Memorization and Reasoning of Semantic Network Using Paired-Associate Mapping of Patterns

Shouichi Suzuki †

Abstract

In knowledge engineering, there is a semantic network as one method expressing knowledge. Inheritance of properties is possible for a semantic network, and it is a directed graph which consists of nodes and links. Memorizing a semantic network and reasoning there are usually performed by manipulating symbols.

A method of memorizing a semantic network and reasoning can be performed by pattern processing is presented here. In other words, it is shown that memory action and reasoning of a semantic network can be performed using model-construction operator T , similarity-measure function SM , and a paired-associate mapping A suggested by S,Suzuki.

The five advantages are as follows :

(1) The concept and the property memorized by adoption of pattern expression have been easily expressed by the linear-predictive-coefficient vector of the pronounced sound..

(2) The concept memorized and the property became strong to noise by adoption of a pattern model.

(3) In two phases of memorizing and reasoning made by the association, it became strong to noise by adoption of the paired-associate mapping.

(4) In the reasoning phase by association, it became strong to noise by adoption of the similarity-measure function.

(5) It can simply perform adding to a semantic network new triple that consists of concept 1, property, and concept 2 by easy correction of the paired-associate mapping and the similarity-measure function.

Key words

semantic network　　Reasoning processing by the patterns　　paired-associate mapping
pattern model　　similarity-measure function

1. Introduction

There are (1*) a procedural representation and (2*) a declarative expression for expressing knowledge in knowledge engineering. The expression by program language LISP is typical of the procedural representation and the expression by program language Prolog is one example of declarative expression. It divides roughly into the latter declarative expression, and there are four kinds [6]:

(1#) Predicate calculus which adopted Prolog etc. as an expression language.

(2#) Production rule.

(3#) Semantic network.

(4#) Frame. □

These declarative expressions are usually realized as processing which manipulates a sequence of symbols [7].

Inheritance of propertiese is possible for a semantic network, and it is directed graph which consists of a set of nodes and links. This semantic network to which a node expresses a concept and the link expresses the relation between nodes is a collection of the triples , for example ,

(2$) <Taro IS-A human>, etc.

expressing the knowledge

(1$) Taro is a human .

Taro, a human being, etc. express a concept and IS-A expresses the property, the attribute, the relation, etc. If a triple

(3$) <human can thinking>

expressing the knowledge "human can think" is also memorized, the so-called " inheritance of property " is possible in a semantic network in the form of that

(4$) <Taro can thinking> is reasoned from <Taro IS-A human> and <human can thinking>.

If the recursive neural network of Elman (Elman, J.L.) that consists of four layers of

(1%) an input layer (2%) a hidden layer (3%) an output layer (4%) a primary memory layer which memorize the past activation situation only at the 1 unit time of the hidden layer provisionally, and have connection of feedback to the hidden layer are prepared is employed, it is clear that the grammar structure of complicated English containing a relative clause can be learned as an activation pattern (a distributed representation) [8], and a semantic network can be memorized. However, by this neural network, in order to attain the reasoning which can inherit an property attribute, another framework is needed.

In this paper, the semantic network which is a set of triples is memorized, and reasoning there is now realized as one associative processing of a pattern pair not different from symbol-sequence processing in which a symbol-sequence is manipulated.

In the semantic network of this paper, a triple <human IS-A animal> is memorized as a if-then-rule

(1&) if <human IS-A> then <human animal>.

When the element of a semantic network is for example,“<human IS-A animal>”and <human IS-A animal> is memorized as two pattern pairs <human IS-A> and <human animal>, reasoning expressed with modus ponendo ponens

Fact : <human IS-A>> holds.

Rule : It is <human animal> if it is <human IS-A>.

Conclusion : <human animal> ∴ <animal>

about a declarative knowledge <human IS-A animal> is processed in the form of

"pattern association of the pattern pair <human animal> is carried out from the pattern pair <human IS-A>". □

Five acquired advantages (1@) (5@) brought about by four adopted fundamental design plans

. Human, IS-A, animal, etc. are expressed not as symbol sequences but as 16-dimensional vectors of the real linear prediction coefficients of voice waveforms of the person who uttered these

. Furthermore, the patterns are prepared in the form where noise is removed from a voice waveform by asking for the model of a 32-dimensional linear prediction coefficient vector

. Reasoning using a modus ponendo ponens is realized by work of association

. Memory similar to the obtained reasoning result is looked for by application of a method of maximum similarity which is as follows and is checked by the semantic network implemented in the Java language :

(1@) The concepts and the attributes memorized by adoption of pattern expression have been easily expressed by the linear prediction coefficient vectors of the pronounced sounds.

(2@) The concept and the attribute which are memorized by adoption of a pattern model are strong to noise.

(3@) A memory scene and a recall scene are strong to noise by adoption of the paired associate mapping.

(4@) A recall scene is strong to noise by adoption of the similarity-measure function.

(5@) When it must have to add new triples to an old semantic network , the old memorizing / reasoning function is maintained about this addition and this addition can be simply performed by easy correction of the paired-associate mapping and the similarity-measure function. □

With this five advantages (1@) (5@), this research will offer the technology used as the foundation required to build a full-scale semantic network system.

Although the conventional semantic memory model is realized by sign sequence processing or the neural network [6] [10], this semantic memory model adopts pattern processing of a theory (SS theory [1] [4]) suggested by S.Suzuki, is realized, and this realization becomes clear from three points (a), (b), (c):

(a) In order to use instead of the pattern φ the pattern model $T\varphi$ which can remove the noise which may be in a pattern φ , the model-construction operator T of SS theory was used.

(b) The paired-associate operator A [5] by thinking over the paired-associate matrix [9] suggested by T.Kohonen in the n-dimensional Euclidean space R^n to which the method of minimum square might be applied was obtained in general separable abstract Hilbert space making full use of the composition technique of SS theory ,and the operator A was used as a paired-associational inferencer.

(c) In order to inspect whether the associative output from an associative inferencer has a semantic relation of being similar to one of the contents throats of memory, the similarity-measure function SM of SS theory was used. □

In addition, this function of memorization and reasoning on a semantic network is equipped with the performance beyond the function of memorization and reasoning of the conventional semantic network system done by symbol processing about noise-proof nature. It is because the three following character (1@),(2@) and (3@) is materialized fundamentally in this system:

(1@) The modeling conclusion expression (6) of the model-construction operator T
(2@) The orthonormal expression (8) type of the similarity-measure function SM
(3@) The interpolation expression (A.11) of the paired-association operation A □

2. Design method of semantic network system

The method of designing the semantic network system as one of the semantic memory models is briefly explained by this chapter.

2.1　Expression vector of a pair of symbol sequence

A semantic network can be expressed as set

<A[j] R[j] B[j]>, $j = 1 \sim r$ (1)

of triples.

About symbols in which it is different from each other in the set

A[j],R[j],B[j], $j = 1 \sim r$ (2)

of symbol sequences , all are pronounced and we ask for the voice waveform. The voice waveform in the 1-dimensional section $\{x \mid x = -p \sim +p\}$ of such a symbol sequence is set to be $f(x)$. The 1-dimensional section $\{x \mid x = -p \sim +p\}$ is mostly chosen as the middle portion in the utterance section. The linear prediction coefficient vector

$\vec{C}(f) = col(C_1(f) \quad C_2(f) \quad \cdots \quad C_q(f))$ (column vector)

of f according to Appendix D is asked for. Here, since the orthogonal system of the expression (A.8) of Appendix B is obtained, inequality

$r \le 2q$

must be materialized. In this semantic network system,

$r = 17$, $q = 16$

were set up. Henceforth, $\vec{C}(f)$ may be simply written to be $\vec{C}$.

As opposed to the k-th triple

<A[j] R[j] B[j]> (3)

in a set of triples of expression (1),the　if condition then action　type knowledge representation unit

if <A[j] R[j]> then <A[j] B[j]> (4)

is considered. Expression of the paired-sequence <A[j] B[j]> for a symbol showing a condition shall be a $2q$-dimensional real vector, should connect the q-dimensional real value vector of B[j] after the q-dimensional real value vector of A[j], and shall have been obtained. We asks for the same $2q$-dimensional real vector also from the sequence <A[j] B[j]> for a symbol showing an action.

2.2　The noise removable model $T\varphi$ of Pattern φ

It is considered that the expression vector of paired $j(=1 \sim r)$-th symbol sequence <A[j] B[j]> is an element(pattern) of $2q$-dimensional Euclidean space R^{2q}. The inner product (φ,η) of R^{2q} and the norm $\|\varphi\|$ are defined by the expression (A.5). The $k(=1 \quad 2q)$-th component $(\varphi)_k$ in this expression (A.5) is defined by the expression (A.3). It asks for the model(pattern model) $T\varphi_j$ of φ_j according to 1. of an appendix. The adopted mapping T which is called a model-construction operator is explained by its appendix 1.

Generally, there is the following two fundamental character in the pattern (pattern model) $T\varphi \in R^{2q}$ called the noise removable model of the pattern $\varphi \in R^{2q}$ of expression (A.1):

(1#) (T absorbs multiplying by the any positive constant;cone property)

$\forall \varphi, T(a \cdot \varphi) = T\varphi$ for any positive real number a (5)

(2#) (Idempotent property of T ; conclusion of modeling)

$\forall \varphi, T(T\varphi) = T\varphi$ (6)

□

Two above-mentioned matters (1#) and (2#) shows the following matters (1$) and (2$)respectively:

(1$) The model of the pattern multiplied by positive arbitrary constants is the model of the original pattern which has not been multiplied by the constants.

(2$) The model of a model is the original model. □

There is character in which small noise is removable from the pattern φ of expression(A.1) in mapping T because it can be understood from

(3#) (Noise removable nature)

$$(T\varphi)_k = 0 \quad \text{if} \quad e_-(k) < \frac{a_k}{\max_{l=1 \quad n} |a_l|} < e_+(k)$$

being materialized. $(T\varphi)_k$ is the $k(=1 \sim 2q)$ -th component of expression (A.3) of B of expression $T\varphi$ here. Moreover, it is as follows if one pattern

$$\eta = col\left(b_1 \quad b_2 \quad \mathbf{L} \quad b_k \quad \mathbf{L} \quad b_{2q}\right)$$

is introduced and explained now

(4#) (Each model being equivalent which follows from removing deformation of patterns)

About all the $k(=1 \sim 2q)$, if either in three conditions

(4#_1) $e_-(k) < \frac{a_k}{\max_{l=1 \quad n} |a_l|} < e_+(k) \wedge e_-(k) < \frac{b_k}{\max_{l=1 \quad n} |b_l|} < e_+(k)$

(4#_2) $\frac{a_k}{\max_{l=1 \quad n} |a_l|} \le e_-(k) \wedge \frac{b_k}{\max_{l=1 \quad n} |b_l|} \le e_-(k)$

(4#_3) $\frac{a_k}{\max_{l=1 \quad n} |a_l|} \ge e_+(k) \wedge \frac{b_k}{\max_{l=1 \quad n} |b_l|} \ge e_+(k)$

are materialized, the equivalent nature

$$T\varphi = T\eta$$

of two models will be realized. That is,it is shown that there is character to remove modification of a pattern in mapping T , that is to say, to extract character common to two patterns φ and η. □

When a set of triples of expression (1) is given, in a trial-and-error method, it is quite difficult to ask for patterns which corresponds to each symbol sequence of expression It has become clear by the method using the linear prediction coefficient of two paragraphs 2.1 and 2.2 for this to be possible, without applying a trial-and-error method to almost.

2.3 Paired associate operator A

The paired-associate operator A of expression (A.9) which recollects the model $T\eta_j$ of action pair symbol sequence <A[j] B[j]> is constructed from model $T\varphi_j$ of condition pair symbol sequence <A[j] R[j]> about if-then-rule of expression (4) for any $j(=1\sim r)$ according to appendix 2. This construction is a function of memory of a semantic network.

There is an interpolation property of expression (A.11) in the function of this memory. This interpolation property has guaranteed that the symbol sequence <A[j] B[j]> can be called correctly from the arbitrary symbol sequences <A[j] R[j]> for any $j(=1\sim r)$.

In addition, stage A_{k+1} in the middle of construction is extension of A_k, and there is interpolation property of expression (A.12) also in stage A_k in the middle of construction.

The addition of triple <A[r+1] R[r+1] B[r+1]> with this new fact guarantees becoming easy.

2.4 Similarity-measure function SM

Each $\omega_j(j=1 \quad r)$ of expression (A.16) of Appendix C is set to be the $2q$-dimensional real-number-value vector $T\eta_j$ showing the action pair <A[j] B[j]> in triple <A[j] R[j] B[j]> of expression (3). Namely

$$\omega_j = T\eta_j,\ j=1\sim r \tag{7}$$

will set up.

The similarity-measure function $SM(\varphi,\omega_j)$ which gives the grade to which pattern φ resembles pattern ω_j is defined like expression (A.15) of appendix 3. .There is the three following properties ( ), ( ) and ( ) in the constructed similarity-measure function SM :

( ) (orthonormality)

$$\forall i,\forall j, SM(\omega_i,\omega_j)= \begin{cases} 1\ \mathbf{L}\ i=j \\ 0\ \mathbf{L}\ i\neq j \end{cases} \tag{8}$$

( ) (probability condition,normalization)

$$\forall\varphi, \sum_{j=1}^{r} SM(\varphi,\omega_j)=1. \tag{9}$$

( ) (invariance under mapping T)

$$\forall\varphi,\forall j\in\{1,2,\mathbf{L},r\}, SM(T\varphi,\omega_j)=SM(\varphi,\omega_j). \tag{10}$$

□

The nomalized inner product $nip(T\varphi,T\omega_j)$ of the form of expression (A.14) is not adopted as a degree of similar between two patterns φ,ω_j because $nip(T\omega_i,T\omega_j)=0(i\neq j)$ is not necessarily materialized. That is, it is because the separation between patterns is not good unlike the following (1%).

Three expressions (8), (9), and (10) bring about respectively the effect

(1%) ω_i can completely disassociate from ω_j about all the different j from i

(2%) If one certain degree $SM(\varphi,\omega_i)$ of similar about ω_i becomes maximum 1, the degree $SM(\varphi,\omega_j)$ of similar about ω_j with all different j from i will become the minimum value 0 and

(3%) A pattern model $T\varphi$ can identify with the original pattern φ from a viewpoint of the similarity

3.2.3　The check of the paired-associate mapping A being equipped with interpolation character

About expression (A.10) of appendix 2., it was checked as a matter of fact that

$$\|(\Delta A_j)(T\varphi_k)\| < 10^{-12}, k = 1 \quad j, j = 1 \quad r-1$$

is materialized. It was checked that expression (A.11) having the interpolation character is materialized in the form where inequality

$$\|A_k(T\varphi_j) - T\eta_j\| < 10^{-11}, j = 1 \sim k, k = 1 \quad r \quad \therefore \quad \|A(T\varphi_j) - T\eta_j\| < 10^{-11}, j = 1 \sim r$$

is filled.

It was checked by the above that all the 17 triples (1#)　(17#) are memorized correctly.

3.3　Examination in a recall scene

It was checked that theorem 1 holds good, i.e., all of 17 triples (1#)　(17#) are memorized correctly, and are correctly reasoned including the inheritance of properties.

For example, it sees and needs about (9#) <　　can　　　>.

We will examine whether B is reasoned from A.

For example, the voice waveform　　　is shown in Fig. 2.

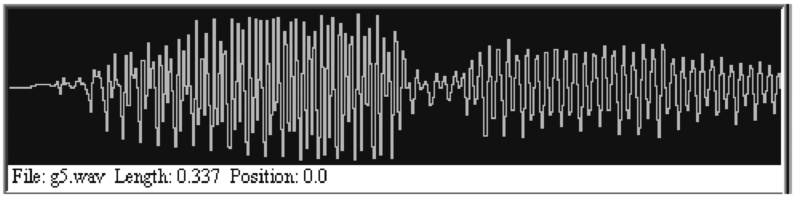

Fig. 2　A voice waveform of uttered“bird”.

In Table 1, the q-dimensional linear prediction vectors of two voice waveforms　　, can　 are expressed as "

$(\varphi)_{k-1} =$ The k-th component C_k ($k = 1$　16) of the q-dimensional linmear prediction vector of voice waveform

and

$(\varphi)_{k+15} =$ The k-th component C_k ($k = 1$　16) of the q-dimensional linmear prediction vector of voice waveform　can

".

Time required for reasoning, a display, and transfer of a reasoning result by a sound etc. was 5812mm second.

In addition, equation (A.12) of appendix 2. being materialized is checked ,that is,although the 1st to k-th triples are correctly reasoned from $A_k (1 \le k \le m)$, the $k+1$-th to r-th triples are not reasoned correctly.

4. Conclusion

The old semantic network system has been made by symbol-sequence processing. In this paper, the original method of and memorizing the semantic network and performing reasoning there by pattern associative processing was proposed exceeding the performance by symbol-sequence processing.It was shown that it can be carried out using the model-construction operator T, the paired-associate mapping A and the similarity-measure function SM proposed by S.Suzuki. It was checked that many original advantages in the introduction are equipped in the semantic network system implemented in the Java language, when the contents of operation are examined. As the method of memorizing the triple < IS-A > for example, on a semantic network

(1) "Three q-dimensional linear prediction coefficient vectors of three voice waveforms which are obtained by utterring respectively with , IS-A, and were used "

is originality of this research. Next, it is also originality of this research to have transposed

(2) triple < IS-A > to if-then-rule if < IS-A> then < >.

(3) "System asks for the pattern model of the $2q$-dimensional linear prediction coefficient vector of < IS-A> and < >, and a paired-associate operator which can recall the latter model from the former model is designed"

was presented here.It is also originality of this research.These 3 originalities make it easy to design a semantic network, and also make it easy to add new triple to the designed semantic network .

Moreover, for example, the following (4),(5) and (6) as a method of reasoning < > from < IS-A> was proposed:

(4) It asks for the model of the $2q$-dimensional linear prediction coefficient vector of < IS-A>, and asks for the $2q$-dimensional vector pair-associated from this model.

Then,

(5) the set of the $2q$-dimensional linear prediction coefficient vector of <1st term 3rd term> of triple containing < > is prepared.

(6) What resembles this $2q$-dimensional vector most is taken out from this set by applying the maximum method of similarity. □

For this reason (6), this designed semantic network will be equipped with proof against noise. It will have predominancy to the semantic network designed by mere symbol processing.

The used pattern model form T, the pair-associate operator A, and the similarity-measure function SM are original with this research which both have not appeared in an old pattern information processing theory except SS theory proposed by S.Suzuki.

Moreover, about noise-proof nature it came to have a performance beyond "the function of memorizing and reasoning" of the conventional semantic network system done by symbol-processing because (1&),(2&),

and (3&) of Section 2.6 act effectively fundamentally.

For example, the system is designed so that the inputted symbol sequence, a reasoning result, etc. may be told also with a sound. Furthermore, about triple < IS-A >, it is desirable to change to the system which gives directions with the sound " IS-A" and can answer with the sound " ". We think that basic technology for that was secured enough.

References

[1] Shoichi Suzuki（ ）: "recognition engineering（ ）（Quantum theory of recognition）", Kashiwa Shobo publishing company（ ）, Feb.1975

[2] Shoichi Suzuki（ ）: "a new mathematical principle of a neural network（ ）（A new-generalized error-backpropagation neural network with hidden layers that can learn an input-output relation from examples,based on a new mathematical point of view）", Modern Literature publishing company（ ）, Sept.1996

[3] Shoichi Suzuki（ ）: "mathematical general solution of a pattern recognition problem（ ）（A newly-generalized means of solving pattern-recognition problems in the employ of categorical-membership knowledges,based on a new mathematical point of view）", Modern Literature publishing company, June 1997

[4] Shoichi Suzuki（ ）: "new deployment of a recognition intelligence information theory（ ）（A newly-developed intelligence informatics about pattern-recognition that contains a potential theory of categorical-membership knowledges） ", Modern Literature publishing company, Aug.1998

[5] Shoichi Suzuki（ ）: "A general solution of " a paired associate problem and its pseudoinverse problem" by the help of a sequence of input-output pairs"（" " " "）,Information and Communication Studies of the Faculty of Information and Communication （Bunkyo University）（ （ ））,no.30,pp.81-137,Jan.2004

[6] Ikuo Tahara（ ）: "The basic knowledge of artificial intelligence（computer science university lecture 20）（ ）", Kindai Kagaku Sha publishing company（ ）, Feb.1999

[7] Shintani Toramatsu（ ）: "a guide to intelligence programming by Java（Introduction to Knowledge Programming with Java）（Java ）", Corona Publishing company（ ）, Oct.2002

[8] Mori Kazuwo（ ）, Takashi Tsuzuki（ ）, Tkasi Kusumi（ ）: "understanding of the heart by the simulation of brain -connectionist model and psychology- （ - -）", Kitaoji Shobo publishing company（ ）, Jun.2001

[9] Teuvo Kohonen（T. ）:" Associative memory-for information engineering and psychology - （Associative Memory A System-Theoretical approach）（ - -）", translation by Kazuo Nakatani（ ）, Saiensu-Sha publishing company （ ）, Aug.1980

[10] Manfred Spitzer（M. ）: "the soul in a brain,i.e.,a network （map which a neural network

draws)(Geist im Netz)(( ))", translation by Toshiya Murai( ) and Hirosi Yamagishi( ), Shin-yo-sha publishing company( ), Nov.2001

[11] Shoichi Suzuki( ): "Generalization of canonical forms called corresponding models of patterns (( ))",Information and Communication Studies of the Faculty of Information and Communication (Bunkyo University) (( )),no.32,pp.169-218,Jan.2005

[12] Yasuhisa Niimi( ) : "speech recognition (information science lecture E-19-3)((E 19 3))", Kyoritsu shuppan publishing company( ), and Oct.1979

[13] Shoichi Suzuki( ): "A corresponding entropy model of a pattern( )",The Transactions of the Institute of Electronics,Information and Communication Engineers( ), vol.J77-D ,no.11,pp.2220-2238,Nov.1994

Appendices

1. The adopted model-construction operator T which brings about the canonical form of a pattern(a corresponding model of a pattern)

There are the following three kind (1#), (2#) and (3#)in the mapping T which generates the pattern model $T\varphi$ used as instead of pattern φ :

(1#) Covariant forms under unitay coordinate transformations [13]

(2#) Invariant forms under unitay coordinate transformations [1], [2], [13]

(3#) What approximates the amplitude of a pattern with the limited or infinite values □

The general form of such mapping T is synthetically studied by reference [11].

The mapping T adopted in this semantic network system is the 3rd thing (3#), and is equipped with the effect of removing small noise. This mapping T is explained below.

About the $2q$-dimensional real number value vector

$$\varphi = col\left(a_1 \quad a_2 \quad \cdots \quad a_k \quad \cdots \quad a_{2q}\right) \text{(column vector)} \qquad \text{(A.1)}$$

, the noise removable model

$$T\varphi = col\left(c_1 \quad c_2 \quad \cdots \quad c_k \quad \cdots \quad c_{2q}\right) \qquad \text{(A.2)}$$

is defined as follows. Henceforth, $a_k, c_k (k = 1 \sim 2q)$ may be expressed as

$$(\varphi)_k \equiv a_k, (T\varphi)_k \equiv c_k (k = 1 \sim 2q) \quad . \qquad \text{(A.3)}$$

Two threshold value vectors

$$\vec{e}_- = col(e_-[1] \quad e_-[2] \quad \cdots \quad e_-[2q]), \vec{e}_+ = col(e_+[1] \quad e_+[2] \quad \cdots \quad e_+[2q]) \qquad \text{(A.4)}$$

which fill inequality

$$-1 \le e_-\left(i\right) < e_+\left(i\right) \le +1, i = 1 \quad 2q$$

are prepared.

Zero 0-calculation rule

$$\frac{a_k}{\max_{\lambda=1 \quad 2q}\left|a_\lambda\right|} = 0 \qquad if \qquad \max_{\lambda=1 \quad 2q}\left|a_\lambda\right| = 0$$

is prepared. The k-th ingredient $c_k = (T\varphi)_k$ of $T\varphi$ is defined as

will be materialized. Therefore, the equation

$$A(T\varphi_j) = T\eta_j,\ j = 1 \sim r \tag{A.11}$$

which shows an interpolation property is realized. In addition, there is interpolation property

$$A_t(T\varphi_j) = T\eta_j,\ 1 \le j \le t \tag{A.12}$$

also in A_t in the middle of composition. Moreover, Expression

$$A_t = A_1 + \sum_{k=1}^{t-1} (\Delta A_k),\ 2 \le t \le r$$

is materialized.

3. The adopted similarity-measure function SM which brings about the work which measures the degree of similar between pattern models

Below, the similarity-measure function SM in the Appendix 2 of reference [4] is explained.

Let all the n -dimensional real vectors that express pair <concept1, concept2> about triple <concept1,relation,concept2> be

$$\omega_j,\ j = 1 \sim r \cdot \tag{A.13}$$

Similarity-measure function $SM(\varphi, \omega_j)$ which gives the grade to which pattern φ resembles Pattern ω_j is defined as follows.

First, normalized inner product $nip(T\varphi, T\eta)$ is defined as

$$nip(T\varphi, T\eta) = \begin{cases} \dfrac{(T\varphi, T\eta)}{\|T\varphi\| \cdot \|T\eta\|} & if \quad \|T\varphi\| \cdot \|T\eta\| \ne 0 \\ 0 & if \quad \|T\varphi\| \cdot \|T\eta\| = 0 \end{cases} \tag{A.14}$$

.Next, non-negative quantity $S(\varphi, \omega_j)$ is defined as

$$S(\varphi, \omega_j) = -\frac{1}{2} \cdot \log_e [1 - |nip(T\varphi, T\omega_j)|^2]$$

and the total

$$S(\varphi) = \sum_{j=1}^{r} S(\varphi, \omega_j)$$

is defined. At this time, $SM(\varphi, \omega_j)$ is defined as

$$SM(\varphi, \omega_j) = \begin{cases} \dfrac{S(\varphi, \omega_j)}{S(\varphi)} & if \quad S(\varphi) > 0 \\ \dfrac{1}{r} & if \quad S(\varphi) = 0 \end{cases} \tag{A.15}$$

.