A Computational Model of the Metaphor Generation Process

Keiga Abe (abe@nm.hum.titech.ac.jp)
Kayo Sakamoto (sakamoto@nm.hum.titech.ac.jp)
Masanori Nakagawa (nakagawa@nm.hum.titech.ac.jp)
Graduate School of Decision Science & Technology, Tokyo Institute of Technology
2-12-1, Ohkayama, Meguro-Ku, Tokyo, 152-8552, Japan

Abstract

The purpose of this research was to construct a computational model of the metaphor generation process. In order to construct the model, first, the probabilistic relationship between concepts and words was computed with a statistical analysis of language data. Secondly, a computational model of the metaphor generation process was constructed with results of the statistical analysis of language data. The results of the simulation were examined from a comparison with metaphors that participants had generated. Finally, a third-party rating of the metaphors the model generated was conducted.

Introduction

Metaphor understanding and generation processes are very important aspects of language study. However, most cognitive studies of metaphor focus on the metaphor understanding process(Lakoff & Johnson, 1986; Glucksberg & Keysar, 1990; Kusumi, 1995), while studies of the metaphor generation processes are relatively few. The purpose of this study is to construct a computational model which generates a "A like B" style metaphor process. In the case of "A like B" sytle metaphors, A is called the "vehicle", and B is called the "topic".

In a previous study, Kusumi(2003) showed that belief or experience affects the metaphor generating process, using a metaphor generation task dealing with the concept of love. Hisano(1996) studied the relationship between the impression of the topic and that of generated metaphors, using a metaphor generation task where the categories of topic and vehicle were limited. However, these studies were limited to a few concepts or categories. It is not clear whether the results are applicable in the case of other concepts. In order to examine the applicability of the studies, the experimenter must conduct a metaphor generation task with a huge number of concepts. It is impossible to cover large scale language knowledge using only a psychological experiment, because psychological experiments require expensive time and labor.

In order to solve this problem, a statistical analysis of language data was used to represent large scale human language knowledge stochastically. Applying statistical analysis, a stochastic language knowledge structure can be automatically constructed without subjective judgement. In this study, a statistical analysis of language data was conducted and a computational model of the metaphor generation process was constructed based on the results of the statistical analysis. After that, a psychological experiment was conducted to examine the validity of the model.

Probabilistic representation of meaning

In previous studies, practical methods to compute the probabilistic relationship between concepts and their words, between words and words have been developed. For example, LSA(Landauer & Dumais, 1997) assumes semantically similar words occur in common contexts. In LSA, text data are represented as a matrix in which each row stands for a unique word and each column for a text passage or other context. Each cell stands for the frequency with which the word of its row appears in the passage denoted by its column. After that, LSA applies singular value decomposition (SVD) to the matrix, as follows:

$$S = U_k \Sigma_k U_k^{'}. \tag{1}$$

Using this method, the meaning of words can be represented in the coordinate of a vector space. Furthermore, semantic similarities between words and words are represented by the cosine distance of vectors.

However, LSA can not treat functional words(for example, "the", "a", "is"). Generally, functional words occur in various contexts with high occurrence frequency. Such cooccurrence between content words and functional words do not necessarily reflect semantic relation. In order to avoid this problem, LSA has to set a strong weight to high occurrence frequency words. or omit low occurrence frequency words, However such a weighting method is likely to be subjective and ad-hoc.

PLSI(Hofmann, 1999) is a probabilistic model for the relationship between concepts and words based on the idea of LSA. PLSI assumes that latent semantic classes c's mediate the probability of cooccurrence between documents d's and words w's. In PLSI, the probability of cooccurrence between a document d and a word w, $P(d, w)$ is represented by the following equation:

$$P(d, w) = \sum_{c \in C} P(d|c)P(w|c)P(c), \tag{2}$$

where $P(d|c)$ stands for the conditional probability of a document d, given a latent semantic class c, $P(w|c)$ stands for the conditional probability of w, given c, and $P(c)$ stands for the probability of c. Applying this probailistic representation, PLSI does not need particular weighting to word, because the weight of a word w is included in the probability of cooccurrence $P(d,w)$.

There are other methods for the probabilistic representation of meaning of words; Pereira(1993) proposed a computational method for the probabilistic representation of the relationships between nouns and verbs. Kameya & Sato(2005) provided another statistical model based on PLSI to represent the relationship between words and words in Japanese. The model assumes that the cooccurrence probability of a word "n_i" and a word "a_j", $P(n_i, a_j)$ is computed by the following formula;

$$P(n_i, a_j) = \sum_k P(n_i|c_k)P(a_i|c_k)P(c_k), \qquad (3)$$

where $P(n_i|c_k)$ means the conditional probability of n_i, given c_k which indicates a latent semantic class assumed in this model. Parameters in the model, $P(c_k), P(n_i|c_k)$ are estimated to be the values that maximaizes the likelihood of co-occurrence data measured from the language corpus, with the EM algorithm.

In this model, the meanings of words are represented as a probability distribution of $P(a_i|c_k)$ or $P(n_j|c_k)$. Furthermore, Kameya & Sato's model can represent a semantic similarities between words and words as KL-divergence. This model was used for computational models of high order cognitive processes, for example, the metaphor understanding process(Terai & Nakagawa, 2005). This model can also be applied to the metaphor generation process using this probabilisitic representation of meaning.

In this study, first, a probabilistic representation of language knowledge was constructed, by applying Kameya & Sato's model to a language corpus taken from the Japanese newspaper, the Mainichi-Shinbun, over a period of 10 years (1993-2002). One of the main reasons for using this Japanese newspaper is the fact that it is read by a wide range of Japanese readers. The language corpus consisted of 2783 adjectives and 14000 nouns. The probabilistic representation consisted of 50 latent semantic classes. Some examples of the result are shown in Table 1. Table 1 shows the rank order of conditional probabilities of c_i, given noun n_j, or conditional probability of c_i, given adjective n_j. The rank order of nouns and adjectives suggests that the latent semantic class represents a conceptual category about "infant" or "art". While the names of the latent semantic classes were applied by the authors for the practical convenience, this naming has no effect on the results of the simulation discussed below. The 2783 adjectives and 14000 nouns are classified by 50 latent semantic classes.

Table 1: Results of statistical analysis of language data.

"infant" latent semantic class			
nouns	$P(C\|N)$	adjectives	$P(C\|A)$
grandchild	0.8077	young	0.9711
girl	0.7184	fine	0.891
son	0.6996	lovable	0.8701
character	0.6753	mild-mannered	0.8549
child	0.6721	slight	0.8469
sister	0.6665	docile	0.7986
baby	0.6328	small	0.7906
sleeping face	0.6231	slender	0.779
body	0.6204	innocent	0.7626
initial cry	0.6143	tragic	0.7596

"art" latent semantic class			
nouns	$P(C\|N)$	adjectives	$P(C\|A)$
harmony	0.7564	mild	0.932
tune	0.7465	witty	0.931
amazement	0.7333	noble	0.9161
merody	0.7073	plain	0.9115
singing voice	0.6946	heroic	0.894
lyric	0.6792	fresh	0.8933
strain	0.6571	flowing	0.8655
poetry	0.6509	massive	0.8606
landscape	0.6508	elegant	0.8553
mid-age	0.6466	hard	0.855

Metaphor generation model

In this study, it is assumed that the metaphor generating process is a kind of word association between base words (vehicle) and target words (topic). The association process can be represented as a cooccurrence relationship between words and words in Kameya & Sato's model. Furthermore, it is assumed that the latent semantic class c_k as a high order semantic category in human-being's concepts, and conditional probability $P(a_j|c_k)$, $P(n_i|c_k)$ in Kameya & Sato's model as relationship strengths between the semantic category and adjectives or nouns.

Based on the above assumption, a computational model that trasforms adjective-modified nouns (for example, "young, innocent, and fine character") into "A like B" style metaphors (for example, "the character is like a child") was constructed. The model consists of the three layers below(Figure 1):

input layer: Each node in this layer corresponds to a word which constructs the phrases for metaphors.

hidden layer: Each node in this layer corresponds to a latent semantic class c_k in Kameya & Sato's model assumed as a high order semantic category of human-being's concepts.

output layer: In this layer, each node corresponds to the word for the vehicle of a metaphor.

In this model, weights of links between each layer are determined with conditional probability $P(a_j|c_k)$, $P(n_i|c_k)$.

According to the model, metaphor generation is processed in the following steps:

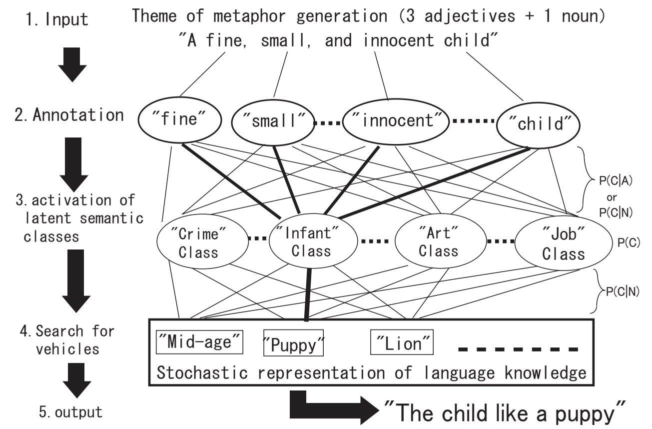

Figure 1: The image of metaphor generation model

1. When a phrase for metaphors is input to the model, the model runs a syntactic analysis of the phrase, and decomposes the phrase to adjectives and nouns.

2. Binary values are assigned to nodes in the input layer. The value 1 is assigned to the nodes corresponding to the adjectives or nouns, while the value 0 is assigned to the other nodes.

3. Activations of the input layer are transferred to the hidden layer. The activation value of node i in the hidden layer, u_i is computed as follows:

$$w^1_{ij} = P(c_i|n_j), \tag{4}$$

$$s_i = \Sigma_j w^1_{ij} \cdot n_j, \tag{5}$$

$$u_i = \frac{1}{1 + exp^{-s_i}}. \tag{6}$$

In these equations, w^1_{ij} is the conditional probability corresponding to the weight of the links between the input layer and the hidden layer. Applying a sigmoid function in equation (5), even though such a large value is used for a specific node, the final activation value does not become larger than 1.

4. In the output layer, each node receives the activation transferred from the hidden layer. The activation value of each node o_l is computed with the equations as follows:

$$w^2_{il} = P(c_i|n_l), \tag{7}$$

$$v_l = \sum_i u_i \cdot w^2_{il}, \tag{8}$$

$$o_l = \frac{1}{1 + exp^{-v_l}}. \tag{9}$$

In these equations, w^2_{il} is the conditional probability corresponding to the weight of the links between the input layer and the hidden layer. In the model, it is assumed that the activation value of each node of the output layer represents the probability of the word being represented by the node as the vehicle of the metaphor.

In this study, a probabilistic representation of language knowledge was constructed by applying Kameya & Sato's model to a language corpus. After that, a metaphor generation model with probabilistic representation of language knowledge was constructed.

Simulation

In order to evaluate the model, simulations were conducted using three types of input phrases. Each input phrase consists of a noun with three adjectives. Each word of the input phrases were selected at random from top ten words according to rank order of conditional probabilities $P(C|N)$ or $P(C|A)$.

1 class input: This type consists of nouns and adjectives which are strongly related to the same latent semantic class. For example, in the case of the input phrase "young, innocent, and fine character", all words are strongly related to the "infant" latent semantic class.

2 classes input: This type consists of adjectives strongly related to one latent semantic class and a noun related to another latent semantic class. For example, in the case of the phrase "excellent, admirable, and famous son", the adjectives are strongly related to the "Job" latent semantic class, and the noun is strongly related to the "infant" class.

4 classes input: This type consists of words strongly related to separate latent semantic classes independently. For example, In the case of the phrase "small, elegant, and disconsolate nobility", each word is strongly related to the "infant", "art", "emotion" and "job" classes, respectively.

In this simulation, the activation of output values concerning input phrases was computed . After that, the top 20 words were considered as metaphors the model generated. Results of the simulation are shown in Tables 2,3,4.

According to the model, in the case of 1 class input, all words of each input phrase activate a certain specific class. In this case, the metaphors the model generated are concrete and easy to imagine. On the other hand, in the case of 2 classes input, the input phrase activates two latent semantic classes. The model then generates intermediate words between the two classes. Therefore, the metaphors the model generated are a little ambiguous compared to the case of 1 class input. In the case of 4 classes input, the metaphors the model generated are less easy to visualize compared to the metaphors from 1 class or 2 classes input.

For the comparison with these models' output, a metaphor generation task was conducted for 22 native Japanese speakers. In this task, participants generated "A like B"style metaphors from 3 input phrases. Those phrases presented to participants were the same input phrases that were used for the model simulation

Table 2: Metaphors the model generated from the input phrase "young, innocent and fine character"

order	a character like a "XXX"	output value
1	grandchild	0.5928
2	girl	0.583
3	son	0.5809
4	child	0.5777
5	sister	0.5772
6	baby	0.5731
7	sleeping face	0.5721
8	body	0.5719
9	baby's first cry	0.5711
10	character	0.5685
11	physical frame	0.5641
12	young man	0.561
13	boy	0.5592
14	daughter	0.5587
15	old folks	0.5585
16	infant	0.5575
17	appearance	0.5571
18	entrepreneurial spirit	0.5563
19	eldest-son	0.5551
20	second son	0.5545

Table 3: Metaphors the model generated from the input phrase "excellent, admirable, and famous son"

order	the son like a "XXX"	output value
1	academic	0.5728
2	surgeon	0.5657
3	human resource	0.5599
4	artist	0.5598
5	nobility	0.5559
6	painter	0.5551
7	soldier career	0.552
8	forerunner	0.5502
9	sense	0.5501
10	old man	0.5485
11	artist of calligraphy	0.548
12	military commander	0.547
13	flower	0.5462
14	student	0.545
15	general	0.5439
16	shrine	0.5436
17	heated battle	0.5435
18	engineer	0.5433
19	musician	0.5423
20	mis-thrown pitch	0.5415

Table 4: Metaphors the model generated from the input phrase "small, elegant, and disconsolate nobility"

order	the nobility like a "XXX"	output value
1	mind-set	0.5505
2	expression	0.5476
3	scream	0.5468
4	passion	0.5454
5	singing voice	0.5421
6	harmony	0.5417
7	mentality	0.5413
8	tune	0.5412
9	amazement	0.54
10	lost point	0.5394
11	grand child	0.5389
12	melody	0.5388
13	appearance	0.5383
14	lyric	0.5381
15	manner	0.538
16	landscape	0.5371
17	girl	0.5369
18	poetic state of mind	0.5369
19	strain	0.5368
20	ring	0.5367

above. Participants were asked to generate as many metaphors as possible in 5 minutes. The results of the task are shown in Tables 5,6,7. In the metaphor generation task of the input phrases "young, innocent and fine character" and "excellent, admirable, and famous son", most participants generated the same metaphors as the model did with high output value (For example, "a character like a child", "a grandchild like a academic"). Some of the metaphors the participants generated didn't consist of the same metaphors the model generated. However, participants do not always generate good metaphors. There is a possibility that participants generated nonsense metaphors, while the model generated good metaphors the participants did not. Therefore, in the next section, a third-party rating of the metaphors both the participants and the model generated was conducted.

Rating

In this section, a third-party rating of the metaphors both the model and the participants generated was conducted.

Method

raters: In this evaluation, 13 college students participated. All raters were native Japanese speakers.

materials: Metaphors participants evaluated consist of three groups.

Model's metaphors: This group consists of metaphors the model generated, and human participants did not. Three metaphors were chosen

Table 5: Metaphors participants generated from the phrase "young, innocent and fine character"(*:matched with model output).

"young, innocent and fine character"		
order	the character like a "XXX"	number of answers
1	child*	16
2	puppy	13
3	sun	7
4	boy, flower, cat	3
5	infant*, girl*, glass, hamster, fireworks	2
6	moon, ball, air, sky, strawberry, straight line, wind, doll, puffball, the color of yellow, summer, budworm, yarn, typhoon	1

Table 6: Metaphors participants generated from the phrase "excellent, admirable, and famous son"(*:matched with model output).

"excellent, admirable, and famous son"		
order	the son like a "XXX"	number of answers
1	academic*	8
2	sun	6
3	diamond, teacher	4
4	god	3
5	military commander*, proffessor, ball,angle, president, adult, king	2
6	govemment official, top, father, forerunner*, sample, doctor, music, elite, witster, thinker, poet, padre, monk, star···	1

Table 7: Metaphors participants generated from the phrase "small, elegant, and disconsolate nobility"

"small, elegant, and disconsolate nobility"		
order	"nobility like a XXX"	number of answers
1	cat	5
2	aristocrat, dame,gem, grandmother, inkstick, rich folk,fallen leaves	2
3	moon, rose, flower, diamond, chocolate, doll, neckrace, perl, chesil, rainbow, angel, dead tree···	1

from each input phrase, so this type consists of 9 metaphors (3 phrases x 3 metaphors).

Partcipants' metaphors: This group consists of metaphors the human participants generated, and those the model did not. Three metaphors were chosen from each input phrase, so this type consists of 9 metaphors(3 phrases x 3 metaphors).

Matched metaphors: This group consists of metaphors both the human participants and the model generated. This type consists of 6 metaphors because there were no matched metaphors from the input phrase "small, elegant, and disconsolate nobility" (2 phrases x 3 metaphors).

participants were shown these metaphors with the materials used for genenerating these metaphors.

procedure: Metaphors were presented without informing the raters as to who generated it. Raters rated the metaphors by 3 types of scales of 1 point to 7 point.

adequacy: In this scale, the more adequate the expression of material, the higher the score is.

ease of visualization: In this scale, the more easily visualized the metaphor is, the higher the score is.

amusingness: In this scale, the more amusing the metaphor is, the higher the score is.

novelty: In this scale, the more novel the metaphor is, the higher the score is.

Results

Table 8 shows the result of the rating. In this analysis, the average scores of each type of metaphors were compared, by each input phrase. For the comparison, the average scores on each scale were computed, by each type of metaphor in the input phrase.

In comparison with other cases using Bonferroni's method, the metaphors of the input phrase "young, innocent and fine character", the matched metaphors gained high evaluation score on the scales of adequacy ($F(2,24) = 37.667, p < 0.01$) and ease of visualization ($F(2,24) = 50.665p < 0.01$). On the other hand, the model metaphors gained significantly high evaluation scores on the scale of novelty compared to the human metaphors ($F(2,24) = 7.866, p < 0.01$).

The metaphors of the input phrase "excellent, admirable, and famous son", the matched metaphors gained higher evaluation scores on the scales of adequacy ($F(2,24) = 4.791, p < 0.01$) and ease of visualization than the model metaphors ($F(2,24) = 5.576, p < 0.01$). On the scale of novelty, the scores of model metaphors were significantly higher than the others ($F(2,24) = 5.473, p < 0.01$).

In comparison with the model metaphors of the input phrase "small, elegant, and disconsolate nobility" with the t-test, the human metaphors gained higher evaluation scores on scales of adequacy($t(12) = -6.434, p <$

0.01) and ease of visualization than model metaphors ($t(12) = -6.08, p < 0.01$). On the scale of novelty, the scores of the model metaphors were significantly higher than the human metaphors ($t(12) = 5.505, p < 0.01$).

Table 8: Third party evaluation of metaphors generated by the participants and model

"young, innocent and fine character"

	model	human	matched
adequacy	3.49	2.74	5.59
ease of visualization	3.46	3.00	5.72
amusingness	3.54	3.18	3.23
novelty	4.08	3.51	2.79

"excellent, admirable, and famous son"

	model	human	matched
adequacy	2.51	3.51	3.59
ease of visualization	3.15	4.00	4.31
amusingness	3.82	3.62	3.85
novelty	5.13	4.46	4.15

"small, elegant, and disconsolate nobility"

	model	human
adequacy	1.79	3.87
ease of visualization	1.44	4.08
amusingness	3.85	3.79
novelty	5.64	3.74

Discussion

In this study, a statistical analysis of language data was conducted, and a computational model of the metaphor generation process was constructed, based on the results of the statistical analysis. The central focus of this study is the application of a new statistical method, as a probabilistic version of LSA, to the construction of the computational model.

Futhermore, the simulation of the model was conducted, and the results were compared with metaphors human participants generated. From a comparison between model output and participants answers in the case of 1 class and 2 classes input phrases, most of the participants generated metaphors which had high output values in the model. Furthermore, in third-party rating of metaphors in the case of 1 class input phrases, the metaphors the model generated were more highly evaluated compared to those generated by the participants. This result suggests that the model might generate good metaphors participants overlooked.

However, in comparison to metaphors from 4class input phrases, the model did not match participants' metaphors at all. Furthermore, in the third-party rating of metaphors in the case of 4class input phrase, on the scales of ease of visualization and adequacy, the scores of metaphors the model generated were significantly lower than participants' metaphors. In other words, the model generates nonsense metaphors in the case of 4class input phrases. These sharp differences may reflect cognitive mechanisms that the model does not possess. For 4 classes input phrases, participants could generate metaphors which were adequate and easyily visualize. There is a possibility that participants generated abstract images, which intermediate words of input phrase moderately. Alternatively, participants might focus on emergent features. In this case, emergent features are assumed to be features which become salient only if specific words are well combined. In this study, participants might use emergent features of input phrases to generate metaphors.

The future challenges of this study are to clarify a mechanism of internal evaluation in generating metaphors. Metaphor generation is a kind of divergent thinking. Therefore it does not necessarily have a single or correct answer. However, participants have an internal evaluation method for discriminating the metaphors they generate. There is a possibility that this internal-evaluation mechanism filters out nonsense metaphors.

References

Glucksberg, S & Keysar, B. (1990). Understanding metaphorical comparisons: Beyond similarity, *Psychological Review*, **97**,3-18.

Hisano, M.(1996). Topic no Tokusei ga Hiyu Seisei ni Oyobosu Eikyou, *Proceedings of 38th Annual Conference of Japanese Association of Educational Psychology.* 449.(in Japanese).

Hofmann,T.(1999). Probabilistic latent semantic indesing, *Proceedings of the 22nd Annual ADM Conference on Research and Development in Information Retrieval (SIGIR'99)*, 50-57.

Kameya. Y & Sato. T.(2005). Computation of Probabilistic Relationship between Concepts and their Attributes Using a Statistical Analysis of Japanese Corpora. *Proceedings of Symposium on Large-Scale Knowledge Resources*, 65-68.

Kusumi, T.(1995). *Hiyu No Syorikatei To Imikouzou*, Kazama Shobou.(in Japanse).

Kusumi, T.(2003). Hiyu Seisei wo Sasaeru Shinnen to Keiken: Ai no Hiyu no Haigo ni Aru Renai Kihan to Keiken,*Proceedings of 66th Annual Meeting of Japanese Psychological Association* 811.(in Japanse).

Lakoff & Johnson. (1980). *Metaphors We Live By.* The University of Chicago Press.

Landauer, T.K. & S.T. Dumais. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, **104**(2):211-240,

Pereira, F., Tishby, N., and Lee, L.(1993). Distributional Clustering of English Words, *The Proceedings of 31st Meeting of the Association for Computational Linguistics*, 183-190,

Terai, A. & Nakagawa, M. (2005) A development method of a metaphorical search engine, *World Conference on Educational Multimedia, Hypermedia & Telecommunications*, 1298-1303.