

# Developing Japanese WordNet Affect for Analyzing Emotions

Yoshimitsu Torii<sup>1</sup>   Dipankar Das<sup>2</sup>   Sivaji Bandyopadhyay<sup>2</sup>   Manabu Okumura<sup>1</sup>

<sup>1</sup>Precision and Intelligence Laboratory, Tokyo Institute of Technology, Japan

<sup>2</sup>Computer Science and Engineering Department, Jadavpur University, India

torii@lr.pi.titech.ac.jp, dipankar.dipnil2005@gmail.com

sivaji\_cse\_ju@yahoo.com, oku@pi.titech.ac.jp

## Abstract

This paper reports the development of Japanese *WordNet Affect* from the English *WordNet Affect* lists with the help of English *SentiWordNet* and Japanese *WordNet*. Expanding the available synsets of the English *WordNet Affect* using *SentiWordNet*, we have performed the translation of the expanded lists into Japanese based on the synsetIDs in the Japanese *WordNet*. A baseline system for emotion analysis of Japanese sentences has been developed based on the Japanese *WordNet Affect*. The incorporation of morphology improves the performance of the system. Overall, the system achieves average *precision*, *recall* and *F-scores* of 32.76%, 53% and 40.49% respectively on 89 sentences of the Japanese judgment corpus and 83.52%, 49.58% and 62.22% on 1000 translated Japanese sentences of the *SemEval 2007* affect sensing test corpus. Different experimental outcomes and morphological analysis suggest that irrespective of the google translation error, the performance of the system could be improved by enhancing the Japanese *WordNet Affect* in terms of coverage.

## 1 Introduction

Emotion analysis, a recent sub discipline at the crossroads of information retrieval (Sood *et al.*, 2009) and computational linguistics (Wiebe *et al.*, 2006) is becoming increasingly important from application view points of affective computing.

The majority of subjective analysis methods that are related to emotion is based on textual keywords spotting that use specific lexical resources. *SentiWordNet* (Baccianella *et al.*, 2010) is a lexical resource that assigns *positive*, *negative* and *objective* scores to each *WordNet* synset (Miller, 1995). Subjectivity wordlist (Banea *et al.*, 2008) assigns words with the strong or weak subjectivity and prior polarities of types *positive*, *negative* and *neutral*. Affective lexicon (Strapparava and Valitutti, 2004), one of the most efficient resources of emotion analysis, contains emotion words. To the best of our knowledge, these lexical resources have been created for English. A recent study shows that non-native English speakers support the growing use of the Internet<sup>1</sup>. Hence, there is a demand for automatic text analysis tools and linguistic resources for languages other than English.

In the present task, we have prepared the Japanese *WordNet Affect* from the already available English *WordNet Affect* (Strapparava and Valitutti, 2004). Entries in the English *WordNet Affect* are annotated using Ekman's (1993) six emotional categories (*joy*, *fear*, *anger*, *sadness*, *disgust*, *surprise*). The collection of the English *WordNet Affect*<sup>2</sup> synsets that are used in the present work was provided as a resource in the "Affective Text" shared task of *SemEval-2007* Workshop.

The six *WordNet Affect* lists that were provided in the shared task contain only 612 synsets in total with 1536 words. The words in each of the six emotion lists have been observed to be not more than 37.2% of the words present in the corresponding *SentiWordNet* synsets. Hence, these six lists are expanded with the synsets retrieved from the

<sup>1</sup> <http://www.internetworldstats.com/stats.htm>

<sup>2</sup> <http://www.cse.unt.edu/~rada/affectivetext/>

English *SentiWordNet* (Baccianella *et al.*, 2010). We assumed that the new sentiment bearing words in English *SentiWordNet* might have some emotional connotation in Japanese even keeping their part-of-speech (POS) information unchanged. The numbers of entries in the expanded word lists are increased by 69.77% and 74.60% at synset and word levels respectively. We have mapped the synsetID of the *WordNet Affect* lists with the synsetID of the *WordNet 3.0*<sup>3</sup>. This mapping helps in expanding the *WordNet Affect* lists with the recent version of *SentiWordNet 3.0*<sup>4</sup> as well as translating with the Japanese *WordNet* (Bond *et al.*, 2009). Some affect synsets (e.g., 00115193-a *huffy, mad, sore*) are not translated into Japanese as there are no equivalent synset in the Japanese *WordNet*.

Primarily, we have developed a baseline system based on the Japanese *WordNet Affect* and carried out the evaluation on a Japanese judgement corpus of 89 sentences. The system achieves the average *F-score* of 36.39% with respect to six emotion classes. We have also incorporated an open source Japanese morphological analyser<sup>5</sup>. The performance of the system has been increased by 4.1% in average *F-score* with respect to six emotion classes.

Scarcity of emotion corpus in Japanese motivated us to apply an open source google translator<sup>6</sup> to build the Japanese emotion corpus from the available English *SemEval-2007* affect sensing corpus. The baseline system based on the Japanese *WordNet Affect* achieves average *precision, recall* and *F-score* of 83.52%, 49.58% and 62.22% respectively on 1000 translated test sentences. The inclusion of morphological processing improves the performance of the system. Different experiments have been carried out by selecting different ranges of annotated emotion scores. Error analysis suggests that though the system performs satisfactorily in identifying the sentential emotions based on the available words of the Japanese *WordNet Affect*, the system suffers from the translated version of the corpus. In addition to that, the Japanese *WordNet Affect* also needs an improvement in terms of coverage.

The rest of the paper is organized as follows. Different developmental phases of the Japanese *WordNet Affect* are described in Section 3. Prepa-

ration of the translated Japanese corpus, different experiments and evaluations based on morphology and the annotated emotion scores are elaborated in Section 4. Finally Section 5 concludes the paper.

## 2 Related Works

The extraction and annotation of subjective terms started with machine learning approaches (Hatzivassiloglou and McKeown, 1997). Some well known sentiment lexicons have been developed, such as subjective adjective list (Baroni and Vegnaduzzo, 2004), English *SentiWordNet* (Esuli *et al.*, 2006), Taboada's adjective list (Voll and Taboada, 2007), SubjectivityWord List (Banea *et al.*, 2008) etc. Andreevskaia and Bergler (2006) present a method for extracting *positive* or *negative* sentiment bearing adjectives from *WordNet* using the Sentiment Tag Extraction Program (STEP). The proposed methods in (Wiebe and Riloff, 2006) automatically generate resources for subjectivity analysis for a new target language from the available resources for English. On the other hand, an automatically generated and scored sentiment lexicon, *SentiFul* (Neviarouskaya *et al.*, 2009), its expansion, morphological modifications and distinguishing sentiment features also shows the contributory results.

But, all of the above mentioned resources are in English and have been used in coarse grained sentiment analysis (e.g., *positive, negative* or *neutral*). The proposed method in (Takamura *et al.*, 2005) extracts semantic orientations from a small number of seed words with high accuracy in the experiments on English as well as Japanese lexicons. But, it was also aimed for sentiment bearing words. Instead of English *WordNet Affect* (Strapparava and Valitutti, 2004), there are a few attempts in other languages such as, Russian and Romanian (Bobicev *et al.*, 2010), Bengali (Das and Bandyopadhyay, 2010) etc. Our present approach is similar to some of these approaches but in contrast, we have evaluated our Japanese *WordNet Affect* on the *SemEval 2007* affect sensing corpus translated into Japanese. In recent trends, the application of mechanical turk for generating emotion lexicon (Mohammad and Turney, 2010) shows promising results. In the present task, we have incorporated the open source, available and accessible resources to achieve our goals.

<sup>3</sup> <http://wordnet.princeton.edu/wordnet/download/>

<sup>4</sup> <http://sentiwordnet.isti.cnr.it/>

<sup>5</sup> <http://mecab.sourceforge.net/>

<sup>6</sup> <http://translate.google.com/#>

### 3 Developmental Phases

#### 3.1 WordNet Affect

The English *WordNet Affect*, based on Ekman’s six emotion types is a small lexical resource compared to the complete *WordNet* but its affective annotation helps in emotion analysis. Some collection of *WordNet Affect* synsets was provided as a resource for the shared task of *Affective Text* in *SemEval-2007*. The whole data is provided in six files named by the six emotions. Each file contains a list of synsets and one synset per line. An example synset entry from *WordNet Affect* is as follows.

*a#00117872 angered enraged furious infuriated maddened*

The first letter of each line indicates the part of speech (POS) and is followed by the *affectID*. The representation was simple and easy for further processing. We have retrieved and linked the compatible *synsetID* from the recent version of *WordNet 3.0* with the *affectID* of the *WordNet Affect* synsets. We have searched each *WordNet Affect* synset in *WordNet 3.0*. If a matching *WordNet 3.0* synset is found, the *WordNet 3.0 synsetID* is mapped to the *WordNet Affect affectID*. The linking between two synsets of *WordNet Affect* and *WordNet 3.0* is shown in Figure 1.

<p><b>WordNet Affect:</b> <i>n#05587878 anger choler ire</i> <i>a#02336957 annoyed harassed harried pestered vexed</i></p> <p><b>WordNet:</b> <i>07516354-n anger, ire, choler</i> <i>02455845-a annoyed harassed harried pestered vexed</i></p> <p><b>Linked Synset ID with Affect ID:</b> <i>n#05587878</i> ↔ <i>07516354-n anger choler ire</i> <i>a#02336957</i> ↔ <i>02455845-a annoyed harassed harried pestered vexed</i></p>
--

Figure 1: Linking between the synsets of *WordNet Affect* and *WordNet*

#### 3.2 Expansion of WordNet Affect using SentiWordNet

It has been observed that the *WordNet Affect* contains fewer number of emotion word entries. The six lists provided in the *SemEval 2007* shared task contain only 612 synsets in total with 1536 words. The detail distribution of the emotion words as

well as the synsets in the six different lists according to their POS is shown in Table 1. Hence, we have expanded the lists with adequate number of emotion words using *SentiWordNet* before attempting any translation of the lists into Japanese. *SentiWordNet* assigns each synset of *WordNet* with two coarse grained subjective scores such as *positive* and *negative* along with an *objective* score. *SentiWordNet* contains more number of coarse grained emotional words than *WordNet Affect*. We assumed that the translation of the coarse grained emotional words into Japanese might contain more or less fine-grained emotion words. One example entry of the *SentiWordNet* is shown below. The POS of the entry is followed by a *synset ID*, *positive* and *negative* scores and synsets containing sentiment words.

*SentiWordNet:*

*a 121184 0.25 0.25 infuriated#a#1 furious#a#2 maddened#a#1 enraged#a#1 angered#a#1*

Our aim is to increase the number of emotion words in the *WordNet Affect* using *SentiWordNet*, both of which are developed from the *WordNet*. Hence, each word of the *WordNet Affect* is replaced by the equivalent synsets retrieved from *SentiWordNet* if the synset contains that emotion word. The POS information in the *WordNet Affect* is kept unchanged during expansion. A related example is shown in Figure 2. The distributions of expanded synsets and words for each of the six emotion classes based on four different POS types (*noun N*, *verb V*, *adjective Adj.* and *adverb Adv.*) are shown in Table 1. But, we have kept the duplicate entries at synset level for identifying the emotion related scores in our future attempts by utilizing the already associated *positive* and *negative* scores of *SentiWordNet*. The percentage of entries in the updated word lists are increased by 69.77 and 74.60 at synset and word levels.

#### 3.3 Translation of Expanded WordNet Affect into Japanese

We have mapped the *affectID* of the *WordNet Affect* to the corresponding *synsetID* of the *WordNet 3.0*. This mapping helps to expand the *WordNet Affect* with the recent version of *SentiWordNet 3.0* as well as translating the expanded lists into Japanese using the Japanese *WordNet* (Bond *et al.*, 2009).

Emotion Classes	WordNet Affect Synset (S) and Word (W) [After SentiWordNet updating]							
	N		V		Adj		Adv	
	S	W	S	W	S	W	S	W
Anger	48 [198]	99 [403]	19 [103]	64 [399]	39 [89]	120 [328]	21 [23]	35 [50]
Disgust	3 [17]	6 [21]	6 [21]	22 [62]	6 [38]	34 [230]	4 [5]	10 [19]
Fear	23 [89]	45 [224]	15 [48]	40 [243]	29 [62]	97 [261]	15 [21]	26 [49]
Joy	73 [375]	149 [761]	40 [252]	122 [727]	84 [194]	203 [616]	30 [45]	65 [133]
Sadness	32 [115]	64 [180]	10 [43]	33 [92]	55 [129]	169 [779]	26 [26]	43 [47]
Surprise	5 [31]	8 [28]	7 [42]	28 [205]	12 [33]	41 [164]	4 [6]	13 [28]

Table 1: Number of POS based Synsets and Words in six *WordNet Affect* lists before and after updating using *SentiWordNet*

<p><b>Linked Affect word:</b>  <i>n#05587878</i> ←→ <i>07516354-n anger cholera ire</i></p> <p><i>SentiWordNet</i> synsets containing “anger”:  <i>07516354-n anger, ire, cholera</i>  <i>14036539-n angriness, anger</i>  <i>00758972-n anger, ira, ire, wrath</i>  <i>01785971-v anger</i>  <i>01787106-v see_red, anger</i></p> <p><i>SentiWordNet</i> synsets containing “cholera”:  <i>07552729-n fretfulness, fussiness, crossness, petulance, peevishness, irritability, cholera</i>  <i>05406958-n cholera, yellow_bile</i></p> <p><b>Expanded Affect word:</b>  <i>n#05587878</i> ←→ <i>07516354-n anger cholera ire</i>  <i>14036539-n angriness anger</i> <i>00758972-n anger ira, ire wrath</i> <i>01785971-v anger</i>  ... <i>05406958-n cholera</i></p>
--

Figure 2: Expansion of *WordNet Affect* synset using *SentiWordNet*

As the Japanese *WordNet*<sup>7</sup> is freely available and it is being developed based on the English *WordNet*, the synsets of the expanded lists are automatically translated into Japanese equivalent synsets based on the *synsetIDs*. The number of translated Japanese words and synsets for six affect lists are shown in Table 2 and Table 3 respectively. The following are some translated samples that contain word as well as phrase level translations.

*07510348-n surprise* → 愕き, 驚き  
*07503260-n disgust* → むかつき, 嫌悪  
*07532440-n unhappiness, sadness* → 不仕合せさ, 哀情, 悲しみ, 不幸せさ, 不幸さ...

<sup>7</sup> <http://nlpwww.nict.go.jp/wn-ja/index.en.html>

*07527352-n joy, joyousness, joyfulness* → ジョイ, 愉楽, うれしいこと, 慶び, うれしさ, 歓び, 悦楽, 歓, 嬉しさ, 欣び, 楽しいこと...

Emotion Classes	Translated <i>WordNet Affect</i> list in Japanese (#Words)			
	N	V	Adj	Adv
Anger	861	501	231	9
Disgust	49	63	219	10
Fear	375	235	334	104
Joy	1959	1831	772	154
Sadness	533	307	575	39
Surprise	144	218	204	153

Table 2: Number of POS based translated word entries in six Japanese *WordNet Affect* lists

Emotion Classes	Japanese <i>WordNet Affect</i> list			
	Trans (#Syn)	Non-Trans (#Syn)	Translated Morphemes	
			(#W)	(#P)
Anger	254	159	1033	450
Disgust	57	24	218	97
Fear	146	74	615	315
Joy	628	238	2940	1273
Sadness	216	97	846	519
Surprise	112	25	456	216

Table 3: Number of translated (*Trans*) and non-translated (*Non-Trans*) synsets (*Syn*), words (*W*) and phrases (*P*) in six Japanese *WordNet Affects*.

### 3.4 Analyzing Translation Errors

Some *SentiWordNet* synsets (e.g., 00115193-a *huffy, mad, sore*) are not translated into Japanese as there are no equivalent synset entries in the Japanese *WordNet*. There were a large number of word combinations, collocations and idioms in the Japanese *WordNet Affect*. These parts of synsets show problems during translation and therefore manual

translation is carried out for these types. Some of the English synsets ('07517292-n *lividity*') were not translated into Japanese. But, an equivalent gloss of the word '*lividity*' that is present in the Japanese *WordNet* is "*a state of fury so great the face becomes discolored*". One of the reasons of such translation problems may be that no equivalent Japanese word sense is available for such English words.

## 4 Evaluation and Analysis

We have evaluated the lexical coverage of the developed Japanese *WordNet Affect* on a small emotional judgment corpus and *SemEval 2007* affect sensing corpus.

### 4.1 Evaluation on Judgment Corpus

The judgment corpus that is being developed by the Japan System Applications Co. Ltd.<sup>8</sup> contains only 100 sentences of emotional judgments. But, this corpus is not an open source till date. We have evaluated our Japanese *WordNet Affect* based baseline system on these 100 sentences and the results for each of the six emotion classes are shown in Table 4. We have also incorporated an open source morphological analyzer<sup>9</sup> in our baseline system.

The algorithm is that, if a word in a sentence is present in any of the Japanese *WordNet Affect* lists; the sentence is tagged with the emotion label corresponding to that affect list. But, if any word is not found in any of the six lists, each word of the sentence is passed through the morphological process to identify its root form which is searched through the Japanese *WordNet Affect* lists again. If the root form is found in any of the six Japanese *WordNet Affect* lists, the sentence is tagged accordingly. Otherwise, the sentence is tagged as non-emotional or *neutral*. The average *F-Score* of the baseline system has been improved by 4.1% with respect to the six emotion classes. Due to the fewer number of sentential instances in some emotion classes (e.g., *joy*, *sadness*, *surprise*), the performance of the system gives poor results even after including the morphological knowledge. One of the reasons may be the less number of words and synset entries in some *WordNet Affect* lists (e.g., *fear*). Hence, we have aimed to translate the Eng-

lish *SemEval 2007* affect sensing corpus into Japanese and evaluate our system on the translated corpus.

Emotion Classes (#Sentences)	Judgment Corpus (in %)		
	Before Morphology [ <i>After Morphology</i> ]		
	Precision	Recall	F-Score
<i>Anger</i> (#32)	51.61 [ <b>64.29</b> ]	50.00 [ <b>68.12</b> ]	50.79 [ <b>66.14</b> ]
<i>disgust</i> (#18)	25.00 [ <b>45.00</b> ]	5.56 [ <b>10.56</b> ]	9.09 [ <b>17.10</b> ]
<i>fear</i> (#33)	NULL		
<i>joy</i> (#3)	3.45 [ <b>8.08</b> ]	66.67 [ <b>100.00</b> ]	6.56 [ <b>14.95</b> ]
<i>Sadness</i> (#5)	NULL		
<i>surprise</i> (#9)	6.90 [ <b>13.69</b> ]	22.22 [ <b>33.33</b> ]	10.53 [ <b>19.41</b> ]

Table 4: Precision, Recall and F-Scores (in %) of the system per emotion class on the Judgment corpus by including and excluding morphology.

### 4.2 Evaluation on Translated SemEval 2007 Affect Sensing Corpus

The English *SemEval 2007* affect sensing corpus consists of news headlines only. Each of the news headlines is tagged with a valence score and scores for all the six Ekman's emotions. The six emotion scores for each sentence are in the range of 0 to 100. We have considered that each sentence is assigned a single sentential emotion tag based on the maximum emotion score out of six annotated emotion scores. We have used the Google translator API<sup>10</sup> to translate the 250 and 1000 sentences of the trial and test sets of the *SemEval 2007* corpus respectively. The experiments regarding morphology and emotion scores are conducted on the trial corpus. We have carried out different experiments on 1000 test sentences by selecting different ranges of emotion scores. The corresponding experimental results are also shown in Table 5. Incorporation of morphology improves the performance of the system. On the other hand, it is observed that the performance of the system decreases by increasing the range of Emotion Scores (ES). The reason may be that the numeric distribution of the sentential instances in each of the emotion classes decreases as the range in emotion scores increases.

<sup>8</sup> <http://www.jsa.co.jp/>

<sup>9</sup> <http://mecab.sourceforge.net/>

<sup>10</sup> <http://translate.google.com/#>

Emotion Classes	Japanese Translated <i>SemEval 2007</i> Test Corpus (in %)					
	Before Morphology [ <i>After Morphology</i> ]					
	Emotion Score (ES) $\geq 0$			Emotion Score (ES) $\geq 10$		
	Precision	Recall	F-Score	Precision	Recall	F-Score
<i>Anger</i>	61.01[ <b>68.75</b> ]	18.83[ <b>31.16</b> ]	28.78[ <b>42.88</b> ]	44.65[ <b>52.08</b> ]	25.54[ <b>33.32</b> ]	32.49[ <b>40.35</b> ]
<i>disgust</i>	79.55[ <b>85.05</b> ]	8.35[ <b>16.06</b> ]	15.12[ <b>27.01</b> ]	40.91[ <b>41.46</b> ]	9.89[ <b>18.07</b> ]	15.93[ <b>24.97</b> ]
<i>Fear</i>	93.42[ <b>95.45</b> ]	10.26[ <b>16.77</b> ]	18.49[ <b>28.52</b> ]	77.63[ <b>81.82</b> ]	13.32[ <b>21.42</b> ]	22.74[ <b>34.03</b> ]
<i>Joy</i>	69.07[ <b>72.68</b> ]	57.03[ <b>80.30</b> ]	62.48[ <b>76.29</b> ]	53.89[ <b>55.61</b> ]	56.50[ <b>96.22</b> ]	55.17[ <b>70.40</b> ]
<i>sadness</i>	83.33[ <b>84.29</b> ]	10.58[ <b>19.54</b> ]	18.77[ <b>31.67</b> ]	67.78[ <b>69.87</b> ]	11.78[ <b>19.88</b> ]	20.07[ <b>30.86</b> ]
<i>surprise</i>	94.94[ <b>94.94</b> ]	7.84[ <b>13.65</b> ]	14.48[ <b>23.99</b> ]	72.15[ <b>74.58</b> ]	8.25[ <b>15.87</b> ]	14.81[ <b>26.30</b> ]
Emotion Score (ES) $\geq 30$			Emotion Score (ES) $\geq 50$			
<i>Anger</i>	21.38[ <b>28.12</b> ]	39.08[ <b>62.45</b> ]	27.64[ <b>38.59</b> ]	6.92[ <b>10.42</b> ]	57.89[ <b>78.02</b> ]	12.36[ <b>18.26</b> ]
<i>disgust</i>	2.27[ <b>5.04</b> ]	3.70[ <b>6.72</b> ]	2.82[ <b>6.15</b> ]	NIL	NIL	NIL
<i>Fear</i>	44.74[ <b>56.82</b> ]	16.67[ <b>28.76</b> ]	24.29[ <b>38.45</b> ]	21.05[ <b>29.55</b> ]	17.98[ <b>31.26</b> ]	19.39[ <b>30.79</b> ]
<i>Joy</i>	31.48[ <b>33.42</b> ]	56.86[ <b>97.08</b> ]	40.52[ <b>50.53</b> ]	12.04[ <b>24.98</b> ]	61.32[ <b>87.66</b> ]	20.12[ <b>39.10</b> ]
<i>sadness</i>	37.78[ <b>69.86</b> ]	15.60[ <b>25.31</b> ]	22.08[ <b>37.22</b> ]	13.33[ <b>23.07</b> ]	12.12[ <b>22.57</b> ]	12.70[ <b>18.71</b> ]
<i>surprise</i>	17.72[ <b>20.34</b> ]	8.14[ <b>18.56</b> ]	11.16[ <b>20.35</b> ]	3.80[ <b>8.50</b> ]	7.50[ <b>12.50</b> ]	5.04[ <b>10.11</b> ]

Table 6: Precision, Recall and F-Scores (in %) of the system per emotion class on the translated Japanese *SemEval 2007* test corpus before and after including morphology on different ranges of Emotion Scores.

### 4.3 Analysis of Morphology

Japanese affect lists include words as well as phrases. We deal with phrases using Japanese morphology tool to find affect words in a sentence and substitute an affect word into its original conjugated form. One of the main reasons of using a morphology tool is to analyze the conjugated form and to identify the phrases. For example, the Japanese word for the equivalent English word ‘*anger*’ is “怒る (*o ko ru*)” but there are other conjugated word forms such as “怒った(*o ko tta*)” that means ‘*angered*’ and it is used in past tense. Similarly, other conjugated form “怒っていた (*o ko tte i ta*)” denotes the past participle form ‘*have angered*’ of the original word ‘*anger*’. The morphological form of its passive sense is “怒られる (*o ko ra re ru*)” that means ‘*be angered*’. We identify the word forms from their corresponding phrases by using the morpheme information. For example, the phrase “怒られる (*o ko ra re ru*)” consists of two words, one is “怒ら (*o ko ra*) that is in an imperfective form and other word is “れる (*re ru*) which is in an original form. The original form of the imperfective word 怒ら (*o ko ra*) is “怒る (*o ko ru*)”. It has been found that some of the English multi-word phrases have no equivalent Japanese phrase available. Only the equivalent Japanese words are found in Japanese *WordNet*. For exam

ple, the following synset contains a multi-word phrase ‘*see-red*’. Instead of any equivalent phrases, only words are found in Japanese *WordNet*. *01787106-v anger, see -red* → 怒る, 憤る, 立腹

## 5 Conclusion

The present paper describes the preparation of Japanese *WordNet Affect* containing six types of emotion words in six separate lists. The automatic approach of expanding, translating and sense disambiguation tasks reduces the manual effort. The resource is still being updated with more number of emotional words to increase the coverage. The sense disambiguation task needs to be improved further in future by incorporating more number of translators and considering their agreement into account. In future we will adopt a corpus-driven approach for updating the resource with more number of emotion words and phrases for extending the emotion analysis task in Japanese.

## Acknowledgments

The work reported in this paper is supported by a grant from the India-Japan Cooperative Programme (DST-JST) 2009 Research project entitled “Sentiment Analysis where AI meets Psychology” funded by Department of Science and Technology (DST), Government of India.

## References

- Andreevskaia A. and Bergler Sabine. 2007. CLaC and CLaC-NB: Knowledge-based and corpus-based approaches to sentiment tagging. *4th International Workshop on Semantic Evaluations (SemEval-2007)*, pp. 117–120, Prague.
- Baccianella Stefano, Esuli Andrea and Sebastiani Fabrizio. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *In Proceedings of the 7th Conference on Language Resources and Evaluation*, pp. 2200-2204.
- Banea, Carmen, Mihalcea Rada, Wiebe Janyce. 2008. A Bootstrapping Method for Building Subjectivity Lexicons for Languages with Scarce Resources. *The Sixth International Conference on Language Resources and Evaluation (LREC 2008)*.
- Baroni M. and Vegnaduzzo S. 2004. Identifying subjective adjectives through web-based mutual information. *Proceedings of the German Conference on NLP*.
- Bobicev Victoria, Maxim Victoria, Prodan Tatiana, Burciu Natalia, Anghelus Victoria. 2010. Emotions in words: developing a multilingual WordNet-Affect. *CICLING 2010*.
- Bond, Francis, Hitoshi Isahara, Sanae Fujita, Kiyotaka Uchimoto, Takayuki Kuribayashi and Kyoko Kanaoka. 2009. Enhancing the Japanese WordNet. *7th Workshop on Asian Language Resources, ACL-IJCNLP 2009*, Singapore.
- Das Dipankar and Bandyopadhyay Sivaji. 2010. Developing Bengali WordNet Affect for Analyzing Emotion. *23rd International Conference on the Computer Processing of Oriental Languages (ICCPOL-2010)*, pp. 35-40, California, USA.
- Ekman Paul. 1992. An argument for basic emotions, *Cognition and Emotion*, 6(3-4):169-200.
- Esuli, Andrea. and Sebastiani, Fabrizio. 2006. SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining, *LREC*.
- Hatzivassiloglou V. and McKeown K. R. 1997. Predicting the semantic orientation of adjectives. *35th Annual Meeting of the ACL and the 8th Conference of the European Chapter of the ACL*, pp. 174–181.
- Miller, A. G. 1995. WordNet: a lexical database for English. *In Communications of the ACM*, vol. 38 (11), November, pp. 39-41.
- Mohammad, S. and Turney, P.D. 2010. Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. *Proceedings of the NAACL-HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, LA, California, 26-34.
- Neviarouskaya, Alena, Prendinger Helmut, and Ishizuka Mitsuru. 2009. SentiFul: Generating a Reliable Lexicon for Sentiment Analysis. *International Conference on Affective Computing and Intelligent Interaction (ACII'09)*, IEEE, pp. 363-368.
- Sood S. and Vasserman, L. 2009. ESSE: Exploring Mood on the Web. *3rd International AAAI Conference on Weblogs and Social Media (ICWSM) Data Challenge Workshop*.
- Strapparava Carlo and Valitutti, A. 2004. Wordnet-affect: an affective extension of wordnet, *In 4th International Conference on Language Resources and Evaluation*, pp. 1083-1086.
- Strapparava Carlo and Mihalcea Rada. 2007. SemEval-2007 Task 14: Affective Text. *45th Annual Meeting of Association for Computational Linguistics*.
- Takamura Hiroya, Inui Takashi, Okumura Manabu. 2005. Extracting Semantic Orientations of Words using Spin Model. *43rd Annual Meeting of the Association for Computational Linguistics*, pp.133-140.
- Voll, K. and M. Taboada. 2007. Not All Words are Created Equal: Extracting Semantic Orientation as a Function of Adjective Relevance. *In Proceedings of the 20th Australian Joint Conference on Artificial Intelligence*. pp. 337-346, Gold Coast, Australia.
- Wiebe Janyce and Riloff Ellen. 2006. Creating Subjective and Objective Sentence Classifiers from Unannotated Texts. *International Conference on Intelligent Text Processing and Computational Linguistics*, Mexico City, pp. 475–486.