Identifying and Classifying Subjective Claims

Namhee Kwon, Eduard Hovy, Liang Zhou
USC Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292
{nkwon,hovy,liangz}@isi.edu

Stuart W. Shulman
University of Pittsburgh
121 University Place, Suite 600
Pittsburgh, PA 15260
shulman@pitt.edu

ABSTRACT
To understand the subjective documents, for example, public comments on the government's proposed regulation, opinion identification and classification is required. Rather than classifying documents or sentences into binary polarities as in much previous work, we identify the main claim or assertion of the writer and classify it into the predefined classes of opinion (attitude) over the topic. For the classification of the claims, we automatically build a list of multi-word subjective expressions by extending a small set of seed words, using automatically generated paraphrases from machine translation corpus. Our supervised machine learning method shows significant improvement over the baseline both in identification and classification of claims.

Categories and Subject Descriptors
I.2.7 [**Natural Languages**]: Text analysis

General Terms
Algorithms, Experimentation, Languages

Keywords
Electronic Rulemaking, Public Comments, Text Annotation, Opinion

1. INTRODUCTION
U.S. regulatory agencies are required to announce the proposed regulation and respond to the public comments on it before finalizing the regulation. For some controversial regulations, a huge volume of comments from stakeholder interest groups, lobbyists, lawyers, and citizens are collected, and to read all comments is a big burden for rule-writers. This requires a supportive tool to summarize and classify comments or to search important and substantive comments. Since the comments are mainly opinion-oriented arguments about the regulation, the identification and classification of main subjective claims would help rule-writers to preview and summarize the comments.

There has been a resurgence of interest in subjectivity analysis recently, including detecting subjectivity and classifying polarities for a document, sentence, or phrase. Although much research has focused on finding subjective expressions and classifying them, the readers are interested in what the writer claims or insists about the topic in many applications. More specifically, rule-writers want to know if the comments support or oppose the regulation.

In this paper, instead of classifying the whole document or recognizing small pieces of subjective expressions, we identify conclusive sentences showing the author's attitude to the main topic and classify them to polar classes. This approach has many benefits: first, it can detect multiple opinions within a single document; second, we can focus on the main opinion of the document, showing the purpose of writing, rather than detecting all (unrelated) subjective expressions. We applied a supervised machine learning method to identify claims using sophisticated lexical and structural features and classify them by their attitude to the topic: *support*, *oppose*, and *propose a new idea*.

A second novel point of our work is in the polarity classification. We adopt the basic approach of much previous work to utilize the subjective expressions obtained from annotation or bootstrapping from the seed word lists. However, we automatically build multi-word subjective lists by extending a small set of seed words from the paraphrases automatically obtained from a machine translation corpus. It is important to note that each subjective phrase is a proper semantic unit of polar expression, rather than an arbitrary length of multiple words (*n*-gram).

Our experiment was performed on the public's comments about the Environmental Protection Agency (EPA)'s proposed emission standard rule on hazardous pollutants. This is an extension of our previous work on multidimensional text analysis [7], but focuses on classifying claims. We apply polarity classification after near-duplicate detection [21][22] over the public comments.

In the rest of this paper, we first briefly introduce previous work on opinion analysis. Next, we describe our main claim identification, and explain the claim classification. The classification includes the process of building a depository of polar expressions and machine learning model to integrate related features as well as subjective expressions. The experimental results and conclusion follow.

2. Related Work
Sentiment Analysis the goal to understand subjectivity expressed in text. Recently, much research has been done on sentiment analysis, including detecting subjectivity [18][23], identifying opinion strength or intensity [19], and classifying semantic orientation (positive or negative characteristics) of words [14], phrases [20], sentences [6][23], or documents [12][23].

While most research has concentrated on classification of polarity (e.g., "positive" and "negative", sometimes including "neutral" and "both"), Pang and Lee in [13] adopted the rating scales of

The 7th Annual International Conference on Digital Government Research '06, May 21–24, 2006, San Diego, CA, USA.

one to five "stars" for movie reviews, and Mishne attempted mood classification of blogs into labels such as "amused", "happy", "bored", and "depressed" [9].

Most approaches search for the subjective expressions appearing in text. Especially for the polarity classification of a sentence or a phrase, people search for clue instances of polar expression that are either obtained from hand annotation or extended from a seed list by checking word co-occurrences [14][23] or semantic relations defined in WordNet [6].

Since subjective clues are mostly defined at the word level, the more important problem is how to combine the individual polar clues in a phrase or a sentence. Yu and Hatzivassiloglou [23] used the average log-likelihood score of positive and negative word frequencies, and similarly Kim and Hovy in [6] applied the mean of the weights of polar expressions and also tried several heuristic rules including the multiplication of polarities (1 for positive and -1 for negative).

While these approaches focused only on subjective expressions, Wilson et al. [20] attempted to detect the *contextual polarity* of a phrase in context, by also using sentence, structure, and document information represented in 28 complicated features.

3. Main Claim Identification

In many applications, a single document contains multiple subjective expressions. Some of them are about the main topic of interests while others are not directly related to the main topic. Rather than focusing on all subjective expressions from documents, we attempt to find the central idea of the documents. Finding the main claim is comparable to the task of discovering the "thesis statement" in essays, defined as "the sentence that explicitly identifies the purpose of the paper or previews its main ideas" in [2], but our goal is, more specifically, to identify the main subjective claims of the document. This process is identical to the "root" identification of the argument structure analysis suggested in our previous work in [7].

Since multiple claims can exist, we define the problem as a binary classification of each sentence as either claim or not, given all sentences in a document. We use many syntactic and semantic features in a supervised machine learning framework as follows:

- **Unigram, bigram, word's lexeme**: Since the documents are from a single domain, frequent *n*-gram can help find claims. We also expect that these features detect popular patterns or (adverbial) phrases.
- **Subjectivity**: We count positive and negative words as defined in the General Inquirer [5].
- **Position**: Especially in well-written texts, the main claim is highly related to the position in the text. We indicate a position with three values:
    - *Relative paragraph position*: the position of paragraph that includes the given sentence. Since the number of paragraphs is different for each document, the position is represented as a relative position scaled to the interval [0,1].
    - *Sentence position in a paragraph*: the order of the sentence in a paragraph in the region of [1,*n*] where *n* is the total number of sentences in a paragraph.
    - *Relative sentence position in a paragraph*: the sentence order represented as a relative position in a paragraph, scaled to the interval [0,1].
- **Subhead**: From the sentence parse produced by Charniak's parser, the main predicate is obtained by taking the headword of the sentence[1]. For the parent of the predicate (verb phrase (VP) for verb predicate), the direct children nodes are selected and the sequence of headwords of each child node is used. E.g., for the sentence "I strongly urge you to withdraw the proposed rule and return to a path of requiring power plants to reduce their emissions of mercury by 90 percent by 2008", the subhead is identified as "urge you to".
- **Subtopic "policy"**: Since the arguments in this domain are about the proposed regulation, we assume the main claim will probably mention the regulation. We therefore create a separate process to identify the subtopic "policy", which focuses on a particular policy in the regulation. We implement an SVM classifier using lexical features including unigrams, bigrams, and synonyms of the word, and obtain an F-measure of 0.71 [7]. 11% of sentences whose topic is "policy" are indeed main claims in the training set.

4. Claim (Attitude) Classification

Having identified the main claims of the documents, we classify the claim into predefined classes of attitudes. In the following example excerpted from the training set, the sentences shown in *italics* are identified as claims and classified into their polarities:

> The previous use of cap and trade methods for SOX removal is a good idea, and will work under the new mercury regulations. *I support the cap and trade method, even if it may produce hot spots where more mercury settles ...*
>
> The current plan proposed by President Bush lacks toughness when preventing mercury to be emitted into our air. *The regulations set to 30% and 70% reduction by 2010 and 2018, respectively, are too lenient on power plants.* Since the maximum available control technology can reach 70 to 90% mercury removal from stack gases if the removal is done efficiently, the regulations should be stricter.

Instead of classifying claims into traditional polarities of positive, negative, neutral, we classify them in terms of their attitude to the topic, which we define as "support" or "oppose" for the given topic (regulation), as well as another class "propose a new idea" for claims not directly stating an opinion about the main topic but proposing some new idea, as shown in the following example:

> *We request that you extend the comment period either until June 30, or until 30 days after the completion and public availability of any new analysis, whichever is later.*

[1] The headword is determined by Collins's heuristic rules in http://people.csail.mit.edu/mcollins/papers/heads.

As a resource of polarity detection, we first build positive and negative word (phrase) lists using a small set of seed words. The process of obtaining and extending the seed list is described below.

4.1 Subjective Clues

Each word can be characterized with many attributes; one of them is positive and negative semantic orientation. For example, “good”, “honest”, and “happy” have positive orientation, and “bad”, “disturb”, “violence” have negative orientation. This semantic orientation is often used as a basis to recognize the polar opinion in sentence or document.

However, the semantic orientation is not always clear, since a single word can show different polarities depending on its sense in context. For instance, “economical” can be used in positive context, but sometimes it does not show a polarity (neutral). We therefore have to be careful in determining the semantic orientation of words. A large set of positive and negative words may introduce much noise (include ambiguous polar expressions); on the other hand, using fewer polar words may miss many useful expressions.

We start with a small set of relatively clear polar (positive and negative) expressions and extend the list to multi-word phrases containing the contextual polarity, to obtain less ambiguous and more comprehensive lists. As a seed word list, we obtain polar words defined in General Inquirer [5]. General Inquirer is a manually developed and publicly available dictionary defining various properties of words including positive and negative polarity. Among 8,720 lexicons (11,788 senses), 1,622 lexicons (1,915 senses) are annotated as positive and 1,992 lexicons (2,291 senses) as negative. Example positive and negative words are shown in Table 1.

Polarity	Example Words
Positive	absolve, abundance, accentuate, accept, approval, capability, celebrate, complement, expert, familiar, friendly, glad, graceful, harmony, healthy, merciful, miraculous
Negative	abhor, absence, accident, admonish, agitation, breakdown, bribe, cancel, chaotic, condemn, exterminate, fearsome, manipulate, mediocre, quarrel, resentful

Table 1. Sample positive and negative entries in General Inquirer

4.2 Paraphrases

Using the subjective clues described above, we next obtain multi-word subjective lists, since the presence of subjectivity and its polarity are often determined by the way the positive and negative clues are combined in context rather than by their frequency standing isolated as a single word. A simple example is a negation where the polarity is inverted, such as “not good”, “no doubt” and “nobody likes”. Sometimes, a word can work as a subjective clue or as an intensifier (e.g., “a great future” vs. “a great disappointment”). It is even harder to determine the polarity when the expression is long (e.g., “there is no reason to believe” vs. “there is no doubt”; “the beautiful background of the monument” vs. “the background of a stagnant economic condition”).

We extend the repository of positive and negative expressions using paraphrases obtained from a machine translation corpus. Our method to automatically construct a large domain-independent paraphrase collection is based on the assumption that two different phrases of the same meaning may have the same translation in a foreign language. Phrase-based Statistical Machine Translation (SMT) systems analyze large quantities of bilingual parallel texts in order to learn translational alignments between pairs of words and phrases in two languages [11]. The typical sentence-based translation models make word/phrase alignment decisions probabilistically by computing the optimal model parameters with application of the statistical estimation theory. This alignment process results in a corpus of word/phrase-aligned parallel sentences from which we can extract phrase pairs that are translations of each other.

We ran the alignment algorithm from [10] on a Chinese-English parallel corpus of 218 million English words, available from the Linguistic Data Consortium (LDC). Phrase pairs are extracted by following the method described in [11] where all contiguous phrase pairs having consistent alignments are extraction candidates. Using these pairs we built paraphrase sets by joining together all English phrases that have the same Chinese translation. We speculate the resulting paraphrase table is of high quality relying on the fact that both the alignment models and phrase extraction methods have been shown to produce very good results for SMT. (The output is available at [24].)

Given the paraphrase collection derived from word and phrase alignment as explained, we assign the polarity by bootstrapping from the positive and negative words in General Inquirer [5].

Using this list of extended polar expressions, we determine the contextual polarity of a given phrase or sentence without additional polarity annotation.

This extended repository seems similar to *n*-gram approach, but different in finding proper units (boundaries) of polar expression rather than the arbitrary length of *n*-grams. The extended list not only increases the coverage of the polar expressions but also detects the contextual polarity of phrases, so that we can improve the polarity classification by simply applying this paraphrase list at the phrase and at the sentence level.

4.3 Assigning Polarities to Paraphrases

We assign the polarity to the cluster of paraphrases (all phrases within a cluster share the same meaning) by referring the positive and negative words defined in General Inquirer [5] [2]. When one or more paraphrases in a cluster is annotated as positive and no paraphrase as negative in General Inquirer, the entire cluster is assigned as positive, and similarly for negative polarity. If both positive and negative paraphrases co-exist in one cluster, the

[2] We do not perform sense disambiguation. If at least one sense has the polarity of positive or negative, we assume the polarity for all senses of the lexeme. When a lexeme has both positive and negative polarities for different senses (15 entries found in total), we assign “neutral” to the lexeme.

cluster is defined as neutral[3]. For each classified cluster, the paraphrases composed of only stopwords are excluded.

Table 2 shows the example clusters annotated with polarity. The word in bold font denotes the lexeme in General Inquirer. All phrases in *Cluster 1* are extended from the word "help" defined in General Inquirer, and "we had" is removed because "we" and "had" are stopwords. We obtain 4,592 positive clusters (24,031 phrases) and 3,629 negative clusters (17,893 phrases) by the process.

Cluster1: *Positive*
service will not only assist | will not only facilitate | it would not only help | not only assists the | not only been useful | will not only enable | would not only help | will not only help | not only helps to | not only enables | it will help to | not only helps | not only would | not only help | will help to | have helped | would help | will help | can help | helps to | it helps | enables | ~~we had~~ | **helps**
Cluster 2: *Negative*
characterized by a lack of tangible progress | lacking in anything worth mentioning | two immediate concerns of the people | have little to write home about | nothing worth taking note | nothing worth mentioning | promoting it development | lacking in intensity | not worth mentioning | nor merits to speak | lack of progress | **inadequate**

Table 2. Sample paraphrase clusters with polarity

4.4 Classification Features

Although we have built opinion-holding paraphrases, we need even more features to classify the claim sentence in terms of opinion or attitude over the main topic (the given regulation in our experiment). Many syntactic and semantic features, as well as polar expressions we obtained, are used, which are integrated in a machine learning framework.

- **Positive & Negative words**: The positive and negative words defined in General Inquirer and their accumulated frequencies are used. For example, from the sentence of "We also oppose the proposal to allow toxic mercury credit trading.", "oppose" has negative polarity and "allow" and "credit" has positive polarity.
- **Positive & Negative phrases**: The positive and negative phrases extended from General Inquirer using paraphrases obtained from machine translation corpus are used. For example, "I believe there is no reason that the original stipulation to remove 90% of mercury by 2008 should be altered." does not contain any polar words defined in General Inquirer, but "there is no reason" is found from the polar phrase list. We find the longer polar expressions first and then consider shorter one for the remaining parts. For example, we count "evil forces" as a unit, but we do not count "evil" and "forces" separately although they are defined in our collection.
- **Main predicate**: The headword of the sentence is identified from the parsed data. In the above example sentence, "believe" is a main predicate.
- **FrameNet frame**: As a way to generalize the main predicate, we find possible frames of the main predicate defined in FrameNet [1]. For the main verb "believe", four possible frames of *Awareness*, *Certainty*, *Trust*, and *Religious_belief*, are obtained.
- **Subcategorization**: The parsing rule that expands the parent of the main predicate, verb phrase (VP) for a verb predicate, is obtained.
- **Unigram, bigram, trigram**: The traditional *n*-gram features are applied to find useful subjective expressions or topical information.

We use only the *positive* and *negative* words or phrases, and do not include clue words for the "propose a new idea" category. The decision for that class depends on lexical information of bigrams, trigrams, and main predicates, combined with positive and negative polarity.

4.5 Human Annotation

To provide training and evaluating material, manual annotation was performed on randomly selected documents from our domain. The documents were assigned to four coders; every document was annotated by at least two coders. However, because of low inter-coder agreement, we only included documents whose annotation by two particular coders showed high enough agreement. This resulted in 119 documents in total, of which 78 documents were annotated in parallel.

To assess the agreement between coders, we first compute Cohen's Kappa coefficient [4] that is generally used for assessing rater agreement as a chance-corrected measure.

$$Kappa = \frac{P(A) - P(E)}{1 - P(E)}$$

where *P(A)* is the agreement ratio and *P(E)* is the chance agreement.

As another measure for the inter-coder agreement, we compute F-measure from *precision* and *recall* when we consider one annotation as correct, which means the average ratio of matching instances over the annotated instances. F-measure is formulated as follows:

$$F - measure = \frac{\frac{m}{C1} * \frac{m}{C2} * 2}{\frac{m}{C1} + \frac{m}{C2}} = \frac{m * 2}{C1 + C2}$$

where *C1* and *C2* are the total instances identified by each coder respectively, and *m* is the number of matching instances.

[3] Since the automatically generated paraphrase collection contains errors, this simple binary decision for the cluster outperforms the probabilistic model using the frequencies of individual phrases.

Since our focus is recognizing the main claims from all other sentences, F-measure as in an information retrieval task can provide a valid measure for the agreement.

Category		Value
Number of files		78
Average number of sentences per file		30
Average number of words per file		761
Claim	Kappa	0.62
	F-measure	0.68
Opinion & Ratio (support/oppose/propose) (%)		7/59/34
Opinion	Kappa	0.80
	F-measure	0.80

Table 3. Human agreement

Table 3 shows the inter-coder agreement on identifying and classifying the main claims. The claim identification was difficult even for human, since some of documents were hard to understand due to complicated and professional contents or poorly-written structures.

To see if the classification categories are similar to the polarities of positive and negative in other previous work, we also created a separate annotation, into "positive", "negative", and "neutral", based on the contextual polarity of subjective expressions (not based on the opinion over the main topic). The comparison between these two different annotation styles is shown Table 4. The first two rows show the kappa agreement for each annotation style, and the last row shows the agreement between the two different annotation styles.

While it shows much overlap between "positive" and "support", and between "negative" and "oppose", the neutral expressions introduced ambiguity since many of neutral expressions were annotated as "support" or "oppose" the regulation.

Classification type	Kappa
Claim type (support/oppose/propose)	0.80
Polarity (positive/negative/neutral)	0.69
Claim type: Polarity	0.61

Table 4. Comparison on opinion categories

5. Support Vector Machine vs. Boosting

As a framework to interpolate diverse features, we performed and compared Support Vector Machine (SVM) classification and a boosting algorithm. SVM [17] is a machine learning method widely used in classification or regression problems. It finds a hyperplane that separates the positive and negative training examples with a maximum margin in the vector space. For a data set $\{(x_1,c_1), (x_2,c_2),...,(x_n,c_n)\}$ where the point x_i is a p-dimensional vector and the c_i is a class that x_i belongs to, the problem is to minimize $|w|$ subject to the following:

$$c_i(w \cdot x_i - b) \geq 1, \quad 1 \leq i \leq n$$

For testing, the decision is made by the following rule:

$$\hat{c} = \begin{cases} 1, & if \;\; w \cdot x + b \geq 0 \\ -1, & if \;\; w \cdot x + b \leq 0 \end{cases}$$

As another supervised machine learning method, we applied a boosting algorithm [8] implemented in BoosTexter [14]. Boosting is a meta-algorithm to perform supervised machine learning, incrementally improving the weak rule by iteration. BoosTexter combines many simple, moderately inaccurate rules into a single accurate rule, by sequential training where each rule is tweaked in favor of the instances misclassified by the preceding rules. BooxTexter supports multi-class classification problem allowing more than one class for an instance, with textual, discrete or continuous features.

6. Experiments

The experiment was conducted in two steps: claim identification and classification.

For claim identification, 10-fold cross validation on 119 annotated documents was done with SVM and BoosTexter using the features described in Section 4.4. We implemented a baseline system selecting the first sentence as "claim", which is a baseline often used in text summarization research.

Table 5 shows the performance of our system, compared with the baseline system and human agreement. F-measure based on human annotation is provided, where *Human* shows inter-human agreement and *Baseline*, *System1*, and *System2* show the agreement with one specific human coder. Our system using BoosTexter shows significant improvement over the baseline and valuable performance compared to human agreement.

Human	Baseline	System 1 (SVM)	System 2 (BoosTexter)
0.68	0.19	0.52	**0.55**

Table 5. Result (F-measure) on claim identification (p<0.01)

Next, for claim classification, 240 claim sentences were extracted from the 119 annotated documents and 5-fold cross validation was performed. Since most sentences were annotated as "oppose", the baseline was implemented by simply assigning "oppose" for all instances. Because of the high prevalence for "oppose", it is hard to train and evaluate; however, our system outperforms the baseline for both 2-classes and 3-classes (shown in Table 6).

Classes	Human	Baseline	System (BoosTexter)
2 (support/oppose)	1	0.86	**0.91**
3 (support/oppose/propose)	0.80	0.59	**0.67**

Table 6. Result (F-measure) on opinion classification (p<0.01)

7. Analysis

So far, we have shown that our system performs comparably to a human's understanding. Although readers want to recognize the main claim of text, the low agreement between human coders implies that the identification task is difficult. Our assumption about the difficulty is that the documents are poorly written. We plan to classify documents into structured or not, and then provide different analysis for each case in future work. Further, for well-formed documents, we will define a more detailed model including claim, request, and related reasons, and investigate a method to automatically identify them.

Our multi-word subjective clues automatically built from the machine translation corpus were good indicators of the polar classification. However, we still need to combine multi-word expressions in a sentence to determine the whole polarity. We are also interested in developing a sophisticated method to combine the polar phrases in a sentence or paragraph by investigating syntactic and structural information from text. Further, the subjective paraphrase collections can be utilized in various ways. It can be used for more refined classes of subjectivity, not limited to binary polarity, using various seed lists.

8. Conclusion

We have described a model to understand the subjective public comments on the government's proposed regulation. We identify the main subjective claims and classify them using lexical and syntactic features, showing significantly higher performance than the baseline. Especially, a large repository of subjective words and phrases is automatically built from the paraphrases, which can be applied to other domains or classifications of different categories. Our claim identification and classification provides clearer understanding of documents by detecting more important opinions in terms of the regulation. We will investigate a method to refine the classes of identification and to provide more flexible analysis in future work.

9. ACKNOWLEDGMENTS

The researchers wish to acknowledge the EPA for providing the datasets on which this is report is based. This work was supported by NSF grants IIS-0429293 and IIS-0429360.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

REFERENCES

[1] Baker, C.F., Fillmore, C.J., and Lowe, J.B. The Berkeley FrameNet Project. In *Proceedings of COLING-ACL*. Monteral, Canada, 1998.

[2] Burstein, J., Marcu, D., Andreyev, S., and Chodorow, M. Towards Automatic Classification of Discourse Elements in Essays. In *Proceedings of ACL-01*. Toulouse, France, 2001.

[3] Charniak, E. A Maximum-Entropy Inspired Parser. In *Proceedings of NAACL-2000*. Seattle, WA, 2000.

[4] Cohen, J. A Coefficient of Agreement for Nominal Scales. *Education and Psychological Measurement*. 43(6):37—46. 1960.

[5] General Inquirer. *http://www.wjh.harvard.edu/inquirer/*.

[6] Kim, S. and Hovy, E. Determining the Sentiment of Opinions. In *Proceedings of COLING*. Geneva, Switzerland, 2004.

[7] Kwon, N., Shulman, S., and Hovy, E. Multidimensional Text Analysis in eRulemaking. In *Proceedings of 7th National Conference on Digital Government Research*. San Diego, CA, 2006.

[8] Meir, R. and Ratsch, G. An Introduction to Boosting and Leveraging. *Advanced Lectures on Machine Learning*. Springer-Verlag New York, Inc., 2003.

[9] Mishne, G. Experiments with Mood Classification in Blog Posts. In *Proceedings of the 1st Workshop on Stylistic Analysis of Text for Information Access at SIGIR*. Salvador, Brazil, 2005.

[10] Och, F.J. and Ney, H. A Systematic Comparison of Various Statistical Alignment Models, *Computational Linguistics*, 29(1):19-51, 2003.

[11] Och, F.J. and Ney, H. The Alignment Template Approach to Statistical Machine Translation, *Computational Linguistics.* 30(4), 2004.

[12] Pang, B., Lee L., and Vaithayanathan, S. Thumbs up? Sentiment Classification using Machine Learning Techniques. In *Proceedings of EMNLP*. Philadelphia, PA, 2002.

[13] Pang, B. and Lee L. Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with respect to Rating Scales. In *Proceedings of ACL-05*. Ann Arbor, MI, 2005.

[14] Schapire, R. and Singer, Y. BoosTexter: A Boosting-Based System for Text Categorization. *Machine Learning*. 39(2/3):135-168, 2003.

[15] Shulman, S.W. e-Rulemaking: Issues in Current Research and Practice. *International Journal of Public Administration* 28: 621-641. 2005.

[16] Turney, P., and Littman, M. Measuring Praise and Criticism: Inference of Semantic Orientation from Association. *ACM Transactions of Information Systems (TOIS)*. 21(4):315-346, 2003.

[17] Vapnik, V. N. *The Nature of Statistical Learning Theory*. Springer, 1995.

[18] Wiebe, J., Wilson, T., Bruce, R, Bell, M., and Martin, M. Learning Subjective Language. *Computational Linguistics.* 30(3):277-308, 2004.

[19] Wilson, T., Wiebe, J., and Hwa, R. Just How Mad are You? Finding Strong and Weak Opinion Clauses. In *Proceedings of AAAI-04*. San Jose, CA, 2004.

[20] Wilson, T., Wiebe, J., and Hoffmann, P. Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In *Proceedings of HLT-EMNLP*. Vancouver, Canada. 2005.

[21] Yang, H. and Callan, J., Near Duplicate Detection for eRulemaking. In *Proceedings of the Sixth National Conference on Digital Government Research*, Atlanta, GA. 2005.

[22] Yang, H., Callan, J., and Shulman, S., Next Steps in Near-Duplicate Detection for eRulemaking. In *Proceedings of the 7th National Conference on Digital Government Research*, San Diego, CA. 2006.

[23] Yu, H. and Hatzivassiloglou, V. Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences. In *Proceedings of EMNLP-03*. Sapporo, Japan, 2003.

[24] Zhou, L., Lin, C., Munteanu, D.S., and Hovy, E. *http://www.isi.edu/~liangz/DEMO/PARA/*, 2006.