Semantic Classification and Weight Matrices Derived from the Creation of Emotional Word Dictionary for Semantic Computing

Minsu Ko
Division of Web Science Technology, School of Computing
Korea Advanced Institute of Science and Technology
Daejon, Republic of Korea
Email: ryan0802@kaist.ac.kr

***Abstract*—This paper introduces a general creation method for an emotional word dictionary (EWD) which contains a semantic weight matrix (SWM) and a semantic classification matrix (SCM) which will be used as an efficient foundation for opinion mining. These two matrices are combined into a single n by 7 matrix called as a classification and weight matrix (CWM) in a machine-processable format. Such a matrix would also have applications in the field of semantic computing.This paper also details investigations which were performed in order to gather information on the efficiency of using CWM based on categorizing synonymous relations and frequencies. The multilingual extensibility of the EWD will benefit semantic processing of opinion mining as a generic linguistic resource which has an emotional ontology structure and linked data.**

I. Introduction

In recent times, the number of internet documents has increased exponentially due to the availability of easy creation methods and instant publication, both of which help users express their ideas. In particular, online reviews are useful data for use in opinion mining because they not only reveal rich sentiments but also have the potential to affect the future choices of other users. Therefore, many active research projects attempt to analyze the emotions and opinions of online reviews and documents.

The polarity of a document as a basic unit is determined by comparing it against a list of emotional words. A number of studies analyzing the sentiment of texts have tried to create dictionaries which automatically classify these words, but these studies did not utilize a numerical vector with continuous values corresponding to the word meaning. This study contributes not only the notion of a categorization of emotional words, but also the vectorization of each word's attributes; an approach which is reusable for other NLP applications.

These idea started as a way to use real values for word meaning which had been categorized and graded by both a specified criterion and the word's role containing the various semantic relationships between words. Even though we mainly focus on Korean as a source language, the method is language-independent, since the seed list of synset and the creation algorithm are purposed.

Sentiment analysis and metasearch of a software agent based on the EWD in a semantic web has generally been expected to produce good results because of EWD's delicate features and relational properties. Numerous studies exist regarding the polarity of words but, assuming the decision making process of humans to be a computable function of the relevant data gathered, our approach attempts to emulate this process by performing computations on the available data. Thus, the data collected here may also have applications in future machine learning projects. We created the CWM for each word to incorporate gradability into the feature focus (SWM) and then examined the input text for emotional process identification (SCM). The resulting data entry could effectively be used to analyze the semantic orientation of a word and to supply this information to an analyzing system.

II. Related Work

Research on identifying the polarity of expressions has received increased interest in the past few years but work regarding automatic sentiment classification dictionaries is currently limited because research relating to the creation of dictionaries has received relatively little attention. In this section we will take a brief look at some of the available research which uses linguistic resources and dictionary creation.

One approach to determine word polarity is to use a linguistic resource. This method allows one to predict the polarity of a text using the relations of synonyms and antonyms found in a thesaurus. [1] used distance measurements on the syntactic category of adjectives to develop WordNet-based measurements for the semantic orientation of adjectives. This method depended on the hypothesis that all synonyms had the same polarity. [2] employed WordNet to check the relations between synonyms and antonyms. [3] proposed probabilistic models to estimate the strength of a word's polarity. Using synonyms found in WordNet. [4], [5] gave polarity values to words based on the WordNet gloss corpus.

Appraisal taxonomies classify an appraisal group using its emotional adjectives [6] in accordance with appraisal theory [7]. This method evaluates the attitudes that appear in the text and examine how the words deal with human relationships. [7] described an attitude, with regards to the evaluation of feelings, as having three parts: the individual's way of feeling including the emotional responses, the decision of action, and the evaluation of a particular object. Attitude is the most important aspect of an appraisal because it reveals the very essence of the intended sentiment.

III. Construction of the Emotional Word Dictionary

A. Emotional Word Dictionary

The framework of dictionary building was fundamentally based on statistical and mathematical approach using linguistic resources and corpora with several conceptual motivations.

The creation of an EWD stems from the following two motivations: **a**. Can the intersection of meaning of emotional words be found in a single language society? **b**. Can synonyms be graded relative to each other? That is, this was the result of attempting to construct a form of computerized processing to extract all of the emotional words in a text and find the shared parts of meanings in each synset. Polarity synsets are classified by their synonymous relations and exist inter-independently. These synsets have n words and can be used to search the vocabulary of the emotional word dictionary and find the assigned polarity values at a certain scale.

The EWD (ver 1.2) is now designed to be a language resource which can be formalized as a opinion mining database with reusablility in semantic webs. The database structure of an EWD consists of 12 columns. (SID, WID, W, AT, PR, Dom, ONT, P, PreVal, TFIDF, ZT, SV)

B. Conceptual Foundation of Emotional Word Dictionary

1) Semantic Classification Matrix (SCM): The concept of semantic classification matrix is at the core of using an EWD for classification method. The SCM consists of a n by 4 matrix which contains a quadruple representing the statistical and vectorized information of each word as a semantic classification feature.

Semantic classification feature (SCF) is defined as a quadruple with categorical information corresponding to each word: SCF = (AT, PR, ONT, P), where AT is the attitude type assigned to synset according to White's classification, PR is the degree of word's prototypicality, ONT is the information corresponding to the mikrokosmos (μK) ontology, and P is the polarity information. Each feature is stored as a vector in the database and it can be grouped by SID order. The links of the ATs compose an ontological structure which has semantic concept nodes which can be evaluated by the distance information between n-words. The role of P is to specify AT, and the hierarchical property of PR builds a binary asymmetrical top-down tree structure of the input text, when a system recalls SCFs. AT is a naive discriminator, but it has a significant role in the problem of disambiguation. ONT exists to provide future usability expanding the attitude tree of EWD to a large ontology.

@SYNSET_ID	@WORD_ID	@Word	@POS	@ATTITUDE_TYPE	@PIDX	@DOMAIN_INFO
SPredN014	W01_SPredN014	헛되	ADJ	appreciation_VALUATION	0	MOVIE
SPredN014	W02_SPredN014	부질없	ADJ	appreciation_VALUATION	1	MOVIE
SPredN014	W03_SPredN014	덧없	ADJ	appreciation_VALUATION	2	MOVIE
SPredN014	W04_SPredN014	무상	NNS	appreciation_VALUATION	3	MOVIE
SPredN014	W05_SPredN014	허무하	ADJ	appreciation_VALUATION	4	MOVIE
SPredN014	W06_SPredN014	허황	NNS	appreciation_VALUATION	5	MOVIE
SPredN014	W07_SPredN014	맥빠지	ADJ	appreciation_VALUATION	6	MOVIE
SPredN014	W08_SPredN014	뭣하	ADJ	appreciation_VALUATION	7	MOVIE
SPredN014	W09_SPredN014	희망없	ADJ	appreciation_VALUATION	8	MOVIE
SPredN014	W10_SPredN014	새되	ADJ	appreciation_VALUATION	9	MOVIE

@ONTOLOGY_INFO	@POL	@PREDEF_VAL	@TFIDF	@Z_TRANSFORM	@SEMANTIC_VALUE
UTILITY-ATTRIBUTE	negative	9	0.405155	-1.499385498	-1.03405
UTILITY-ATTRIBUTE	negative	9	1.012888	0.500614502	-1.34575
UTILITY-ATTRIBUTE	negative	8	0.675387	-0.0231734196	-1.748
UTILITY-ATTRIBUTE	negative	8	1.125645	1.9768265804	-1.98805
UTILITY-ATTRIBUTE	negative	7	0.47657	-1.0	-2.08055
UTILITY-ATTRIBUTE	negative	6	0.469679	-1.5047339397	-2.53405
UTILITY-ATTRIBUTE	negative	6	0.704518	-0.4952660603	-2.65605
UTILITY-ATTRIBUTE	negative	5	0.604932	-1.3323300684	-3.0467
UTILITY-ATTRIBUTE	negative	5	1.209865	0.6676699316	-3.3743
UTILITY-ATTRIBUTE	negative	4	1.446055	1.0	-3.92065

Fig. 1: Example of the EWD DB Structure

2) Semantic Weight Matrix (SWM): The concept of a semantic weight matrix is at the core of using an EWD for computation method. The SWM consists of an n by 3 matrix which contains a triple representing the statistical and vectorized information of each word as a semantic weight feature.

A semantic weight feature (SWF) is defined as a triple with numerical information corresponding to each word: SWF = (TFIDF, ZT, SV), where TFIDF is the term frequency - inverse document frequency, ZT is the value for Gaussian distribution function to gain the probability belonging to the synset, and SV is the semantic value which shows the grading form and contains the numerical meta information of emotional words' gradation. Each feature is stored as a vector in the database and it can be grouped by SID order. TFIDF is the basic feature of a statistical interpretation for text, ZT is the normalized value of each emotional word, and SV is the overt feature for vectorization of texts.

3) Prototypicality and Prototype Meaning: [8] proposed that the structure of prototypical categories take the form of a radial set of clustered and overlapping meanings. This would imply that each synset can be represented in such a way as to allow one to find the core meaning and apply it to a function of ontology class mapping.

$$Word_{core} = min(\{|V_n||V_n \in \{SV_i\}, n \in N\}) \qquad (1)$$

Prototype meaning can be used to identify the center of each synset and also be adopted to select adequate emotional words for building a language-extensible list. If every word has its own properties and they all have a certain shared property, then this becomes the center of the synset and acts as a prototype which is assigned to each word. The prototype meaning mentioned here is not related to the set of necessary and sufficient conditions in classical categorization theory. Ontological questions about the prototypicality of words are proposed as psychological objects. Although it's intangible, it's still observable in the idea that the meaning of words conforms to a minimum intersection.

4) Synonymous Relations and Synsets: The emotional word dictionary is designed to hold lists of all of the emotional words. Due to the importance of consistency, the building process of synsets required a set of robust and unified criteria. Positive/Negative synsets were produced for all entries made from thesauri and chunk lists obtained from corpora. Synonymous relations are defined here as several words having close practically-related usages. If two or more words have the same meaning then they are in a synonymous relationship with each other. Additionally, we define that sharing the same meaning infers sharing the same prototype meaning. [9] also proposed that sharing the same truth value

was the very definition of the term ‘same meaning’ and the truth value relies on the usage of a word in context. A synset composed of emotional words ultimately has a counterpart with an opposite polarity which it may be concatenated with.
Theorem (Requirements of Synset) :
$Synset_i = \{m_j | m_j \in S_i, S_i \cap S_{i\pm1} = \varnothing, i, j \geq 1\}$
1. Associative law : $\forall m_i, m_{i+1}, m_{i+2} \in S_i$: $(m_i * m_{i+1}) * m_{i+2} = m_i * (m_{i+1} * m_{i+2})$
2. Identity element : $\exists e \in S_i : m_i * e = e * m_i = m_i$
3. Inverse element : $\exists x \in S_i : m_i * x = e$

Each element in the same grade for each synset may be graded in different dimension but the SVs obey a unified criterion. Each SV present in a synset means a point of data type real exists on the same continuous scale. The process of grading values in a synset is done by keeping the following definitions:

[**Def. 1**] If and only if two words belong to other synsets and their intersection is the empty set, i.e., they are in relation of independent sets, two grades with the same index in different dimensions are completely irrelevant.
[**Def. 2**] The synset is an open set. Any new element ‘m_j’ will be added to an existing synset. Each element in a synset has the possibility of addition/deletion because the meaning of the words iterates the process of extinction, transformation, and creation with the flow of time.
[**Def. 3**] We can find a number of the infinite empty places between the points of SVs on the scale, but the realization of definition 2 relies on this definition. Thus, the possible size of a single synset can be figured out as follows (2), because a synset has a half scale of integer.

$$n(Synset_i) \leq O(2^{\aleph_0}) \tag{2}$$

C. *Process of Building Emotional Word Dictionary*

1) Seed Words from Domain Corpus - Initial Phase Example: The seed list of emotional words in the emotional word dictionary began with a corpus of movie reviews. The seed words were manually extracted from the Cine21 [19] Movie Review Corpus (229,192 reviews, 2,047,110 words) based on frequency of occurrence. The expansion process to include colloquial forms frequently used in the same domain was later done using the Naver Movie [20] Review Corpus. (for line 1)

Assuming domain-specific sentimental word identification, we chose the domain as movie reviews containing dense sentimental expressions to be the source of a high proportion of the emotional words in the emotional word dictionary. Basic emotions tended to be identifiable as a function of the word frequency-order in a corpus. This means the information gathered through the sentimental expressions can be classified under several subclasses of emotion. The process to create lists of synonyms was done semi-automatically by following the two phases. After the completion of this process and until the lists satisfied the definitions above, they are considered to be a synset.

[**Phase 1**] Words were sorted by frequency-order. This was used to determine the coreness of an emotion among countless words. Low frequency words tended to be variations on standard form. Thus, they were lined up with high frequency words, the core of a certain emotion category.
[**Phase 2**] The synonym lists were still domain-specific after the phase 1. Although the lists were extracted from a large-sized corpus, they unexpectedly still had a deficit of general emotional words. To ensure a generic dictionary, emotional words from Korean thesauri [18] were added to the lists through the result of a synonym search and developer’s intuition. Each search result had to be reclassified for our work.

Algorithm : Abstract Process of Building EWD
1: import seed_list
2: import corpus
3: **for all** e in seed_list do # e in seed_list = {w, SCF}
4: build expanded S_i
5: **for all** s in S_i do
6: **if** count(s)>0 in corpus
7: calculate SWF then
8: apppend to S_i
9: **if** count(s)=0 in corpus
10: calculate $V_{estim.}$ from S_i then
11: apppend to S_i
12: duplicate SCF # Update CWM
13: align PR in SCF
14: **end for**

2) Extensions as a General Linguistic Resource: The synsets were still potentially lacking some emotional word entries which could prevent them for being useful as a general linguistic resource. Phase 2 was performed in an iterative manner to add previously unknown emotional words. (line 4) Our proposal that a representative of synset can automatically be selected for a word based on the nearest coreness reveals the important problem of ambiguity which was previously discussed in both [5] and [10]. We employed the idea of a practical treatment to deal with the ambiguity problem and expected that the slot of semantic values in existence would have more branches regarding the usage and observations with topic analysis to solve this problem.

First, emotional words from semantic classes in the Sejong Noun Dictionary[1] were added to their related synsets. The semantic classes of the SND made the process of identifying and extracting each emotional word’s polarity easier than using a raw corpus. Second, the emotional words corresponding to KOLON (the Korean Lexicon Ontology) [11] were also manually checked and added to the synsets.

3) Building Classification and Weight Matrices: The CWMs consist of two main parts, categorical (SCM) and numerical information (SWM) with the purpose of providing reusability in semantic classification and opinion mining.

The feature set of SWM, SWF, contains statistically reusable data as described above. Each feature of SWF is continuously calculated, TFIDF (3), ZT (4) and SV(8). (line 5-13)

4) Grading of Semantic Values and Allocation: We set the five Gaussian distribution functions at each polarity scale as a model template, and each synset found its adequate model

[1]The Sejong Dictionary is one of outputs of the 21^{st} Sejong Project started in 1998 with a 10-year plan by Korean government.

by their frequency-based information. Thus, we have 10 sub-distributions of a single synset at 2-dimensional vector space and the models will have n-duplications when a new synset is created. We assumed that if a synset was integrated from 0 to 1 on the negative and positive infinite timeline, it would recover its ideal shape of the model (6).

$$TFIDF_{i,j} = \frac{wfs_{i,j}}{\sum_k wfs_{k,j}} \times ln(\frac{|D|}{tfs}) \tag{3}$$

$$\tilde{d_i} = \frac{d_i - E(d)}{\sigma_d} \tag{4}$$

$$E(d) = \frac{1}{N}\sum_{i=1}^{N} d_i \tag{5}$$

$$N(x|\mu,\sigma^2) = \int_0^1 S_i(x)dx \tag{6}$$

Considering that the meaning of a word is contextually limited and the semantic relation has more priority than its lexical description, the separation of the practical application from its defined meaning is reasonable. Supposing that the decision making process and the linguistic intuition of writers is reflected within a corpus, we focused on the idea that comparing the TFIDF values of each emotional word revealed meaningful differences among emotional words. In other words, statistical results denoted an internal decision by random people in a single language society. Each word is set to have its semantic value according to a normalized probability distribution function. The calculation process for these values consisted of three stages as follows (line 7-8):

$$\sigma_d = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N}(d_i - E(d))^2\right)} \tag{7}$$

$$SV_i = V_{grade_i} + V_{PDF_i} \tag{8}$$

$$N(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{9}$$

[**Stage 1**] First, the polarity synset was recognized as a line, and the element with the most meaningful TFIDF value was selected as the representative. The words were ranked according to their already existing grades. Each word of the same rank was sorted by the weight of its TFIDF. The initial grading scale was created by converting the grades assigned by the reviewers to a positive scale of [0.5, 10.5], or a negative scale of [-0.5, -10.5] to allow for calculations using normalized values.

[**Stage 2**] Next, the normalized values were automatically allocated to the initial grade of emotional words followed by the standard normalized distribution(9)'s ZT(4) of TFIDF(3). As a result, all of the grade points had a continuous distribution with a uniform range.

[**Stage 3**] Finally, the SVs of the two stages above were distributed in the range of [-5.5, 5.5] which was compressed to half the size of its previous size during stage 2.

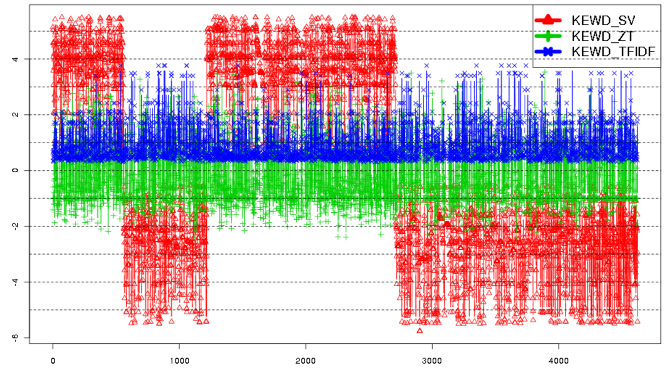

Fig. 2: Red dots (triangles, KEWD_SV) are scattered at four areas. A1, A2, A3, A4 by left-to-right order. A1 and A2 mean the polarity synsets of predicates, A3 and A4 mean the polarity synsets of nouns. They are like a mirror image to each other.

D. Strategies to Allocate SV of Null Frequency Words

The list expansion of emotional words using general linguistic resources aimed to create a domain-independent dictionary. Unfortunately, when a null frequency word in the corpus is found in the word list, the SV cannot be calculated. The SV essentially relies on frequencies and existing grade information so even estimation is impossible due to the lack of trace information. However, this problem had to be overcome in order to allow the emotional word dictionary to be used as a more general resource. Generally accepted smoothing techniques were not appropriate for this situation since they are related to n-gram models and this system not only lacked information about n-gram but also required entries to be independent from existing word sequences.

$$V_{estim.} = E(SV_i) + |min(\{V_{PDFi}|i \in N\})| \tag{10}$$

$$Cond._{estim.} = \{SV_i|SV_i \in Synset_j\} \tag{11}$$

Therefore, a back-up plan was created to help estimate values for null frequency words, based on the idea of coreness. (line 9-11) First, the word in question already belonged to a synset. The minimum SV of hapax legomenon in a synset could be used as a good base for estimation. That is, the function (10) can be applied to the function for SV (8). Using (10), the estimated SV, $V_{estim.}$, was achieved by taking a mean of the SVs in a synset and minimum value of the probability density function.

If we use this function to check the estimated SV of the synset, $S_{PredN014}$, as an example, we find $V_{estim.} = -2.485135$. This situation forms a significant proportion (approximately 10%) of the entries in the emotional word dictionary. As the corpus size is periodically increased, the estimated values will also be recalculated to reflect the changes to the data used in the estimation process.

E. Multilingual Extensibility

We designed our approach to underline the multilingual extensibility that can create emotional word dictionaries in any language. The matters of language alternation and the input size of word list were basically independent variables because the methodology and techniques of creation were combined into a single process like a compositional function.

1) KEWD and EEWD: Korean emotional word dictionary 1.0 [13] was a semantic dictionary with a number of emotional words. Generality was ensured by the use of general linguistic resources. Information about Korean parts-of-speech existed as a unit of morpheme because of the attributes of agglutinative languages. Therefore, the unit of entry was based on a morpheme instead of a chunk.

KEWD 1.1 was created using the 1.0 version as a base, and KEWD 1.2 has adopted the new concept of CWM, subdividing the attitude type. All of the synsets were matched one-to-one with word classes in KOLON in order to introduce the possibility creating a partial ontology with connective information for semantic web.

A test version for an English emotional word dictionary (EEWD) was performed to verify the extensibility using the IMDb [22] corpus (1.62 times the size of Cine21). MPQA [12] subjectivity lexicon which is a part of OpinionFinder was selected as seed words and each word expanded to a synset based on WordNet. This provided developers the advantages of both time management and the technical consitant procedure.

2) Reusable Seed List of EWD: The basic seed list of EWD needs to be set commonly for reusability in any language. If developers have no common list, the EWD of multi-languages cannot be only compatible, but also mutual-interchangeable with a metasearch of the emotion between multi-languages. Thus, we distribute a reusable seed list of EWD on our webpage. We set a threshold $\theta(p, k = 2)$ to build a basic seed corresponding to 413 synsets. (926 emotional words)

$$p = \underset{x \in N}{\operatorname{argmin}}(PR_i) \tag{12}$$

F. Investigation of Applicability Aspects

This paper mainly focuses on the detailed description of a methodology for dictionary creation and its components, but we will now briefly investigate the applicability of these

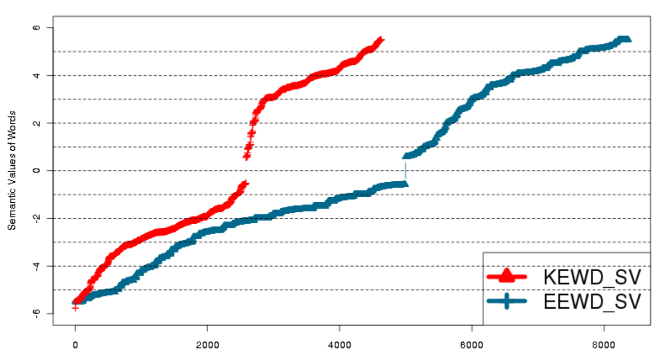

Fig. 3: Linear Distribution of the SVs of KEWD and EEWD : The size of EEWD word list is roughly over two times more than KEWD, but they show similar line shapes, the scalability aspects and the relation of multilingual extensibility of emotional word dictionary.

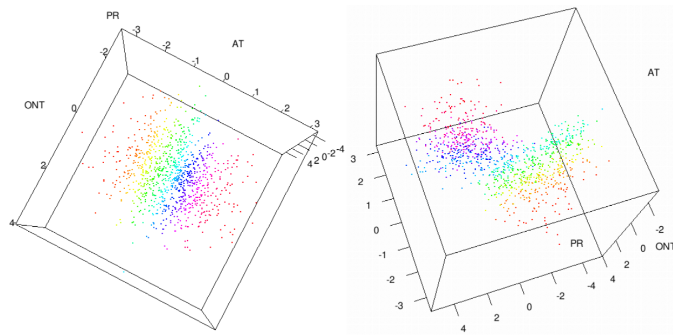

Fig. 4: Significant Distribution - *on the axes of AT and PR*: The density of points is going to be high at the center of the cube, low at the edge of the cube. *on the axes of AT and PR*: The coordinates of the points are proportionally increased from the zero point in general, i.e., the emotional words belonging to affect are most prototypical.

dictionaries as linguistic resources to demonstrate guidelines for reuse.

1) Automatic Rating with SWM: The statistical characteristics of SWM have actualized the practical use of an automatic analysis of the continuous scale of expressions and determine the grade of input text. One available example of NLP application is the ARSSA [21] (The Automatic Rating System for Sentiment Analysis) [13]. They use the SV of SWM as a basic feature to classify the grades of review texts using linguistic features, machine learning and SVM classifiers. The vectorization of word meanings contribute to a useful feature set of emotional expressions for use with any data mining technique.

2) Emotion Detection and Interpretation with SCM: Emotion detection and interpretation are related with the theme of metasearch in semantic web. The feature of SCM is expected to be adopted due to its relational characteristics. SCM has 24 ATs, 243 ONTs and 336 AT-ONT relations. (P-AT-ONT relations are double sized.) The ATs compose a kind of partial emotional ontology themselves, but we have to consider the AT-ONT relations for a description of an attribute of AT and a portable outlink to a large ontology structure. ONT-AT relations are also considered to find the attitude of an explicit description. This interactive structure contributes to the detection and metasearch of the emotional expressions and the emotion interpretation of n-language.

IV. Evaluations and Experiments

A. Calculation and Expansion Methods for Matrices

We now have the basic features for language processing, but expansion method is also required for human-like compositional computation. Concatenated rules, 13 conjunctives, and negation markers which change a sentiment orientation are selected as the additional factors for calculations. Each word which is extracted from the text has its own features but the context of an emotional word highly affects the whole meaning. We assumed that any adverbs or demonstrative adnouns occuring before emotional words must be considered

	AT	PR	ONT	P	TFIDF	ZT	SV
AT	31.13202980	**2.0492813**	NA	**0.58419084**	0.07858809	0.19193711	**1.7692351**
PR	**2.04928129**	41.7656189	NA	0.21968370	0.37982676	**0.78712409**	**2.4939756**
ONT	NA	NA	NA	NA	NA	NA	NA
P	**0.58419084**	0.2196837	NA	14.76884895	**0.63277199**	0.07090735	**3.2629574**
TFIDF	0.07858809	0.3798268	NA	**0.63277199**	0.42933202	**0.54974620**	-0.4643019
ZT	0.19193711	**0.7871241**	NA	0.07090735	**0.54974620**	1.20760817	-0.2135022
SV	**1.76923508**	**2.4939756**	NA	**3.26295736**	-0.46430193	-0.21350216	12.6219321

TABLE I: Covariance Matrix of CW combination: Significant relationships are marked in bold face. To determine the dividing point of significance, we normalized the covariance values with significance level 0.01. The confidence interval is between ± 0.4710414, but we reversed the interval to find the significant values of relationship, cutting the left side. $\bar{X} \geq \theta + \frac{2.58}{\sqrt{n}}$

to be a single unit as a concatenated segment weighting to SWF.

Four basic rules of concatenation and two sub rules are combined to catch the segment. For matrix computation, we have combined SCM and SWM to one n by 6 CWM without ONT. ONT is too insignificant to be adopted for calculation and it is just used for AT specification. However, it will definitely be recalled for the expansion to a large ontology system as a linking factor. From the ANCOVA results in table 1, we can delimit multiple features which have a high discriminating power for our experiments. The selection of AT-PR-P feature combination is adequate for testing of polarity detection, and the selection of AT-PR-SV for testing of grading the levels of emotion.

In the tests with AT-PR-P and AT-PR-SV, the P or SV roles will act as a discriminator which determines the polarity or the grade. PR has hierarchical structure roles which affect the weighting feature of P/SV. We multiplied the weight 1-0.01886792 as a shrinking factor to P/SV every step from 0 to n, because the maximum depth of PRs is 53. PR also affects the weight of ATs as it represents the coreness. We can expect the contrast markings of emotional areas.

Rules of Concatenated Segments	
B1:	$BasicRule_{concat}$
	$BC_{R01} : md_{[0,1]} + (md\|a*)_{[0,1]} + nc(a\|s)*$
	$BC_{R02} : md_{[0,1]} + (md\|a*)_{[0,1]} + pa$
	$BC_{R03} : md_{[0,1]} + (md\|a*)_{[0,1]} + pv$
	$BC_{R01} : md_{[0,1]} + (md\|a*)_{[0,1]} + pa + nc(a\|s)*$
S1:	$SubRule_{pa}$
	$pa_{R01} : nc + xn$
	$pa_{R02} : nca + xpv$
	$pa_{R03} : ncs + xpa$
	$pa_{R04} : (nc) * +(pa\|px) * +exm$
	$pa_{R05} : nc + jcm$
	$pa_{R06} : nc + jc + pa$
S2:	$SubRule_{a}$
	$a_{R01} : pv + ecs$
	$a_{R02} : pa + xa$

B. Experiments

1) Experimental Set-up and Exemplifications: Although this paper focuses on the detailed methodology of dictionary creation and the verification as a machine-processable resource for semantic computing, we will show some brief approaches of application tests to suggest some practical examples. The detection of emotional words depends on the registration list of EWD database.

In this section, some experiments are proposed to examine the generality and validity of the EWD. For these experiments, 1000 book reviews from Aladdin bookstore [23], which were not previously in the EWD, were selected at 100 reviews per grade. Basic test sets (1000 reviews) were randomly divided into 10 subsets to help ensure a robust result and cross validation. Each subset of a grade contained 50 reviews at the same grade and 50 at the others. The results of each experiment were compared with the existing author's rating. Two limited matrices were utilized in the test version of semantic computing in this paper, but the unlimited matrices offer additional features which may be useful in different types of applications.

<Experimental Set-up for Polarity Prediction and Grading> : The discriminating power of the CWM will be proved in four experiments. Each AT represents the attitude of semantic segment and PR discriminates the contrast level of AT and roles identically in all cases.

$$P_{\omega}^{i,j} = P_{init}^{i,j} \times (1-\omega)^{PR} \quad (13)$$

$$SV_{\omega}^{i,j} = SV_{init}^{i,j} \times (1-\omega)^{PR} \quad (14)$$

[**Exp_P&P : AT-PR-P#Polarity Prediction**] : Determining polarity of input text using AT-PR-P matrix.
⇒ We proved the possibility that AT-PR-P matrix can be used in a polarity prediction system (thumps up/down). First, the Ps of the detected emotional words were used in polarity prediction. The weighting function (13) is used to modify the initial P input values.
[**Exp_P&SV : AT-PR-SV#Polarity Prediction**] : Determining polarity of input text using AT-PR-SV matrix.
⇒ We proved the possibility that AT-PR-SV matrix can be used in a polarity prediction system (thumps up/down). First, the SVs of detected emotional words were used in polarity prediction. The weighting function (14) is used to modify the initial SV input values.
[**Exp_G&P : AT-PR-P#Grading**] : Determining grades of input text using AT-PR-P matrix.
⇒ We proved the possibility that AT-PR-P matrix can be used in grading system. First, the Ps of detected emotional words were used as basic grading values. The weighting function (13) is used to modify the initial P input values.
[**Exp_G&SV : AT-PR-SV#Grading**] : Determining grades of input text using AT-PR-SV matrix.
⇒ We proved the possibility that AT-PR-SV matrix can be

Exp \ F1 score	1-Fold	2-Fold	3-Fold	4-Fold	5-Fold	6-Fold	7-Fold	8-Fold	9-Fold	10-Fold
Exp_A1	.852	.832	.83	.819	.802	.824	.832	.838	.83	.84
Exp_A2	.92	.913	.912	.89	.859	.88	.894	.905	.894	.91
Exp_G&P	.86	.845	.832	.83	.819	.83	.85	.87	.85	.86
Exp_G&SV	.94	.935	.935	.91	.905	.895	.90	.905	.88	.92
Evaluation$_{human}$ - **Exp_A3**	.945	.928	.831	.742	.715	.795	.829	.878	.921	.951

TABLE II: Experimental results

used in grading system. First, the SVs of detected emotional words were used as basic grading values. The weighting function (14) is used to modify the initial SV input values.

Exemplifications of Sample Representations	
Author's Rating	★★☆☆☆
Extracted Matrix	(11 7 -0.8751698 -2.629185)
(AT-PR-P$_{\omega}^{i,j}$)-SV$_{\omega}^{i,j}$)	(16 7 -0.8751698 -2.951309)
Context of Sample Phrase	***cilwu/nca ha/xpv ko/ecx***$_{conj}$ ithallia/nq ey/jca tayha/pv n/exm ***cisik/nc i/jc eps/pa*** ese/ecs kuleh/pa nci/ecs
Translation	it is ***boring*** and maybe we ***don't have knowledge*** on Italy
Author's Rating	★★☆☆☆
Extracted Matrix	(16 12 -0.795664 -2.696191)
(AT-PR-P$_{\omega}^{i,j}$)-SV$_{\omega}^{i,j}$)	(3 2 -0.9626202 -2.487651)
Context of Sample Phrase	penyek/nc ul/jc ***cal/a mosha/px n/exm*** key/nb i/jcp nci/ecs maintu/nc mayp/nc silcen/nc pwupwun/nc i/jc eps/pa m/exn i/jc ***aswiwum/nc***
Translation	it is maybe ***bad translation***. ***I'm afraid that*** i have no mind map
Author's Rating	★★☆☆☆
Extracted Matrix	(2 6 0.892 -4.493405)
(AT-PR-P$_{\omega}^{i,j}$)-SV$_{\omega}^{i,j}$)	(16 9 -0.8424561 -3.016456)
Context of Sample Phrase	***nemwu/a kitay/nca ha/xpv*** esses/efp ***na/ecs***$_{conj}$ ***pyello/nc*** i/jcp ess/efp m/exn ./s.
Translation	Did i ***expect too much***? It was ***not good***.
Author's Rating	★★★☆☆
Extracted Matrix	(16 2 -0.9626202 -1.521469)
(AT-PR-P$_{\omega}^{i,j}$)-SV$_{weighted}^{i,j}$)	(16 4 0.9266376 3.317872)
Context of Sample Phrase	***kantan/ncs ha/xpa ciman/ecs***$_{conj}$ al/pv nun/exm kwukki/nc ka/jc nao/pv myen/ecs ***hungmiiss/pv*** e/ecx ha/px pnita/ef
Translation	It's ***naive but*** it's ***interesting*** whenever a flag which is known is refered

<*Additional Experiments*> : Additional experiments were conducted for comparison with the main experiments. The first two additional experiments represented a counter method of polarity prediction in previous studies and the last experiment was an intuitive evaluation of human participants.

[**Exp_A1**] : Determining polarity through Delta TFIDF weights.

⇒ TFIDF weights simply represented the statistical significance, not the polarity. Colloquial texts often consist of only 40 character long documents. This created problems with TFIDF if a word appeared only once per document. Thus, Delta TFIDF [14] was used to determine the polarity of a word. Its function was inversely modified to better suit our experiment.

[**Exp_A2**] : Determining polarity through SVM classifier estimations in ARSSA.

⇒ ARSSA used all of the features, including the words, conjunctions, negatives, negators, and syntactic structures, to calculate values of each text and automatically determine a grade of each value with trained data.

[**Exp_A3 - Human Evaluation**] : Human Evaluation which is entirely dependent upon the rater's intuition.

⇒ All the experiments above were compared with human test data. We asked three subjects to give a grade to each review on a 1 to 10 discrete scale without additional information of the semantic values or the original grades. Fleiss' kappa between the three raters was $\kappa = -0.154$.

2) Experimental Results and Discussions: We conducted the experiments both ways, using ROC curves for polarity prediction test and F1 scores for grade prediction test as the evaluation measures of analysis. Some patterns of the experiments showed more consistent than the human evaluation.

Polarity prediction tests (Exp_P&P, Exp_P&SV) with CWM in figure 5 returned highly accurate AUC values. This shows that these features are adequate for polarity prediction and the criterion has a discrimination sensitivity.

The results of the four experiments in table 2 were generally better than those from previous works [13], [15] due to the style of book reviews being relatively simple when compared to other types of reviews, such as movie reviews. Sarcasm, irony, or requisite world knowledge were significantly less prevalent in this domain when compared to movie reviews. Exp_A1 returned an average F1 score of 0.8258. This result was considered quite respectable since it relied on a wholly statistical approach without any regard for any other features or considerations. Exp_A2 returned an F1 score of 0.897 on average. Some rules of the system and the trained classifier control the calculation. It can be accepted that the SV(only) can be successfully adapted to application systems as semantic feature weights. Exp_G&P returned an average F1 score of 0.8446. Ps also have a certain discriminating power for grading because Ps are hierarchically weighted by shrinking factor ω derived from PR, but the global weighting is limited to the grade prediction in a degree. Exp_G&SV returned an average F1 score of 0.9125 on average. This result is much higher than previous experiment's score.

This result proved the possibility of semantic computing using the EWD across different domains for sentiment analysis and the applicability of term weighting results. This approach was more robust than human evaluation and is guaranteed to be a useful resource for NLP. However, we were also met with two potential pitfalls at the limit line of natural language, i.e. pragmatic or idiomatic expressions which were mentioned above and disambiguation in a retrieving process. How can a natural language system using computation method

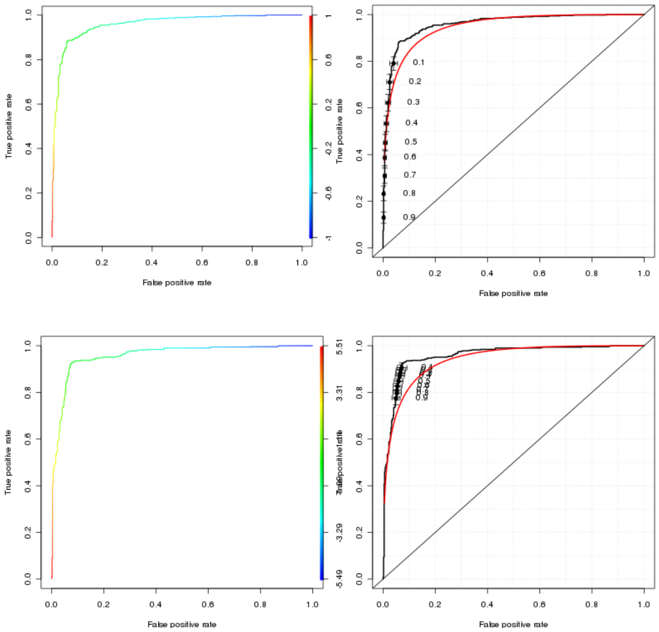

Fig. 5: ROC(Receiver Operating Characteristic) curve of Exp_P&P(top row) and Exp_P&SV(bottom row) : The confidence interval which is calculated by bootstrapping the observations and prediction is *True* and the number of bootstrap samples is *100*. AUC(area under the ROC curve)s of Exp_P&P(top row) and Exp_P&SV are 0.9587494 and 0.957363.

semantically overcome this kind of problem? [17] showed a computational dictionary for idioms (Phraseo-Lex) which contained the notion that *partially compositional idioms* consist of both meaningful and meaningless components. The meaningful components in these idioms can inspire the methodological reusability of semantic computing.

Then, how can one disambiguate emotional word senses in a retrieving process? We are frequently exposed to a number of words which share the same forms but have different meanings and usages. The basic forms of emotional words are identical at the morphological level. This is the limitation of the experiments above and therefore one must approach this problem from a different standpoint. We expect an NLP application adopting some form of WSD (word sense disambiguation) will help and have future plans for further research.

V. Conclusion and Future Work

Creating a dictionary for identifying sentiment orientation has recently been attempted at many research institutions in two main streams. One approach was to make a certain ontological structures using a category for emotions on their subjective basis. The other approach was to calculate some weights from statistical methods and estimate the polarity of text based on an objective basis. The former approach was paradoxically too categorical to understand the fuzziness of emotions and feelings while the latter was too statistical to reflect the human decision making process. We focused on how appropriately the two approaches could be combined for creating any emotional word dictionary in a subject-independent manner and how an extensible generalized model for multilingual dictionary creation could be built for reuse in semantic computing.

The semantic CWMs of the emotional word dictionary were created to address these problems using categorization and statistical methods. Validity and generality were proved through a series of experiments and we now conclude that the emotional word dictionary is able to be used as a basic feature set for other NLP applications involving analyzing or grading documents. We will continue to investigate the new applicability using the EWD to analyze the emotional expressions and represent the relations of web data and the boosting method for the values of the matrices. The language-specific side of lexical level will also be considered in our future work.

References

[1] J. Kamps, M. Marx, R.J. Mokken, M. de Rijke, *Using wordnet to measure semantic orientations of adjectives* In: Proceedings of LREC., 2004.
[2] B. Liu, M. Hu, J. Cheng, *Opinion observer: Analyzing and comparing opinions on the web*, In: Proceedings of WWW., 2005.
[3] S. Kim, E. Hovy, *Determining the sentiment of opinions*, In: Proceedings of COLING, pages 1367–1373, 2004.
[4] A. Esuli, F. Sebastiani, *SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining*, In Proceedings of LREC-2006, The fifth international conference on Language Resources and Evaluation, Genova, Italy, 24-25-26 MAY 2006.
[5] S. Baccianella, S. Esuli, F. Sebastiani, *SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining*, In: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), European Language Resources Association (ELRA), Valletta, Malta, 2010.
[6] C. Whitelaw, N. Garg, S. Argamon, *Using Appraisal Taxonomies for Sentiment Analysis* In: Paper session IR-8 (information retrieval): sentiment and genre classification, 2005.
[7] J.R. Martin, P.R.P. White, *The language of evaluation : appraisal in English*, Palgrave Macmillan, 2005.
[8] D. Geeraerts, *Prospects and Problems of Prototype Theory* In: Linguistics, 27:587-612, 1989.
[9] A.G. Miller, W.G. Charles, *Contextual correlates of semantic similarity* In: Language and Cognitive Processes, Volume 6, Number 1, 1991.
[10] A. Gliozzo, *Semantic Domains and Linguistic Theory*, In: Proceedings of the LREC 2006 workshop, 2006.
[11] H. Shin, *KOLON(the KOrean Lexicon mapped onto the Mikrokosmos ONtology): Mapping Korean Words onto the Mikrokosmos Ontology and Combining Lexical Resources*, In: Eoneohag, Volume 56:159-196, 2010.
[12] T.A. Wilson, *Fine-grained subjectivity and sentiment analysis: recognizing the intensity, polarity, and attitudes of private states*, Doctoral Dissertation, University of Pittsburgh, 2008.
[13] M. Ko, H. Shin, *Grading System of Movie Review through the Use of An Appraisal Dictionary and Computation of Semantic Segments* In: Korean Journal of Cognitive Science, Volume 21, Number 4., 2010.
[14] J. Martineau, T. Finin, *Delta TFIDF: An Improved Feature Space for Sentiment Analysis*, In: Third AAAI International Conference on Weblogs and Social Media, SanJose CA, 2009.
[15] M. Ko, *Automatic Classification and Grading System of Movie Reviews based on the Appraisal Dictionary* Master's Thesis, Seoul National University, 2010.
[16] A. Nourbakhsh, C. Khoo, J. Na, *A Framework for Sentiment Analysis of Political News Articles*, In: Paper presented at the annual meeting of the International Communication Association, TBA, Montreal, Quebec, Canada, May 22, 2008.
[17] R. Dormeyer, I. Fischer, *Building Lexicons out of a Database for Idioms*, In: Proceedings of the First International Conference on Language Resources and Evaluation (Granada/Spain), 1998.
[18] U. Choi, K. Kim, *Development of Korean Wordnet and Thesaurus* Natmal Corporation, 2010.
[19] http://www.cine21.com
[20] http://movie.naver.com
[21] http://clab.snu.ac.kr/arssa
[22] http://www.imdb.com
[23] http://www.aladin.co.kr