IPSJ Transactions on Bioinformatics Vol.6 29–34 (July 2013)

[DOI: 10.2197/ipsjtbio.6.29]

Database/Software Paper

NegFinder: A Web Service for Identifying Negation Signals and Their Scopes

Kazuki Fujikawa[1,a)] Kazuhiro Seki[1,b)] Kuniaki Uehara[1,c)]

Received: February 8, 2013, Accepted: April 23, 2013, Released: July 10, 2013

Abstract: More and more biomedical documents are digitally written and stored. To make the most of the rich resources, it is crucial to precisely locate the information pertinent to user's interests. An obstacle in finding information in natural language text is negations, which deny or reverse the meaning of a sentence. This is especially problematic in the biomedical domain since scientific findings and clinical records often contain negated expressions to state negative effects or the absence of symptoms. This paper reports on our work on a hybrid approach to negation identification combining statistical and heuristic approaches and describes an implementation of the approach, named NegFinder, as a Web service.

Keywords: scope boundaries, heuristics, supervised classification, API

1. Introduction

More and more biomedical documents, including academic articles and clinical records, are digitally written and stored, where it is important to accurately find documents and/or information pertinent to user's needs. One of the obstacles in finding information in natural language text (free text) is negated and uncertain expressions, which reverse or obscure the semantics of a sentence or clause. This is especially problematic in the biomedical and clinical domains since scientific findings or clinical records often include negated and/or uncertain expressions to state negative effects revealed by experiments or the absence of symptoms from medical examination, such as "hydroxylated estrogens do *not* activate cAMP/PKA" and "*no* fever." According to Szarvas et al. [1], 13.5% of the sentences in biological paper abstracts and 6.6% of the sentences in clinical records have negated expressions. Ignoring such expressions degrades the quality of information access and may lead to false conclusions. However, accurately identifying negated/uncertain expressions is not trivial. Negative words, such as "not," do not always make negated expressions and a negation scope may extend beyond typical phrase boundaries, such as a comma and adverb as in "The prior odds ratio (Oprior) is difficult to estimate because we do [not know all the true interactions, even for a small subset of proteins]." (PMID: 17615067), where the negation scope is indicated by square brackets. Given the importance and challenge of the problem, a number of studies have been made on the identification of negated/uncertain expressions, which would improve the performance of biomedical knowledge processing, including information retrieval, information extraction, and text data mining [2].

This study expands on the previous work [3] by incorporating syntactic information through manually constructed rules, and provides a Web service to allow the users to annotate negated expressions with their documents via its RESTful API.

The remainder of the paper is organized as follows: Section 2 summarizes the related work, and Section 3 describes our approach in detail. Section 4 presents the results of our experiments on three benchmark data sets. Section 5 describes the Web API of our system, NegFinder. Section 6 concludes the present paper with a brief summary.

2. Related Work

There is much work in the identification of negated and/or uncertain expressions. They can be roughly categorized into manually crafted rule-based approaches [4], [5], [6], [7] and supervised classification-based [8], [9], [10], [11]. The former looks at lexico-syntactic patterns peculiar to negated expressions so as to spot them. Exhaustively making such patterns or rules is costly and may require domain knowledge. The latter takes advantage of machine learning techniques and classifies whether each token of input text is a negated expression. The performance of the approach largely depends on the choice of a model, features, and the values of hyper parameters. Also, quality and size of training data make significant difference. In the following, we summarize a representative work for each category of the approaches.

Huang et al. [12] proposed an approach based on hand-crafted grammatical rules. They manually analyzed POS of negation and frequent tokens around negation in 30 radiology reports and defined heuristic rules to categorize negation patterns. Let us consider the following sentence as an example: "There is no evidence of cervical lymph node enlargement," whose parse tree is shown in **Fig. 1**. The negated phrase (NegdPhr) of this sentence is identified by the following procedure.

---

[1] Kobe University, Kobe, Hyogo 657–8501, Japan
[a)] fujikawa@ai.cs.kobe-u.ac.jp
[b)] seki@cs.kobe-u.ac.jp
[c)] uehara@kobe-u.ac.jp

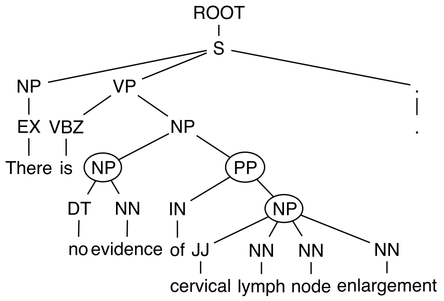


Fig. 1 A parse tree of “There is no evidence of cervical lymph node enlargement.”

(1) Locate the noun phrase (NP) with a head from a small set of nouns such as “evidence” and modified by a token “no,” “without”, or “absent.”
(2) Locate the prepositional phrase (PP) headed by “of” or “for” following the above NP.
(3) Extract the NP under the above PP, which contains the negated phrase (NegdPhr).

Morante et al. [3] developed a meta learning approach to identifying negation scopes. The approach takes a cascading, two-step classification procedure. In the first phase, an input sentence is compared against a list of negation signals: *absence, absent, fail, failure, impossible, lack, loss, miss, negative, neither, never, no, none, nor, not, unable*, and *without*. Tokens matching any negation signal in the list are classified as a negation signal. A supervised classifier is then applied to other tokens to classify each as the beginning, inside, or outside of a negation signal. As a classifier, Morante et al. used IGTree [13], instance-based learning algorithm with tree structure. An instance (token) was represented by the following features: word form, part-of-speech (POS), chunk IOB tags of one token to the left and to the right, and word form of the second token to the left and to the right. Here, a chunk is a group of adjacent words which function as a single grammatical unit (e.g., a noun phrase) in a sentence, and chunk IOB tags are three distinct tags (I, O, B), indicating the inside (I), outside (O), and beginning (B) of a chunk, respectively. These features were extracted from the output of the GENIA Tagger [14]. In the second phase, a pair of a negation signal and a token in the same sentence was treated as an instance. In other words, the number of instances in a sentence is the number of negation signals found in the sentence multiplied by the number of tokens in the sentence. Each instance (a token associated with a particular negation signal) was classified as the first token (F), last token (L), or neither of a negation scope.

In classification, three different classifiers using different algorithms, IGTree, support vector machine (SVM) [15], and conditional random field (CRF) [16], were employed. The feature set to represent an instance was similar to the ones used in identifying negation signals, including lemma, word forms, POS, and IOB tags, but more extensive. Then, a meta learner implemented by CRF assigned each instance with a class (first, last, or neither) using as features the classification results of the three classifies in addition to the similar set of features used by the classifiers. Finally, the scope of each negation signal was determined by a set of post-processing rules, such as “If one token is classified as F (first), and no token is classified as L (last), the negation scope is concluded between F and the end of the sentence.” These rules are needed since a classifier may predict none or multiple F’s and L’s for one negation signal.

Although both approaches were shown effective on real-world data, they both have limitations. A shortcoming of Morante et al. [3]’s approach is that negation scopes are identified based only on the limited context; only two tokens before or after the token in question are considered. According to our preliminary study, this sometimes leads to the grammatically odd solutions. Huang et al. [12]’s approach has limitation in adaptability. Their rules are extensive but may be effective only in limited domains since their rules contain domain-specific words, such as “mammographic” and “area.” It is laborious to manually create such extensive rules in different domains.

Junhui et al. [17] proposed another approach using simplified shallow semantic parsing. They regarded a negation signal as a predicate while mapping the negation scope into several constituents which are deemed as arguments of the negation signal. Their experiments on the BioScope corpus showed the effectiveness of the approach, achieving better Percentage of Correct Scope (PCS) scores than Morante et al. [3]. However, they focused only on negation scope identification without dealing with negation signal detection. It is known that slight decrease in perfomance in negation signal detection degrades the combined performance of the negation signal and scope identification [3]. Moreover, to make a complete negation scope identification system available to public as attempted in the present work, both negation signals and their scopes need to be considered.

3. A hybrid Approach

3.1 Overview

This section describes our proposed approach to negation identification combining supervised classification and parsing to address the issues mentioned in the previous section. The approach is composed of three phases: identification of negation signals, identification of negation scopes, and adjustment of negation scope. The first two phases are based on supervised classifiers, IGTree, similarly to Morante et al. [3], [11], and the last phase is based on a heuristic rule using grammatical parsing. Each phase is described in the following sections.

3.2 Identification of Negation Signals

The first step toward identifying negated expressions is to identify negation signals. Negation signals are words implying negation, such as “no” and “not.” There are roughly two approaches to the identification of negations signals, namely dictionary-based and supervised classification-based. A dictionary-based approach compiles a set of negation signals in advance and exhaustively searches an input text for the signals. On the other hand, a supervised classification-based approach uses training data annotated with negation signals and learns a model to identify negation signals based on a given learning algorithm. We adopt the latter because of its advantages over the former that no dictionary is necessary, which improves the applicability of the approach to other domains, and that the local context can be easily taken into account as features.

In classification, each token in an input is classified as the beginning of a negation signal (FIRST), inside (INSIDE), or outside (OUTSIDE). Considering the previous work, we used the following features to represent each instance (token). For each feature, an example for the third token "no" in a sentence, "there is no evidence of cervical lymph node enlargement," is presented in the parentheses.

- Raw word and root form. (no, no)
- POS and chunk IOB tag. (DT, B-NP)
- Root form, POS, and chunk IOB tag of one token to the left and to the right. (be, VBZ, B-VP, evidence, NN, I-NP)
- Root form of the second token to the left and to the right. (there, of)

These features can be extracted from publicly available NLP tools, such as the GENIA Tagger [14].

3.3 Identification of Negation Scopes

Each token in an input is paired with its nearest negation signal detected in the previous phase in the same sentence and forms an instance for this phase. Each instance is classified as the beginning of a negation scope (START), end of the scope (END), or neither (NEITHER). The feature set used to represent an instance follows Morante et al.'s work [3] and is summarized below. The sixth token "cervical" in the sentence, "There is no evidence of cervical lymph node enlargement," is used as an example below.

- Features regarding a detected negation signal
  - Raw word. A multi-word negation signal is hyphenated. (no)
  - The relative position (PRE, POST, or SAME) of the token in question with respect to the negation signal (POST).
  - Distance to the token in question counted as the number of words (3).
  - Whether or not the token is a negation signal (FALSE).
- Features regarding the token to be classified
  - Raw word and root form, POS, and chunk IOB tag (cervical, cervical, JJ, B-NP).
  - Root form, POS, and chunk IOB tag of one token to the left and to the right (of, IN, B-PP, lymph, NN, I-NP).
  - Root form of the second token to the left and to the right (evidence, node).
- Features regarding a chunk containing the token to be classified
  - The first and last token in the chunk (cervical, enlargement).
  - Sequence of the tokens in the chunk (cervical-lymph-node-enlargement).
  - Sequence of the POS tags in the chunk (JJ-NN-NN-NN).
  - The first and last token, hyphenated all tokens, and hyphenated all POS tags of two chunks to the left and two chunks to the right (of, of, of, IN; no, evidence, no-evidence, DT-NN). Note that there are only preceding chunks in this particular example.

3.4 Adjusting Negation Scope

The earlier two phases together could identify negation scope but suffer from the fact that they do not consider grammatical structure of input sentences. Our preliminary experiment revealed that the accuracy of scope identification is worse at the end (right-most boundary of the scope) than at the beginning (left-most boundary). In further analysis, it was found that the incorrectly identified right-most boundaries were often grammatically invalid (e.g., a boundary was located in the middle of a phrase). Given these observations, we adjust the end of a scope boundary considering the grammatical structure of the input sentence.

In essence, we locate the right-most boundary of a negation scope of a detected negation signal by tracing back the parse tree from the beginning of the scope (detected as "START" in the previous phase), such that the right-most boundary is the last (right-most) descendant node of the highest ancestor node which contains the beginning (START) as the first (left-most) descendant node. Algorithm 1 shows the pseudo-code for adjusting a negation scope, where Parent(x) and Children(x) functions return the parent and children of x, respectively.

Algorithm 1 Adjusting negation scope

```
Input: parse tree $T$, beginning of a negation scope $s$
Output: end of the negation scope $e$
$n \leftarrow s$
$C \leftarrow$ Children(Parent($n$))
while $n$ is the left-most node in $C$ do
  $n \leftarrow$ Parent($n$)
  $C \leftarrow$ Children(Parent($n$))
end while
while $C \neq \emptyset$ do
  $C \leftarrow$ Children($n$)
  $n \leftarrow$ right-most node of $C$
end while
$e \leftarrow n$
```

For illustration, **Fig. 2** shows the parse tree of a sentence, "PMA treatment and not retinoic acid treatment of the U937 cells acts in inducing NF-KB expression in the nuclei." (PMID: 1984449). In the parse tree, the correct negation scope is indicated by the dashed box, which is a sequence of the child nodes of the NP node indicated by the circle. The supervised classification approach described in the previous sections detects "not" as START and (incorrectly) "nuclei" as END of a negation scope.

The negation scope of the negation signal "not" (circled) in this sentence is located through the following procedure: First, we focus on the beginning of the scope, "not," detected in the previous steps. Then, we look at the child nodes of the parent node ("RB") of "not." As the children of "RB" is only "not" and thus the left-most, we shift our focus to the parent, "RB." By repeating these steps, we trace back to "NP" indicated by a circle. Note that the parent of the NP (which is also NP right under ROOT) no longer has the NP as its left-most child and is not being traced back. Since the circled NP's right-most leaf node is "cells," the token is identified as the end of the negation scope.

4. Performance

4.1 Experimental Settings

To demonstrate the effectiveness of the approach, we con-

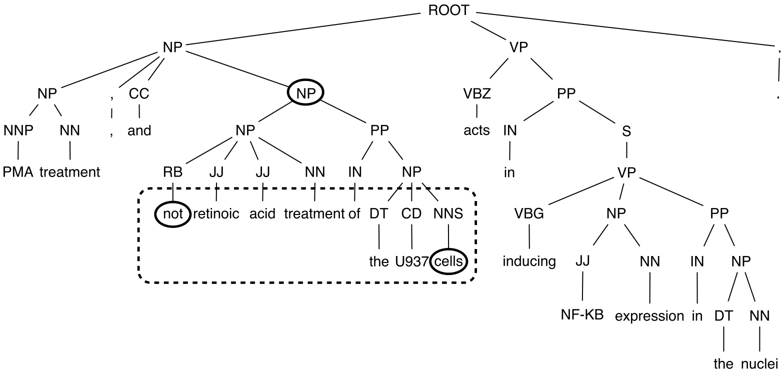


Fig. 2 An example parse tree illustrating the procedure to adjust scope boundary.

Table 1 Results of the identification of negation scopes.

Approach	Corpus	Precision	Recall	F-score
Baseline	Abstracts	79.5	71.5	75.3
	Full text	69.9	47.1	56.2
	Clinical	90.4	85.0	87.6
	Overall	79.9	67.9	73.0
Proposed	Abstracts	83.0	71.4	76.8
	Full text	73.0	54.5	62.4
	Clinical	89.9	82.3	85.9
	Overall	82.0	69.4	75.0

ducted experiments on the BioScope Corpus [1] in which each sentence is annotated with information about negation and speculation. This corpus consists of three data sets: 1,273 biological paper abstracts, 9 biological full-text, and 1,954 clinical free texts.

For identification of negation/uncertain signals and initial identification of the beginning of their scopes, our approach uses supervised classification and determines the end of the scopes by heuristics considering grammatical structure. As a classifier, we adopted IGTree [13]. To obtain the set of features of an instance and a parse tree of an input, we used GENIA Tagger [14] and Stanford Parser [18], respectively. We used recall, precision, and F-score as evaluation metrics both in negation/uncertain signal identification and scope identification.

4.2 Results

We applied the supervised classification-based approaches described in Sections 3.2 and 3.3 to to identify negation signals and their scopes, respectively. Subsequently, the negation scopes were adjusted as described in Section 3.4. **Table 1** summarizes the results, where the results of the classification-based approaches are shown as “baseline,” which roughly corresponds to the previous work [3] using supervised learning approaches. The “overall” results show macro-averaged precision, recall, and F-score of the results of the three data sets.

Except for the clinical domain, the performance (in F-score) of the negation scope identification improved for both abstracts and full text by 2.0% and 11%, respectively. Overall, the improvement was statistically significant at the $p = 0.01$ level by a randomized test [19]. Although the result indicates the effect of Algorithm 1, the decrease of the performance in the clinical domain should be noted. Our investigation suggested that it may be due to the large proportion of ungrammatical sentences in the clinical domain. In fact, only 35% of the sentences in clinical records have verbs as compared with 95% and 90% in the biomedical abstracts and full text, respectively. Algorithm 1 takes advantage of syntactic information and thus would not be effective for ungrammatical input. More in-depth analysis and discussion of the proposed approach are reported elsewhere [20].

It should be mentioned that Morante et al. [3] reported better F-scores than ours shown above except for the clinical domain. This is presumably due to the difference in performance for negation signal detection. They reported F-scores of 97.5, 96.1, and 98.8 for abstracts, full text, and clinical, respectively, whereas ours, based on Morante et al.’s approach, were found to be 89.8, 78.7, and 94.2. Although the difference needs to be investigated, detection of negative signals and their scopes are independent processes and thus one could expect similar improvement over stronger baseline resulted from improved negative signal detection.

5. RESTful API

Our negation identification system, named NegFinder, can be used via RESTful API[*1]. Through the API, users can specify input/output formats and the targeted domain (training data used for learning classification models) to accommodate their needs. A Web demo system[*2] is also provided to quickly test the system’s functionality as shown in **Fig. 3**. The following sections describe the request parameters and response fields of the API.

5.1 API Request Parameters

A request to the Web service is made as an HTTP URL in the following form:

`http://www.ai.cs.kobe-u.ac.jp/˜NegFinder/api/?PARAMETERS.`

As is standard in URLs, all parameters are separated using an ampersand (&) character. **Table 2** shows the list of parameters and their possible values.

Users can submit an input text with the `sentence` parameter or `src` parameter. The former receives the value as a sentence in which users would like to identify negation scopes, and the latter receives the contents of the specified file (URL) as input. Either parameter is required for a valid request. The `train_type` parameter allows one of the following training data,

---

[*1] http://www.ai.cs.kobe-u.ac.jp/˜NegFinder/

[*2] http://www.ai.cs.kobe-u.ac.jp/˜NegFinder/demo/

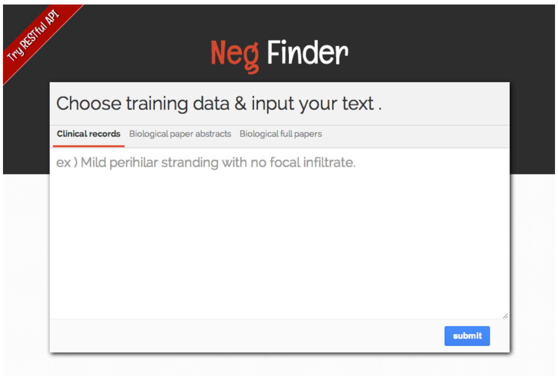


Fig. 3 A screenshot of the Web demo system.

Table 2 Request parameters of API.

Parameters	Descriptions
`sentence`	Specify an input sentence. Required unless `src` is provided.
`src`	Specify an URL for an input file. Required unless `sentence` is provided.
`train_type`	Specify one of following target domain, "`clinical_records`," "`abstracts`," or "`full_papers`"
`output_type`	Specify an output format, either "`json`" or "`text`."

```
{
  status: "OK",
  results: [
    {
      result_annotated_scope: "There is
        <neg_scope>no evidence of cervical lymph
        node enlargement</neg_scope> . ",
      result_annotated_signal: "There is
        <neg_signal>no</neg_signal> evidence
         of cervical lymph node enlargement . ",
      isNegation: true
    }
  ]
}
```

Fig. 4 An example of JSON response for "There is no evidence of cervical lymph node enlargement."

`clinical_records`, `abstracts`, or `full_papers`. If the parameter is not provided, the system uses `clinical_records` as the default. The `output_type` parameter allows either `json` or `text` as the output format.

5.2 API Response Fields

Figure 4 shows an example JSON response for "There is no evidence of cervical lymph node enlargement." It contains two root elements, `status` and `results`.

The `status` field contains the status of the request, and may contain debugging information to help users track down why the request failed. This field has four types of values: `OK`, `INVALID_REQUEST`, `INTERNAL_ERROR`, and `SRC_FILE_NOT_FOUND`. `OK` indicates that no error occurred and negation scopes were successfully identified. `INVALID_REQUEST` indicates that request parameters were not correct. A possible reason is that a required parameter is missing. `INTERNAL_ERROR` indicates that our server has a temporal problem. `SRC_FILE_NOT_FOUND` indicates that the system could not download user's requested file.

The `results` field has an array of sentences, each containing three types of information: `result_annotated_scope`, `result_annotated_signal`, and `isNegation`. The former two show the input sentence annotated with information about negation scopes and signals, respectively. The last type indicates whether the sentence has negated expressions.

6. Conclusion

This paper reported on our work to develop a hybrid approach to identifying the scope of negated and uncertain expressions by cascading supervised classification-based and grammatical rule-based approaches. Specifically, the rule took advantage of syntactic structure of an input sentence and adjusted the right-most boundary of a negation scope, which was difficult to identify by a classification approach alone with limited local context. Through the evaluative experiments on the BioScope Corpus composed of three data sets (i.e., abstracts, full text, and clinical records), it was shown that the performance of negation identification was improved on average. In addition, we implemented the system as a Web service for public use. Through the API, users can send an HTTP request to annotate their own text with negation signals and scopes and easily deploy the functionality to build a larger system.

While our Web service may be beneficial to researchers and practitioners, the target language is currently limited to English due to the fact that the approach is language-dependent. For future work, we would like to extend the service to other languages, specifically, Japanese by exploiting the Japanese clical corpus recently released for the NTCIR MedNLP task[*3].

Acknowledgments This project is partially supported by the Kayamori Foundation grant #K23-XVI-363.

References

[1] Szarvas, G., Vincze, V., Farkas, R. and Csirik, J.: The BioScope corpus: Annotation for negation, uncertainty and their scope in biomedical texts, *Proc. Workshop on Current Trends in Biomedical Natural Language Processing*, pp.38–45 (2008).

[2] Hersh, W.: *Information retrieval: A health and biomedical perspective*, Springer Verlag, New York (2009).

[3] Morante, R. and Daelemans, W.: A metalearning approach to processing the scope of negation, *Proc. 13th Conference on Computational Natural Language Learning*, pp.21–29 (2009).

[4] Apostolova, E., Tomuro, N. and Demner-Fushman, D.: Automatic extraction of lexico-syntactic patterns for detection of negation and speculation scopes, *Proc. 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Vol.2, pp.283–287 (2011).

[5] Ballesteros, M., Francisco, V., Díaz, A., Herrera, J. and Gervás, P.: Inferring the scope of negation in biomedical documents, *Proc. 13th Computational Linguistics and Intelligent Text Processing*, pp.363–375 (2012).

[6] Chapman, W., Bridewell, W., Hanbury, P., Cooper, G. and Buchanan, B.: A simple algorithm for identifying negated findings and diseases in discharge summaries, *Journal of Biomedical Informatics*, Vol.34, No.5, pp.301–310 (2001).

[7] Harkema, H., Dowling, J., Thornblade, T. and Chapman, W.: ConText: An algorithm for determining negation, experiencer, and temporal status from clinical reports, *Journal of Biomedical Informatics*, Vol.42, No.5, pp.839–851 (2009).

[8] Agarwal, S. and Yu, H.: Biomedical negation scope detection with conditional random fields, *Journal of the American Medical Informatics Association*, Vol.17, No.6, pp.696–701 (2010).

[9] Councill, I., McDonald, R. and Velikovich, L.: What's great and what's not: Learning to classify the scope of negation for improved sentiment analysis, *Proc. Workshop on Negation and Speculation in*

[*3] http://mednlp.jp

Natural Language Processing, pp.51–59 (2010).

[10] Cruz Díaz, N., Maña López, M., Vázquez, J. and Álvarez, V.: A machine-learning approach to negation and speculation detection in clinical texts, *J. Am. Soc. Inf. Sci. Technol.*, Vol.63, No.7, pp.1398–1410 (2012).

[11] Morante, R., Liekens, A. and Daelemans, W.: Learning the scope of negation in biomedical texts, *Proc. 2008 Conference on Empirical Methods in Natural Language Processing*, pp.715–724 (2008).

[12] Huang, Y. and Lowe, H.: A novel hybrid approach to automated negation detection in clinical radiology reports, *Journal of the American Medical Informatics Association*, Vol.14, No.3, pp.304–311 (2007).

[13] Daelemans, W., Van Den Bosch, A. and Weijters, T.: IGTree: Using trees for compression and classification in lazy learning algorithms, *Artificial intelligence review*, Vol.11, No.1-5, pp.407–423 (1997).

[14] Tsuruoka, Y., Tateishi, Y., Kim, J., Ohta, T., McNaught, J., Ananiadou, S. and Tsujii, J.: Developing a robust part-of-speech tagger for biomedical text, *Proc. 10th Panhellenic Conference on Informatics* (2005).

[15] Burges, C.J.C.: A tutorial on support vector machines for pattern recognition, *Data mining and knowledge discovery*, Vol.2, No.2, pp.121–167 (1998).

[16] Lafferty, J.D., McCallum, A. and Pereira, F.C.N.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data, *Proc. 18th International Conference on Machine Learning*, pp.282–289 (2001).

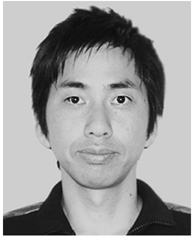
[17] Li, J., Zhou, G., Wang, H. and Zhu, Q.: Learning the Scope of Negation via Shallow Semantic Parsing, *Proc. 23rd International Conference on Computational Linguistics*, pp.671–679 (2010).

[18] Klein, D. and Manning, C.: Accurate unlexicalized parsing, *Proc. 41st Annual Meeting on Association for Computational Linguistics*, pp.423–430 (2003).

[19] Chinchor, N.: The statistical significance of the MUC-4 results, *Proc. 4th Conference on Message Understanding*, pp.30–50 (1992).

[20] Fujikawa, K., Seki, K. and Uehara, K.: A hybrid approach to finding negated and uncertain expressions in biomedical documents, *Proc. 2nd International Workshop on Managing Interoperability and CompleXity in Health Systems*, pp.67–74 (2012).

Kazuki Fujikawa was born in 1988. He received his B.E. in engineering from Kobe University. His research interests include machine learning, natural language processing, and data mining. He is currently an M.E. student in the Graduate School of System Infomatics, Kobe University.



Kazuhiro Seki received his Ph.D. in information science from Indiana University, Bloomington. His research interests are in the areas of natural language processing, information retrieval, machine learning, and their applications to intelligent information processing and management systems. He is currently an associate professor in the Graduate School of System Informatics at Kobe University, Japan.

Kuniaki Uehara received his B.E., M.E. and D.E. degrees in information and computer sciences from Osaka University, Japan. He was an assistant professor in the Institute of Scientific and Industrial Research at Osaka University and was a visiting assistant professor at Oregon State University. Currently, he is a professor in the Graduate School of System Informatics at Kobe University, Japan. His is conducting research in the areas of machine learning, data mining, and multimedia processing. He is a member of the Japanese Society for Artificial Intelligence, Japan Society for Software Science and Technology, and AAAI.

(Communicated by *Tetsuo Shibuya*)