Self-organizing Conceptual Map and Taxonomy of Adjectives

Noriko Tomuro
DePaul University, School of Computer Science,
Telecommunications and Information Systems,
Chicago, IL 60604 USA
tomuro@cs.depaul.edu

Kyoko Kanzaki, Hitoshi Isahara
National Institute of Information and
Communications Technology,
Soraku-gun, Kyoto, 619-0289 Japan
{kanzaki,isahara}@nict.go.jp

Abstract

This paper presents a comprehensive, corpus-based study of adjectival concepts. An adjectival concept is a semantic class of adjectives, for example, "adjectives which express feeling". We represent adjectival concepts by abstract nouns (such as "feeling"), and extract instances where abstract nouns are modified by adjectives (such as "happy feeling") from a large corpus of Japanese newspaper articles. Then we automatically construct a two-dimensional map of adjectival concepts by using the Kohonen Self-organizing Map (SOM). Not only does this SOM map effectively visualize the concept space of adjectives, it also readily allows us to extract a taxonomy of adjectival concepts. Using the SOM map and the taxonomy extracted, we are able to investigate the breadth of various adjectival concepts and the inter-relations between them.

Keywords: Self-organizing Map; SOM; Conceptual Map; Adjectival Concepts; Adjectival Concept Taxonomy.

Introduction

Adjectives are notoriously polysemous. The function of adjectives is to modify or elaborate the meanings of nouns. By being modified by adjectives, nouns will come to bear specific values for their attributes. For example, "warm soup" is (a bowl of) soup which has the value "warm" (moderately hot) for its temperature attribute. On the other hand, adjectives seem to take on different meanings, or change its meaning, depending on the nouns they modify as well. For example, when "warm" modifies the noun "person", the meaning of "warm" is psychological (as versus physical in the case of "warm soup"), elaborating the personality or the way the person deals with others. Traditionally, meanings of a polysemous word are enumerated in dictionaries. For example, WordNet (Miller 1990), a large on-line dictionary and thesaurus, lists ten senses for "warm", four of which are shown below.

1. (having a comfortable degree of heat): "a warm body", "a warm climate"
2. (psychologically warm): "a warm greeting", "a warm personality"
3. ((color) inducing the impression of warmth): "warm reds and yellows"
4. (having warmth or affection): "a warm embrace", "a warm glance"

How to distinguish different senses of polysemous words is a difficult problem for lexicographers for any part of speech. However for adjectives, the problem is more salient because the change in the meaning, or *meaning shift*, seems quite flexible and occurs dynamically when they are combined with different nouns (Murphy & Andrew 1993). A good example would be metaphor – a figurative use of a word which tries to apply its (original) concept to a new domain, for instance "dark conversation" (a discussion on grim topics, or to talk pessimistically).

In linguistics, there is a large body of work on adjectives, although the attention they have received is much less than that for nouns and verbs. Adjectives are usually considered in the context of nouns which they modify, and thought to add only auxiliary information to the nouns. In computational linguistics or Natural Language Processing (NLP), there exist only few other works which tackled adjectives specifically. Adjectives are generally difficult to incorporate in NLP applications because of their polysemous nature, as mentioned above. A critical issue is how to represent adjectives so that the dynamic interactions between nouns and adjectives are modeled and facilitated computationally. Most of the recent work on adjectives in NLP has focused either on specific applications (e.g. classifying documents according to semantic orientation or subjectivity (Hatzivassiloglou & McKeown 1997; Wiebe 2000)), or on specific types of adjectives (e.g. event adjectives (Lapata 2001), and gradable adjectives (Hatzivassiloglou & Wiebe 2000)).

In this work, we investigate *adjectival concepts* from a comprehensive point of view. In particular, we attempt to visualize the conceptual space for adjectives in a two-dimensional map, and investigate the relations between various concepts. Adjectival concepts are essentially *semantic classes* of adjectives, for instance "adjectives which express feeling". By meaning shift (including metaphors), many adjectives take on different meanings, thereby belonging to several concepts. Conversely, a given concept includes adjectives whose meanings have extended from their original/core meaning, thereby forming the *breadth* of the concept. Fellbaum describes the breadth of different groups of adjectives in (Fellbaum, Gross, & Miller 1993, pg. 32) as:

> Adjectives expressing evaluations (good/bad, desirable/undesirable) can modify almost any noun; those expressing activity (active/passive, fast/slow) or potency (strong/weak, brave/cowardly) also have wide ranges of applicability. Other adjectives are strictly limited with respect to the range of their head nouns (mown/unmown; dehiscent/indehiscent).

We aim to investigate the breadth and relations between adjectival concepts, and construct a model for the conceptual space of adjectives from real data. To this end, we collect instances of adjectives modifying *abstract nouns* from a large corpus of Japanese newspaper articles. Here, abstract nouns are used as the *semantic class labels* of adjectives. For example, the noun "emotion" is a class label for a group of adjectives which express human emotion, such as "happy" and "sad". Then we represent each of those abstract nouns by a vector of adjectives which co-occurred with the noun, and automatically construct a two-dimensional map using the Kohonen Self-organizing Map (SOM) (Kohonen 1984; 1995). The resulting map is a map of adjectival concepts, represented by abstract nouns and visualized on a two-dimensional plane. From this map, we also extract a *taxonomy* of adjectival concepts. This taxonomy is based on the subsumption relation between the concepts, and represents the inter-relations among them in finer details and in a hierarchical manner. Such an adjectival concept taxonomy would be useful and of interest to various areas of study, for example as a basis in analyzing adjectival meaning shift in linguistics, as a potential cognitive model for adjectives in cognitive science, and as a lexical resource in various NLP and other applications including the Semantic Web.

Abstract Nouns as Adjectival Concepts

In this work, we use abstract nouns to represent adjectival concepts based on the following past studies. Takahashi (1975) investigated various linguistic clues which indicated meronymy (the part-whole relation) in Japanese sentences. He found that, in many phrases of the form "X ga Y" where X is a noun, Y is an adjective and "ga" is a post-positional subject marker, if X is an abstract noun, it corresponds to the semantic category of Y. For example,

A.	"Yagi	wa	<u>seishitsu</u>	ga	otonashii."
	(goat)	tm	(nature)	sm	(gentle)
	"The nature of goats is gentle."				
B.	"Zou	wa	<u>hana</u>	ga	nagai."
	(elephant)	tm	(nose)	sm	(long)
	"The nose of an elephant is long."				

where tm stands for topic-marker, and sm stands for subject-marker post-positions in Japanese. In A, "seishitsu (nature)" is an abstract noun, while in B, "hana (nose)" is a concrete noun. Also in A, "gentle" is an adjective which expresses and elaborates the concept of "nature", while in B, "long" is a property/part of "nose" and there is no concept elaboration relation between them. This concept elaboration relation can also be identified by omitting the noun – for A, we can omit "seishitsu (nature)" and say "Yagi wa otonashii (Goats are gentle)" without changing the meaning of the sentence, while for B, we cannot omit "hana (nose)" and say "Zou wa nagai (Elephants are long)" without changing the meaning. Similarly, Nemoto (1969) argued that expressions such as "<u>iro</u> ga akai (the color is red)" and "<u>sokudo</u> ga hayai (the speed is fast)" as a kind of repetition of meaning, or tautology. Schmid (2000) called such abstract nouns *shell nouns*: a class of (non-referencial) abstract nouns which function as conceptual shells for complex pieces of information (i.e., contents) that are elaborated by other words or clauses in a text. Other examples of abstract nouns and adjectives which elaborate the concepts include:

kimochi (feeling)	–	ureshii (glad), kanashii (sad)
kibo (scale)	–	ookii (large), chiisai (small)

In this work, we used a dataset which we had constructed in our previous work (Isahara & Kanzaki 1999; Kanzaki *et al.* 2004). This dataset contained a large number of examples where abstract nouns were modified by adjectives. They were extracted from real data: a large corpus consisting of articles from a total of 42 years' worth of Japanese newspapers (11 years of Mainichi Shinbun, 10 years of Nihon Keizai Shinbun, 7 years of Sangyou Kinyuu Ryuutsu Shinbun, and 14 years of Yomiuri Shinbun). In addition to the form "AbN ga Adj" described above, we also extracted phrases of the form "Adj AbN" (e.g. "yasashii seishitsu (gentle nature)"). Then we manually removed instances where the abstract noun and the adjective were not in the concept elaboration relation (Isahara & Kanzaki 1999). There were a total of 361 abstract nouns which co-occurred with 4 or more adjectives. The total number of adjectives was 2374.

In the data, we represented each abstract noun by a feature vector, where a feature was an adjective co-occurred with the noun, and the value was the *pointwise mutual information* (Manning & Schutze 1999) computed from the co-occurrence counts extracted from the corpus. Mutual Information (MI) is based on information theory, and has been used in many NLP tasks such as clustering words (e.g. (Lin & Pantel 2002)). Let x be a noun and y be an adjective. The pointwise MI between x and y, denoted $I(x, y)$, is defined as:

$$I(x, y) = log\frac{p(x, y)}{p(x)p(y)}$$

where $p(z)$ is the probability of a word z (a noun or an adjective) in the corpus (i.e., the total number of times z occurred in the corpus divided by the total number of times all words occurred), and $p(x, y)$ is the joint probability of x co-occurred with y (i.e., the total number of times x co-occurred with y in the corpus divided by the total number of times all words occurred). MI essentially works as a way to put weights on the frequencies (in fact, weights which are inverse to the frequencies). It is similar to the TFiDF term weighting scheme (Salton & McGill 1983) used in information retrieval; the difference is that MI in effect places heavier weights on words that occurred less frequently.

A Self-organizing Map of Adjectival Concepts

Using the data described in the previous section, we produced a map of adjectival concepts using an algorithm called Kohonen Self-organizing Map (SOM) (Kohonen 1984; 1995). The SOM was originally inspired by the way in which various human sensory stimuli are neurologically mapped into the brain such that spatial or other relations among stimuli resemble the spatial relations among the neurons (Kohonen 1996). Apart from neuroscience, the SOM is also widely used as an exploratory data analysis and visualization tool in many practical applications, including image processing, NLP, process control and economic analysis. In this work, we use the SOM as a tool, rather than as a model of human brain.

The SOM is a neural network model, also an unsupervised learning method where input instances are projected onto a grid/map of nodes arranged in an n-dimensional space in such a way that relative distances between the instances are preserved. Input instances are usually high-dimensional data, and the map is usually two-dimensional (i.e., n = 2). Thus, the SOM essentially reduces the dimensionality of the data, and can be used as an effective tool for data visualization – akin to Multidimensional Scaling (MDS).

The SOM can also be utilized for clustering. Every node in a map represents a cluster, and is associated with a reference vector of m-dimension, where m is the dimension of the input instances. During learning, input instances are mapped to a map node whose (current) reference vector is the closest to the instance vector, and the reference vectors are gradually smoothed so that the differences between the reference vector and the instance vectors mapped to the node are minimized. This way, instances mapped to the same node form a cluster, and the reference vector essentially corresponds to the centroid of the cluster.

SOM maps are self-organizing in the sense that input instances that are similar are gradually pulled closer during learning to be assigned to nodes that are topographically close to one another on the map. The mapping from input instances to map nodes is one-to-one (i.e., an instance is assigned to exactly one node), but from map nodes to instances, the mapping is one-to-many (i.e., a map node is assigned with zero, one or more instances).

Constructing A SOM Map

To produce a SOM map, we used a tool called SOM_PAK (Kohonen *et al.* 1996). The input data was the set of 361 abstract nouns defined by the 2374 co-occurring adjectives as features, as described in the previous section. In the data, we also added an extra, artificial instance "TOP" which had a frequency value of 1 for all features (and re-computed the MI's). This "TOP" instance was intended to represent "the most abstract noun of all". According to Caraballo & Charniak (1999), the generality/specificity of nouns is determined fairly accurately by the number of their co-occurring adjectives. Thus, a noun modified by all adjectives represents the most abstract noun possible (in the given dataset).

Then during learning, we specified the "TOP" to be assigned to the center node in the map (utilizing the option available in SOM_PAK). By forcing "TOP" at the center of the map, our first intention was to reduce the topographical variability of the maps generated depending on the parameter settings (such as the initial values of the reference vectors). Another purpose was to fix the orientation of the map and obtain one in which the nouns are laid out hierarchically in the general-to-specific order from the center in all directions. Such a map was highly probable based again on the findings by (Caraballo & Charniak 1999) – because nouns which are fairly abstract are modified by a large number of adjectives, they are similar to "TOP", therefore likely to be mapped near the center node. Other, more specific nouns would be placed further away from the center, because their similarities to "TOP" are low. Then the resulting map would likely be one where hierarchies of nouns are radially extending from the center in all directions, each of which descends to a specific adjectival concept. Thus, not only can such a map show the similarity/horizontal relations between nouns (mapped to near-by nodes or regions on the map), it can also show the hierarchical/vertical relations (mapped radially from the center) at the same time.

The SOM Map

Figure 1 shows the map obtained.[1] We used a two-dimensional grid of 45 x 45 nodes. Each node is associated with a reference vector of length 2374, corresponding to the number of features (co-occurring adjectives) for a given noun. The map shown was obtained after 5 million iterations of the SOM algorithm.

In the map, "TOP" is indicated by a solid black rectangle at the center. We also circled groups of nodes whose reference vectors are significantly close to each other. We call such groups of nodes *tight clusters*. In SOM, some nodes are mapped with multiple nouns which are very similar. However, some nouns which are similar might have been mapped to different nodes, because the algorithm's self-organization is sensitive to the parameter settings. In order to account for this, we extracted tight clusters whose average cosine coefficients (Salton & McGill 1983) between the reference vectors in the cluster were significantly high (where we used the threshold of greater than 0.96). The cosine between two reference vectors was computed as standard: for given vectors f and g, the cosine coefficient between them, denoted $cos(f, g)$, is:

$$cos(f, g) = \frac{\sum_i f_i \times g_i}{\sqrt{\sum_i f_i^2 \times \sum_i g_i^2}}$$

where v_i is the value for the ith feature in the vector v.

The tight clusters extracted in this way included, for example, {"keijyo (shape)", "keitai (form)", "katachi (shape)"} (indicated with A in Figure 1), {"kuukan (space)", "okuyuki (depth)", "shizen (nature)", "nagame (view)"} (indicated with B), and {"ikioi (velocity)", "sokudo (speed)"} (indicated with C).

To evaluate the tight clusters, we compared them with the noun *synonyms* recorded in "Bunrui Goihyo" – a com-

[1] We will discuss the lines drawn in the map in the section "Taxonomy of Adjectival Concepts".

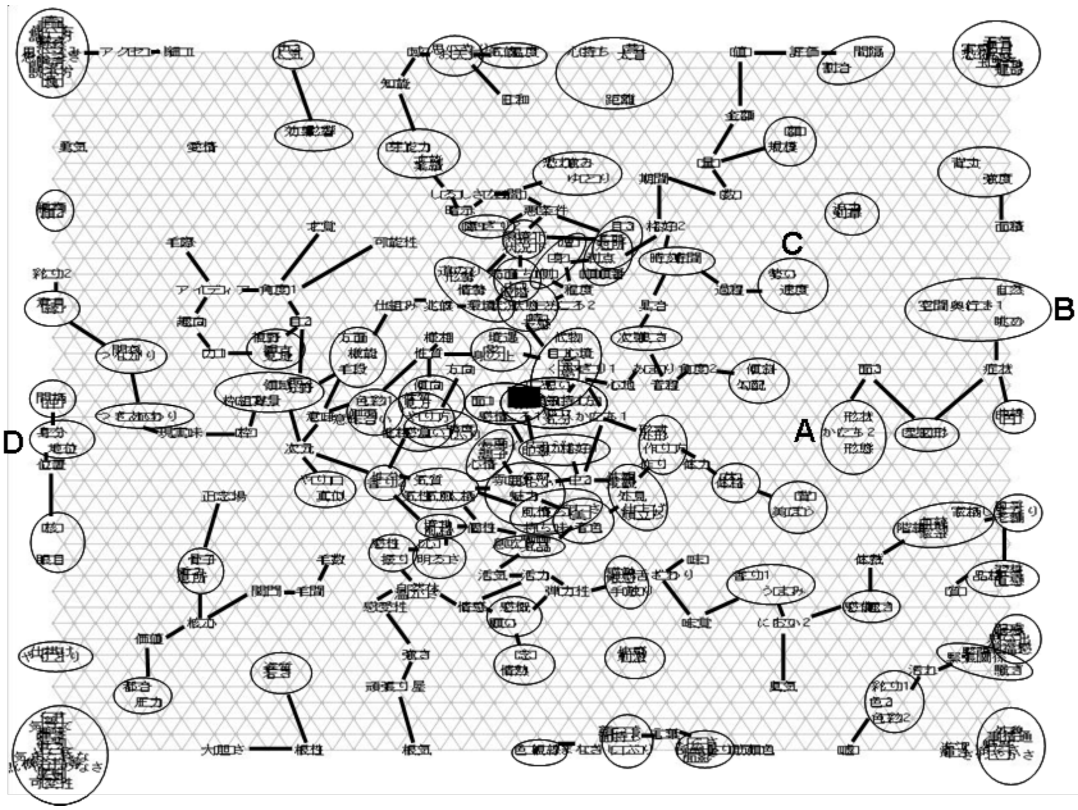

Figure 1: The SOM Map of Abstract Nouns

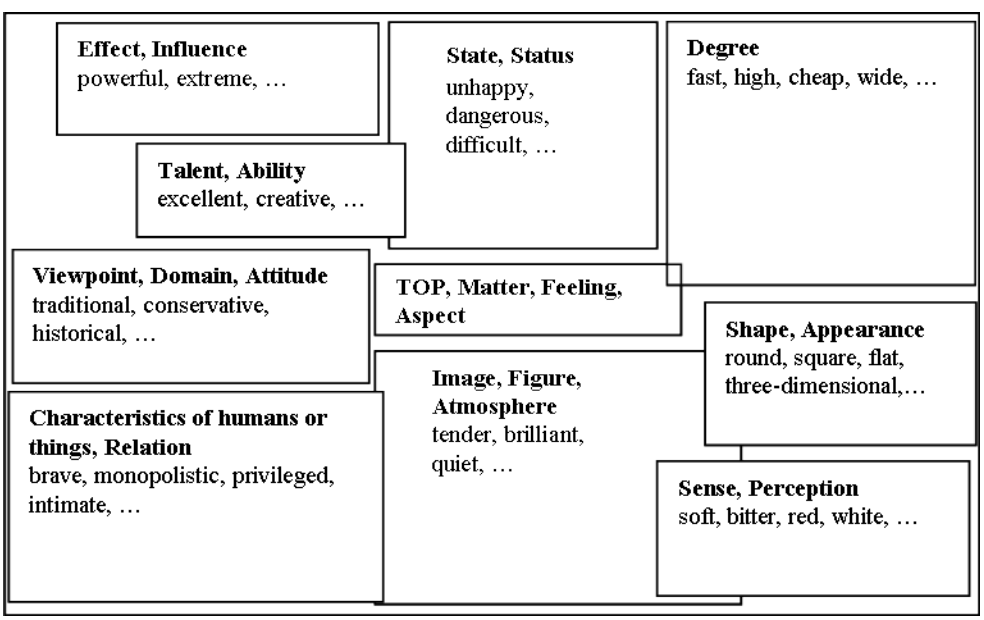

Figure 2: Conceptual Area Map of Adjectives

prehensive Japanese thesaurus (compiled manually by lexicographers) which contained over 100,000 synonym sets for various parts of speech.[2] Our inspection revealed that, of the total 88 tight clusters we extracted, 46 of them consisted of words which were assigned the same semantic category (thus synonyms) in the thesaurus. This means the precision was 52 %. Although our abstract nouns were represented only by modifying adjectives (rather than words which appeared in a wider context surrounding the nouns, such as verbs or other nouns), it seemed the modifying adjectives were good enough features to define abstract nouns, and moreover our SOM implementation was able to capture the similarity between abstract nouns fairly well. Furthermore, the precision suffered somewhat due to the coverage of the thesaurus. For example, our method extracted a cluster {"mibun (social status)", "chii (social position)"} (indicated with D in the Figure). Although those words seem very similar, they were not recorded as synonyms in "Bunrui Goihyo".

As a note, we also computed the recall. Words that were included in our 88 tight clusters appeared in a total of 114 categories in "Bunrui Goihyo". Thus, the recall was 77 %.

Conceptual Area Map

Using those tight clusters as guides, we segmented the map into ten rough, coarse-grained adjectival concepts based on our linguistic intuitions. Figure 2 shows the segmented map (annotated with English translations). In each area, some representative adjectives which occurred with the nouns in the area are also shown. As we had expected, nouns which are extremely general and abstract, such as "matter", "feeling", are placed very close to "TOP" at the center of the map. Indeed in the data, "koto (matter)" co-occurred with 1594 (out of 2374) adjectives. Note that the area with "TOP" is not really a conceptual area; it is rather a collection of nouns which are considered the top-most concept in their respective hierarchies (except for "koto (matter)", which is extremely general). We will discuss the hierarchies obtained from the map in the next section. Also as you can see in the map, tight clusters in this area overlap quite a bit (where the same word belongs to several clusters), whereas other tight clusters become more separated towards the edges. This implies the concepts at an extremely high level are very much mingled. We are planning to investigate this hypothesis in psychological experiments in our future work.

Other nine concept areas are placed surrounding the center. From this map, we can identify the relative relations between the concept areas. Although the precise relations in the original data in the high dimensional space are not entirely captured in the low-dimensional map, we can infer the closeness of concepts from neighboring areas.[3] For example, the area "Sense, Perception" is adjacent to the areas "Shape, Appearance" and "Image, Figure, Atmosphere". This corresponds well to our intuitions: a sensation resulting from physically experiencing something through perception (e.g. touching something 'soft') can be metaphorically used to describe things that have the appearance or give the impression of the sensation. Since those concept areas are immediate neighbors on the map, we can infer that the relations between them are quite strong in our cognition. On the other hand, "Sense, Perception" and "Effect, Influence" are mapped far from each other. Although perception may cause effect, we can infer from the map that their relation is not as strong. Similarly, we see in the map that "Characteristics of humans or things" is more strongly related to "Image" than to "State". Although concepts exist in continuum, we can easily see their relative closeness by visualizing them on a map.

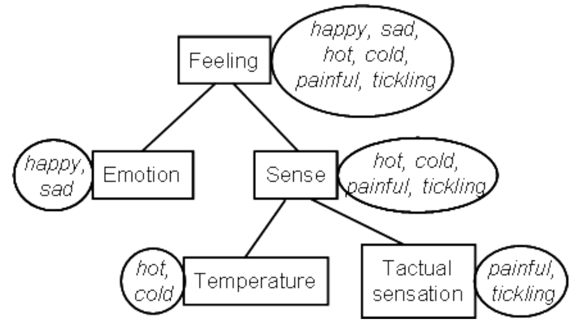

Figure 3: Example Adjectival Concept Taxonomy

Taxonomy of Adjectival Concepts

As we mentioned in the previous section, in constructing a SOM map of abstract nouns, we had expected the nouns to be distributed hierarchically in the general-to-specific order from the center on the map. That would create a taxonomy of abstract nouns. But in our case, each abstract noun represents a semantic concept for a group of adjectives. Also here, abstract nouns which are general are modified by more number of adjectives than specific nouns. Thus, a *taxonomy* of adjectival concepts is a structure organized hierarchically according to the *set subsumption* relation – a set of adjectives associated with nouns at higher levels subsume those associated with nouns at lower levels.

Figure 3 shows an example adjectival taxonomy. It is a hypothetical example, given for the purpose of illustrating and clarifying what we mean by an adjectival concept taxonomy. In the figure, the adjectival concepts are indicated in square boxes, and the adjectives which belong to the given concept are indicated in a circle next to the concept. As you can see, the adjectives are in the subsumption

[2] http://www.kokken.go.jp/english/en/publications/Bunrui_Goihyo.html

[3] For the map obtained, we computed the *topographic error*: a measure which indicates how well the map preserved the topology of the input data. The topographic error essentially measures how much the map is "twisted" – the error is large if there are many input instances whose first best-matching unit (BMU) and the second BMU are not adjacent to each other on the map (http://koti.mbnet.fi/~phodju/nenet/SelfOrganizingMap/Theory.html). Our map showed the topographic error of zero, indicating that the topology of the input data is accurately preserved on the map.

relation. By organizing the adjectival concept space based on subsumption, an adjectival concept taxonomy can represent the *breadth* of each concept and the ways concepts relate to each other in a hierarchical manner. For example in Figure 3, we can see that adjectives which express "feeling" range widely: some are related to (or extended from) physical perception ("sense"), and others are related to psychological state ("emotion"). In turn, there are many kinds of physical perception, obtained through various perceptual experiences such as touching ("tactual sensation") and feeling the heat/coldness ("temperature").

An adjectival concept taxonomy would be useful and of interest to various areas of study. For example in linguistics, it could be used as a basis in analyzing metonymy and metaphor of adjectives. In cognitive science and psychology, taxonomical relations could be used as a potential cognitive model for adjectives and tested by experiments, for instance to study how humans acquire adjectives (Mintz 2005). In NLP, a taxonomy of adjectives could be used in a variety of applications such as Information Extraction, Question-Answering and Machine Translation, in a similar way as noun and verb taxonomies are used. For instance, in comparing two sentences for similarity, the similarity between the adjectives (obtained by computing the distance using the adjective taxonomy) can be included in addition to the noun and verb similarities to improve the accuracy of the overall sentence similarity. Another potential application would be the Semantic Web – using the adjective taxonomy as additional meta-data, the input data could be tagged with more meta-level labels, thus become more descriptive and allow more precise inferences.

However, there have been very little work in linguistics, cognitive science or NLP which tried to view adjectives hierarchically. In conventional dictionaries, adjectives are grouped into a flat list of categories, for example, "color adjectives", "emotion adjectives", "relative adjectives" etc. In various wordnets built in recent years (e.g. English WordNet (Miller 1990), EuroWordNet (Vossen 1998)), nouns and verbs are organized in taxonomies, while adjectives are simply grouped into a small number of coarse categories (descriptive and relational adjectives). One exception is Spanish WordNet (Soler 2004), which proposed taxonomical structures for adjectives by mapping them to existing concept ontologies. GermaNet (Hamp & Feldweg 1997) does organize adjectives in hierarchies, but only lexically and at lower levels (e.g. happy → merry → amusing); at the top level however, adjectives are grouped into fourteen semantic classes (e.g. perceptional, spatial, weather-related, mood-related), and they are not organized into taxonomy. Also, dictionaries and wordnets are compiled manually by lexicographers. Our work, on the other hand, aims to construct a semantic hierarchy of adjectives automatically from the corpus data. We also extract the taxonomy from a SOM map, thereby adding another dimension to the two-dimensional conceptual map.

Extracting A Taxonomy

To extract a taxonomy from a SOM map, we identified map nodes and tight clusters which are in a parent-child relation

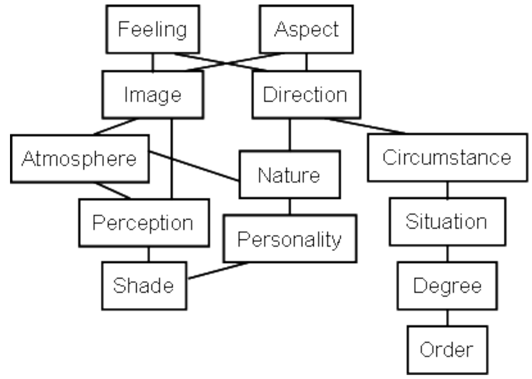

Figure 4: Taxonomy for "kibishii (tough/hard/strict)"

and connected them. We determined the parent-child relation by using two measures: cosine and entropy. Cosine is a measure of similarity. For two nodes to be in the parent-child relation, they must be at least fairly similar. Therefore, pairs of nodes whose reference vectors have a relatively high cosine coefficient are potential candidates. However, cosine alone is not sufficient, because it does not indicate the hierarchical information. Entropy (Shannon 1948) is a measure of impurity used in information theory. For a given random variable X, the entropy of X is large if the distribution of the values X takes is uniform, while if the distribution is rather skewed, the entropy value is small. This property serves our purpose quite well because, in our data, nouns which are general co-occurred with a large number of adjectives, thus map nodes assigned with general nouns yield larger entropies. Caraballo & Charniak (1999) also reported that entropy was the most effective indicator of the specificity of nouns when their features were pre-nominal modifiers (adjectives, verbs and nouns which modify a noun and appear before the noun). Indeed, entropy has been used often recently in building domain ontologies (e.g. (Ryu & Choi 2006)). We computed the entropy of a reference vector v, denoted $H(v)$, of a map node as follows.

$$H(v) = -\sum_i p(v_i) \times log_2 p(v_i)$$

where v_i is the value for the ith feature, and $p(v_i)$ is its probability in v. By using these two measures, we determined that a map node n1 is a (direct) parent of n2 if their cosine coefficient is greater than a pre-specified threshold and n1's entropy is larger than that of n2. For a tight cluster, the centroid of the reference vectors of the map nodes assigned to the cluster was used as the representative vector.

The Taxonomy

The taxonomy obtained is indicated by bold lines connecting nodes and tight clusters in Figure 1. As you can see, the branches are descending from the center of the map in directions roughly corresponding to the conceptual areas we segmented based on our intuitions (Figure 2). We used the threshold of cosine ≥ 0.6 in determining the parent-child relation, so some nodes/clusters were not connected whose similarities to any other nodes were less than the threshold. However the overall taxonomy is clearly observable.

One notable thing about this taxonomy is that the top is not a single node/cluster. Rather, it is a cloud of nodes and tight clusters surrounding "TOP" at the center of the map, consisting of nouns which are extremely general and abstract. Those nodes and clusters were not connected to each other by the parent-child relation; instead they are densely overlapping, as we mentioned in the previous section. This suggests that the highest level of the adjectival taxonomy might be a cloud of extremely abstract concepts which are vague and indistinguishable in our cognition.

Figure 4 shows a part of the taxonomy from Figure 1 where the adjective "kibishii (tough/hard/strict)" appeared as a modifier (e.g. "kibishii kanji (tough feeling)", "kibishii houkou (tough direction)"). From this taxonomy, we can see that adjectival concepts are not mutually exclusive between branches; the taxonomy is a graph, rather than a tree, so some branches are shared. And the way nodes are connected is quite complex. For example, "image" subsumes "perception" directly as well as through "atmosphere", and "nature" and all of its descendants are subsumed by "feeling" and "aspect" through different paths.

We can also observe the breadth of various adjectival concepts from this taxonomy. For example, we can see that "image" is a very broad concept – there are many kinds of image, including atmospheric image (such as "quiet image"), perceptual image (such as "soft image"), and personality image (such as "brave image"). In contrast, the concept of "order" is narrow; there are not so many kinds of order, and indeed in our data there were only 27 adjectives (out of 2374) which modified "junban (order)", as compared to "imeji (image)" which was modified by 871 adjectives.

Finally, we can also observe the relative closeness of the adjectival concepts from this taxonomy. For example, "perception" is much closer to "nature" than to "circumstance". That correlates well with our intuitions: the sensations perceived through physical experiences (such as "soft", "bitter") can describe the nature of the things being perceived metonymically (i.e., without changing the domains of the nouns modified), but perceptual adjectives modify situations only metaphorically (e.g. "soft situation", "bitter situation"; requiring the transfer of the domains which the adjectives modify).

Conclusions and Future Work

In this paper, we visualized the adjectival concept space based on the data collected from corpus, and built a conceptual model. The results, both the conceptual map and the concept taxonomy, showed that the highest level of the adjectival concept space is a cloud of extremely abstract concepts. Also the breadth of the concepts create a complex, overlapping structure. Our findings provide insights on various properties and behavior of adjectives, and will be useful in many areas and applications which deal with adjectives.

For future work, an immediate task would be to apply the techniques we developed in this work to other kinds of data (not newspaper articles) and other languages (such as English) to verify the validity of our approach. We are also planning to conduct psychological experiments to see how our SOM map correlates with the Hyperspace Analogue to Language (HAL) model of memory (Audet & Burgess 1999). Although our conceptual map is a model produced by a Machine Learning method and not cognitively motivated, it would be very interesting to compare the two models. Another thing we would like to investigate is how to incorporate our adjectival concepts in a lexical semantic theory, in particular, the Generative Lexicon Theory (GL) (Pustejovsky 1995). In GL, the phenomenon of adjectival polysemy is explained by a mechanism called *Selective Binding*: when an adjective is combined with a noun, a generative process takes place where the adjective as a function (or predicate) selects a specific facet of the noun to produce a plausible interpretation. We are planning to investigate the applicability of our adjectival taxonomy in this generative selection process.

Finally, we are also planning to use the data to cluster adjectives instead of abstract nouns. By clustering adjectives, we will obtain groups of adjectives which have similar *meaning extension patterns*. For example, adjectives which express both temperature and personality, such as "warm", "cold" and "cool", will be grouped in the same cluster. Then by organizing those clusters hierarchically (based again on the subsumption relation), we will be able to observe the patterns of meaning extension. For instance, a pattern `temperature-personality` may be subsumed by another pattern `temperature-personality-color` (e.g. "warm temperature", "warm personality", "warm color"). Such a hierarchy will be extremely useful in investigating how various adjectival meanings extended – polysemy, metonymy and metaphor of adjectives.

References

Audet, C., and Burgess, C. 1999. Using a High-dimensional Memory Model to Evaluate the Properties of Abstract and Concrete Words. In *Proceedings of the Cognitive Science Society*, 37–42.

Caraballo, S., and Charniak, E. 1999. Determining the Specificity of Nouns from Text. In *Proceedings of the joint SIGDAT conference on Empirical Methods in Natural Language Processing (EMNLP)*, 63–70.

Fellbaum, C.; Gross, D.; and Miller, K. 1993. Adjectives in WordNet. In *Five Papers on WordNet*, included in a WordNet delivery.

Hamp, B., and Feldweg, H. 1997. GermaNet - a Lexical-Semantic Net for German. In *Proceedings of ACL workshop on Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*. On GermaNet adjectives: `http://www.sfs.uni-tuebingen.de/lsd/Adj.html`.

Hatzivassiloglou, V., and McKeown, K. R. 1997. Predicting the Semantic Orientation of Adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics (ACL/EACL-97)*, 174–181.

Hatzivassiloglou, V., and Wiebe, J. 2000. Effects of Adjective Orientation and Gradability on Sentence Subjectivity. In *Proceedings of the 18th International Conference on Computational Linguistics (COLING-2000)*.

Isahara, H., and Kanzaki, K. 1999. Lexical Semantics to Disambiguate Polysemous Phenomena of Japanese Adnominal Constituents. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, 489–496.

Kanzaki, K.; Yamamoto, E.; Ma, Q.; and Isahara, H. 2004. Construction of an Objective Hierarchy of Abstract Concepts via Directional Similarity. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING-2004)*, 1147–1153.

Kohonen, T.; Hynninen, J.; Kangas, J.; and Laaksonen, J. 1996. SOM_PAK: The Self-organizing Map Program Package. Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, Finland. `http://www.cis.hut.fi/research/som_lvq_pak.shtml`.

Kohonen, T. 1984. Self-organization and Associative Memory. In *Springer series in Information Science, Vol. 8*. Berlin: Springer.

Kohonen, T. 1995. *Self-organizing Maps*. Berlin: Springer.

Kohonen, T. 1996. The Self-organizing Map: A Possible Model of Brain Maps. *Medical & Biological Engineering & Computing* 34:5–8.

Lapata, M. 2001. A Corpus-based Account of Regular Polysemy: The Case of Context-sensitive Adjectives. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics (NAACL)*, 63–70.

Lin, D., and Pantel, P. 2002. Concept Discovery from Text. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING-2002)*, 768–774.

Manning, C., and Schutze, H. 1999. *Foundations of Statistical Natural Language Processing*. The MIT Press.

Miller, G. 1990. WordNet: An Online Lexical Database. *International Journal of Lexicography* 3(4). WordNet available at `http://wordnet.princeton.edu/`.

Mintz, T. 2005. Linguistic and Conceptual Influence on Adjective Acquisition in 24- and 36-Month-Olds. *Developmental Psychology* 41(1):17–29.

Murphy, G. L., and Andrew, J. M. 1993. The Conceptual Basis of Antonymy and Synonymy in Adjectives. *Journal of Memory and Language* 32:301–319.

Nemoto, K. 1969. The Combination of Nouns and Adjectives with "ga-case". *Language Research for Computers* 63–73. in Japanese.

Pustejovsky, J. 1995. *The Generative Lexicon*. The MIT Press.

Ryu, P.-M., and Choi, K.-S. 2006. Taxonomy Learning using Term Specificity and Similarity. In *Proceedings of the 2nd Workshop on Ontology Learning and Population, held at COLING/ACL-2006*, 41–48.

Salton, G., and McGill, M. 1983. *Introduction to Modern Information Retrieval*. McGraw Hill.

Schmid, H.-J. 2000. *English Abstract Nouns as Conceptual Shells*. Germany: Mouton de Gruyter.

Shannon, C. E. 1948. A Mathematical Theory of Communication. Technical Report 27(3), Bell Systems.

Soler, C. 2004. Extension of the SpanishWordNet. In *Proceedings of the 2nd Global WordNet Conference*, 213–219.

Takahashi, T. 1975. Various Phases Related to the Part-whole Relationship Investigated in Sentences. *Studies in the Japanese Language* 103:1–16. in Japanese.

Vossen, P. 1998. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Dordrecht: Kluwer Academic Publishers.

Wiebe, J. 2000. Learning Subjective Adjectives from Corpora. In *Proceedings of the 17th National Conference on Artificial Intelligence (AAAI-2000)*, 735–740.