

LINGUISTIC MODEL BASED ON THE GENERATIVE TOPOLOGICAL INFORMATION SPACE

Seiichi Uchinami and Yoshikazu Tezuka
Faculty of Engineering, Osaka University
Yamada-Kami, Suita, Osaka, 565 JAPAN

Based on the *structuralism*, we propose a generative semantic model which has a topological information space generative grammar as basic rules. In this model a semantic map which is called a topological information space, is generated by the grammar, and the space can express implications and similarities among concepts. In the syntax, a syntactic generative grammar is defined based on the space grammar, and a mapping from the map to the language is defined. The mapping is composed of two mappings: one is a meaning affix mapping Φ which maps a conceptual area in the space to a token in the language, and the other is an operator mapping Ψ which maps a generative rule in semantics to a rewriting rule in syntax. By these mappings, a derivation tree in semantics is mapped to a derivation tree in syntax, and vice versa.

An algebraic system on the space is defined, and an algebraic system on the sentences is derived by the (Φ, Ψ) -mappings. English will be analyzed according to this model and the algebraic systems on them.

Finally an information processing model is described based on the model. The information processing in a natural language is carried out in the following steps: recognizing the inputs, parsing, interpreting, deducing, updating them, and outputting. These processes are discussed in details.

1. INTRODUCTION

Linguistics has been investigated from three aspects: *syntax*, *semantics*, and *pragmatics*.

As for syntax, many studies have been done since Chomsky presented a generative grammar based on *structuralism*, and splendid accomplishments have been made.

As for semantics, some reports have been published. In the *standard theory* and the *Case grammar*, syntactic rules are considered to be powerful enough to describe and manage meanings, but it is insufficient to describe meanings completely with the discrete and branching rules, because meanings have continuous properties.

In the *stratificational grammar* and the *generative semantics*, semantic components seem to be more important than in the standard theory, because semantic rules are used basically. But they have the similar defects as the standard theory, because these rules are discrete and branching ones.

Lakoff and McCawley's *generative semantics* model are not enough to describe meanings, because their semantic rules are treated with the same form of syntax.

In general, their notations adopt classifications based on the categories or binary classifications. So it has the following weak points: lack of expressive power of similarity and distance, and ordering of property value of concepts.

We have proposed to adopt a high dimensional semantic map (i.e. a *topological information space* complementing for these defects.

It is said that when one wants to express one's intention, not a sentence but a concept is formed in his brain at

first, then it is translated into a sentential form, feeding back the difference between the concept and the sentence translated, and at last it is expressed as a speech or a writing.

Therefore we have proposed the model that the concept which we want to utter is formed at first as a specified area in the semantic information space, next it is transformed into a phrase-marker (P-marker) in the deep structure, and at last it is transformed in the sentential form in the surface structure being fed-back in each stage.

2. OUTLINE OF THE LINGUISTIC ACTIVITY MODEL

There are two ways in describing meanings: one is connotational semantics; in this case meanings of sentences in a language A are described by sentences in a language B, for example, English is translated into Japanese.

The other is the operational semantics; in programming language, input data are computed and the results are output. So the meanings of a programming language is expressed as an operator that prescribe input-output relation.

In our model these two semantics are considered as follows. As for the connotational semantics, we take the topological information space as the language that describes meanings.

In this model we express the meanings of a natural language in the following way: we locate the concepts in the semantic map, i.e. the topological information space as follows; if a concept A implies a concept B, we assign them as the area B contains the area A, and if a similarity (or distance) between a concept A and B is greater than a similarity (or distance) between a concept A and

C, we locate the area B nearer (or farther) the area A than C.

As for the operational semantics, there are two kinds of operators. One is responses or reactions against the received message, the other is updatings or rearrangements of the memory. The former is carried out by responding with a speech, writing, or actions. The latter is carried out by updating the database.

The model has four description levels:

- (1) *Expression level* (utterances, writings, and actions)
- (2) *Surface level* (sentences)
- (3) *Deep level* (phrase-markers)
- (4) *Semantic level* (information space).

In the expression level, one's intention is expressed through one's organs.

The surface level is the level of the language which we use in the daily life, and the set of sentences compose elements in this level.

In the deep level, each element is a generator of a surface description, and is in the form of P-marker, i.e. so called a derivation tree.

In the semantic level, the concept is expressed as an area or a file of areas in the information space, where the information space is something like a semantic map, and is made up of the neighbourhood system or the metric space, and has topological properties, which can express distances, similarities and inclusions so on.

In the functional point of view, the information space is composed of a *meaning subspace* and an *estimating subspace*, where the *meaning subspace* expresses topological relations among concepts, the *estimating subspace* expresses various estimated values in the process of management. The estimating subspace has an activity axis, an occurrence probability axis and so on.

There are three levels of processing on them as follows:

- (1) *Expression processing*
(expression level \leftrightarrow surface level \mathfrak{D}),
- (2) *Syntactic processing*
(surface level \leftrightarrow deep level \mathfrak{D}),
- (3) *Semantic processing*
(deep level \leftrightarrow semantic level \mathfrak{D}).

The expression processing is to utter, write or act, and in the opposite direction to recognize the received signs. The syntactic processing is mappings from the deep level to the surface level and vice versa, in other words, mappings from P-markers to the sentences and the mappings among P-markers.

A *syntactic structure* consists of two level elements and the algebraic structures on them.

The semantic processing is mappings from the semantic level to the deep level and vice versa, and mappings among the semantic level, in other words map-

pings from domains in the information space to P-markers and vice versa, and the mappings among domains in the information space. A *semantic structure* consists of two level elements and the algebraic structures on them.

The correspondence between the semantic structure and the syntactic structure is as follows. There are two mappings which connect two structures, one is a *meaning affix mapping* Φ that maps an area in the space to a word in the syntactic grammar, the other is an *operator mapping* Ψ that maps an operator on the space to a production rule in the syntactic grammar.

The algebra on the syntactic and the semantic structure are defined.

3. TOPOLOGICAL INFORMATION SPACE

In the syntax-generative grammar, it is of no use to introduce relations among sentences, because there are no semantic relations among sentences which are generated by the grammar. But in the semantic-generative grammar, not only elements but also a topology over them should be generated. Because interpreting meanings of sentences, distances, similarities and inclusions among concepts which are expressed by points or areas, are very important.

In an algebraic point of view, a topological space is defined by an ordered pair (X, T) , where X is a support (set of points), and T is a topology on X .

The whole topological space is generated as a dictionary, or a thesaurus, or the universe of knowledge. In conversation with someone or a description of one's concept, a particular area is pointed out to communicate one's intention with others. Then we need two kinds of production rules: one is the *universe generative rules* which generate the whole topological space, i.e. elements set and a topology on them, and the other is the *utterance generative rules* which generate distinct conceptual areas.

In terms of sorts of topology, we investigated topology-types which are required to express properties of meanings of about one thousand Chinese characters for ordinary use. Following types of topology are listed up;

1) *Linear number axis*. In this type, a property value is expressed in a numerical value such as an integer or a real number, e.g. a weight axis: 150 lbs.

2) *Cyclic number axis*. Property values are expressed in numerical values in a given closed interval, and both end points are identified, i.e. a number sequence is defined cyclically, e.g. a time axis; one o'clock, two o'clock, .. twelve o'clock, one o'clock, ..

3) *Ordinal axis*. Property values can not be expressed in numerical expression,

but property values are lined up in order along its strength and weakness, e.g. a sensation of warmth axis; cold cool, comfortable, warm, hot, heated.

4) *Cyclic ordinal axis*. Property values are lined up in order cyclically, for example, a season axis; spring, summer, autumn, winter, spring,...

5) *More general topology*. Property values can not be expressed in type 1)-4), and are given weaker topology as near-and-far relations or similarity relations among elements, for example, chromaticity diagram on color; yellow, red, blue, green, etc.

6) *Simple set*. There is no topology on elements, so only a medley of elements constitutes a set, e.g. a set of terms which expresses aspects or units.

We'll formalize the model that can cover all sorts of topology listed above.

A topological information space generative grammar (TISG) consists of three levels of production rules which are restricted in the rewriting order.

Def.1 Space Constitution Grammar : G_1
 $G_1 = \langle V_{N1}, V_{T1}, P_1, S_1 \rangle$ (1)

where $V_1 = V_{N1} \cup V_{T1}$, $V_{N1} \cap V_{T1} = \phi$ (2)

$D = \{ \cdot, \#, | \}$ (3)

$L^\Delta(V, D) = \{ (\alpha\beta)^*\beta \mid \alpha, \beta \in V, \delta \in D \}$ (4)

V_1 is a finite set (vocabulary), V_{N1} and V_{T1} are called a nonterminal vocabulary, a terminal vocabulary respectively. S_1 is a subset of V_{N1} (initial vocabulary), and P_1 is a finite set of production rules (rewriting rules) of the form: $P_1 : A \rightarrow \omega$ (5)

where $A \in V_{N1}$, $\omega \in L^\Delta(V_1, D)$.

D is a space constitutor set, and consists of *direct product*: \cdot , *connected sum*: $\#$, and *direct sum*: $|$.

Def.2 Space Property Grammar : G_2
 $G_2 = \langle V_{N2}, V_{T2}, P_2, S_2 \rangle$ (6)

where $V_2 = V_{N2} \cup V_{T2}$, $V_{N2} \cap V_{T2} = \phi$.

property production rule P_2 is a finite set of production rules of the form:

$P_2 : A \rightarrow \omega$, where $A \in V_{N2}$, (7)

$\omega \in \{ a, \dot{a}, aa, aa^{-1}, aba^{-1}b^{-1}, aba^{-1} \}$, $a, b \in V_{T2}$. (8)

A grammar G_2 generates properties of subspaces. Sorts of properties are as follows:

- (i) *Line axis* (a)
- (ii) *Cyclic axis* (a)
- (iii) *Sphere axis* (Sphere) (aa^{-1})
- (iv) *Projective axis* (aa)
(Projective plane)
- (v) *Torus axis* (Torus) ($aba^{-1}b^{-1}$)
- (vi) *Boundary axis* (aba^{-1})
(Surface with boundary)

where type (iii)-(vi) are of polyhedral forms of manifold.

Def.3 Topology Prescription Grammar: G_3
 $G_3 = \langle V_{N3}, V_{T3}, P_3, S_3 \rangle$ (9)

where $V_3 = V_{N3} \cup V_{T3}$, $V_{N3} \cap V_{T3} = \phi$.

P_3 consists of two parts, one is P_{3T} which generates the universe, the other is P_{3A} which generates a specified area to point out a concept.

$P_{3T} : A \rightarrow \omega$ (10)

$\omega \in \{ \text{topological spaces of each types} \}$

$P_{3A} : A \rightarrow \omega$ (11)

$\omega \in \{ \text{areas of topological spaces} \}$

G_3 generates the topology of each subspace, then following types of topologies are adopted adaptively: $T_0, T_1, T_2, R, T_3, CR, T, N, T_4, T_5, T$, Compact, Metric, Uniform, Separable^V and so on.

Def.4 The Topological Information

Space Generative Grammar : G_I
 $G_I = \langle G_1, G_2, G_3 \rangle$, (12)

where G_1, G_2 and G_3 are sub-grammars defined above. And the relations among three levels are as follows:

$S_1 = S, S_2 = V_{T1}, S_3 = V_{T2}$.

Def.5 The Topological Information
Space : $I, I = I(G_I)$ (13)

The topological information space generated by G_I is written as eq.(13).

Def.6 The Syntax Generative Grammar: G_S
 $G_S = \langle V_N, V_T, R, S \rangle$ (14)

where S is an initial vocabulary, and subset of V_N , R is a set of phrase structure rewriting rules.

The other side V is divided into two: $V = V_p \cup V_g$, where V_p is *significant vocabulary*, and V_g is *grammatical vocabulary*. An intersection of V_p with V_g is not always null.

A generation tree, i.e. derivation tree of a sentence is called a *P-marker* (*phrase marker*), and the set of all P-markers generated by G is written as $P = P(G)$. The terminal string extracted from the treetop of the P-marker is called a *sentence*, and the set of all sentences generated by G , is written as $L(G)$, and is called the *language* generated from the grammar G .

The P-marker sets and the set of all sub-P-markers gotten by decomposing P are called *universe of P-markers*, and written as \hat{P} .

Def.7 The Semantic Algebraic System: \mathcal{U}_I
The semantic algebraic system \mathcal{U}_I is defined as

$\mathcal{U}_I = \langle I, I_I, E_I \rangle$ (15)

where I is the information space defined by G_I , and I_I, E_I are the interior and exterior operators on I respectively.

Generally speaking interior operators are functions $I^n \rightarrow I$ ($n \geq 2$), and exterior operators are functions $I \times \Omega \rightarrow I$, where Ω is an exterior operator.

Def.8 The Syntactic Algebraic System: \mathcal{U}_S
The syntactic algebraic system is defined as

$\mathcal{U}_S = \langle P, I_S, E_S \rangle$ (16)

where \hat{P} is a set of P-markers generated by G , and I_S and E_S are the interior and exterior operators on \hat{P} .

Def.9 The meaning affix mapping : Φ
 $\Phi' : \hat{D} \rightarrow V_p^+$, (17)

where \hat{D} is the set of all areas in the space I. Φ' assigns an area (a concept in semantic map) to a token or string of tokens (words) in the syntactic level. We interpret a sentence through used-rewriting rules, so in order to avoid ambiguity we extend V_p^+ to \hat{P} , i.e.

$\Phi : \hat{D} \rightarrow \hat{P}$ (18)

Def.10 The operator mapping : Ψ
 $\Psi : O_I \rightarrow O_S$ (19)

where $O_I = P_i \cup \Omega \cup I_I \cup E_I$ (20)

$O_S = R \cup V_G \cup I_S \cup E_S$ (21)

Ψ assigns the semantic operator in the semantic space to the syntactic operator in the syntax.

4. THE SEMANTIC ALGEBRAIC SYSTEM IN ENGLISH

We will show the semantic algebraic system in English briefly. Fig.1 shows the outline of G_I .

[1] The organization of the support in English --- organization of the meaning subspace ---

The meaning subspace is used to express the topology among concepts. It consists of three subspaces: *intellectual, feeling, and ordering subspaces.*

The intellectual subspace is a field to represent, state and judge. The feeling subspace is a field to express our actions, emotions and senses. The ordering subspace is a field to express the ordering of events or the informations.

(1) The intellectual subspace

In English there are four typical intellectual subspaces: {1} nouns or noun phrases, {2} adjectives or adjective phrases, {3} adverbs, {4} verbs. These four subspaces are written as I_N , I_A , I_D , and I_V respectively.

They constitute the support of the system, but some of adverbs correspond to the operators, for example, "very". These four subspaces are composed to the higher level by nesting each other.

(2) The feeling subspace

The feeling subspace is divided into

P_1 : <information space> \rightarrow <meaning ss>•<estimating ss>
 <meaning subspace> \rightarrow <intellectual ss>•<feeling ss>•<ordering ss>
 <estimating subspace> \rightarrow <reliability ax>•<activity ax>•<occurrence probability ax>•<...> ..
 <intellectual subspace> \rightarrow <noun ss>|<adjective ss>|<adverb ss>|<verb ss>
 <feeling subspace> \rightarrow <action driving ss>•<emotional ss>•<sense ss>
 <action driving subspace> \rightarrow <reply ss>•<action ss>
 <emotional subspace> \rightarrow <state of emotions ss>•<degree of emotions ss>
 <ordering subspace> \rightarrow <events ordering ss>•<causality ordering ss>
 <events ordering subspace> \rightarrow <seasons>|<months>|<times>|<...>|...

P_2 : <reliability axis> \rightarrow I, ...
 <seasons> \rightarrow S, ... (where ss means "subspace", and ax means "axis")

P_{3T} : I \rightarrow [0,1], ... P_{3A} : I \rightarrow x, (0 \leq x \leq 1)
 S \rightarrow <spring> <summer> <autumn> <winter> S \rightarrow <spring>|<summer>|<autumn>|<winter>

Fig.1 A part of the rewriting rules in the information space generative grammar

three subspaces: *action driving, emotional and sense subspaces.*

(i) The action driving subspace

This subspace expresses forms of actions and action-driving potentials. The forms of actions are imperative, prohibition, request, hope, desire, petition, question, doubt, obligation, permission, causative, voluntary and intention so on.

This subspace is divided into the *reply* and the *action subspaces*, in the functional point of view. The reply subspace expresses " *answer by language* ", and the action subspace expresses the *actions except utterance*, e.g. eat.

The words are mapped from the areas in the space by the meaning affix mapping Φ and the operator mapping Ψ as follows: if one receives words which arouse actions, one finds *form of actions* in this subspace by Φ^{-1} , and finds *driving potentials* by Ψ^{-1} .

The meaning affix mapping Φ is defined on the following types of words: verbs which arouse actions (for example, ask, order etc.), imperative and interrogative forms of verbs, and auxiliary verbs of prohibition and wish so on.

The meaning affix mapping designates only " *forms of actions aroused* ", and " *the content of actions* " is designated by the area in I. For example, imperative sentence " \overset{v} Read " is managed as follows. The form of action is designated " *ordinary imperative* ", and an action " *to read* " is designated as a corresponding area in I_V subspace.

The operator mapping Ψ has following functions. As for verbs listed above, the aroused degree of the action is designated. And if there is an adverb that modifies a verb, and its adverb expresses the degree, Ψ is designated by the degree expressed by the adverb, for example, a word " *absolutely* " designates " *extreme* " degree.

(ii) The emotional subspace

The emotional subspace is used to express emotions of a speaker and a hearer, and to modify a criterion level of decision of taking action, and this subspace does not arouse immediate actions.

This subspace is divided into two: the state of emotions and the degree of emotion's subspaces.

Usually the area is computed both on the subject and the indirect object. Sorts of emotions are joy, anger, pity, fear, love, hatred and sadness etc.

Φ is computed by adjectives, verbs and nouns that express emotions, and interjections. And Ψ is determined by the degree of words listed above, adverbs, and emphasizing styles.

(iii) The sense subspace

Kinds of senses are the senses of sight, hearing, smell, taste, touch, and internal organs. These senses are awake by direct excitations and by particular words in conversation. In case of words, there are two types as follows:

{1} The words which designate a certain sense area in the space, for example, dazzling, hungry. These words are mapped by Φ^{-1} into the sense subspace.

{2} The words which activate the potential of a certain area in the sense estimating subspace. For example, a word "rain" arouses the sense of cold and chilly in Japan. The sense potential is raised by these words, and the degree is computed by Ψ^{-1} .

The stimuli from the five organs of sense are gained by actual experiences, but the stimuli by words request to image the same experience as a speaker.

(3) The ordering subspace

This subspace expresses ordering series of domains in time or in logic. Usually the ordering axis is a real time axis of $(-\infty, \infty)$, and is divided into three areas: past, present, and future. But present is the relative value, and has meaning of only a time base.

(i) Ordering the events

This designates the occurrence time of the contents described in the intellectual and feeling subspaces.

There are two kinds of time assignment, one is simple: past, present, or future. The other is composite: perfect, simultaneous, or delay. The composite assignment designates ordering in time or in logic among clauses or events.

(ii) Ordering the causality

As for logic, input clauses are classified into two groups: one is a group of assumption or causes, the other is a group of conclusion or effects. And a suitable causality type is assigned to each clause.

[2] OPERATORS IN ENGLISH

(1) The intellectual subspace

(i) The exterior operators The exterior operators modify or restrict the conceptual area in a specified axis, and there are two types as follows: the value-shift type and the quantifier type. For example, let $u \in V_p^+$, $v \in V_G^+$, and d :

an area in the space, ω : exterior operator on the space, if $\Phi^{-1}(u)=d$ and $\Psi^{-1}(v)=\omega$, then the area corresponding to a phrase "uv" is formed by " ωd " in the space.

There are four types in the exterior operators as follows:

- (a) Restricting to limit, e.g. very
- (b) Settlement of range,
 - (i) Extension of range, e.g. all
 - (ii) Reduction of range, e.g. small
- (c) Inversion of range, e.g. unpleasant
- (d) Shift or limit the value of reliability in the estimating subspace, e.g. infer.

Sorts of limiting are conclusion, presumption, hearsay, degree, possibility, question, negation & emphasis.

(ii) The interior operators

There are four levels in the interior operators as follows: words, phrases, clauses and sentences. Let $a, c \in V_p^+$, $b \in V_G^+$, $\Phi^{-1}(a)=\alpha$, $\Phi^{-1}(c)=\gamma$, $\Psi^{-1}(b)=\beta$, $\alpha, \gamma \in D$, $\beta \in \{\text{binary operators over the space}\}$. Then $\delta = \Phi^{-1}(abc) = \Phi^{-1}(a) \cdot \Psi^{-1}(b) \cdot \Phi^{-1}(c) = \alpha\beta\gamma$.

{1} The interior operators common to each level. (space) (operators) (syntax) synthesis of areas enumeration in, by

- Sum: parallel logical sum
ex. red and white tulips,
- Selection: selection exclusive or
ex. either you or I,
- Product: series logical product
ex. beautiful and healthy.

{2} The interior operators in words

Compound words in syntax are corresponding to these operators. Usually the operator is designated by a rewriting rule in the syntax, and V_G does not appear in the compound words. This operator limits the area expressed by the main word in the compound word by the area expressed by the subordinated word. For example, the conceptual area of "letter paper" is the restricted area of "paper" for a letter.

{3} The interior operators in phrases

These operators resemble to the operators in the words. But in this level, not only rewriting rules but also V_G such as "the monkey in a cage" designates an operator too. In case of noun phrases there are two kernels: V_p and one V_G , and these are mapped to one V_p area. In case of adjective phrases or adverb phrases such as "in a cage", there are one kernel V_p and one V_G , so the lack of V_p causes that these phrases are mapped to unary operators in the space.

A phrase has a case, and each phrase is attached a pair of the concept and its case.

{4} The interior operators in clauses

The process of composing clauses from words and phrases is corresponding to the unification of some conceptual areas as follows. There are three main

concepts: subject, object and action. And there are three types of unifications as follows;

**Mutual dependency*: composition of transitive verb with direct object, and composition of direct object with indirect object.

**Uni-directional dependency*: composition of the subject and verb (phrase).

**Mutual independency*: composition of two clauses by a coordinate conjunction.

{5} The interior operators in sentences

There are two methods of composition by a coordinate junction or nesting.

(2) The feeling subspace

This information does not always appear, but if it appears it designates the state of feeling between the sender and the receiver.

(3) The ordering subspace

(i) The exterior operators

A shift or limitation of the value in the ordering axis is designated by the inflection of verbs and tense.

(ii) The interior operators

Conjunctions designate interior operators. The operators are classified into subjunctive, indicative and imperative groups in the modal point of view. Also operators are divided into copulative, alternative, adversative, disjunctive and illative conjunctive groups in the conjunctive point of view.

5. INFORMATION PROCESSING MODEL ON THE INFORMATION SPACE MODEL

Our communication is carried out bi-directionary. On sending we compose a concept which we want to communicate others, next point out a corresponding area in the topological information space, map an area or areas to tokens in the syntax by the meaning affix mapping ϕ or operator mapping ψ , form the P-marker in the deep level, get the sentences, and at last speak or write sentences.

Descriptions of entities and relationships are managed as follows:

(1) when we quote one entity, or when we express one entity's properties, a particular area in the information space is pointed out.

(2) when we express two or more entities and relationships among them, a direct product of two or more areas of entities, and an area of the relationship is pointed out. Both entity-subspace and a relationship-subspace are topological information spaces. Then the composed space become a topological space too.

On receiving we listen or read the sentences, analyze them lexically, parse them, get the P-marker in the surface level, transform it to the deep P-marker, map it to the derivation tree of the topological information space generative grammar, find the corresponding areas, compose the areas and relationship among them, interpret the message, update the memory, and respond the required actions.

These two processes are inverse relation, so we here describe in detail these processes in the latter case.

We depict processing steps in Fig.2.

[1] Lexical analysis and Parsing

In this step syntactic processing is carried out. There are four recognition levels of syntax as follows:

(1) Recognizing phonemes & characters

The received phonemes or characters are checked whether they belong to permissible phonemes or characters.

(2) Lexical analysis Input strings are divided by the word delimiters, and recognized as words. Words are checked whether they are contained in the token-dictionary or not.

(3) Parsing & Deriving P-markers

Delimiters of sentences are searched and each sentence is parsed. If ambiguity occurs in parsing, all feasible P-markers are listed up.

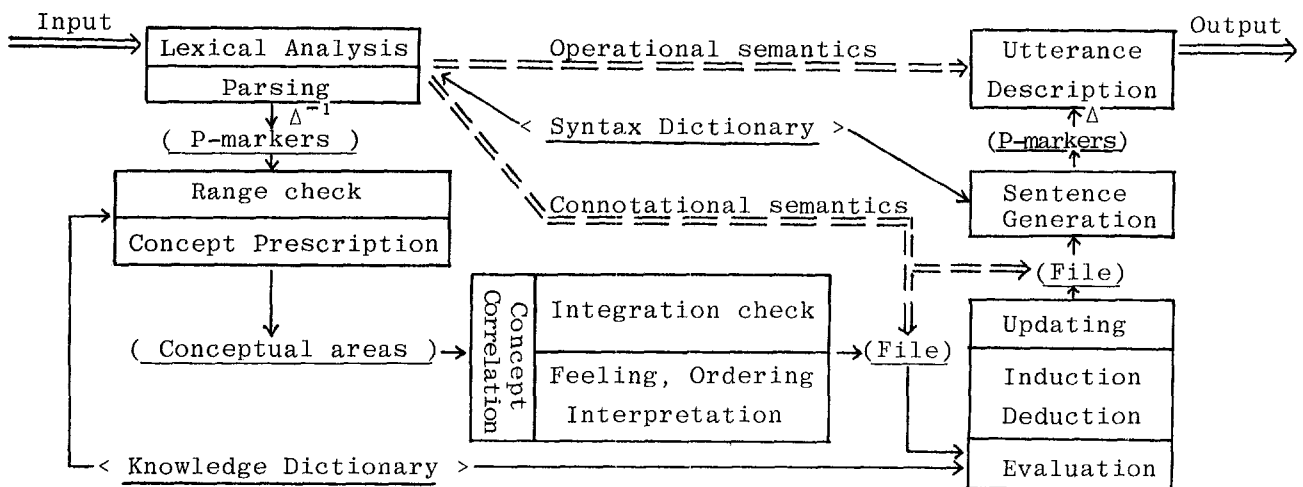


Fig.2 The Information Processing Model

(4) Reduction of P-markers to generator P-markers The parsed P-marker in the surface level is transformed into a generative P-marker in the deep level.

If recognition is failed in each step, the degree of potential for question is increased. And if the degree of potential for question exceeds the threshold, a question to resolve the ambiguity is formed, and is uttered.

[2] Range check

The value of property of the input data is checked whether it is out of range of permissible interval or not, i.e. range check of property value is carried out. Because in the information space, the domain of property value is defined by the topological information space generative grammar. This process is divided into two sub-processes: (i) the validity judgement of the conceptual area that is mapped from the words or phrases in the input sentences, i.e. the check for grammatical sentences, (ii) meta-description of a virtual world, i.e. the definition of words or assumption of property values etc. "A" is defined to be "B" is interpreted that the range of $\Phi(B)$ is adjusted to $\Phi(A)$ by modifying the characteristic functions of A for the virtual view.

[3] Intellectual interpretation

The object is definitely prescribed to be in such a situation, and is managed as follows.

(1) The intellectual prescription

In this step the focussed conceptual area is pointed out to interpret the communicated intention, and the area is mapped from the words, phrases, clauses or sentences. This means the specialization of universal predicates. An element of direct product of I_A, I_N, I_V, \dots etc. are pointed out.

(2) Correlation

The correlation between subject and object in the sentence is clarified, and the relationship among entities are investigated, and the cases of each word is made clear; there are following cases; *Agentive, Factitive, Objective, Dative, State, Action arousing, Emotional, Feeling, Time, Cause, Result, Degree, Locative and Goal cases* so on. As for cases we don't make verbs special treatment as the dependency model, so verbs are assigned cases, too.

The correlated tuples of conceptual areas are filed with cases.

[4] Integration check

In bussiness database systems, the conceptual schema for the enterprise is analyzed at set-up-time, and the most suitable schema is selected, and is not time-variant. And the correspondence between the external schema and the conceptual schema is preset, so the input data is integrated to the conceptual

schema automatically.

On the contrary, in the academic use of the database or in our thinking, the universe of knowledge is not fixed, and is growing according to the input data from the circumstances.

So we introduce *View (Aspect)* for integrating the same object in different records. A group of experimental data, a group of sentences in a book, etc., have the same view. This view is different from the view in the relational model.

A group of fragmentary data with the same view is checked whether they can be integrated to one object if they express the same object. A group of fragmentary data, which has the view that contains or is contained by the input data view, is checked whether they can be integrated to one object or not.

[5] Feeling and Ordering interpretation

(1) The feeling interpretation

In the action arousing subspace, whether action is required or not is examined, where the action means responses by hands, speech or writing so on. There are emotional and feeling subspaces. As for the emotional-state, the conditions of the sender and the receiver are judges by particular words concerning emotions. As for the emotional state, the same judgement is executed.

(2) Ordering information interpretation

For described events, before-and-after relations on time, and cause-and-effect relations are judged. These informations are used in deduction and induction of theorems from axioms, and in management of subjunctive mode.

[6] Evaluation of the received contents

After the grammatical valuation of input sentences and semantic interpretation in the space, following managements are carried out: evaluation of communicated contents, and judgement of logical validity. These managements are executed by inferences using deductions, inductions and arithmetic operations on areas in the space. By these processing, confirming the validity on the communicated contents, and drawing new conclusions from inputs and knowledge, are executed.

(1) Arithmetic operations

Arithmetic operations are used in deduction, induction and other managements. There are following property-types:

- \hat{D} : a power set of areas,
- \hat{H} : a power set of hyperplanes,
- \hat{V} : a power set of values,
- \hat{E} : a power set of estimating value.

And following kinds of operations exist.

$$\hat{D} \times \hat{H} \times \hat{V} \times \hat{E} \rightarrow \hat{D} \times \hat{H} \times \hat{V} \times \hat{E}$$

(2) Evaluation of Declaration

The input informations are checked whether they contradict with the data that are already input and confirmed.

(3) Deduction and Induction

(i) Deduction

The question is rewritten to the questions on the space. Using the algebraic properties, the equation is transformed and solved. In case that the algorithm is cataloged, a suitable procedure is searched and triggered. Also suitable procedures are triggered if the estimating value exceeds the threshold in each process.

In case that the algorithm is not cataloged, or the input is not a question, deductions from the input is executed. We communicate each other with some view, and the received information is interpreted in that view, i.e. if the received data contradicts with the same view data, the input data must be checked whether it is right or not. And if a contradiction is not deduced, the valuable data deduced from the input are arranged and stored. In these processes, the deduction is executed by syllogisms or resolution.

(ii) Induction

The induction is executed as follows: at first some special predicates in the same view are selected, common variables are searched, universal special predicates are reconstructed by choosing the domain appropriately from the existential special predicates. These existential predicates are not always valid, but if the contradiction is not found by checking the domain in the information space feeding back the result, the predicates are regarded to be correct probably, and are stored as a hypothesis, and output. Where induction is executed on the data which have the same view, and the predicates are selected from the predicates which have the same properties in the same view data.

This operation is useful in creative thinking, for example, from the data gathered, common properties are extracted, and the induction is executed to the predicates which contain the common properties. As a result, from the experimental data or the observational field data, some valuable hypotheses are generated.

(4) Updating the Memory or Database

The received message are parsed and mapped to the semantic level, interpreted and deduced. As a result new facts or new informations are clarified. These data are stored in the database. There are two ways of updatings: one is a permanent amendment: if the data is a principle one, modifications of the information space are needed, so the information space generative grammar is reconstructed, and the corresponding amendments of files are carried out. The other is a temporary modification in an interim file.

[7] Composition of Intention & Utterance

If the question is ambiguous or the answer is not deducible from the database

and inputs, the action-driving potential to fill up the information insufficiency is elevated, then the question to resolve the ambiguities is composed and uttered. If the question is correct and the answer is deducible from the dictionary and inputs, the answer file in the semantic level is translated into the P-marker, and is uttered. If the input is not an interrogation, then the database is updated, and no answer or agreeable responses are made. In every case, the intention is described in file form in the semantic level, i.e. tuples of conceptual areas with cases. The file is recomposed to the derivation tree in the syntactic level by ϕ and ψ , transformed to the surface description, and uttered.

6. CONCLUSION

A generative semantics model based on the topological space generative grammar was proposed. This model has more descriptive power than other models because both implications and similarities can be generated in the basic level. The aim of this model is to describe and manage the semantics algebraically, and for that purpose English was arranged based on the model.

The question-answering process in a natural language was clarified. In this model not only a concept but also estimation of the information is considered. Then the motivation for next actions can be interpreted. Based on the model, we are now implementing a research-DBMS. Using it we can built up and test hypotheses by KJ-methods and so on.

REFERENCES

1. Chomsky, N. "Aspect of the Theory of Syntax", MIT Press, 1965
2. Fillmore, J.C. "The Case for Case", In Universals in Linguistic Theory. Bach, E; Harms, T.R. eds. Holt, Rinehart & Winston, 1968, pp1-90.
3. Jackendoff, R.S. "Semantic Interpretation in Generative Grammar", MIT Press, 1972.
4. CODASYL Dev. Committee, "An Information Algebra Phase I Rep.", C.ACM 5-4, April 1962, pp190-207.
5. Uchinami, S., Tezuka, Y. and Kasahara, Y. "A Model of Linguistic Information Algebra", Tech. Rep. of Osaka Univ. 1970, 20(962), pp711-723.
6. Tsurumi, Y., Uchinami, S. and Tezuka Y. "Information Retrieval using the Meaning Space", Trans. of IECE of Japan, 1972 May, 55D5, pp307-314.
7. Uchinami, S. and Tezuka, Y. "On the Formal Properties of the Topological Information Space Generative Grammar & its Space", Trans. of IECE of Japan, 1980 Jan., 63D1, pp79-86