

形容詞で表される属性概念の分類と 図形パターンの自然言語理解†

岡 田 直 之††

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系の組織的な解明に取り組んでいる。本稿はその第3報で、とくに形容詞で表される属性概念の分類を行うとともに、一連の概念調査で得た成果を図形パターンの自然言語理解へ応用する。初めに、第1報と第2報で提案した属性概念のクラスならびに各クラスを特徴づける構造的、内容的性質に基づいて、分類のためのアルゴリズムを作成する。次にそのアルゴリズムに従って日常の言語生活でほぼ十分とされる約630の属性概念を実際に分類する。その結果形容詞で表される属性概念の系の定性的かつ定量的性質が明らかになる。図形パターンの自然言語理解は、動図形における事象の意味を自然言語で表現するシステム SUPP を基礎に行う。まず、分類したすべての属性概念データを SUPP のソース部分に登録し機械可読化する。次に登録したデータのなかから10組の要素的属性概念を選び、第2報の概念分析に基づいてそれらを意味辞書に組み込む。同時に各概念に対応する視覚的特徴を図形パターンから抽出するためのサブプログラムも作成する。実験により拡張された SUPP が図形パターンのもつ属性も理解できることが確認される。

1. ま え が き

近時、異なった情報源である自然言語および図形を計算機に統一的に意味理解させようとする社会的気運が高まっている。そのためわれわれは日本語の形容詞や形容動詞で表される属性概念の系の組織的な解明に取り組んでいるが^{5), 6)}、その方針は次のとおりである。

- 1) 属性概念を分類分析する際の基礎として、属性概念を対象の間の“差”に関する概念と見なしている。
- 2) 概念の抽象化過程を背景にして属性概念を要素的な概念とそれらから導ける連結・合成概念とに大別している。
- 3) 要素的な概念については、外的な（すなわち現実世界の）物理的対象あるいは内的な心理的対象との対応に注目し、構造および内容を明確にしようとしている。
- 4) 連結・合成概念については、それらが要素的な概念からどのようにして導かれるかに注目し、構造および内容を明確にしようとしている。
- 5) 日常の言語生活でほぼ十分とされる形容詞、形容動詞を対象とすることにより、属性概念の系全体の定性的かつ定量的性質を明らかにしようとしている。

第1報では方針1と3に従って形容詞で表される要素的な概念の性質について議論した⁵⁾。また第2報では方針2と4に従って形容詞で表される属性概念の抽象化過程について議論した⁶⁾。引き続き本稿では、方針5に従って形容詞で表される属性概念の組織的な分類を行うとともに、一連の概念調査で得られた成果を図形パターンの自然言語理解に応用する。本稿は自然言語および図形を総合的に処理するうえで、大規模な知識ベースを作成するのに有効な基礎資料を提供するとともに、統一的な意味理解システムに対し一つのひな形を提案するものである。

2. 属性概念の分類

2.1 分類の対象と項目

まず分類の対象であるが、分類語彙表は日常の言語生活でほぼ十分と見られる約3万3千語を収録したソースである⁷⁾。それらの語は体、用、相およびその他の四つの類に大分類され、さらに各類は同義、類義の関係で細かく分類されている。ここでは相の類に登録されているすべての形容詞（約540語）の表す概念を分類の対象としよう。語は一般に多義をもつが、これについては分類語彙表の分類項目の意図している意味、岩波国語辞典に記述されている意味⁸⁾、多くの用例を集めた文献⁸⁾に見いだされる意味を参考にし、それらで不足している場合は筆者の考察に基づいて補足する。

次に分類の項目について、第1報および第2報で得た属性概念のクラスとそれらの特徴づける構造的および

† Conceptual Classification of Attributes Represented by Adjectives and Natural Language Understanding of Picture Patterns by NAOYUKI OKADA (Department of Information Science and Systems Engineering, Faculty of Engineering, Oita University).

†† 大分大学工学部組織工学科

び内容的性質を以下にまとめる。

属性概念

要素的概念 A_p

基準概念† A_{ps}

構造の型 (S_p): 7種類 (第1報の3.2節)

内容のカテゴリ (C_p): 19種類 (第1報の4.2節)

類似概念† A_{pr}

構造の型, 内容のカテゴリ: A_{ps} に同じ。

連結・合成概念 A_c

複合概念 A_{ca}

結合規則 (S_{ca}): 3種類 (第2報の4.1節)

内容: 要素的概念の内容とそれらの結合関係

複合概念 B_{cb}

構造の型: A_{ps} に同じ。

内容のカテゴリ (C_{cb}): 14種類 (第2報の4.2.2項)

派生概念 A_{cd}

接辞 (S_{cd}): 2種類 (第2報の4.3節)

派生情報 (C_{cd}): 10種類 (第2報の4.3節)

以上のうち†印は2.2節で述べる。また S_i および C_i の詳細は、〔 〕内に第1報あるいは第2報で述べている節を示しているのので、そちらを参照されたい。

2.2 分類のアルゴリズム

事象概念を分類する際一つの分類アルゴリズムを提案した^{*)}。以下、処理手順に関してはほぼそのアルゴリズムに準拠し、処理基準に関して2.1節の構造および内容的性質 (S_i および C_i) を用いたアルゴリズムを示す。

アルゴリズム

対象とする属性概念の集合を A とする。

1) 派生概念の分類

A のなかから、まず S_{cd} の接辞をもつすべての語を抽出する。次にそのおのおのについて語基がそれ自身で意味をなし、かつ接辞のもたらす情報が C_{cd} の内容を満たすかどうかチェックする。もしいずれの条件をも満たせばクラス A_{cd} のメンバとする。

メンバとしない例。“うやうやしい”は一見“一しい”という接尾辞をもつ派生概念に見えるが、“うやうや”だけでは意味をなさないの派生概念とはしない。

2) 複合概念Aの分類

$A - A_{cd}$ のなかから、まず2語で構成されるすべての複合語を抽出する。次に各複合語について要素であ

る語がそれ自身で意味をなし、かつ結合関係が S_{ca} を満たすかどうかチェックする。もしいずれの条件をも満たせばクラス A_{ca} のメンバとする。

メンバとしない例。“水臭い”は一見“水が臭い”という、 S_{ca} の構文的結合に則った概念のように見える。しかし内容においては文字どおりの化学的性質を示していないので、複合概念Aと見なさない。

3) 複合概念Bの分類

$A - A_{cd} - A_{ca}$ のおのおのについて、もしその内容が C_{cb} のいずれかのカテゴリを満たせば、クラス A_{cb} のメンバとする。さらに A_{cb} を基準概念の構造の型 S_p に準じて細分類する。

4) 類似概念の分類

$A - A_{cd} - A_{ca} - A_{cb} (= A - A_c)$ のなかにもし内容の類似した概念があれば、まずそれらをグループ化する。次に各グループにおいて“基準概念”を一つ選び、残りは“類似概念”とする。すべてのグループから類似概念を取り出し、クラス A_{pr} のメンバとする。

メンバとする例。“煙い”に対する“煙たい”、“塩辛い”に対する“しょっぱい”、“おいしい”に対する“うまい”など。

ただし一つの概念が類似概念であるとともに、他に對して基準概念となることはないものとする。

5) 基準概念の分類

$A - A_c - A_{pr}$ のおのおのについてその構造が S_p かつ内容が C_p で把握できるかどうかチェックし、いずれの条件をも満たせばクラス A_{ps} のメンバとする。

以上1)~5)において必要なら一つの概念を複数の項目に分類するものとする。

2.3 分類の結果

前述のアルゴリズムに従って実際に分類を行った^{*}。表1~6にその結果を示す。表1~3は連結・合成概念の分類結果、表4~6は統計データである。要素的概

表1 複合概念Aの分類

Table 1 Classification of complex concepts A.

構造	複合概念A
XX I	—
XX II	細長い, 青白い, 浅黒い, 甘酸っぱい, 暑苦しい, 悪賢い, ...
XX III - I	毛深い, 欲深い, 力強い, 程よい, 肌寒い, 目ざとい, 身軽い, きまり悪い, ...
XX III - II	縁遠い

* 分類に際し本学大学院生渡辺国房君 (現在三菱電機勤務) の協力を得た。ここに謝意を表す。

表 2 複合概念 B の分類

Table 2 Classification of complex concepts B.

内容	構造	複 合 概 念 B
10	I	何気ない, 賢い, さとい, ...
	III	肩身が広い, うやうやしい
	IV	親しい, 心安い, 気安い, 気まずい
	VI	心強い, 頼もしい, ゆかしい, おこがましい, 片腹いたい, 気忙しい, ...
11		—
12	I	めでたい
13	I	まわりくどい, 口汚ない, 口はばったい
	III	うるさい, やかましい, 口やかましい
14.0	I	草深い, せちがらい
14.1	I	乳臭い, あどけない, みずみずしい, ...
14.2	I	いかがわしい, するい, あくどい, ...
14.3	III	肩身が広い, うやうやしい, うるさい, やましい, 優しい, 水臭い, ...
	IV	親しい, むつまじい, 心安い, ...
15	I	偉い (1), 賢い, おとなしい, 涙もろい, 凜々しい, 図太い, 仰々しい, しかつめらしい, さもしい, がめつい, だらしない, そそっかしい, 深い, しぶとい, ...
	III	優しい, なれなれしい, よそよそしい, 水臭い, きびしい, ...
	IV	親しい, むつまじい, 心安い, ...
	V	うまい, まずい, つたない, うとい
	VI	心強い, 頼もしい, ゆかしい, おこがましい, 片腹いたい, 気忙しい, ...
16	I	忙しい, あわただしい, ...
17.0	I	貴い (1), くだらない, 高い, 安い
	I	ぼろい, がめつい, あくどい, ...
18		—
19	I	貴い (2), 偉い (2), いやしい, ...
	VI	ゆかしい, おこがましい, 恐れ多い, ...
20, 21		—
22.0	I	貧しい, つましい, お寒い
22.1	I	花々しい, みずぼらしい, しがたない
	VI	晴れがましい
23	I	物々しい, 重苦しい

念の分類結果についてはすでに第 1 報で示した。紙面の都合により, 表 1~3 では一部だけ表示している。すべてを分類した結果については文献 4) を参照されたい。また表 1~6 において内容のカテゴリおよび構造の型をそれぞれアラビア数字およびローマ数字で表している。同じ理由によりやむをえず数字のみを示しているが, それらの意味は第 2 報を参照されたい。

表 3 派生概念の分類

Table 3 Classification of derivative concepts.

内容	構造	派 生 概 念
50.0	L I	手ごわい, 手ひどい, 手ぬるい, ...
50.1	L I	真新しい, 真近い, 真白い
50.2	L II	かわいらしい, 憎らしい, 汚ならしい, 男らしい, 長たらしい, むごたらしい, 若々しい, 重々しい, ...
51.0	L I	小ざかしい, 小憎らしい, ...
51.1	L II	おしつけがましい, 差出がましい
52		—
53.0	L II	見やすい, わかりやすい
53.1	L II	住みにくい, 言にくい, 堪えがたい, やみがたい, 息苦しい, ...
53.2	L II	掃ぎない, 限りない, 遠慮ない, ...
54~57		—
58.0	L I	生暖かい, 生ぬるい, 生白い, ...
58.1	L I	薄暗い, 薄赤い, 薄ら寒い, ...
58.2	L III	白っぽい, 湿っぽい, 飽きっぽい, ...
58.3	L II	古くさい, 面倒くさい, 陰気くさい, ...
59		喜ばしい, 悩ましい, 好ましい, 誇らしい, 狂わしい, うらやましい, ねたましい, 腹立たしい, 呪わしい, ...

表 4 複合概念 B の分布

Table 4 Distribution of complex concepts B.

内容	構 造							合計
	I	II	III	IV	V	VI	VII	
10	6		2	4		14		26
11								0
12	1							1
13	3		3					6
14	13		22	6				41
15	47		14	6	4	14		85
16	5							5
17	11							11
18								0
19	6					5		11
20								0
21								0
22	6					1		7
23	2							2
合計	100	0	41	16	4	34	0	195

2.4 検 討

表 1~6 でとくに注目される点を検討しておこう。表 1 を見ると各複合概念 A において二つの概念が単純な規則できれいに結合している。表 2 では言語的形態

表 5 派生概念の分布

Table 5 Distribution of derivative concepts.

内 容	構 造		
	L I	L II	合計
50	14	31	45
51	3	2	5
52			0
53		15	15
54~57			0
58	10	13	23
59		29	29
合 計	27	90	117

表 6 形容詞で表される属性概念の分布

Table 6 Distribution of attribute concepts represented by adjectives.

ク ラ ス	語/概 念		
	相異なる語数	相異なる概念数	延べの概念数
連結・合成			
派生概念	116	116	117
複合概念A	30	30	30
複合概念B	126	128	195
要素的			
類似概念	88	88	88
基準概念	178	272	305
合 計	538	634	735

が単純語、複合語あるいは派生語（に準ずる語）と雑多である。しかし内容的には各カテゴリによりかなり整理されていることがわかる。複合概念Bでは複数のカテゴリにまたがるものが多い。表4ではカテゴリ15（態度・性格・能力）が目立つ。人はとりわけ人の属性に興味を示すものと理解される。構造的には最も単純なI型が際立っている。表6は形容詞で表される属性概念全体の定量的性質を示している。相異なる概念数は約630で、そのうち基準概念が約270である。要素的概念は外的な物理的対象あるいは内的な心理的対象と直接対応しているが、そのような対象の世界はこれら約270の基準概念で覆いうることを示している。

上記分類結果を得るまでには2.2節のアルゴリズムに従って数度の見直しを行った。もし第1報および第2報で示した、属性概念のクラスを特徴づける構造的または内容的特徴によって把握できない属性概念が存在するなら、最終的に $A-A_c-A_p-A_{p_1}$ が空でなくなる。幸いそのような属性概念は見当たらず、多少例

外的な場合でも各クラスにおけるその他の項目の範囲内で処理できた。

以上により、第1報および第2報の議論の妥当性が確認されるとともに、形容詞で表される属性概念の系の定性的かつ定量的性質が把握できた。

3. 図形パターンの自然言語理解

3.1 SUPP の概要

時間的に変化する図形パターンを受け入れ、そのなかで生じている事象の意味内容を理解し、事象の間の類似性を推論し、その結果を日本語と英語で記述するシステムとしてすでにSUPPを作成している^{2),3)}。その概要を図1に示す。以下では、図形内の属性をも理解できるようSUPPを拡張する。

3.2 分類・分析データの知識ベース化

3.2.1 分類データの登録

2.3節で分類したすべての属性概念をSUPPのシソーラス部門に登録し、機械可読化する。各概念に割り付けられたレコードの構成を図2に示す。

3.2.2 分析データの登録

シソーラス部門に登録される属性概念のうち、第2報で分析した約200の要素的概念についてはSUPPの意味辞書および構造辞書にも組み込みたい。しかしすべての要素的概念を組み込むのは以下に示すサブプログラムの作成が困難である。すなわち認識に際して音、熱など画像入力装置以外の入力装置あるいはセンサを必要とする要素的概念が約70あるが、これらは

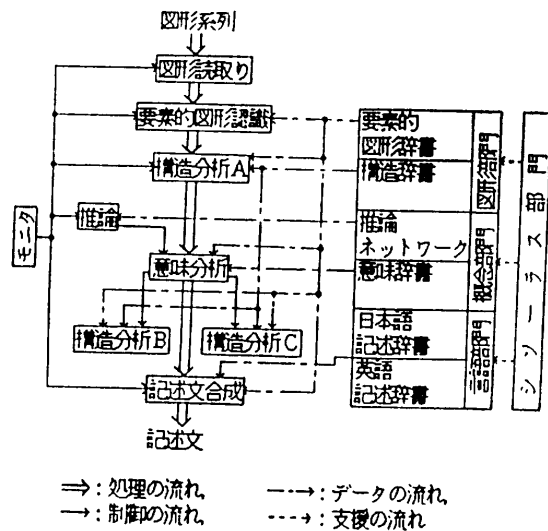


図 1 SUPP の概要

Fig. 1 Overall schema of SUPP.

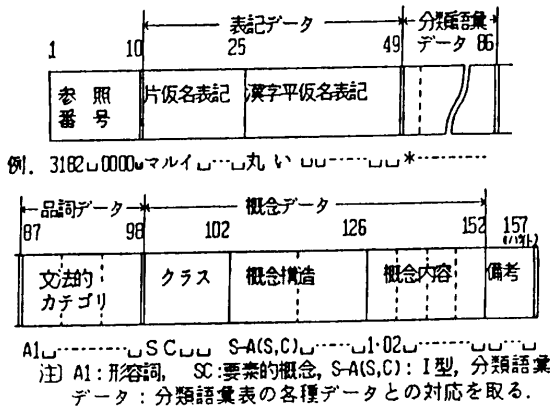


図 2 シソーラスのデータ構造
Fig. 2 Data structure of the thesaurus.

SUPPで取り扱う範囲を越えている。残りが画像処理の対象となるが、SUPPの要素的図形認識は現在のところ比較的簡単な2次元の線画像50種類なので、さらに対象が限られる。結局、現在の要素的図形認識で視覚的特徴の抽出可能なものとして、とりあえず内容的カテゴリが場所、形および質に属する10組を試行的に組み込む。

まず第2報で分析した概念的特徴を参考にして、その特徴に連合される視覚的特徴が図形内でどのような形態をとるかを考察する。次に文献2)で述べた図形処理の手法を用いて、入力図形からその視覚的特徴を抽出するためのサブプログラムを作成し、事象概念の場合と同様に一括して構造辞書に組み込む。以下に組み込んだサブプログラムの名称、対応概念および処理内容を示す。

- KT(1) 高い(1)/低い(1)
- sとcの間の、基準面からの垂直な距離の差を取る。
- KH(1) 深い(1)/浅い(1)
- sとcの間の、(環境の)表面または出入口からの距離の差を取る。
- KN(2) 長い(2)/短い(2)
- sとcの間の、おのおのの長軸に沿った長さの差を取る。
- KT(2) 高い(2)/低い(2)
- sとcの間の、おのおのの向きに沿った長さの差を取る。
- KT(3) 高い(3)/低い(3)
- sとcの間の、画面での垂直方向の長さの差を取る。
- KH(3) 深い(3)/浅い(3)

sとcの間の、表面または出入口から底または奥までの長さの差を取る。

- KR(1) 広い(1)/狭い(1)
- sとcの間の面積の差を取る。
- KF(2)* 太い(2)/細い(2)
- sとcの間の、おのおの短軸に沿った長さの差を取る。
- KO(1)* 大きい(2)/小さい(2)
- sとcの間の面積の差を取る。
- KM* 重い/軽い
- sとcの間の重さ(=面積×密度)の差を取る。

注 1. *とcは属性概念の構成要素で、それぞれ主体、比較の対象を示す。
2. *印のKF(2)については面積、またKO(1)とKMについては体積の計算が必要であるが、2次元図形であるためそれぞれ長さと同面積で置き換えている。

最後に上記サブプログラムの名称を一部に用いて、10組の要素的概念を事象概念におけると同様に性質リストとして意味辞書に組み込む(表7参照)。

3.3 階層的知識と理解過程

第2報で知識の系を生データ、知覚的特徴、概念的特徴、要素的概念および連結・合成概念の五つのレベルに分け、理解の過程を制御を伴った解釈の過程とする考えを示した。SUPPの理解過程は物としての要素

表 7 意味辞書—追加概念
Table 7 Semantic dictionary—added concepts.

概念	性質		
	A	s	c
高い(1)	KT(1)*	MONO	MONO
深い(1)	KH(1)*	MONO	MONO
長い(2)	KN(2)*	HOSONAGA	HOSONAGA
高い(2)	KT(2)*	HOSONAGA, MUKI-V	HOSONAGA, MUKI-V
高い(3)	KT(3)*	MONO	MONO
深い(3)	KH(3)*	KUBOMI	KUBOMI
広い(1)	KR(1)*	HEIMEN	HEIMEN
太い(2)	KF(2)*	HOSONAGA	HOSONAGA
大きい(2)	KO(1)*	RITTAI	RITTAI
重い	KM*	MONO	MONO

注 1. 性質欄は全体として各概念の構造が第1報で述べたA(s; c)型であることを示す。
A欄: たとえばKT(1)*はサブプログラムKT(1)が処理結果として正の値を返すことを示す。
s, c欄: たとえばHOSONAGA, MUKI-Vは、sおよびcがそれぞれ“細長い”および“垂直方向の向き”をもつことを示す。
2. 紙面の都合で低い(1), 浅い(1)などの反対概念を省略している。

表 8 シソーラスの一部
Table 8 A part of the thesaurus.

1	ハンコウ	1	ヒヨウキ	1	ガイネン	1	コウゾウ	1	サイヨウ	1
1	3503 0012	1	カンタ ² カイ	1	SC	1	S-A(S,C)	1	1.09	1
1	3503 0013	1	アタマシイ	1	SC	1	S-A(S,C)	1	1.09	1
1	3503 0050	1	* タカイ	1	SC	1	S-A(S,C)	1	1.09	1
1	3503 0051	1	* ヒクイ	1	SC	1	S-A(S,C)	1	1.09	1
1	3503 01001	1	* ヤカマシイ	1	SC	1	INTR S-A(S,C)	1	1.09	1
1	3503 01002	1	* ヤカマシイ	1	SC	1	TRAN S-A(S,O,C)	1	0.011 & 1.09	1
1	3503 0101	1	* ウルツイ	1	SC	1		1		1
1	3503 0102	1	ワカ ² シイ	1	DC	1	D-SFX	1	59	1
1	3503 0103	1	ソウゾウ ² シイ	1	DC	1	D-SFX	1	50.2	1

注) INTR: 自動詞的屬性概念, TRAN: 他動詞的屬性概念, S-A(S,O,C): VI型, DC: 派生概念, D-SFX: 接尾辞, *: 使用頻度に関する情報

の図形を認識する過程²⁾と事としての属性や事象を理解する過程³⁾とに分かれるが、後者の概要を次に示す。

1) 読み取られた生データレベルの図形パターンに対し、要素的図形の認識後前記サブプログラムによる構造分析を行うが、これは知覚的特徴レベルの処理に相当しよう。

2) 知覚的特徴の抽出が終わると、用いたサブプログラムの名称に注目する。各サブプログラムの名称は第2報で述べた“第一種の概念的特徴”に相当するが、このことは1)により自動的に知覚的特徴が概念的特徴に連合されたことになる。

3) 次に要素的 개념レベルでの解釈が行われる。たとえば要素的図形 P_1 と P_2 の間に性質 $KT(3)^+$ が見出され、かつ要素的図形辞書(省略)のチェックにより P_1 と P_2 がそれぞれ MONO という性質を満たすなら、表7により“ P_1 が P_2 より高い[3]”が解釈される。

4) さらに推論ネットワーク(省略)により類似性という観点からいくつかの要素的 개념が関連づけられるが、これは連結・合成概念レベルでの処理といえる。

以上において実際の処理順序はもっと複雑である。また SUPP における制御は推論プログラムと意味分析プログラムがある程度の機能を有している。

3.4 実験と検討

3.4.1 実験*

SUPP は FACOM 230-60 で原型が開発されたが³⁾、その後 FACOM 230-38S に移植された。プログラムの大きさは当初約5,400ステートメント(FORTRAN)であったが、逐次改良され、現在は約6,500ステートメントになっている。今回の属性概念にかかわる部分

* 実験に際し(元)本学基礎情報工学研究室三浦愛子助手の協力を得た。ここに謝意を表す。

は、そのうち約500ステートメントである。

初めにシソーラス部門に登録された属性概念データの一部を表8に示す。登録に際しては作業を容易にするため会話型の支援プログラムを作成した。先に述べたように SUPP の認識しうる要素的図形は50種類であるが、それらを種々組み合わせて20の図形パターンを作成し、入力した。図3に入出力の一例を示す。図3の処理においてCPUタイムは3分29秒で、うち2分38秒が要素的図形認識に費やされた。要素的図形は拡大・縮小、回転、反転および平行移動の図形変換を施されており、また手書き等による雑音も許されている。

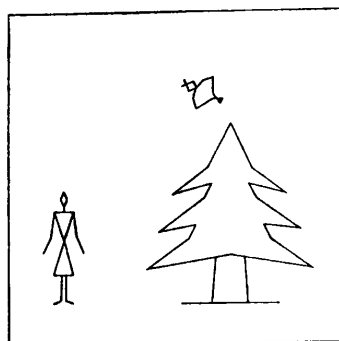
出力文-1では比較の対象 c の表現が省略されている。第1報で6種類の省略の型を示したが、出力文-1の型はそのなかの“平均”に相当する。すなわち要素的図形辞書のモデルが平均的な c を与えていると見るのである。出力文-2では第2報におけるきめの細かい概念分析が“タカイ/ヒクイ”の多義の上に反映されている。“オンナがトリ(1)よりタカイ”は少し理解しにくい、が、“オンナ”と“トリ(1)”の画面における垂直方向の長さが比較された。出力文-1, 2では英語文も出力されている。日本語と英語では形容詞の概念が異なるので、日本語の属性概念を最もよく表している英語の形容詞を選んで英語記述辞書を作成した。

他の図形パターンについてもほぼ同様の出力が得られ、SUPP が期待どおり拡張されたことが確認できた。

3.4.2 検討

(1) SUPP の能力

SUPP の核心は知覚的特徴レベルから要素的 개념レベルにかけての、図形に関する知識と自然言語に関する知識のインタフェースの部分である。この部分が従



(a) 入力図形

- 注 1. "オンナ"を主体とすることが外部から指示された。
 2. 出力文-1においては、入力図形の"オンナ"が要素的図形辞書のモデルと比較された。
 3. *印は、以下のサブプログラムが構造分析に用いられたことを示す。
 *: KT [1], **: KT [2], ***: KT [3]

- 1-J) オンナ カ* ナカ*イ。
 1-E) THE WOMAN IS LONG.
 2-J) オンナ カ* タカ*イ,**
 2-E) THE WOMAN IS TALL.**
 3-J) オンナ カ* オオ*キイ。
 3-E) THE WOMAN IS BIG.
 4-J) オンナ カ* フ*タイ。
 4-E) THE WOMAN IS THICK.
 5-J) オンナ カ* オ*モイ。
 5-E) THE WOMAN IS HEAVY.
 (b) 出力文-1

- 1-J) オンナ カ* キ ヨリ ヒ*クイ,***
 1-E) THE WOMAN IS LOWER THAN THE TREE,***
 2-J) オンナ カ* キ ヨリ ヒ*クイ,*
 2-E) THE WOMAN IS LOWER THAN THE TREE*
 3-J) オンナ カ* キ ヨリ サ*イ*イ。
 3-E) THE WOMAN IS SMALLER THAN THE TREE.
 4-J) オンナ カ* キ ヨリ オ*モイ。
 4-E) THE WOMAN IS HEAVIER THAN THE TREE.
 1-J) オンナ カ* トリ(1) ヨリ タ*カ*イ,***
 1-E) THE WOMAN IS HIGHER THAN THE BIRD(1),***
 2-J) オンナ カ* トリ(1) ヨリ ヒ*クイ,*
 2-E) THE WOMAN IS LOWER THAN THE BIRD(1).*
 3-J) オンナ カ* トリ(1) ヨリ オ*オ*キイ。
 3-E) THE WOMAN IS BIGGER THAN THE BIRD(1).
 4-J) オンナ カ* トリ(1) ヨリ オ*モイ。
 4-E) THE WOMAN IS HEAVIER THAN THE BIRD(1).
 (c) 出力文-2

図3 図形パターンの自然言語理解

Fig. 3 Natural language understanding of a picture pattern.

来個別に取り扱われた自然言語理解とパターン認識とを融合させる役割を果たす。事象概念や属性概念に関する丹念な調査は、SUPPのこの方式が実験に用いた少数の事象や属性に対してだけでなく、要素的概念と直接対応するほぼすべての事象や属性に対して有効であることを示唆している。連結・合成概念は要素的概念から導けることも考えあわせると、SUPPは動詞と形容詞で表されるほぼすべての事象や属性を処理する(潜在的)能力をもつと考えられる。

(2) SUPPの拡張性と柔軟性

今回の事象から属性への拡張は意味辞書と構造辞書への項目の追加のみではほぼ実現できた。他の要素的事象/属性概念を追加する場合でもシステムの枠組に大幅な手直しが必要とは考えにくい。また辞書項目の追加の際はやはり概念調査の結果が明確な見通しを与えてくれる。

SUPPは他の研究成果に対しても柔軟である。濃淡画像、3次元画像あるいは動画像に対する進んだ処理技術は構造辞書の個々のサブプログラムに容易に反映

されうる。たとえば3.2.2項の*印の処理内容は、概念の特徴レベル以上の処理あるいは辞書に影響を与えることなく、3次元画像向きに改良されうる。一方種々行われている自然言語理解の研究に対しては生成論的な立場から寄与しうる。知識の系の階層構造はイメージデータからの自然語文の生成を可能にした。加えてシソーラス部門の基準概念は理解や推論を行うためのプリミティブを提供するとともに、シソーラス全体としては連想や推論のネットワーク作成のための基礎データを提供するものである。

4. むすび

第1報および第2報で述べた理論に従って日本語の形容詞で表される属性概念を分類し、これら一連の概念調査で得られた成果を図形パターンの自然言語理解へ応用した。本稿により自然言語および図形の統一的意味理解の研究を一步前進させることができた。今後の課題は形容動詞で表される属性概念についてさらに研究を押し進めることである。

なお本研究の一部は、文部省科学研究費特定研究59218019の補助を受けた。

参 考 文 献

- 1) 岡田直之, 田町常夫: 自然語および図形理解のための非単純事象概念の分析および分類, 電子通信学会論文誌, Vol. 56-D, No. 10, pp. 591-598 (1973).
- 2) 岡田直之, 田町常夫: 図形の意味解釈とその自然語記述—要素的図形認識と構造分析, 電子通信学会論文誌, Vol. J59-D, No. 5, pp. 323-330 (1976).
- 3) 岡田直之, 田町常夫: 動図形の意味解釈とその自然語記述—意味分析, 電子通信学会論文誌, Vol. J59-D, No. 5, pp. 331-338 (1976).
- 4) 岡田直之: 自然言語および図形理解のための形容詞で表される属性概念の分類—非単純概念の場合, 情報処理学会, 自然言語処理研究会資料, 38-1 (1983).
- 5) 岡田直之: 自然言語および図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. 25-31 (1985).
- 6) 岡田直之: 自然言語および図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).
- 7) 国立国語研究所(編): 分類語彙表, 秀英出版, 東京, p. 362 (1964).
- 8) 西尾寅弥: 形容詞の意味用法の記述的研究, 秀英出版, 東京, p. 467 (1972).
- 9) 西尾実他(編): 岩波国語辞典, 岩波書店, 東京, p. 1160 (1975).

(昭和59年7月5日受付)

(昭和59年11月15日採録)