Nagoya International Symposium On Knowledge Information And Intelligent Communication 1989

A Computational Model of Understanding Metaphors

TANAKA Hozumi tanaka@cs.titech.ac.jp

IWAYAMA Makoto iwa@cs.titech.ac.jp

TOKUNAGA Takenobu take@cs.titech.ac.jp

Deptartment of Computer Science

Tokyo Institute of Technology

November 14, 1989

Abstract

We consider understanding metaphors as a transfer process of properties from a Source concept to a Target concept. For example, "a man like a wolf" usually means a man who is vicious, fierce, dangerous and nocturnal, but doesn't mean a man who has four legs, two eyes, and are hairy without a particular context. Thus, only the wolf's properties such as "vicious", "fierce", "dangerous" and "nocturnal" have been transferred from "wolf" (*Source concept*) to "man" (*Target concept*).

This transfer process consists of following two stages. At the first stage, the candidates of properties to be transferred have to be selected. These properties are the typical properties of the Source concept under the current context and are likely to be transferred to the Target concept. At the second stage, how should some of the properties of a Target concept be changed? Some candidate properties may be dropped at this stage.

To select the candidates, we introduce *salience* of a property. Salience represents how typical or prominent the property of the concept is. In our framework, Salience is calculated by the entropy of properties and the property difference among similar concepts. Thus, unlike the models proposed so far, the degree of salience proposed in this paper is not discrete but continuous values. In case of anomalies, the properties with high saliences are not transferred to a Target concept. On the other hand, in case of literal sentences almost all properties with high saliences are transferred to a Target concept. Metaphors are just between anomalies and literal sentences.

We do not consider the influence of the context in our model. How to incorporate the contextual information in our model? It is left as a future work.

1 Introduction

As natural language is a rich source of metaphors, we should consider metaphors in developing a better natural language understanding system. Metaphors have a strong relationship with our conceptual structure that has been acquired through our everyday life [2].

Generally speaking, understanding metaphors is regarded as a transfer process of properties from a Source concept to a Target concept. Human being express and understand a Target concept by means of a Source concept, and extract a new meaning of the Target concept. From a limited number of concepts, we can create and express immense meanings through metaphors.

Consider a metaphor, "a man is a wolf" in which "man" is a Target concept and "wolf" is a Source concept. The meaning of "man" in the sentence is not prototypical but metaphorical one. What is problems of understanding metaphors? The one of the most difficult problems is that we have not yet known how to interpret the Target concept "man" based on the Source concept "wolf." Some salient properties in the Source concept "wolf" seems to be transferred to the Target concept "man" through the metaphor understanding. In other words, we view the Target concept "man" from the Source concept "wolf." We express this by *(Man) \ *(Wolf) which is called *a viewpoint expression* in the paper.

The importance of salient properties in understanding metaphors has been suggested by many researchers [5, 4, 8, 10]. Roughly speaking, salience represents how typical or prominent the property of the concept is. The properties which have high salience in a Source concept are apt to be transferred to a Target concept. However their computational models to calculate salient properties have been very simple. In the paper, we propose a computational model that calculates the salience of properties in the Source concept using more precise computational model. In order to calculate the salience of properties in some concepts, we make use of the measure of entropy of properties and property difference among other similar concepts. Using viewpoint expressions, we will show that not only metaphors but also literal sentences will be processed.

Based on our viewpoint expressions proposed in section 2, we will give a unified view of understanding metaphors, literal sentences and anomalies. In other word, understanding viewpoint expressions implies understanding metaphors, literal sentences, and anomalies. In section 3, we will describe our computational model that calculates the salience of properties. In section 4, we also describe an algorithm that enables us to make change of the salience of properties in a Target concept when metaphors will be understood.

2 Metaphors and Viewpoint expressions

In this section we will describe the importance of viewpoint expressions in understanding metaphors, literal sentences, and anomalies. In the process, calculation of saliences will play a very important role. The notation, *(T) \ *(S) is a *viewpoint expression* where *(T) and *(S) are *a Target concept* and *a Source concept* respectively. The intuitive meaning of *(T) \ *(S) is that *(T) is viewed from *(S). In the following section *(X) stands for a concept with name "X".

Consider the following sentence.

(1) I saw a girl with a telescope.

There is a PP in the above sentence, namely "with a telescope." The sentence (1) is ambiguous and has at least two meanings depending on the location of PP attachment, "a girl with a telescope" and "saw with a telescope." With regard to interpret a concept "telescope," we can extract following two viewpoint expressions, *(Telescope) \ *(Thing) and *(Telescope) \ *(Tool). In both cases, the Target concept *(Telescope) is a subordinate of the Source concept *(Thing) and *(Tool). (See Figure 1)

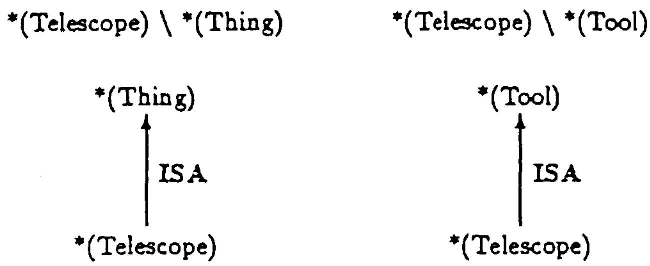

Figure 1. – Viewpoint expressions which hold ISA relation

However, it is not always the case. A viewpoint expression extracted from a metaphor does not hold a superordinate and subordinate relation between a Source and a Target concept. Rather, metaphors make use of another type of viewpoint expressions that violate superordinate and subordinate relations between a Source and a Target concept.

Consider the following metaphor[8].

(2) Mary's cheeks are like apples.

In this sentence, "Mary's cheeks" should be interpreted by the following viewpoint expression, *(Cheek) \ *(Apple), and in consequence "Mary's cheeks" should become round and red. The viewpoint expression *(Cheek) \ *(Apple) is not based on a superordinate and subordinate relation. (See Figure 2) There is a more indirect relation between a Source and a Target concept. We call it "*a metaphorical relation.*"

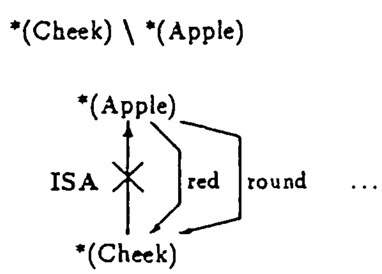

Figure 2. – A viewpoint expression which does not hold ISA relation

How to identify a metaphorical relation between a Source and a Target concept? Many researchers have admitted the importance of high salient properties in a Source concept, which will be transferred to a Target concept[5, 4, 8, 10], but they have deal with saliences based on qualitative discussions. What we need is to develop a computational method that enables us to calculate high salient properties quantitatively. In the sentence (2), the properties of "shape" and "color" of *(Apple) have high saliences that have been transferred to a Target concept *(Cheek). Some properties are not transferred to a Target concept. For example "an apple is a plant," "an apple has a core," "an apple has seeds" are not transferred to *(Cheek), because they are properties with lower salience than that of "shape" and "color." In section 3, we will propose our computational model that calculates the salience of properties.

3 Knowledge representation and salience of properties

In this section we will describe our knowledge representation in understanding metaphors. Using the representation, we will propose a computational model that calculates the salience of properties.

3.1 Knowledge representation

The knowledge representation we will use is the same one as proposed by Smith[7]. His knowledge representation of a concept is composed of a set of properties each of which consists of an attribute with a set of possible values to which frequency is attached. We will give a definition of concepts.

Definition 1 Concepts

Concept C is a set of *properties* S_i.

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

Property S_i is a pair of *an attribute* a_i and *a set of possible values* V_i.

$$S_i = a_i : V_i$$

where V_i is a set of possible values of an attribute a_i to which *frequency* is attached.

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

As all elements in a set of possible values are exclusive each other, the frequency stands for a probability.

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

The most likelihood value (MLV) $v_{i,max}$ is the value which has the highest probability and *the most likelihood property (MLP)* $S_{i,max} = a_i : v_{i,max}$ is a pair of MLV $v_{i,max}$ and its attribute a_i.

Consider the definition of *(Apple) shown in the example 1.

Example 1 Definition of *(Apple)

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

In the example, "$\mathit{color} : \{\mathit{red}\#0.8, \mathit{green}\#0.15, \mathit{brown}\#0.05\}$" is *a property*, and "$\{\mathit{red}\#0.8, \mathit{green}\#0.15, \mathit{brown}\#0.05\}$" is *a set of possible values* of the attribute "*color*" where a number attached to each value is *frequency* of the value. The possible value "*red*" is the most likelihood value (MLV) of the attribute "*color*," and the "*color* : *red*" is called the most likelihood property (MLP).

3.2 Calculation of the salience of a property

In section 2, the importance of saliences in understanding metaphors is explained. In this section, we will describe a computational model of the salience of a property. Smith regards the salience as frequency of a possible value[7]. His method is too simple to calculate salience reasonably. The salience should be determined not only by the frequency of a possible value but also by considering other similar concepts. In our method the salience will be calculated by both entropy of a set of possible values and property difference among other similar concepts.

3.2.1 Entropy of a set of possible values of a property

At first we have to compute the entropy of a property(MLP) which has MLV. Note that as entropy is a measure of randomness of possible values of a property , the smaller the entropy the larger information contained in the property. Therefore MLV with smaller entropy is more salient than the one with larger entropy. For example, consider the following two sets of possible values.

$\mathit{Set}_1 = \{\mathit{red}\#0.6, \mathit{green}\#0.1, \mathit{yellow}\#0.1, \mathit{blue}\#0.1, \mathit{brown}\#0.1\}$

$\mathit{Set}_2 = \{\mathit{red}\#0.6, \mathit{green}\#0.4\}$

Entropy(relative entropy) of Set_1 and Set_2 becomes 0.7627 and 0.9705 respectively. Although the frequency of MLV are both 0.6, MLV of Set_1 should contribute more to the value of salience than that of Set_2, because the former has a smaller value of entropy than the latter.

Definition 2 The amount of information of a set of possible values
Given a set of possible values $V_i = \{v_{i,1}\#w_{i,1}, v_{i,2}\#w_{i,2}, \ldots, v_{i,m}\#w_{i,m}\}$, let define the amount of information $r(v_i)$ 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[3].

Example 2 Calculation of amount of information

Suppose the following four sets of possible values of an attribute "*color*."

$$V_{Color\_of\_Apple} = \{red\#0.8, green\#0.15, brown\#0.05\}$$
$$V_{Color\_of\_Strawberry} = \{red\#0.9, purple\#0.1\}$$
$$V_{Color\_of\_Tomato} = \{red\#0.8, brown\#0.2\}$$
$$V_{Color\_of\_Grape} = \{violet\#0.7, green\#0.2, red\#0.1\}$$

The amount of information for each set becomes,

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

$$r(V_{Color\_of\_Strawberry}) = 0.5310$$

$$r(V_{Color\_of\_Tomato}) = 0.2781$$

$$r(V_{Color\_of\_Grape}) = 0.2702$$

Among the four MLPs, "*color* : *red*" of *(Apple), *(Strawberry), *(Tomato), and "*color* : *violet*" of *(Grape), MLV of *(Strawberry) will contribute most to the value of salience, because it has the largest amount of information in this example.

Note that as the amount of information of a property is determined by the distribution of frequency, the salience will depend on cultures and individual experiences.

3.2.2 Property difference among sibling concepts

Consider a metaphor of sentence (2) in which *(Apple) is a Source concept and a viewpoint expression is *(Cheek) \ *(Apple). In order to calculate amount of information, we only think about a set of possible values in a property. However we have to take into account of other similar sibling concepts if we want to compute saliences. The property difference among sibling concepts seems to contribute to the value of salience, because properties of Source concept with larger property difference are likely to be transferred to a Target concept.

Definition 3 The property difference of MLP

Property difference of MLP $S_{i,max}$ is defined by the following formula.

$$d(*(C), S_{i,max}) = \frac{r(V_i)}{\sum^{*(C_j)\in\{Sibling concepts of *(C)\cup *(C)\}} r'(*(C_j), S_{i,max})}$$

where sibling concepts of *(C) are defined by concepts with a common superordinate concept, and $r'(*(C_j), S_{i,max})$ is obtained by the following formula.

$$r'(*(C_j), S_{i,max}) = \begin{cases} r(V_k) & \text{if there is a } S_{i,max} = S_{k,max} \text{ where} \\ & S_{k,max} \text{ is MLP of } *(C_j) \\ 0 & \text{otherwise.} \end{cases}$$

Example 3 Calculation of property difference

Let us take the following set of concepts each of which belongs to a common superordinate concept *(Fruit).

$$\{*(Apple), *(Strawberry), *(Tomato), *(Grape)\}$$

According to the example 2, we will get,

$$r(V_{Color\_of\_Apple}) = 0.4421$$

$$r(V_{Color\_of\_Strawberry}) = 0.5310$$

$$r(V_{Color\_of\_Tomato}) = 0.2781$$

$$r(V_{Color\_of\_Grape}) = 0.2702$$

For each MLP, we can calculate the property differences.

$$d(*(Apple), color : red) = \frac{0.4421}{0.4421 + 0.5310 + 0.2781} = 0.3533$$

$$d(*(Strawberry), color : red) = \frac{0.5310}{0.4421 + 0.5310 + 0.2781} = 0.4244$$

$$d(*(Tomato), color : red) = \frac{0.2781}{0.4421 + 0.5310 + 0.2781} = 0.2223$$

$$d(*(Grape), color : violet) = \frac{0.2702}{0.2702} = 1$$

It is easy for us to understand that the more the number of sibling concepts is, the smaller the value of property difference. The fact coincides with our intuition because if MLP is a distinguished feature, the number of sibling concepts, which has the same MLP, will be small. For example, as only the concept *(Grape) has MLP "*color : violet*," the property difference will take the maximum value 1. On the other hand, the most likelihood property "*hasa : seeds*" will be the smallest property difference, because all of fruits "have seeds."

3.2.3 The salience of MLP

According to the amount of information and property difference, we can define the salience of MLP.

Definition 4 Salience of MLP
Salience of MLP $S_{i,max}$ of concept $*(C)$ is defined by the following formula,

$$salience(*(C), S_{i,max}) = r(V_i) \times d(*(C), S_{i,max})$$

Example 4 Calculation of salience
Using the results obtained by example 2 and 3, we can compute the saliences of properties with an attribute "*color*."

$$salience(*(Apple), color : red) = 0.4421 \times 0.3533 = 0.1562$$

$$salience(*(Strawberry), color : red) = 0.5310 \times 0.4244 = 0.2254$$

$$salience(*(Tomato), color : red) = 0.2781 \times 0.2223 = 0.06182$$

$$salience(*(Grape), color : red) = 0.2702 \times 1 = 0.2702$$

4 Understanding Metaphors based on Viewpoint Expressions

In this section, we will explain a computational model of understanding metaphors using viewpoint expressions. However, our model described here is so simple that there are a lot of limitations to handle all kinds of metaphors. As an example to understand metaphors based on viewpoint expressions, let us consider the metaphor in section 2.

(2) Mary's cheeks are like apples.

In this sentence, Mary's cheeks should be interpreted by the viewpoint expression $*(\text{Cheek}) \setminus *(\text{Apple})$. At first we have to calculate the salience of properties of the Source concept *(Apple). The Salience of property with an attribute "*color*" has already calculated in example 4. *(Apple) has the highest salience for the property "*color : red*." Using the same method described in section 3, we can obtain the salience for the properties of *(Apple), namely "*color : red*," "*shape : round*," "*texture : smooth*," "*taste : sweet*" and "*hasa : seeds*."

$$salience(*(Apple), color : red) = 0.1562$$

$$salience(*(Apple), shape : round) = 0.1427$$

$$salience(*(Apple), texture : smooth) = 0.1062$$

$$salience(*(Apple), taste : sweet) = 0.1153$$

$$salience(*(Apple), hasa : seeds) = 0.06244$$

As we think that all sibling concepts of *(Apple) has seeds, the salience of property of "*hasa* : *seeds*" gets the smallest value. What properties should be transferred to the Target concept *(Cheek)? Provided that a threshold value of salience is set to 0.1, "*color* : *red*," "*shape* : *round*," "*texture* : *smooth*," and "*taste* : *sweet*" become candidate properties that will be able to be transferred to *(Cheek). However, not all such salient properties will be transferred to *(Cheek). If some of salient properties of *(Apple) does not exist in *(Cheek), they should not be transferred to *(Cheek). For example, a property "*taste* : *sweet*" of *(Apple) should not be transferred to *(Cheek), because there is no attribute of "*taste*" in *(Cheek). Therefore only properties "*color* : *red*," "*shape* : *round*," and "*texture* : *smooth*" of *(Apple) are transferred to *(Cheek) and made those properties highlight in *(Cheek). (See Figure 3)

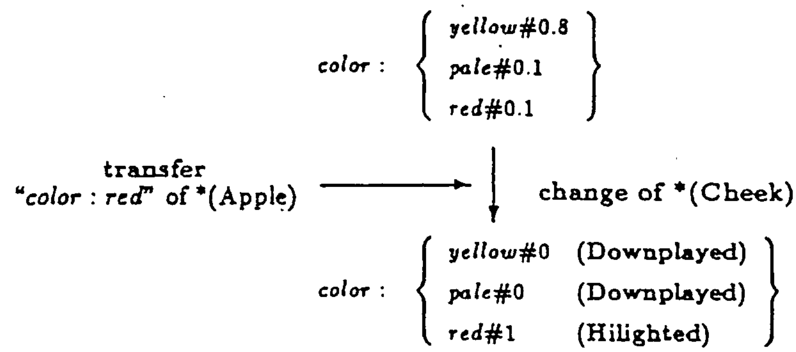

Figure 3. – Transfer process of properties

Finally, we will get the result of understanding *(Cheek) \ *(Apple). (Figure 4) In other words, it is a result of understanding a metaphor based on the viewpoint expression shown above.

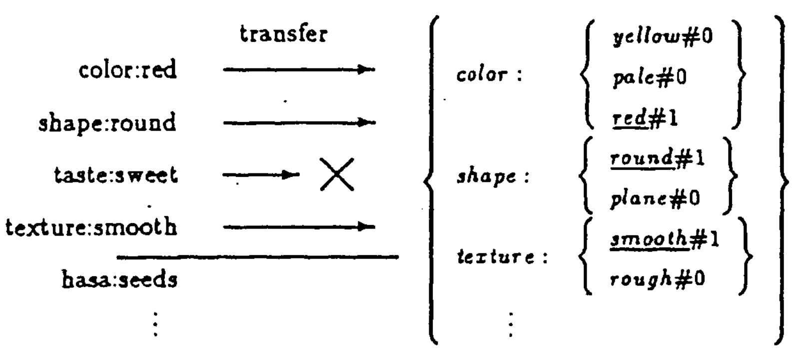

Figure 4. – The result of understanding *(Cheek) \ *(Apple)

5 Conclusion

In this paper, we propose a computational model of understanding metaphors based on viewpoint expressions. We emphasize the importance of salience and give a method how to calculate the value of saliences. We know understanding metaphors need very complex processes in our brain, and we do not know yet exactly how we human being understanding metaphors. Our model proposed in this paper is absolutely impossible to solve all problems of metaphor understanding. We have to make more efforts in this field theoretically and experimentally. Let us pick up some of limitations of our model. (1) How to attach probability to each possible value? (2) How to determine the threshold value of salience? (3) Is the salience affected by contextual information dynamically? (4) Is our definition of calculation property difference reasonable? All these questions are open in our model. However, we think our model proposed in this paper will give us a small step toward better understanding metaphors.

References

[1] D. Fass and Y. Wilks. Preference semantics, ill-formedness, and metaphor. *AJCL*, 9(3–4):178–187, 1983.

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

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

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

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

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

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

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

[9] P. H. Winston. Learning by creating and justifying transfer frames. *Artificial Intelligence*, 10(2):147–172, 1978.

[10] M Yamanashi. *Hiyu to Rikai*. Tokyo Daigaku Shuppankai, Tokyo, Japan, 1988. (in Japanese).