CONCEPTUAL TAXONOMY OF JAPANESE VERBS FOR UNDERSTANDING
NATURAL LANGUAGE AND PICTURE PATTERNS

Naoyuki Okada
Department of Information Science
and Systems Engineering
Oita University
Oita 870-11, Japan

Summary
This paper presents a taxonomy of "matter concepts" or concepts of verbs that play roles of governors in understanding natural language and picture patterns. For this taxonomy we associate natural language with real world picture patterns and analyze the meanings common to them. The analysis shows that matter concepts are divided into two large classes: "simple matter concepts" and "non-simple matter concepts." Furthermore, the latter is divided into "complex concepts" and "derivative concepts." About 4,700 matter concepts used in daily Japanese were actually classified according to the analysis. As a result of the classification about 1,200 basic matter concepts which cover the concepts of real world matter at a minimum were obtained. This classification was applied to a translation of picture pattern sequences into natural language.

1 Introduction
As is generally known, the intellectual activities of human beings are very instructive in higher processing of natural language and picture patterns, especially real world picture patterns. There are three sides to intellectual activity:

1. Recognition and understanding.
2. Thinking and inference.
3. Expression and (intellectual) action.

The system of concepts or knowledge plays an essentially important role in each activity. The base of the system is considered to be placed on those concepts formed by direct association with the real world, which are closely related with both syntactic and semantic structures of natural language. The aim of this paper is to make this system clear from the linguistic viewpoint.

There are two linguistic approaches to the analysis of the system. One is the understanding of the outline of the whole system and the other is the detailed analysis of a small part of the system. Compilation of a thesaurus is considered of the former type. Thesauruses compiled so far, however, are not sufficient for machine processing because of the following:

1. Abstraction processes of concepts
2. Interrelation among concepts
3. Criterion for classification

Concepts of verbs are the core of the system from the linguistic viewpoint. We classify almost all concepts of verbs in daily Japanese by association of natural language with the real world, answering the above-mentioned problems. As for problem 1, a working hierarchy along an abstraction process is constructed in the system. As for problem 2, case frames are shown in "simple matter concept," and connecting relations among elementary matter concepts are shown in "non-simple matter concept." As for problem 3, an algorithm is introduced into the classification.
mantics because they have signs and syntactic relations. In the case of real world picture patterns, however, there exist neither signs nor syntactic relations. Here we observe real world objects named by human beings. If we consider them something like signs, we can think of the syntax, and then the semantics, too. The meanings are common to natural language and picture patterns, although their syntactic structures differ largely from each other.

2.2 Paradigms for Interpretation and Understanding

In order to clarify the notions of interpretation and understanding, first, we propose a working hierarchy of knowledge along the abstraction process, as follows:

- **Level 1** Raw data: Data close to copies of things and events in the real world. Image-like data.
- **Level 2** Data of visual features: Features extracted from raw data.
- **Level 3** Data of conceptual features: Symbolic data associated with visual features. Some of them correspond to Chomsky's syntactic features in the lexicon.
- **Level 4** Concept data: Data obtained by organizing conceptual features. Most data have names as words. In case of the verb they roughly correspond to Minsky's surface semantic frames.
- **Level 5** Interconnected concept data: Networks of concept data. A concept can be interconnected with other concepts from various viewpoints.

The knowledge system is so massive and complicated that it is necessary to make systematic analyses. Since the number of verbs are finite, concepts of verbs at level 4 provide a clue to systematic and exhaustive analyses of knowledge from the linguistic viewpoint.

The concepts of verbs are divided into two large classes: "simple matter concepts" and "non-simple matter concepts." 2,3

3 Simple Matter Concepts

The simple matter concepts are not reduced into any more elementary matter concepts while the non-simple ones are reduced. Most of them are so concrete that they are well analyzed by direct association with the real world.

3.1 Structural Patterns

An object in the real world identified by a verb is called "matter." Unlike things matter does not occur alone. It arises accompanied by things, events, and attributes, which are called "constituents," so this concept can be regarded as the concept of a dynamic or static relation among constituents and be expressed by:

\[ v(s, o, f, c, o_{m}, o_{s}, o_{s}, o_{p}, t, r, \ldots) \] (A)

where each symbol in parentheses represents a constituent specified below.

- **s**: subjective concept
- **o**: objective concept
- **f**: starting point in ac-

Fig. 1 Hierarchy of the knowledge system
tion, or initial state of change

\( o_1 \): finishing or target point in action, or final state of change

\( o_m \): opponent in mutual action

\( o_2 \): standard or reference

\( o_4 \): way or means (including instrument)

\( o_c \): concept which supplements attributive aspects

\( p,t,r,\ldots \): place, time, cause (or reason), ...

Out of these, eight constituents \( s \) through \( o_c \) are obligatory because they are indispensable for the recognition of matter. In Japanese sentences, the obligatory constituents are often accompanied with such postpositional words as \(-ga, -o, -kara, -ni, -to, -n\i, -ds, -de, -ki, -to\). But it is difficult to decide the case of a constituent only by such postpositional words.

The combination of obligatory constituents decides the basic frame of matter concepts. Table 1 was obtained after an elaborate investigation of more than 1,500 simple matter concepts. Two comments must be added to Table 1. First, optional constituents participate fairly freely in matter. Table 1 says nothing about this case. Next, some obligatory constituents are not obligatory in every case.

**Table 1**

| No. | Pattern | Example |
|-----|---------|---------|
| I   | \( v(s) \) | (konoha-ga oohiru) |
| II  | \( v(s, o_f) \) | (otoko-ga te-kara dersu) |
| III | \( v(s, o_t) \) | (toroko-ga yuubinkyoku-ni) |
| IV  | \( v(s, o_m) \) | (torokku-ga basu-to) |
| V   | \( v(s, o_g) \) | (ko-ya yaa-ni) |
| VI  | \( v(s, o) \) | (hanako-ga ringo-o) |
| VII | \( v(s, o, o_f) \) | (kansin-so-ga toomiri-o kurasuma-kara) |
| VIII| \( v(s, o, o_t) \) | (neto-ga kaban-ni kyosaku-o) |
| IX  | \( v(s, o, o_m) \) | (kikushi-ga kanaeatiko-to shing-o) |
| X   | \( v(s, o, o_g) \) | (hito-ga saji-de sate-o) |
| XI  | \( v(s, o, o_c) \) | (hito-ga soyokase-o suanakku) |
| XII | Others   | (Men feel a gentle breeze cool). |

**Table 2**

| No. | Semantic feature | Example |
|-----|-----------------|---------|
| 0:0 | Displacement    | oohiru(fall) | 319 |
| 0:1 | Change in the direction | mukeru (turn) | 54 |
| 0:2 | Deformation A | magaru (bend) | 183 |
| 0:3 | Spiritual change | okoru (get angry) | 128 |
| 0:4 | Sensual change | kanjiru (feel) | 50 |
| 1:00 | Deformation B | yaseru (get lean) | 22 |
| 1:01 | Change in quality | kurasu (rot) | 61 |
| 1:02 | Change in quantity | hiron (decrease) | 35 |
| 1:03 | Optical change | hikaru (flash) | 30 |
| 1:04 | Colour change | akumaru (turn red) | 29 |
| 1:05 | Thermal change | hieru (grow cold) | 34 |
| 1:06 | Change in force and energy | tsuyomaru (intensify) | 53 |
| 1:07 | Vocal change | utau (sing) | 52 |
| 1:08 | Occurrence, appearance and disappearance | anawaremu (appear) | 54 |
| 1:09 | Start, end and stop | tomoru (stop) | 21 |
| 1:10 | Temporal change | hoyamaru (hasten) | 28 |
| 2:0 | Continuation | tsusaku (continue) | 24 |
| 2:1 | State | soibamu (power) | 29 |
| 3:0 | Abstract | motsukatu (base) | 98 |
| 3:1 | Others | tabaru (eat) | 129 |

*Total 1,433*

There are 1,209 different concepts in the classified concepts.

This method was applied to the set of concepts described in Sect. 3.1 and the result is tabulated in Table 2. Here distribution was obtained by the classification of Chapter 5. In Table 2, the first digit 0, 1 and 2 in the classification numbers roughly represent movement, change, and state, respectively.
4. Non-Simple Matter Concepts

Generally, non-simple matter concepts are so abstract in comparison with simple ones that it is hard to show a clear association of natural language with the real world. We emphasize the analysis of how they are composed of simple ones.

4.1 Complex Concept A

If two elementary matter concepts \( v_i \) and \( v_j \) (not necessarily simple ones) are connected according to one of the rules shown in Table 3 and the connected concept is expressed by a Japanese complex word of two verbs for \( v_i \) and \( v_j \), it is called a "complex concept of A." The rules in Table 3 were obtained from the investigation of about 900 matter concepts which consist of two matter concepts and are expressed by a Japanese complex word.

In rule XXI.I, \( v_j \)("deru") is an upper-grade concept of \( v_i \)("afure-deru") and contains the contents of \( v_i \). Rule XXI.I is concerned with the whole and a part of the same matter, while rule XXI.II with two different matters. The former is considered as a special case of the latter in which two matters coincide with each other.

Rule XXI and XXII are logical while rule XXI[I is linguistic. As "cause" is one of the constituents in (A) in Sect. 3.1, XXI may be considered as a part of XXII.

The semantic contents of complex concept A consists of the \( v_i \) and \( v_j \) contents and their connecting relation.

4.2 Complex Concept B

Complex concept B consists of several elementary matter concepts and is usually expressed by a Japanese simple word. However, no general rule can be found to connect elementary matter concepts, so a hierarchical analysis was made for a small number of complex concepts of B as shown in Fig. 2 and Table 4.

From the diachronic point of view, there seems to be a reason why a complex concept of B is expressed by a simple word. The relation among elementary matter concepts can not well be expressed by enumerating each verb as in the case of complex concept A. When one is going to designate matter in the real world without the verb identifying it, one must utter several sen-

---

Table 3 Connecting rules of complex concept A

| No. | Connecting rule | Example | Remark |
|-----|-----------------|---------|--------|
| XXI | Cause and effect | (mizu-ga) afure-deru. | If water overflows, water comes out. |
| XXI.I | Implication | (afure-deru) | |
| XXII | Cause and effect | (dareka-ga watakshi -o) oski-taou. | If someone pushes me, I am thrown down. |
| XXII | Logical product | (sinsja-ga) fushigamu. | Believers kneel down and pray. |
| XXIII | Syntactic connection | (akago-ga) naki-yamu. | That a baby cries stops. |
| XXIII | Relation between s and v | (a baby) cry-stops. | An announcer misses to read his manuscript. |
| XXIII | Relation between o and v | (kano-a-ga genko-o) yomi-ayamaru. | A guard wakes prisoners by knocking them. |

Table 4 An analysis of complex concepts of B

| Complex concept | Relation among elementary concepts | Temporal shift |
|-----------------|------------------------------------|---------------|
| yuzuru(hand over) | [1] | [1] | [1] | [1] | [1] | [1] | [1] | [1] |
| ateaeru(give) | [2] | [2] | [2] | [2] | [2] | [2] | [2] | [2] |
| uru(sell) | [3] | [3] | [3] | [3] | [3] | [3] | [3] | [3] |
| orosu(sell by wholesale) | [4] | [4] | [4] | [4] | [4] | [4] | [4] | [4] |
| okuru(present) | [5] | [5] | [5] | [5] | [5] | [5] | [5] | [5] |
| kasu(lend) | [6] | [6] | [6] | [6] | [6] | [6] | [6] | [6] |
| amukeru(deposit) | [7] | [7] | [7] | [7] | [7] | [7] | [7] | [7] |

\( v_x \): There is no word to represent it.
### Table 5: Surface contents of complex concept B

| No. | Contents                     | Example                      | Dis. |
|-----|------------------------------|------------------------------|------|
| 10  | Spiritual act                |                              |      |
| 10-0| Thought-recognition          | mitome(remember)             | 35   |
| 10-1| Guess-judgement              | asasurau(guesse)             | 25   |
| 10-2| Respect-contempt             | yamamau(respect)             | 18   |
| 10-3| Naughty-flattery             | hikaraku(sport)              | 20   |
| 11  | Academic and artistic        |                              |      |
| 11-0| act                          | oshieru(teach)               | 33   |
| 11-1| Creation                     | arawasu(write a book)        | 11   |
| 12  | Religious act                |                              |      |
| 12-0| Belief                       | moderu(visit a temple or shrine) | 16   |
| 12-1| Celebration-marrige-funeral  |                              |      |
| 13  | Verbal act                   |                              |      |
| 13-0| Praise-blame                 | homereru(praise)             | 12   |
| 13-1| Instigation-banter           | odateru(instigate)           | 12   |
| 14  | Social act                   |                              |      |
| 14-0| Life                         | harasu(live)                 | 26   |
| 14-1| Fostering                    | yashinshu(bring up)          | 26   |
| 14-2| Antisocial-immoral           | makuru(steer)                | 43   |
| 14-3| Promise-negotiation          | suppaokau(break an appointment) | 35   |
| 15  | Conduct-behavior             |                              |      |
| 15-0|                              | aumaru(assume a prim air)    | 25   |
| 16  | Labour-production            |                              |      |
| 16-0| Labour-work                  | tawakomaru(serve)            | 35   |
| 16-1| Agriculture-industry-commerce| okinshu(deal in)             | 49   |
| 17  | Possession                   |                              |      |
| 17-0| Owning-abandonment           | yururu(own)                  | 11   |
| 17-1| Getting and giving-losing    | ataru(give)                  | 55   |
| 17-2| Selling and buying-lending-and-borrowing | kau(buy) | 19   |
| 18  | Investigation-measurement    |                              |      |
| 18-0| Investigation                | shiraberu(investigate)       | 24   |
| 18-1| Measurement                 | hakaru(measure)              | 19   |
| 19  | Domination-personal-affairs  |                              |      |
| 19-0| Domination-obedience         | sabaru(dominate)             | 32   |
| 19-1| Personal affairs             | yotarou(employ)              | 14   |
| 20  | Attack and defense-          |                              |      |
| 20-0| victory and defeat           | somaru(attack)               | 26   |
| 20-1| Attack and defense-victory   | mokaru(defeat)               | 19   |
| 21  | Refuge-escape                |                              |      |
| 21-0|                              | nigeru(escape)               | 22   |
| 22  | Rise and fall:prosperity and decline |                  |      |
| 22-0| Rise and fall                | horobasu(run)                | 11   |
| 22-1| Prosperity and decline       | akearau(prosper)             | 19   |
| 23  | Others                       |                              |      |
| 23-0|                              | moyoumu(hold a meeting)      | 333  |
| Total|                             |                              | 1041 |

### Table 6: Morpheme representing derivative operators

| No. | Derivative information | Example                   | Remark          |
|-----|------------------------|---------------------------|-----------------|
| 50  | Emphasis               | "tori"-shirakaru          | 101             |
| 50-0| Emphasis               | "tori"-shirakaru          | 55              |
| 50-1| Do completely          | odowki-"tori"             | 55              |
| 50-2| Do violently           | shikut-"tori"             | 61              |
| 51  | Respect-politeness     | mosharau                  | 51              |
| 52  | Vulgarity              | mawaru                    | 31              |
| 53  | Poor practice-failure  | nekuru-"tori"             | 2               |
| 53-0| Be ill able to do      | some-"aguru"              | 7               |
| 53-1| Lose a chance to do    | ku-"hagaruru"             | 7               |
| 53-2| Fail to do in part     | (eat-"miss")              | 7               |
| 53-3| Fail to do             | (write-"leak")            | 16              |
| 54  | Repetition-habit       | tot-"kaeru"               | 12              |
| 54-0| Do again               | (ask-"return")            | 29              |
| 54-1| Be used to do          | (eat-"stick on")          | 11              |
| 55  | Start                   | suri-"daaru"              | 9               |
| 55-0| Begin to do            | (rain-"come out")         | 9               |
| 55-1| Be just going to do    | (nasi-"kakeru")           | (say-"hang up") |
| 56  | Completion             | suri-"aguru"              | 23              |
| 56-0| Have finished          | (print-"go up")           | 5               |
| 56-1| Do from the beginning  | yomit-"tawu"              | 3               |
| 56-2| Have completed         | (read-"pass through")     | (do-"accomplish") |
| 57  | Limit                   | noritori-"hagaruru"        | 20              |
| 57-0| Do until the limit      | (climb up-"cream")        | 15              |
| 57-1| Do thoroughly           | ur-"kiro"                 | (sell-"cut")    |
| 58  | Others                  | tomashiro-"garu"          | 209             |

### Table 7: Derivative information

| No. | Derivative information | Example               | Dis. |
|-----|------------------------|-----------------------|------|
| 50  | Emphasis               | "tori"-shirakaru      | 101  |
| 50-0| Emphasis               | "tori"-shirakaru      | 55   |
| 50-1| Do completely          | odowki-"tori"         | 55   |
| 50-2| Do violently           | shikut-"tori"         | 61   |
| 51  | Respect-politeness     | mosharau               | 51   |
| 52  | Vulgarity              | mawaru                 | 31   |
| 53  | Poor practice-failure  | nekuru-"tori"         | 2    |
| 53-0| Be ill able to do      | some-"aguru"           | 7    |
| 53-1| Lose a chance to do    | ku-"hagaruru"          | 7    |
| 53-2| Fail to do in part     | (eat-"miss")           | 7    |
| 53-3| Fail to do             | (write-"leak")         | 16   |
| 54  | Repetition-habit       | tot-"kaeru"            | 12   |
| 54-0| Do again               | (ask-"return")         | 29   |
| 54-1| Be used to do          | (eat-"stick on")       | 11   |
| 55  | Start                   | suri-"daaru"           | 9    |
| 55-0| Begin to do            | (rain-"come out")      | 9    |
| 55-1| Be just going to do    | (nasi-"kakeru")        | (say-"hang up") |
| 56  | Completion             | suri-"aguru"           | 23   |
| 56-0| Have finished          | (print-"go up")        | 5    |
| 56-1| Do from the beginning  | yomit-"tawu"           | 3    |
| 56-2| Have completed         | (read-"pass through")   | (do-"accomplish") |
| 57  | Limit                   | noritori-"hagaruru"     | 20   |
| 57-0| Do until the limit      | (climb up-"cream")     | 15   |
| 57-1| Do thoroughly           | ur-"kiro"              | (sell-"cut")   |
| 58  | Others                  | tomashiro-"garu"       | 209  |

**Total**: 1041
tences. If such necessities often arise and the relationship is conceptualized, it will be efficient to give it a name.

As for semantic contents, elementary matter concepts and their relationship form a surface contents. Approximately 1,000 complex concepts of $B$ were investigated according to the feature extraction method in Sect. 3.2 and the result is tabulated in Table 5.

4.3 Derivative Concept

Some concepts possess a function of deriving a new concept by operating others. Matter concepts derived from operative concepts with both morphemic structures and derivative information as shown in Table 6 and 7 respectively are called "derivative concepts." Table 7 was obtained from the investigation of about 700 matter concepts, most of which are expressed by a complex word and one concept is operative to the other. The derivative information is very similar to the modal information of auxiliary verbs, but it differs in the sense that some matter concepts are operated upon and those operations are fixed.

5 Classification

In order to determine whether analyses in Chapter 3 and 4 are good or not, we classified about 4,700 basic matter concepts in daily Japanese, which are listed in "Word List by Semantic Principles" edited by National Language Research Institute in Japan.

5.1 Algorithm of Classification

An algorithm is introduced into the classification, referring Fig. 3 and 4. The elements or members of $V_B(x=x, y, \cdots)$ are denoted by $V_B(y=1, 2, \cdots)$ and the sum and difference in the set theory are denoted by + and -, respectively.

1) Preprocessing

For each $V_T$ of $V_T$,

1.1) examine whether $V_T$ functions with others or by itself. If it functions with others, then it is excluded from $V_T$.

Example. -goryu

1.2) examine whether there is $V_T(h-i)$ which has the same contents as $V_T$ and is expressed by the same verb as $V_T$. If there is such $V_T$, $V_T$ is excluded from $V_T$.

Let's denote a class of concepts excluded by 1.1) and 1.2) by $V_p$ and let $V_U=V_T-V_p$.

2) Classification of derivative concepts

For each $V_U$ of $V_U$,

2.1) if $V_U$ is expressed by a derivative word, it is classified as a member of term $L$ in Table 6. It is further classified in more detail according to Table 7;

2.2) if $V_U$ is expressed by a complex word of two verbs and one of these verbs is affixal, then it is regarded as a member of term $L'$ in Table 6, and classified in more detail according to Table 7;

2.3) if $V_U$ is expressed by neither a derivative word nor a complex word, but it is regarded as a member of one of the terms in Table 7, it is classified into the term $L'$ in Table 6.

Let this class of concepts thus obtained be $V_D$.

3) Classification of complex concepts of $A$

For each $V_U(i \neq V_B)$ of $V_U$, if $V_U$ is expressed by a complex word of two verbs and each concept functions by itself, it is considered as a complex concept of $A$ and classified according to Table 5.

The class thus obtained is denoted by $V_A$.

4) Classification of complex concepts of $B$

For each $V_B(i \neq V_A)$ of $V_B$, if its contents does not belong to any term in Table 2, it is regarded as a complex concept of $B$. The class thus obtained is denoted by $V_B$, and subject to the following process:

For each $V_B$,

4.1) examine its surface structure and classify it according to Table 1;

4.2) examine its surface contents and classify it according to Table 5.

Let $V_S=V_D+V_A+V_B$ and $V_S=V_U-V_S$.

5) Classification of similar concepts

In class $V_S$ of simple matter concepts, if there is a group with similar contents, choose a concept as the standard, then classify the re-
mainder as similar concepts.

Example. *Korogeru*(roll), *korobu*(roll), *makenobu*(roll), etc. are similar concepts for standard concept *korogaru*(roll).

Counter-example. *Saezuru*(chirp), *hoeru*(bark), *warau*(roar), *inanaku*(neigh), etc. are not similar concepts for standard concept *naku*(cry).

Here, it is assumed that if a certain concept is a standard concept, it is not a similar concept for another standard one at the same time.

The class of similar concepts thus obtained is denoted as $V_s$ and let $V_b = V_s - V_a$.

6) Classification of standard concepts

   6.1) examine its structural pattern and classify it according to Table 1;

   6.2) examine its semantic contents and classify it according to Table 2.

In the above process 2) through 6), one concept can be classified into two or more terms if necessary.

5.2 Results and Discussion

First, let's discuss the relation among the obtained classes along the abstraction process. There are two kinds of abstraction processes:

(i) extracting common features from concepts as follows; bulldog $\rightarrow$ dog $\rightarrow$ animal $\rightarrow$ living thing $\rightarrow$ thing,

(ii) connecting several concepts to form a new concept as shown in complex concept $B$. From the latter viewpoint, the relation among classes is schematized as indicated in Fig. 5.

Fig. 5 Relation among obtained classes

Simple matter concepts ($V_b$) are regarded as the base of matter concepts in the sense that $V_b$ covers the concepts of real world matter at a minimum and every other matter concept is led from $V_b$ by a rule. Two simple matter concepts are connected by a rule and form a little bit abstract concept or complex concept of $A$. Several matter concepts are organized by a fairly complicated rule into a new abstract concept or complex concept of $B$. One of the elementary concepts in a complex concept of $A$ changes its meaning diachronically and becomes a derivative operator. So, the system of Japanese verb concepts has its own nature—although it is a fact that a large part of the system is universal—and is not manipulated at one level.

Next, Table 8 indicates the distribution of all matter concepts. The minute distribution in each class has been shown in Table 2, 5 and 7, respectively. Table 8 is instructive in investigating the human competence in organizing the language system. For example, if class $V_b$ is regarded as "primitive" concepts, number 1,209 of $V_b$ does not side with Schank's classification, but with Minsky's idea. From Table 2, 5 and 7, we can measure the degree of human concern about real world matter. For example, term 0.0 in Table 2 shows human beings are most interested in displacements of objects.

Finally, we consider that every matter concept under consideration was classified satisfactorily supporting our analyses.

6 Translation of Picture Pattern Sequences into Natural Language

As an application of this taxonomy, system SUPP (System for Understanding Picture Patterns) was constructed. The overall system is shown in Fig. 6.

Fig. 6 Organization of the system

6.1 Knowledge System

The knowledge system consists of four components, visual, conceptual, linguistic and thesaurus. The visual component contains models of primitive pictures and syntactic rules, which correspond to level 2 data in Fig. 1. The rules are applied to picture pattern pairs called "before-after" frame pairs.

The conceptual component contains conceptual features, concepts, and networks of concepts,
which correspond to level 3 through 5 data, respectively. A matter concept is expressed by
\[ v : C1C2 \ldots Cn \text{(E1)}d_1(\text{E2}) \ldots d_m(\text{Em}) \] (B)
where each \( C_i \) denotes a feature of matter itself and is associated with a syntactic rule mentioned above. Each \( d_j(\text{E}) \) denotes the case or role of a constituent and must be filled by a specific instance or concept of constituent. Features \( \text{E}_{j1} \ldots \text{E}_{jn} \) specify the conditions its assignment must meet. A network is constructed among similar matter concepts.

The linguistic component consists of dictionaries for the production of Japanese and English sentences. The thesaurus component contains all the classified concepts in Chapter 5 and supports the development of other components.

6.2 Translation Process
A sequence of picture patterns, or two-dimensional line drawings (handwriting is allowed), is input at time \( t_0, t_1, \ldots, t_n \). A picture pattern at \( t_i(0 \leq i \leq n-1) \) is paired with the one at \( t_{i+1} \), and processed as follows:

1) Primitive picture recognition and syntactic analysis
The picture pattern reader is a curve follower that traces line segments by octagonal scanning. The recognizer is based on Evans's matching program for graph-like line drawings but is improved to handle noisy ones.\(^{13}\)

The syntactic analyzer A decomposes the complex picture, in which two or more primitive pictures may intersect or touch each other, and recognizes them according to Gestalt criteria. The syntactic analyzer B performs Boolean operations on quantized primitive pictures to check such a relation as "MAN INSIDE HOUSE." The syntactic analyzer C performs numerical operations on the data such as coordinates and transformational coefficients of primitive pictures.

2) Semantic analysis and inference
The semantic analyzer detects the meaning of matter-centred change in picture pattern pairs by top-down analysis. Suppose that matter \( v(s, o_0) : C1C2 \ldots Cn d_1(\text{E}0) d_2(\text{E}0) \ldots d_m(\text{Em}) \) is directed by the Inference. The analyzer assigns the role of \( s \) to one of the primitive pictures, say \( P_s \), after checking whether \( P_s \) meets \( o_0 \). It assigns the role of \( o_0 \) to another primitive picture \( P_{o_0} \) in the same way. Then it analyzes each \( C_i \) by calling a correspondent sub-program in the syntactic analyzer B or C. If all the analyses end in success, the meaning of \( v(s, o_0) \) is detected. The present Inference makes inferences about all the similar concepts in the network in depth-first order, directing each matter concept at a node to the semantic analyzer.

Finally, the synthesizer produces Japanese and English simple sentences.

6.3 Experiments
All the programs except the picture pattern reader are written in Fortran and run under the OS/VS of the FACOM 230-38S medium-scale computer at Oita University. Running with the syntactic analyzer B and C, the semantic analyzer occupies approximately 200K bytes of core. Memory usage for all the dictionaries except the thesaurus component amounts to approximately 90K bytes.

Fig. 7 and 8 indicate an example of the recognition of a primitive picture and the translation of a picture pattern pair, respectively. It took 47 seconds to recognize "bird[1]" in Fig. 7 and 60 seconds to analyze and infer the meanings of matter after the recognition of primitive pictures in Fig. 8.

Katz and Fodor pointed out the three problems of a semantic theory: (i) Semantic ambiguity, (ii) Semantic anomaly, and (iii) Paraphrase.
As for (i) it is important to enumerate all the readings of the input picture pattern pair. The output sentences in Fig. 8 shows SUPP understands to a fair degree the change in the input. As for (ii) the ability in detecting semantic anomaly is important. SUPP checks it by Eq (B) in Sect. 6.1, but a little bit anomalous sentence 5) or 7) is output because the constructed dictionary of matter concepts is slightly insufficient. As for (iii) the ability in paraphrasing sentences is needed. Output sentences 4) through 7) are an analytical paraphrase of "THE BIRD[1] PERCHES ON THE TREE" although SUPP has no knowledge about "perch."

7 Conclusions

A taxonomy of Japanese matter concepts has been described. It is summarized as follows:

Simple matter concept
- Standard concept
  - structural pattern ; 12 types
  - semantic contents ; 20 features
- Similar concept

Non-simple matter concept
- Complex concept A
  - connecting rule ; 3 types
  - semantic contents ; each contents of elements and their connecting relation
- Complex concept B
  - connecting rule ; different in each concept
  - surface contents ; 14 features
- Derivative concept
  - derivative operator ; 3 types
  - derivative information ; 9 features

This taxonomy has made clear the outline of the system of all matter concepts in daily Japanese, and by SUPP picture pattern understanding research has come closer to natural language understanding research.

Acknowledgement

The author started his investigation of Japanese matter concepts and the development of SUPP some ten years ago when he was at Kyushu University. The author wishes to express gratitude to Prof. T. Tamati of Kyushu University for his kind guidance and material support.

References

[1] Okada, N. and Tamati, T.: Semantic Information of Natural Language and its Extraction and Classification, Trans. IECE, Japan, 52-E, 10, p. 363 (Oct. 1969).
[2] Okada, N. and Tamati, T.: An Analysis and Classification of "Simple Matter Concepts" for Natural Language and Picture Interpretation, Trans. IECE, Japan, 56-D, 9, p. 523-530 (Sept. 1973).
[3] Okada, N. and Tamati, T.: An Analysis and Classification of "Non-Simple Matter Concepts" for Natural Language and Picture Interpretation, Trans. IECE, Japan, 56-D, 10, p. 591-598 (Oct. 1973).
[4] National Language Research Institute (ed.): "Word List by Semantic Principles," Syuëi Syuppan, Tokyo, 1964.
[5] Roget, P. (Brownie, D.C. (ed.)): "Thesaurus of English Words and Phrases," J.M. Dent and Sons Ltd, London, 1971.
[6] Chomsky, N.: "Aspects of the Theory of Syntax," The M.I.T. Press, Cambridge, Mass., 1965.
[7] Minsky, M.: A Framework for Representing Knowledge, in Winston, P.H. (ed.), "The Psychology of Computer Vision," McGrai Hill Book Co., New York, 1975.
[8] Schank, R.C. and Abelson, R.P.: Scripts, Plans, and Knowledge, Proc. 4ICCAI, Tbilisi, Georgia, 1975.
[9] Schank, R.C. and Rieger, C.T.: MARGIE: Memory, Analysis, Response Generation, and Reference on English, Proc. 4ICCAI, Stanford, 1973.
[10] Okada, N. and Tamati, T.: Interpretation of the Meaning of Picture Patterns and its Description in Natural Language—Primitive Picture Recognition and Syntactic Analysis, Trans. IECE, Japan, J59-D, 5, p. 323-330 (May 1976).
[11] Okada, N. and Tamati, T.: Interpretation of the Meaning of Moving Picture Patterns and its Description in Natural Language—Semantic Analysis, Trans. IECE, Japan, J59-D, 5, p. 331-338 (May 1976).
[12] Okada, N.: SUPP: Understanding Moving Picture Patterns Based on Linguistic Knowledge, Proc. 6ICCAI, Tokyo, 1979.
[13] Evans, T.G.: A Program for the Solution of Geometric Analogy Intelligence-Test Questions, in Minsky, M. (ed.): "Semantic Information Processing." The M.I.T. Press, Cambridge, Mass., 1968.
[14] Katz, J.J. and Fodor, J.A.: The Structure of a Semantic Theory, Language, 39, 2, p. 170-210 (1963).