Word Sense Disambiguation in Contextual Dynamic Network Using Associative Concept Dictionary

Jun Okamoto
Keio Research Institute at SFC,
Keio University / 5322 Endo, Fujisawa-shi
2522-8520, Japan
juno@sfc.keio.ac.jp

Shun Ishizaki
Graduate School of Media and Governance,
Keio University / 5322 Endo, Fujisawa-shi
2522-8520, Japan
ishizaki@sfc.keio.ac.jp

Abstract

Many of the Japanese ideographs (Chinese characters) have a few meanings. They should be disambiguated by using their contextual information. Some of the ideographs have a few different pronunciations depending on their meanings. For example, we have an ideograph which has two pronunciations, /hitai/ and /gaku/, the former means a forehead of the human body and the latter means an amount of money or a picture frame. Conventional methods for such disambiguation have been using statistical methods with co-occurrence of words in their context. In this research, Contextual Dynamic Network Model is developed using the Associative Concept Dictionary, which includes semantic relations among concepts/words and the relations are represented with qualitative distances. In this model, an interactive activation model is used in the contextual semantic network where the activation in the network is calculated using the distances. The proposed method can disambiguate word senses by constructing dynamically the contextual semantic network according to the input text including an ambiguous word. The interactive activation model based on the contextual semantic network is evaluated by using a Japanese newspaper, Mainichi shinbun. Its effectiveness is shown by comparison with a Naive Bayes method.

1 Introduction

Word sense disambiguation is one of the difficult problems in natural language understanding by computers because it needs contextual meanings. A lot of previous works for such disambiguation have been using co-occurrence of words in their context. Several machine learning algorithms have been used based on co-occurrence information among words, such as Naive Bayes methods or Support Vector Machine (SVM) (Murata, et.al., 2003). Effectiveness of neural network approaches to the word sense disambiguation has been suggested (Woltz and Pollack 1985; Takahashi, 1995, Tsuzuki and Saito 1991). Not only the neural network architecture but also large-scale machine readable dictionaries were exploited (Veronis and Ide, 1990).

Many of the Japanese ideographs (Chinese characters) have a few meanings. They should be disambiguated by using their contextual information. In our previous paper, we proposed a method to disambiguate homographic ideographs with different pronunciations by using a Contextual Network Model (Okamoto and Ishizaki 2005). In that model, the contextual semantic network architecture is constructed based on the Associative Concept Dictionary (Okamoto and Ishizaki, 2001) including semantic relation and distance information among the concepts. In this paper, we proposed a Contextual Dynamic Network Model, where the contextual semantic network has a structure which changes dynamically depending on the input sentences or closes. If we use the concept dictionary which includes structuralized information about word senses of ambiguous words, we can disambiguate not only homographic ideogram but also various word sense disambiguations by using the Contextual Dynamic Network Model.

Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics, pages 94–103.

2 Associative Concept Dictionary

Background knowledge is crucial for computers to understand the contents of the text as well as its syntactic or shallow semantic information from input texts. The Associative Concept Dictionary (hereinafter referred to as ACD) has been built based on the results of large-scale online association experiments, which many subjects can use simultaneously in a campus network at a campus of Keio University. In these experiments, the stimulus words were fundamental ones chosen from Japanese elementary school textbooks and were presented to human subjects. The subjects were requested to give any number of associated words from the stimulus words at the seven semantic relations, hypernym, hyponym, part/material, attribute, synonym, action and situation.

Table 1 is an example of associated words by a subject where the stimulus word is "辞書" Its pronunciation is /jisyo/ and means a dictionary. The numbers in the table show the duration spent for the concept association by the subjects of which unit is second. They are durations between starting times of the association and their ending times. This temporal information is included in the results of the association experiments. For example, "publication" is a hypernym of "dictionary" while the duration is 7 second. And next associated word is "book" as hypernym of a dictionary.

Table 1: An example of association experiment (Stimulus word is "辞書" In practice, associated words are Japanese)

Semantic relations	{Associated word Response time}
Hypernym	{Publication 7} {Book 12}
Hyponym	{English dictionary 6} {Japanese dictionary 12}
Part / Material	{Entry word 18} {Word definition 33} {Page 38} {Cover page 44}
Attribute	{Difficult 6} {Easy to understand 11} {Pleasant 16}
Synonym	{Encyclopedia 17}
Action	{Read 5} {Investigate 11} {Consult 15} {Search 19} {Buy 29}
Situation	{Library 6} {Book store 27}

All of the associated concepts are, in the ACD, connected to the stimulus words with distances calculated by the following methods. The distances are obtained using a linear programming method (Okamoto and Ishizaki, 2001). It combines the following three parameters linearly: the frequency of the associated concepts (*F*), the association order of the word (*S*), and the duration spent for their association (*T*). Next, two boundary conditions are given such that one is for the shortest distance and the other for a comparatively long distance. By using Simplex method, we obtain the optimum solution for the parameter's coefficients. The first two parameters, *F* and *S*, found significant for the distance calculation and the third parameter *T* zero. Thus, the distance $D(x,y)$ between concepts, x and y, is shown by the following formula:

$$D(x,y)=0.81F+0.27S\,, \qquad (1)$$

where $F=\frac{N}{n+\delta}$, $\delta=\frac{N}{10}-1$ (N≧10),and $S=\frac{1}{n}\sum s_i$.
N denotes a number of the subjects who joined the experiments, and n denotes a number of subjects who input the associated word y with the same semantic relation for a given stimulus word x. Furthermore, δ denotes a factor introduced to limit the maximum value of F to 10, and s_i denotes an order of the association by a subject.

The ACD is built using the quantified distances and is organized in a hierarchical structure in terms of the hypernym and hyponym. Attribute information is used to explain the features of the given word. The synonyms of the stimulus words are also included as well as action concepts and situation concepts.

The conventional concept dictionaries (EDR, 1990) have tree structure for expressing a hierarchical one. Distances between two concepts in the dictionaries are calculated using the number of links between them, whereas the ACD has quantitative distance information between two concepts.

In the association experiment, each stimulus word has 50 subjects who were students at SFC of Keio University. The number of stimulus words is currently 1100. Total number of associated words is about 280,000. And the number of associated words, when the overlapping words are not counted, is about 64,000 words. In Figure1, "chair" is a stimulus word. And "furniture" is a higher-level concept of "chair". The numbers <1> express frequencies of subjects who gave a same associated word, <2> an average of order of association, <3>an average of response time and <4> a conceptual distances.

```
(chair                 <1>  <2>  <3>  <4>
  (hypernym             ↓    ↓    ↓    ↓
    (furniture        0.92 1.02 0.16 1.09)
    (object           0.04 2.50 0.24 7.43))
  (hyponym
    (sofa             0.48 1.92 0.42 1.96)
    (rocking-chair    0.28 1.43 0.59 2.64))
  (part/material
    (wood             0.60 1.20 0.14 1.52))
  (attribute
    (hard             0.46 1.17 0.32 1.82))
  (synonym
    (seat             0.02 1.00 0.15 8.37))
  (action
    (sit down         0.70 1.03 0.15 8.37))
  (situation
    (school           0.30 2.40 0.22 2.78)))
```

Figure1 Concept dictionary description for a stimulus word "chair" (a part of associated concepts are presented. The stimulus word and associated words are originally in Japanese)

3 Word Sense Disambiguation by Contextual Dynamic Network Model

A lot of previous works for the word sense disambiguation have been using co-occurrence of words in their context. In this research, however, we propose a Contextual Dynamic Network Model (hereinafter referred to as the CDNM) to disambiguate word senses by using an interactive activation method in contextual semantic network. The network is constructed by using the ACD, because it includes semantic relations and distance information among the words in the context. In addition, this network is not a static one but dynamic where the network structure changes depending on the context of the paragraphs, sentences, phrases or clauses which include a homographic ideograph.

By using the dynamic network, this method can disambiguate word senses based on words located near the ambiguous words.

3.1 Construction of Contextual Semantic Network

We can use not only influence of word co-occurrence in their context but also that of comparatively rich network with quantitative distances and context information for the word sense disambiguation.

The following steps show a procedure in detail for this network construction.

- Part of speech information for words (nouns, adjectives, adverbs, verbs and so on) in an input sentence is obtained from morphological analysis by using ChaSen, a Japanese morphological analysis software (Matsumoto et. al., 1999).
- A contextual semantic network is constructed by extracting semantic relations among words in the input sentence from the ACD by using the information obtained from the morphological analysis.
- When a noun in the input sentence is included as a stimulus word in the ACD, a contextual semantic network is added which starts from the noun by tracing semantic relation paths until the distance accumulated becomes a certain numerical level.
- When a word in the input sentence is found in associated words of the ACD and a stimulus word for the word is included in the nodes in the constructed contextual semantic network, the semantic relation path from the associated word to the stimulus word is assigned to the network.
- When a stimulus word is a homographic ideograph and its associated words are associated from another homographic ideograph, inhibitory links between the stimulus words and the associated words are added in the network
- Inhibitory links among stimulus words are added in the network when the stimulus words are included in a homographic ideograph.

Figure 2 shows an example of a contextual semantic network based on the ACD and a homographic ideograph in the input sentence. Let an input sentence be "The picture frame of Picasso's picture dropped from the wall, it struck my head and the forehead bled."

In this sentence, "picture frame" and "forehead" are English expressions which correspond to the homographic ideographs "額" in Japanese. This ideograph has two pronunciations /gaku/ and /hitai/. "Wood" is a part relation concept of "picture frame". "Museum" is a situation concept of "picture frame". "Eye" and "nose" are part relation concepts of "face".

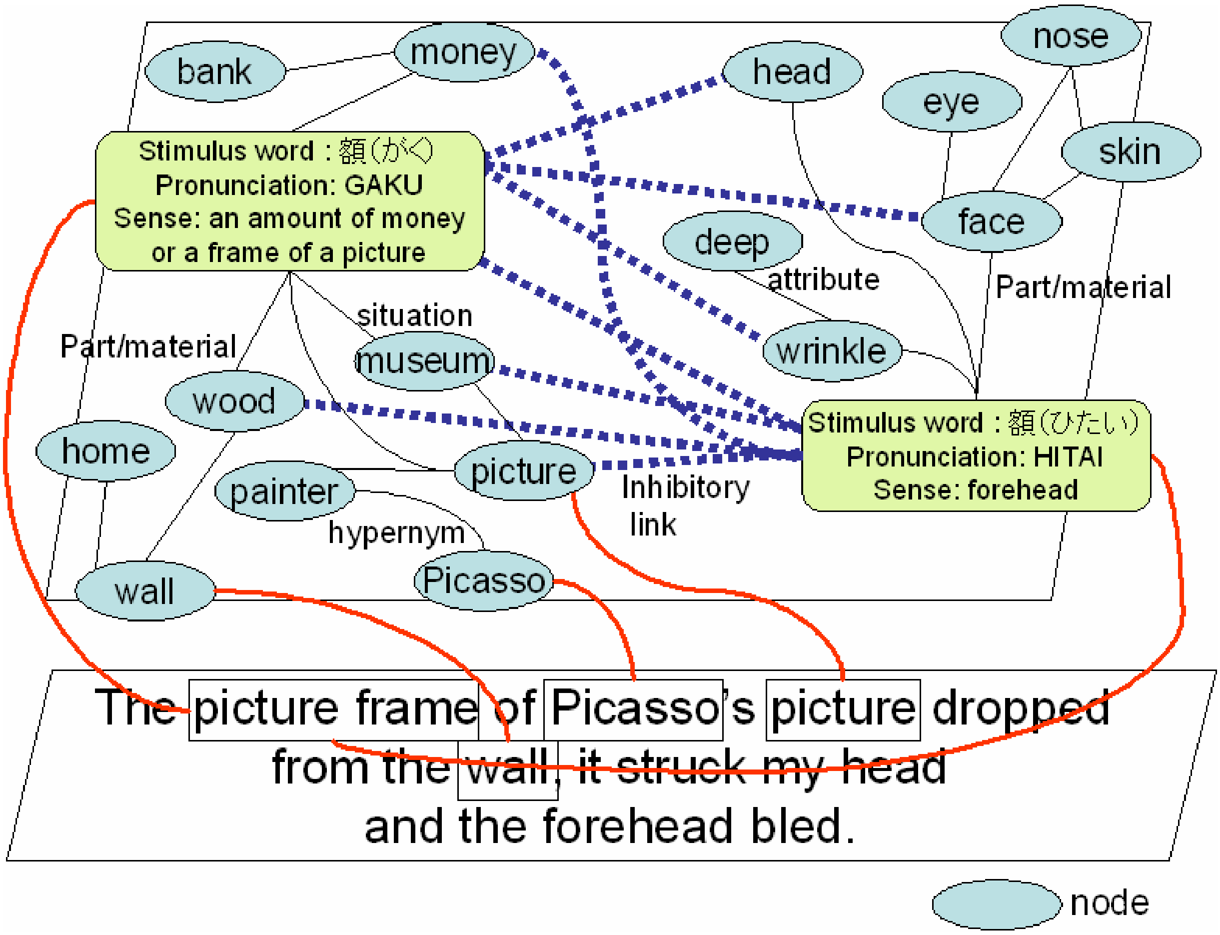

Figure 2 Example of Contextual Semantic Network based on the ACD and the words before and behind one homographic ideograph "picture frame" in an input sentence

"Deep" is an attribute concept for "wrinkle" and "wrinkle" is a part concept for "forehead." "Painter" and "face" doesn't exist in the sentence but is obtained from ACD. Oval shape nodes are added by using ACD and are connected with excitatory links. Square nodes are those of homological ideographs and are connected with inhibitory links. The thick lines connect oval nodes with the word in input sentence. The dotted lines mean inhibitory links. The associated word nodes of a homographic ideograph are connected with other homographic ideographs with inhibitory links.

3.2 Dynamic Reconstruction of Contextual Semantic Network

In this research, an input sentence is divided into multiple segments. For each segment, a contextual semantic network based on the ACD is reconstructed. Human-beings can disambiguate word senses based on two words on the left and two words on the right of the ambiguous word (Choueka and Lusignan, 1985). In our proposed method, when some homographic ideographs are in the sentence, this method can assign appropriate word senses to the given ambiguous word in the sentence depending on its contextual meaning by dividing the sentence into segments. As for the unit of the segments, we can use phrases or clauses which include one ambiguous word in some paragraphs and sentences.

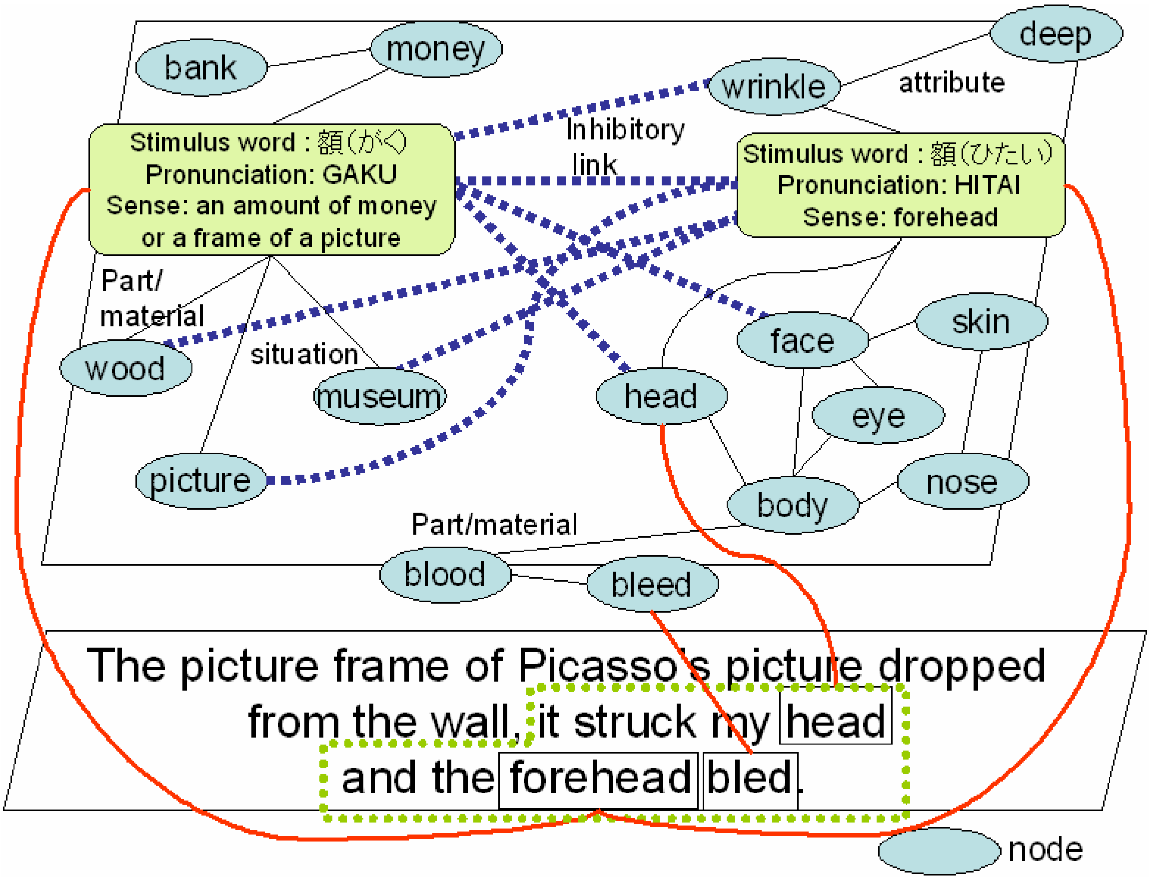

Figure 3 Example of Contextual Semantic Network based on the ACD and the words before and behind the other homographic ideograph "forehead" in an input sentence

Figure 3 shows an example of a contextual semantic network based on the ACD with the other homographic ideograph in the segments of the input sentence. New contextual semantic network is reconstructed by using the second segment in the sentence. The network includes some nodes added by using the previous segment. In addition, the activation values of the node succeed to the values in the previous segment. The polygon with broken lines in the bottom of the figure is an example of the segment in the input sentence.

3.3 Calculation of Node's Activation Value

The activation value of each node is calculated by an interactive activation model in the contextual semantic network. We define the maximum activation level as 1.0. An initial value ($a_i(0)$) is calculated by the following equation.

$$a_i(0) = 1.0 \times S_{ki}, \quad (2)$$

where 1.0 is a normalizing value for S_{ki}, which is a number of node i that appears in sentence k. Next, the new activation value ($a_i(t+1)$) of each node Ni at time $t+1$ is calculated by the following equation (3).

$$a_i(t+1) = a_i(t) - \theta \cdot a_i(t) + \varepsilon_i(t) \ , \quad (3)$$

where the decay parameter θ is assumed to be 0.1 and $\varepsilon_i(t)$ expresses influence of its neighbors at time t. When the neighbors of a node are active, they affect the activation value of the node by excitatory or inhibitory connections, depending on a link between two nodes. Those excitatory and inhibitory influences are combined by a simple equation (4) to yield a net input to the node. Thus, $n_i(t)$ represent the net input to the node by the following the equation.

$$n_i(t) = \sum a_j(t) \Big/ \alpha D_{ij} \ , \quad (4)$$

where $a_j(t)$ denotes an activation value of the node N_j connected with node N_i. α is a constant weight, given by total number of links of the Contextual Semantic Network. D_{ij} denotes a distance between two nodes N_i and N_j. In this paper, The Contextual Semantic Network is constructed by tracing semantic relation paths with accumulating the distance before exceeding the value of 5.0. Therefore, the value of D_{ij} with an inhibitory link is assumed to be -5.0.

When the net input is excitatory, $n_i(t) > 0$, the effect on the node, $\varepsilon_i(t)$, is given by the following equation.

$$\varepsilon_i(t) = n_i(t)[M - a_i(t)] \quad , \quad (5)$$

where M is the maximum activation level of the node and set to 1.0. When the net input is inhibitory, $n_i(t) \le 0$, the effect of the input on the node is given by the following equation.

$$\varepsilon_i(t) = n_i(t)[a_i(t) - m] \quad , \quad (6)$$

where m is the minimum activation level of the node and set to be 0.

4 Experiment for Word Sense Disambiguation by the Proposed Method

For word sense disambiguation, we use the CDNM. This method uses an interactive activation method in a contextual semantic network. Several of Japanese ideographs in the stimulus words in the ACD have a few pronunciations. In the association experiments, such ideographs were presented as stimulus words followed with their pronunciations to avoid ambiguities.

Let an input sentence be "The picture frame of Picasso's picture dropped from the wall, it struck my head, and the forehead bled." In this sentence, "picture frame" and "forehead" are homographic ideographs in Japanese. The sentence is divided into two segments (phrases) "The picture frame of Picasso's picture dropped from the wall" and "it struck my head, and the forehead bled".

At first, we construct contextual semantic network based on the ACD for the first segment "The picture frame of Picasso's picture dropped from the wall". Figure 4 shows activation values of the homographic ideograph for by 20 time cycles. The activation values are probabilities assigned to the word at time t. This homographic ideograph is "額", which have two pronunciations /hitai/ and /gaku/, the former means a forehead and the latter means picture frame. In figure 4, the horizontal axis represents time, and the vertical axis represents activation values.

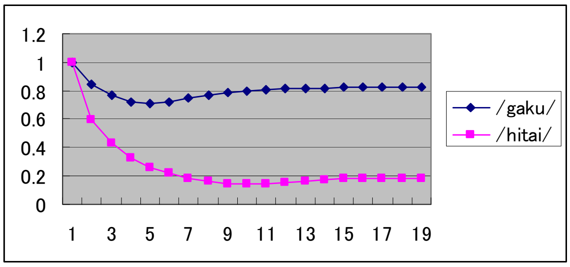

Figure 4 The time course of activation values of selected two nodes "額（がく） /gaku/, picture frame" and "額（ひたい） /hoitai/, forehead" for the first segment

The activation values of "額（がく） /gaku/, picture frame" are higher than the activation values of "額（ひたい） /hitai/, forehead".

Next, we construct a contextual semantic network based on the ACD for the second segment "it struck my head, and the forehead bled". Figure 5 shows an output activation values about homographic ideograph for the simulation through the 20 time cycles. In figure 5, the horizontal axis represents time, and the vertical axis represents activation values.

The activation values of "額（ひたい） /hitai/, forehead" are higher than the activation values of "額（がく） /gaku/, picture frame". We can disambiguate word senses by using words near the homographic ideograph and the interactive activation model.

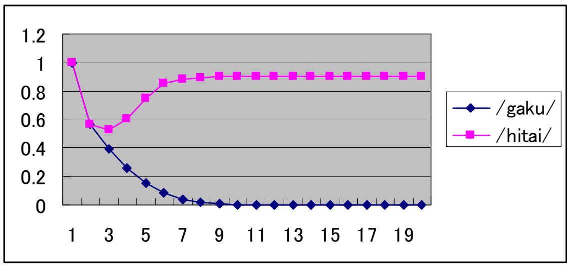

Figure 5 The time course of activation values of selected two nodes "額（がく） /gaku/, picture frame" and "額（ひたい） /hitai/, forehead" for second segment

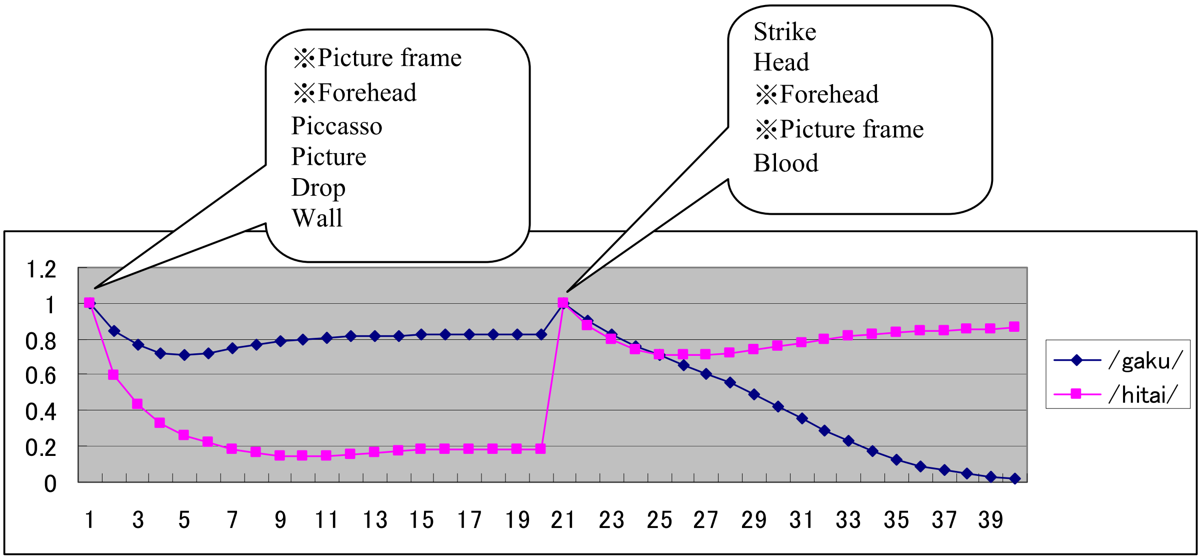

Figure 6 The time course of activation values of selected two nodes "額（がく）/gaku/, picture frame" and "額（ひたい）/hitai/, forehead" for input the sentence "The picture frame of Picasso's picture dropped from the wall, it struck my head, and the forehead bled."

Next, we disambiguate word senses by using the CDNM inputting the sentence "The picture frame of Picasso's picture dropped from the wall, it struck my head, and the forehead bled." The two segments are connected and the activation values of the second segment succeed to the values in the first segment. The initial values of the nodes in the second segment are set to 1.0.

Figure 6 shows activation values of homographic ideograph in the simulation through the 20 time cycles each segment. In figure 6, the horizontal axis represents time, and the vertical axis represents activation values. Several words in two words balloons are speech information for words (nouns, adjectives, adverbs, verbs and so on) in each input segment. In the words balloons, "Picture frame" and "Forehead" are homographic ideograms "額（がく）/gaku/" and "額（ひたい）/hitai/" sequentially.

The contextual semantic network is constructed by using the first segment (picture frame, forehead, Piccasso, picture, drop and wall) in the sentence when the time equals to 1 on horizontal axis. At this point, the initial values of their nodes are set to 1.0. Next new contextual semantic network is reconstructed by using the second segment (strike, head, forehead, picture frame and blood) in the sentence when the time equals to 21 on horizontal axis. The network includes some nodes added by using the previous segment. Here the activation values of the nodes succeed to the values in the first segment and the initial values of their nodes in the second segment are set to 1.0.

The activation values of "額（がく）/gaku/, picture frame" are higher than the activation value of "額（ひたい）/hitai/, forehead" for first segment. Next, the activation values of "額（ひたい）/hitai/" are higher than the activation values of "額（がく）/gaku/, picture frame" for the second segment.

Table 2 shows top ten activation node value for the first segment "The picture frame of Picasso's picture dropped from the wall". "Wood" is a material concept of "picture frame". "Room" is a situation concept of "picture". The activation node values concerning "wood" and "picture" are comparatively high.

Table 2 Top ten activation node values for first segment

Nodes	Activation values
木 -Wood -	0.94
紙 - Paper -	0.92
顔 - Face -	0.91
植物 - Plant -	0.91
壁 - Wall -	0.89
絵本 - Picture book -	0.87
葉 - Leaf -	0.87
絵 - Picture -	0.87
部屋 - Room -	0.87
カレンダー - Calendar -	0.86
皮膚 - Skin -	0.84
森 - Forest -	0.84
根 - Root -	0.84
額（がく） - Picture frame -	0.83
図 - Figure -	0.83

Table3 shows top ten activation node value for second segment "it struck my head, and the forehead bled". "Face" is situation concept of "forehead". "Skin" is part concept of "forehead". "Wounded" is situation concept of "blood". The activation values concerning "face" and "head" are comparatively high. It is thought that the activation value of "wood" is high because the activation value of the "living-thing" is comparatively high.

Table 3 Top ten activation node values for second segment

Nodes	Activation values
木 -Wood -	0.94
血 - Blood -	0.93
顔 - Face -	0.93
植物 - Plant -	0.91
頭 - Head -	0.91
皮膚 - Skin -	0.89
人間 - Human-beings -	0.89
葉 - Leaf -	0.87
額（ひたい） - Forehead -	0.86
森 - Forest -	0.85
傷 - Wounded -	0.84
生物 - Living-thing -	0.84
根 - Root -	0.84
目 - Eye -	0.83
口 - Mouth -	0.81

5 Comparison of Proposed Method and Conventional Method

5.1 Naive Bayes Method for Word Sense Disambiguation

We use a framework of multinomial Naive Bayes text classification for word sense disambiguation. Let s_i, i = 1, 2,…, m, denote a word sense of a homographic ideogram. S is a set of the word senses. Let w_j, j = 1, 2,…, n, denote words appear in the paragraph. We obtain the optimum word sense s which maximizes P($s_i | w_1,..., w_n$) by the following functions.

$$\begin{aligned} s &= \arg\max_{s_i \in S} P(s_i \mid w_1, \cdots, w_n) \\ &= \arg\max_{s_i \in S} P(w_1, \cdots, w_n \mid s_i) P(s_i) \end{aligned},$$

where $P(s_i)$ is a number of paragraphs including s_i divided by total number of the paragraphs. We have a Naive Bayes assumption that words surrounding w_j is independent each other:

$$P(w_1, \cdots, w_n \mid s_i) = \prod_{j=1}^{n} P(w_j \mid s_i) \ .$$

We can determine a probability that it belongs to class s_i by the Bayes' rule:

$$s = \arg\max_{s_i \in S} P(s_i) \prod_{j=1}^{n} P(w_j \mid s_t) \ .$$

Next, we apply the Jeffreys-Perks law (Good, 1965) to solve zero frequency problems. Naive Bayes is a simple and effective method in the statistical machine learning techniques. Despite of its simplicity, this method is effective and often applied to word sense disambiguation.

5.2 Experiment Evaluation and Results

A pair of training and test data from large corpora is used to evaluate the interactive activation model based on the contextual semantic network. The accuracy rate of homographic ideogram's correct pronunciation in all the test data are compared among the interactive activation model and the Naive Bayes method. Several of Japanese ideographs in the stimulus words in ACD have a few pronunciations. In the association experiments, such ideographs were presented as stimulus words with pronunciations to avoid ambiguities by subjects.

A lot of data sets are extracted for the pair of training data and test data including homographic ideograms “額” and “金” from a Japanese newspaper, Mainichi Shinbun (from 1993 to 1995). These data sets include ten words on the left and ten words on the right of the ambiguous words as shown in Table 4. The number of data sets including an ideograph “額” is 1537, and the number of data sets including an ideograph “金” is 3672. The ideograph “金” has two pronunciations /kane/ and /kin/ just like the ideograms “額”, the former means money and the latter means gold, gold medal, Friday or a piece of the Japanese chess.

Next, we labeled correct pronunciations for the homographic ideograms in all the training and test data. Pairs of training data (about 95% of data sets) and test data (about 5% of data sets) are used to evaluate the performance of the proposed method. The evaluation for a word sense disambiguation is designed to check the effectiveness of the CDNM. We use data sets for homographic ideogram “額” which means “forehead” or “picture frame”, and use ones for homographic ideogram “金” which means “gold” or “money”

Table 4 Comparison of test results among the two methods. NB: Naive Bayes Method

	neighborhood words	Our method	NB
An ideograph “額” pronunciation as /hitai/ and /gaku/	10 words	90.4%	80.1 %
An ideograph “金” pronunciation as /kin/ and /kane/	10 words	86%	90.6 %

Table 4 shows accuracy rates of disambiguation for “額” or “金” in test data. The neighborhood words are words on the left or right of an ambiguous word. Our method shows the best score for the two homographs as the correct pronunciations’ ratio in all test data of homographic ideogram “額”. About an homographic ideogram “金”, the data sets of pronunciation /kin/ include a lot of idioms. For example “金の卵 /kin no tamago/” that means a “Golden boy”. Therefore NB shows best score for homographic as the correct pronunciations’ ratio in all test data of homographic ideogram “金” because there are a lot of idioms in test data sets which include neighborhood words of ambiguous words.

It is difficult to disambiguate word sense by using the proposed method. However the correct pronunciations’ ratio is 98% about the test data of pronunciation /kane/.

6 Future Work

In this research, we proposed a method for disambiguation of pronunciations of homographic ideographs by using the CDNM. We made computer simulations to show effectiveness of this model and consistency with empirical data.

The ideograph “額” has two major senses when its pronunciation is /gaku/. One is an amount of money and the other is a picture frame. It is necessary to carry out more experiments to disambiguate the two meanings as well as the two pronunciations. In addition, if we use the concept dictionary which includes structurized information about word senses of ambiguous words, we can disambiguate not only homographic ideogram but also various word senses by using the Contextual Dynamic Network Model.

Acknowledgement

We wish to express our gratitude to the students at SFC, Keio University who were very helpful for the association experiments, and also to the member of Ishizaki Laboratory who helped us to construct and modify the Associative Concept Dictionary.

References

Choueka, Y. and Lusifnan, S., “Disambiguation by short contexts”, Computers and the humanities, Vol. 19, pp 147-157, 1985.

EDR, “EDR Electronic Dictionary Technical Guide”, Japan Electronic Dictionary Research Institute, Ltd., 1990.

Matsumoto, Y., Kitauchi, A., Yamashita, T., Hirano, Y., Matsuda, H. and Asahara, M., “Japanese Morphological Analysis System ChaSen Manual version 2.0 Manual 2nd edition” (in Japanese)., NAIST Technical Report, NAIST-IS-TR99009 Nara, Institute of Science and Technology, 1999.

McCelland J.L. and Rumelhart D.E., “An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings”, Psychological Rev., Vol.88, No.5, pp375-407,1981.

Murata, M., Utiyama, M., Uchimoto, K., Ma, Q. and Isahara, H., “CRL at Japanese dictionary-based task of SENSAVAL-2 -- Comparison of various types of machine learning methods and features in Japanese word sense disambiguation --" (in Japanese)., Journal of NLP Vol.10 No.3, 2003.

Okamoto, J. and Ishizaki, S., "Construction of Associative Concept Dictionary with Distance Information, and Comparison with Electronic Concept Dictionary" (in Japanese)., Journal of NLP Vol.8 No.4, 2001.

Okamoto, J. and Ishizaki, S., "Word Sense Disambiguation Using Contextual Semantic Network Using Associative Concept Dictionary", SNLP 2005, pp205-210, 2005.

Takahashi, N., "Lexical Disambiguation by Layered Neural Networks" (in Japanese)., IPSJ Journal, Vol.36 No.9, 1995.

Veronis, J. and Ide, N. M., "Word Sense Disambiguation with Very Large Neural Networks Extracted from Machine Readable Dictionaries"., Coling '90, 1990.

Waltz, D.L. and Pollack, J.B., "Massively parallel parsing: A strongly interactive model of natural language interpretation", Cognitive Science, Vol.9,pp51-74,1985.