A Method of Calculating the Measure of Salience in Understanding Metaphors

IWAYAMA Makoto and TOKUNAGA Takenobu and TANAKA Hozumi
Tokyo Institute of Technology
O-OKAYAMA, 2-12-1
MEGURO-KU, TOKYO, 152, JAPAN
iwa@cs.titech.ac.jp

Abstract

This paper presents a computaional method of calculating the measure of salience in understanding metaphors. We mainly treat metaphors in the form of "A is (like) B," in which "A" is called *target concept*, and "B" is called *source concept*. In understanding a metaphor, some properties of the source concept are transferred to the target concept. In the transfer process, we first have to select the properties of the source concept that can be more preferably transferred to the target concept. The measure of salience represents how typical or prominent the property is and is used to measure the transferability of the property. By introducing the measure of salience, we have to consider only the high salient properties after the selection. The measure of salience was calculated from Smith & Medin's probabilistic concept[12, 13] according to Tversky's two factors[14]. One is intensity which refers to signal-to-noise ratio; this is calculated from the entropy of properties. The other is diagnostic factor which refers to the classificatory significance of properties; this is calculated from the distribution of the property's intensity among similar concepts. Finally we briefly outline the whole process of understanding metaphors using the measure of salience.

Introduction

Natural language is a rich source of metaphors, and metaphors have strong relationship with the conceptual structure that has been acquired through our everyday life[7]. For this reason, we should consider metaphors to develop a better natural language understanding system.

In this paper, we mainly treat metaphors in the form of "A is (like) B," in which "A" is called *target concept*, and "B" is called *source concept*. We consider the understanding metaphor as a transfer process of properties from the source concept to the target concept. For example, in the case of "A man is a wolf," some properties of "wolf" — "being vicious, dangerous, fierce, etc." are transferred to "man." As a consequence of the transfer, "man's" properties "being vicious, dangerous, fierce, etc." are highlighted.

This transfer process consists of the following three steps. First, we have to select properties of the source concept that can be transferred to the target concept. We call this step *selection step*. Secondly, we have to find the properties of the target concepts which correspond to the properties selected in the selection step. We call this step *mapping step*. Finally, we have to highlight or downplay the properties of the target concept according to the corresponding properties of the source concept. We call this step *variance step*. These steps are very similar to that of analogical reasoning[4]. In this paper, we focus on the selection step and show how this step is achieved.

In the selection step, we have to select the properties of the source concept that can be more preferably transferred to the target concept. We introduce *the measure of salience* that measures the transferability of the property. Generally, the measure of salience represents how typical or prominent the property is. By introducing the measure of salience, we have to consider only the high salient properties after the selection step. There are many properties that play little importance during the whole process of understanding metaphors. With respect to the above example, "wolf's" properties "being vicious," "being fierce," "being dangerous" are high salient properties and are more likely to be transferred to "man." On the other hand, "having two eyes," "having four legs," etc. are low salient properties and cut off at the selection step.

Many researchers have used salience in the process of understanding metaphors[10, 9, 15], but they have not described precisely how salience is calculated. It is necessary to show the foundation which salience was based on and the method of calculating salience based on the foundation. In this paper, we propose a method to calculate the measure of salience from Smith & Medin's probabilistic concept[12, 13] which has a grounding in probability theory. According to Tversky[14], we calculate the measure of salience in terms of two factors. One is intensity which refers to the signal-to-noise ratio; this is calculated from the entropy of properties. The other is diagnostic factor which refers to the classificatory significance of properties; this is calculated from the distribution of the property's intensity among similar concepts. Finally we briefly outline the whole process of understanding metaphors using the measure of salience.

Probabilistic concept and the Measure of Salience

In this section, we describe our method of calculating the measure of salience. First, we briefly review ***probabilistic concept*** that Smith & Medin have proposed[12, 13]. Our measure of salience is calculated based on the probabilistic concept.

Probabilistic Concept

Probabilistic concept is composed of a set of properties, each of which has an attribute with a set of possible values. Probability is attached to each value. We use probabilistic concept as our concept representation form, because:

- Probabilistic concept is not "all or none" concept, and this feature is necessary to calculate the measure of salience as ranging from 0 to 1.
- Once probabilities are attached to attribute's values, they can be treated formally based on probability theory.

We give a definition of concept, which is slight different from Smith & Medin's original definition.

Definition 1 Concept
A *Concept* denoted by *(C) is a set of *properties* S_i.

$$*(C) = \{S_1, S_2, \ldots, S_n\}$$

A property S_i is a pair of an *attribute* a_i and its *possible value set* V_i.

$$S_i = a_i : V_i$$

The possible value set V_i is a set whose element is a pair of a_i's *possible value* $v_{i,j}$ and its *probability* $w_{i,j}$ among the V_i.

$$V_i = \{v_{i,1}\#w_{i,1}, \ldots, v_{i,j}\#w_{i,j}, \ldots, v_{i,m}\#w_{i,m}\}$$

That is,

$$\sum_{j=1}^{m} w_{i,j} = 1$$

Most Likelihood Value(MLV) $v_{i,max}$ is the value with the highest probability among a possible value set V_i, and *Most Likelihood Property(MLP)* $S_{i,max}$ is the pair of attribute a_i and its MLV $v_{i,max}$ and is denoted by $a_i : v_{i,max}$.

Following is an example of *(Apple).

Example 1 Definition of *(Apple)

$$*(\text{Apple}) = \left\{ \begin{array}{ll} \textit{color}: & \left\{ \begin{array}{l} \textit{red}\#0.8 \\ \textit{green}\#0.15 \\ \textit{brown}\#0.05 \end{array} \right\} \\ \textit{shape}: & \left\{ \begin{array}{l} \textit{round}\#0.95 \\ \textit{cylindrical}\#0.05 \end{array} \right\} \\ \textit{texture}: & \left\{ \begin{array}{l} \textit{smooth}\#0.9 \\ \textit{rough}\#0.1 \end{array} \right\} \\ \vdots & \end{array} \right\}$$

"$\textit{color}$: $\{\textit{red}\#0.8, \textit{green}\#0.15, \textit{brown}\#0.05\}$" is a property, and "$\{\textit{red}\#0.8, \textit{green}\#0.15, \textit{brown}\#0.05\}$" is its possible value set, where each real number is the probability of the value. "*red*" is the MLV and "*color* : *red*" is the MLP.

The probability attached to the each value can be understood as the rate of the concept's instances which was observed to have the value. For example, one who has the above representation observed 80% of the "apples" as "red apples."

The Measure of Salience

Each property has a measure of salience which is a real number ranging from 0 to 1. 0 and 1 represents the lowest and the highest salience respectively. The measure of salience represents the typicality of a property and is used in understanding metaphors to decide which properties of a source concept might be more preferably transferred to a target concept. Many researchers have used the measure of salience in the same way as mentioned above[10, 9, 15], but they have not shown the precise method to calculate the measure of salience.

In the field of cognitive psychology, Tversky have pointed out the qualitative nature of salience[14]. Tversky says that salience is determined by two types of factors; intensitive and diagnostic. The former refers to the signal-to-noise ratio and the later refers to the classificatory significance of properties. In the following sections, we show the method of calculating the measure of salience according to Tversky's two factors. Intensitive factor is calculated based on the entropy in information theory, and we call this measure ***the Amount of Information of Property(AIP).*** Another diagnostic factor is calculated based on the distribution of a property's AIP among similar concepts, and we call this measure ***the Difference of Property(DP).***

The Amount of Information of Property(AIP)
The first factor in calculating the measure of salience is the amount of information which a property has. This is calculated by the entropy of a possible value set V_i. Because the entropy is a measure of randomness, the lower the entropy is, the less random a possible value set V_i is, that means V_i has more redundant information. Intuitively, more redundant V_i means that its MLV $v_{i,max}$ occurs more frequent comparing with other values of V_i. It follows that the property S_i with more redundant V_i is the more typical and salient property.

For example, compare the following two properties, $S_1 = color : V_1$ and $S_2 = color : V_2$.

$V_1=\{red\#0.6, green\#0.1, yellow\#0.1, blue\#0.1, brown\#0.1\}$

$V_2=\{red\#0.6, green\#0.4\}$

In these two possible value sets, probability of each MLV "*red*" are the same 0.6. But if we take account of the distribution of all elements in each possible value set, MLV of V_1 occurs more redundantly than that of V_2 because, the degree of concentration of V_1 is higher than that of V_2. In fact entropy (relative entropy) of V_1 (i.e. 0.7627 according to the definition below) is lower than that of V_2 (i.e. 0.9705), and it shows that V_1 is more redundant and has more information than V_2.

Definition 2 The amount of information of property(AIP)
Given a property $S_i = a_i : V_i$, in which

$$V_i = \{v_{i,1}\#w_{i,1}, v_{i,2}\#w_{i,2}, \ldots, v_{i,m}\#w_{i,m}\}$$

the AIP of S_i is denoted by $r(V_i)$ and calculated by the following expression:

$$r(V_i) = 1 - h(V_i)$$

where $h(V_i)$ is

$$h(V_i) = \begin{cases} 0 & \text{if } m = 1 \\ \frac{H(V_i)}{\log_2 m} & \text{otherwise.} \end{cases}$$

$H(V_i)$ is

$$H(V_i) = \sum_{j=1}^{m} w_{i,j} \log_2 \frac{1}{w_{i,j}}$$

$r(V_i)$, $h(V_i)$ and $H(V_i)$ are called "redundancy," "relative entropy," and "entropy" respectively in the information theory[8].

The AIP ranges from 0 to 1 depending on the diversity of a possible value set. If the diversity of a possible value set concentrates on only one value — for example $x : \{a\#1, b\#0, c\#0, d\#0\}$, the property has the highest AIP 1, because all instances has the value a. To the contrary, if the diversity of a possible value set is averaged — for example $x : \{a\#0.25, b\#0.25, c\#0.25, d\#0.25\}$, the property has the lowest AIP 0, because one can not successfully predict which value a instance has.

Example 2 Calculation of AIP
Consider the following three possible value sets which are the value of attribute "*color*" of some fruits.

$V_{Color\_of\_Apple}=\{red\#0.8, green\#0.15, brown\#0.05\}$

$V_{Color\_of\_Strawberry}=\{red\#0.9, purple\#0.1\}$

$V_{Color\_of\_Grape}=\{violet\#0.7, green\#0.2, red\#0.1\}$

Each AIP becomes,

$$r(V_{Color\_of\_Apple}) = 1 - \frac{0.8 \times \log_2 \frac{1}{0.8} + 0.15 \times \log_2 \frac{1}{0.15} + 0.05 \times \log_2 \frac{1}{0.05}}{\log_2 3} = 0.4421$$

$$r(V_{Color\_of\_Strawberry}) = 1 - \frac{0.9 \times \log_2 \frac{1}{0.9} + 0.1 \times \log_2 \frac{1}{0.1}}{\log_2 2} = 0.5310$$

$$r(V_{Color\_of\_Grape}) = 1 - \frac{0.7 \times \log_2 \frac{1}{0.7} + 0.2 \times \log_2 \frac{1}{0.2} + 0.1 \times \log_2 \frac{1}{0.1}}{\log_2 3} = 0.2702$$

The Difference of Property(DP) The second factor in calculating the measure of salience is the difference of a property among similar concepts. It is the distribution of a property's AIP among similar concepts. Intuitively, the more distinguished property from other similar concepts has the higher value of DP and this property becomes higher salient.

Definition 3 The Difference of Property(DP)
Given a concept $*(C)$'s property $S_i = a_i : V_i$ and a set of similar concepts $Sim(*(C))$ including $*(C)$, the DP of the S_i is denoted by $d(S_i)$ and calculated by the following expression:

$$d(S_i) = \frac{r(V_i)}{\sum^{*(C_j) \in Sim(*(C))} \sum^{S_k \in *(C_j)} r'(S_k, S_i)}$$

where $r'(S_k, S_i)$ is calculated by the following expression:

$$r'(S_k, S_i) = \begin{cases} r(V_k) & \text{if } S_{k,max} = S_{i,max} \\ 0 & \text{otherwise.} \end{cases}$$

In this paper, we define the similar concepts as the concepts that has the same parent node in the IS-A hierarchy.

For example, the DP of "color of apple" becomes

$$\frac{\text{The AIP of "color of apple"}}{\sum \text{The AIP of "color of fruit" whose MLP is "color:red"}},$$

because similar concepts of "apple" are "fruits," and the MLP of "color of apple" is "color:red,"

Example 3 Calculation of DP
In this example we calculate the DPs of *(Apple)'s properties, $S_{color\_of\_apple}$ and $S_{shape\_of\_apple}$. Both property's MLPs are "*color : red*" and "*shape : round*" respectively. Figure 1 shows the similar concepts of *(Apple) — these are the child concepts of *(Fruit) — and the AIPs of "color of fruit" whose MLP is "color:red" and the AIPs of "shape of fruit" whose MLP is "shape:round." Blank spaces represent that the concepts do not have the property whose MLP is "*color : red*" or "*shape : round.*". For example the MLP of $S_{color\_of\_lemon}$ is "*color : yellow*" and so the place corresponding to this property is blank space. This is because such a property plays no role in calculation of the DP of $S_{color\_of\_apple}$.

The DPs of $S_{color\_of\_apple}$ and $S_{shape\_of\_apple}$ are calculated as follows.

$$d(S_{color\_of\_apple}) = \frac{0.4421}{0.4421 + 0.5310} = 0.4543$$

$$d(S_{shape\_of\_apple}) = \frac{0.7136}{0.7136 + 0.8586 + 1 + 0.7577 + 0.7577 + 0.1181} = 0.1697$$

Note that because there are less "red fruits" than "round fruits," the DP of "color of apple" has higher score than that of "shape of apple."

		shape:round	color:red
*(Fruit)	*(Apple)	0.7136	0.4421
	*(Strawberry)		0.5310
	*(Watermelon)	0.8586	
	*(Lemon)		
	*(Melon)	1	
	*(Orange)	0.7577	
	*(Banana)		
	*(Grape)		
	*(Pear)	0.7577	
	*(Peach)	0.1181	

Figure 1: *(Apple)'s similar concepts

The DP ranges from nearly equal to 0 to 1. If a MLP of a property is the unique MLP among similar concepts, this property is the most distinguished property and the DP of this property becomes 1. For example the MLP of penguin "*can_fly* : *no*" is in this case, because all other birds have the different MLP "*can_fly* : *yes*." To the contrary, if every similar concepts has the same MLP, the DP of the property which has the MLP become nearly equal to 0. For example, because all fruits have the MLP "*have_seed* : *yes*," the DP of this property of apple becomes nearly equal to 0.

The measure of salience of properties The measure of salience of a property is calculated by the following definition.

Definition 4 The measure of salience of a property The measure of salience of a property S_i is calculated by the following expression:

$$salience(S_i) = r(V_i) \times d(S_i)$$

Because $d(S_i)$ is the rate of $r(V_i)$ occupying among similar concepts, $salience(S_i)$ represent the apparent AIP in similar concepts. For example, if $d(S_i)$ is 1, S_i is the most distinguished property and the apparent AIP is the same as the net AIP (i.e. $r(V_i)$) itself. If $d(S_i)$ is much lower, there are many similar concept that have the same MLP of S_i and the apparent AIP becomes lower than the net AIP.

Example 4 Calculation of the measure of salience Using the results obtained by example 2 and 3, we can calculate the measure of salience of properties $S_{color\_of\_apple}$ and $S_{texture\_of\_apple}$.

$$salience(S_{color\_of\_apple}) = 0.4421 \times 0.4543 = 0.2008$$

$$salience(S_{shape\_of\_apple}) = 0.7136 \times 0.1697 = 0.1211$$

While the net AIP of $S_{shape\_of\_apple}$ (i.e. 0.7136) is higher than that of $S_{color\_of\_apple}$ (i.e. 0.4421), the apparent AIP — that is the measure of salience — of $S_{shape\_of\_apple}$ becomes lower. This is because the effect of the lower DP of $S_{shape\_of\_apple}$.

Using the Measure of Salience in Understanding Metaphors

In this section we describe the overview of our metaphor understanding system AMUSE and explain how the measure of salience is used in AMUSE.

The Overview of AMUSE

The process of understanding metaphors is very similar to analogical reasoning[4]. There are four steps in understanding metaphors. In these four steps, the last three steps correspond to the property transfer process.

1) **Extraction step**: Extracts a Source concept and a Target concept from surface sentence.
2) **Selection step**: Selects the properties of the source concept that are transferable to the target concept. (Calculating salience of source concept's properties)
3) **Mapping step**: Finds the properties of the target concept that correspond to the selected properties in the selection step.
4) **Variance step**: Highlight and downplays the properties of the target concept that are found in the mapping step.

As to analogical reasoning, step 1), step 2), step 3) 4) are corresponding to *Retrieval step*, *Elaboration step*, *Mapping and justification step* respectively[5].

We have focused on the step 2) and shown the method of calculating the measure of salience. We think that the higher salient the property is, the more preferably the property can be transferred.

Now, we show the other steps briefly in understanding the following metaphor.

(1) Mary's cheeks are like apples.

Form this sentence, AMUSE first extract a *viewpoint expression*. This process corresponds to step 1). The viewpoint expression is denoted by *(Target) \ *(Source) and means the *(Target) viewed from *(Source). In this case *(Cheek) \ *(Apple) is extracted.

In the selection step, AMUSE calculates the measure of salience of the properties of *(Apple). For the limitation of space we consider only the following five properties.

$$salience(S_{color\_of\_apple}) = 0.2008$$

$$salience(S_{shape\_of\_apple}) = 0.1211$$

$$salience(S_{taste\_of\_apple}) = 0.1153$$

$$salience(S_{texture\_of\_apple}) = 0.1062$$

$$salience(S_{has\_seed\_of\_apple}) = 0.06244$$

These five properties are rated according to the measure of salience, and AMUSE send the properties to the mapping step in this order. AMUSE has the parameter of threshold which cut off the lower salient properties. In this example if we set the value of threshold to 0.1, then only the first four properties are sent to the mapping step.

The mapping step finds the property of *(Cheek) which corresponds to the property of *(Apple) selected in the previous step. If both property have the same attribute name and the same value which is equal to the MLV of a property of *(Source), AMUSE finds there is a correspondence between both properties. For example, $S_{color\_of\_apple}$ and $S_{color\_of\_cheek}$ are

$$S_{color\_of\_apple} = \{red\#0.8, green\#0.15, brown\#0.05\}$$

$$S_{color\_of\_cheek} = \{yellow\#0.8, pale\#0.1, red\#0.1\}$$

Since these two possible value sets have the same value "*red*" which is the MLV in $S_{color\_of\_apple}$, there is a correspondence between $S_{color\_of\_cheek}$ and $S_{color\_of\_apple}$. Among selected four properties $S_{color\_of\_apple}$, $S_{shape\_of\_apple}$, $S_{texture\_of\_apple}$ and $S_{taste\_of\_apple}$, the $S_{taste\_of\_apple}$ is dropped at the mapping step because there is no corresponding property in *(Cheek). In this case *(Cheek) does not have the attribute "taste."

In the variance step, AMUSE changes the diversity of the properties of *(Cheek) found in the mapping step. For example, in the most simplest version of AMUSE, $S_{color\_of\_cheek}$ shown above is changed to

$$S_{color\_of\_cheek} = \{yellow\#0, pale\#0, red\#1\}$$

The value of $S_{color\_of\_cheek}$ — "*red*" which corresponding to the MLV of $S_{color\_of\_apple}$ is highlighted and the other values — "*yellow*" and "*pale*" are downplayed.

Finally, we get the following representation of $*$(Cheek) $\setminus$ $*$(Apple) as the result of understanding the metaphor (1).

$$*(\text{Cheek}) \setminus *(\text{Apple}) = \left\{ \begin{array}{ll} color: & \left\{ \begin{array}{l} yellow\#0 \\ pale\#0 \\ \underline{red}\#1 \end{array} \right\} \\ shape: & \left\{ \begin{array}{l} \underline{round}\#1 \\ plane\#0 \end{array} \right\} \\ texture: & \left\{ \begin{array}{l} \underline{smooth}\#1 \\ rough\#0 \end{array} \right\} \\ & \vdots \end{array} \right\}$$

In the representation, underlined value is highlighted and others are downplayed according to the high salient properties of *(Apple).

The role of salience in understanding metaphors

There are two advantages to use the measure of salience in understanding metaphors. One is as the measure of preference used in the selection step, which has been described in this paper. The other is as the measure to discern three types of sentences — literal sentences, metaphors and anomalies.

AMUSE is a system which can understand not only metaphors but also literal sentences and anomalies based on the same framework. Many systems have a special device for understanding metaphors, and it is necessary to determine whether an input sentence is a metaphor or not before processing[1, 2]. But there does not exist the clear boundary between metaphors and literal sentences, anomalies. AMUSE does not discern the three types of sentences before processing and all of these are processed in terms of the same transfer process. For example in the case of a literal sentence

(2) I saw the girl with a telescope.,

two viewpoint expressions $*$(Telescope) $\setminus$ $*$(Tool) and $*$(Telescope) $\setminus$ $*$(Thing) are extracted. In the case of a anomaly

(3) Mary's cheeks are like bananas.,

the viewpoint expression $*$(Cheek) $\setminus$ $*$(Banana) is extracted. In the case of the metaphor (1), the viewpoint expression $*$(Cheek) $\setminus$ $*$(Apple) is extracted. In processing these viewpoint expressions, high salient properties of the source concept can be transferred to the target concept as described in the previous section. Considering the percentage of the actually transferred properties among the selected properties in the selection step, almost all properties are transferred in processing $*$(Telescope) $\setminus$ $*$(Tool) and $*$(Telescope) $\setminus$ $*$(Thing). To the contrary, almost all properties are not transferred in processing $*$(Cheek) $\setminus$ $*$(Banana). We define the *Comprehensibility of Viewpoint expression(CV)* as representing this percentage. Following is the definition of the CV.

Definition 5 The Comprehensibility of Viewpoint expression(CV)

The CV of the $*$(A) $\setminus$ $*$(B) is calculated by the following expression.

$$\frac{\sum \begin{array}{c}\text{The AIP of the property of *(B)} \\ \text{which is actually transferred to *(A)}\end{array}}{\sum \text{The AIP of the selected property of *(B) in the selection step}}$$

The CV is the flowing rate of information of the properties in the transfer process. Some properties selected at the selection step are cut off at the mapping step. In literal sentences, the CV of its viewpoint expression is almost 1. In anomalies, the CV of its viewpoint expression is almost 0, because almost all selected properties in the source concept have not corresponding properties in the target concept. So the higher the value of CV a viewpoint expression has, the more literal the viewpoint expression is.

There needs another measure that distinguish metaphors from other types of sentences. The CV measures only the degree of literal or anomalous of viewpoint expressions. From the Ortony's view[10, 9]: in the case of metaphors, transferred properties are high salient in the source concept and low salient in the target concept. In literal sentences, both are high salient. In anomalies, both are low salient. There is an experimental evidence about the Ortony's view[6]. While the CV only considers the properties in the source concept, Ortony's measure considers the properties in both the source concept and the target concept. To formulate the Ortony's measure in AMUSE, we have to make both the mapping step and the variance step more precise. This is out of the scope of this paper.

Conclusion

In this paper, we have proposed a method of calculating the measure of salience.

The measure of salience represents typicality of a property and can be used in various inferences as a measure of preference. This is an aspect of the utility of salience, and the understanding metaphors is one of them. Our metaphor understanding system AMUSE uses the measure of salience to guide which properties can preferably be transferred from a source concept to a target concept.

It is also significant to show the foundation which salience is based on and the method of calculating salience based on the foundation. This is the aspect of how the salience is calculated. To compare with the aspect of the utility, this aspect has not been so enlighten. But both aspects are necessary to accomplish the theory of salience. The measure of salience proposed in this paper is based on the probability attached to attribute's values, and we have shown the precise method how the measure of salience is calculated from the probability. Our measure of salience is based on the entropy in information theory and more formal than other system's score of salience.

But there remains some questions. We have not shown the effect of the context to the measure of salience. There are two relations between contextual information and two factors of the measure of salience. The first one is the diversity change of a possible value set, which causes the change of the first factor — the amount of information of properties. The second one is the variety of selecting similar concepts, which causes the change of the second factor — the difference of properties. Precise analysis of these relations is left as our future work.

Finally, there is a question how the probabilistic concept is constructed. We think Fisher's incremental concept clustering system COBWEB[3] gives us one answer, because COBWEB also uses probabilistic concept as its representation form and constructs probabilistic concepts and their hierarchy incrementally. Fisher proposed the measure of *Category utility* to control the construction of concept hierarchy. Category utility is the measure calculated by two factors — intra-class similarity and inter-class similarity. Roughly speaking, COBWEB incrementally construct concept hierarchy so as to increase intra-class similarity and decrease inter-class similarity. This strategy reflects *basic-level effect* and *typicality effect*[11] observed in human's categorization. Our measure of salience and Fisher's measure of Category utility are similar, that is intra-class similarity and inter-class dissimilarity corresponding to the amount of information and the difference of properties respectively. Therefore we think COBWEB can be incorporated into our metaphor understanding system AMUSE as constructing probabilistic concepts.

References

[1] J. G. Carbonell. Metaphor: An inescapable phenomenon in natural-language comprehension. In W. G. Lehnert and M. H. Ringle, editors, *Strategies for natural language processing*, pages 415–434. Erlbaum, 1982.

[2] D. Fass. Metonymy and metaphor: What's the difference? In *COLING '88*, pages 177–181, 1988.

[3] D. Fisher. Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2:139–172, 1987.

[4] R. P. Hall. Computational approaches to analigical reasoning: A comparative analysis. *Artificial Intelligence*, 39(1):39–120, 1989.

[5] S. T. Kedar-Cabelli. Analogy - from a unified perspective. In D. H. Helman, editor, *Analogical Reasoning*, pages 65–103. Kluwer Academic Publishers, 1988.

[6] T. Kusumi. Effects of categorical dissimilarity and affective similarity between constituent words on metaphor appreciation. *Journal of Psycolinguistic Research*, 16(6):577–595, 1987.

[7] G. Lakoff and M. Johnson. *Metaphors we live by*. Chicago University Press, 1980.

[8] H. Miyagawa. *Information Theory*. CORONA Publishing CO., LTD., Tokyo, Japan, 1979. (in Japanese).

[9] A. Ortony. Beyond literal similarity. *Psychological Review*, 86(3):161–180, 1979.

[10] A. Ortony. The role of similarity in similes and metaphors. In A. Ortony, editor, *Metaphor and Thought*. Cambridge University Press, 1979.

[11] E. Rosch and C. B. Mervis. Family resemblances : Studies in the internal structure of categories. *Cognitive Psychology*, 7:573–605, 1975.

[12] E. E. Smith and D. Medin. *Categories and Concepts*. Harvard University Press, Cambridge, MA, 1981.

[13] E. E. Smith, D. N. Osherson, L. Rips, and M. Keane. Combining prototypes : A selective modification model. *Cognitive Science*, 12(4):485–527, 1988.

[14] A. Tversky. Features of similarity. *Psycological Review*, 84(4):327–352, 1977.

[15] E. J. Weiner. A knowledge representation approach to understanding metaphors. *Computational Linguistics*, 10(1):1–14, 1984.