
- (3) 岡田：形容詞で表される属性概念の分類と图形パターンの向自然言語理解、情報処理学会論文誌、Vol.26, No.3, pp. 505-512 (1985).
- (4) 岡田：自然言語および图形理解のための形容動詞の概念の分類—要素的概念、情報処理学会、自然言語処理研究会資料、50-1 (1985).
- (5) 森岡：現代漢語の成立とその形態、国語と国文学、Vol.24, No.4, pp. 23-47 (1979).
- (6) 野村：三字漢語の構造、国立国語研究所報告、51, pp. 37-62 (1974).
- (7) 田中、水谷、吉田：語と語の関係について、情報処理学会、自然言語処理研究会資料、41-4 (1984).
- (8) 西尾、岩淵、水谷(編)：岩波国語辞典、p. 1160, 岩波書店 (1975).
- (9) 国立国語研究所(編)：分類語彙表、p. 362, 奉美出版 (1964).
- (10) 岡田、田町：自然語および图形解釈のための単純事象概念の分析および分類、電子通信学会論文誌(D), Vol.56-D, No.9, pp. 523-530 (1973).
- (11) 岡田、田町：自然語および图形解釈のための単純事象概念の分析および分類、電子通信学会論文誌(D), Vol.56-D, No.10, pp. 591-598 (1973).