International Journal of Database Theory and Application
Vol.10, No.7 (2017), pp.65-76
http://dx.doi.org/10.14257/ijdta.2017.10.7.06

A Study of Dictionary Based Korean Semantic Role Labeling

Mi-Ran Seok[3], Hye-Jeong Song[1,2], Chan-Young Park[1,2], Jong-Dae Kim[1,2] and Yu-Seop Kim[1,2,*]

[1] Department of Convergence Software, Hallym University, Chun-Cheon, Korea
[2] Bio-IT Research Center, Hallym University, Chun-Cheon, Korea
[3]IT Business DIV. Phill-IT Co. Ltd., An-Yang, Korea
Smr4880@phill-it.com, {hjsong, cypark, kimjd, yskim01}@hallym.ac.kr

Abstract

A semantic role is information used to clarify the role of entities in an event that a sentence describes, including agent, theme, experience, object, and location. Semantic role labeling (SRL) is a process that determines the semantic relation of a predicate and its arguments in a sentence and is an important factor in the semantic analysis of natural language processing, in addition to word sense disambiguation. To date, many manual semantic tagging tasks have been constructed; however, these tasks require a great deal of time and cost. To solve this problem, we propose a method for automatic SRL using frame files included in the Korean version of Proposition Bank (PropBank), which is one of the most widely used corpora. Frame files provide guidelines for PropBank annotators and include a list of framesets, which stand for a set of syntactic frames. First, we select the proper sense of the predicate from among multiple senses of the predicate in the frame files. Senses of the predicate are classified according to the semantic and syntactic properties of the predicate's arguments. We collect the nouns in a sample sentence of a given sense; we also collect all of the nouns that appear in a given sentence. The semantic similarities between the nouns from the sample sentence and the given sentence are measured and the sense with the highest similarity value is selected. The frame information of the selected sense is used for SRL of the given predicate and its arguments.

***Keywords**: Semantic Role Labeling, Natural Language Processing, Proposition Bank, Frame Files, Semantic Similarity*

1. Introduction

Semantic Role Labeling (SRL) is a task that automatically annotates the predicate-argument structure of a sentence with semantic roles [1]. Semantic parsing of sentences is believed to be an important task towards natural language understanding, with immediate applications in tasks such as information extraction and question answering [2].
To date, many manual semantic tagging tasks have been constructed; however, these tasks have required a great deal of time and cost. To solve this problem, we propose a method for automatic SRL using frame files included in the Korean version of Proposition Bank (PropBank), which is one of the most widely used corpora.

Frame files provide the guidelines for PropBank annotators and include a list of framesets, which represent a set of syntactic frames. This study uses semantic similarity to select a semantically suitable frame from several frames. Numerous studies related to semantic similarity have calculated the similarity between concepts using topological similarity. There are two approaches in topological similarity to measuring the semantic distance between concepts. The first approach evaluates similarity using the information

* Corresponding Author

ISSN: 2005-4270 IJDTA

content (called the node-based approach). The second approach evaluates similarity based on conceptual distance (called the edge-based approach) [3].

Information Content (IC) [4-7] is a measure of specificity for a concept. Higher values are associated with more specific concepts, whereas lower values are more general. IC is computed based on the frequency counts of concepts as found in a text corpus [8]. The edge-based approach [9-12] measures the minimal distance between concepts (synsets) in a hierarchical structure. For example, Resnik [13] presents edge-based measures that convert the minimal length between concepts *c1* and *c2*.

To select a frame, this study measured the semantic similarity between the given corpus nouns and sample sentence nouns of the frame using the weighted edge-based method. The mapping information of the selected frame was used for SRL of the given predicate and its arguments. Approximately 90% matched with the manual SRL.

2. Semantic Role Labeling and PropBank

2.1 Semantic Role Labeling

Semantic role labeling is a process that determines the semantic relation of a predicate and its arguments in a sentence. In this relation a word is specified as a predicate and a number of word groups are considered as arguments accompanying the predicate. Those arguments are assigned different semantic categories depending on the roles that they play with respect to the predicate [14]. Below example illustrates this labeling task.

- ✓ **Dan (SUBJ) broke [the laser pointer].**
- ✓ **[The windows (SUBJ)] were broken by the hurricane.**
- ✓ **[The vase (SUBJ)] broke into pieces when it toppled over.**

'SUBJ' means a subjective constituent syntactically and a constituent inside a rectangular parenthesis means an object 'broken'. As shown in the example, an object could be a subjective word or an objective word. SRL should find the relationship between a predicate 'break' and its arguments 'broken' objects in a sentence, regardless its syntactic structure.

[*John*] **threw** a ball to [*Mary*] in the [*park*].

Agent **Recipient** **Location**
or
Causer

The processing of the above sentence should result in the identification of a throwing event involving John as the Agent or Causer of the event, Mary as the Recipient and the ball as the item being thrown. The location of the throwing event, or where it took place, is the park. This description of the event specifies the conceptual relations that the referents of the noun phrases play with respect to the verb. [15]

The goal of SRL is to find semantic role words and expressions automatically. Semantic role labelers are commonly developed using a supervised learning paradigm where a classifier learns to predict role labels based on features extracted from annotated training data [16].

SRL provides a key knowledge that helps to build more elaborated document management and information extraction applications. This shallow semantic level of interpretation has additional uses outside of generalizing question answering, and semantic dialogue systems. One such application is in word sense disambiguation, where the roles associated with a word can be cues to its sense. [17]

2.2. PropBank

Propbank is a corpus where the arguments of each verb predicate are annotated with their thematic roles [18]. PropBank is an annotation of syntactically parsed, or treebanked, structures with "predicate-argument" structures [19]. To generalize the structure of the argument specialized in the predicate, PropBank expresses various arguments for various predicates with numbers and characters; the structure is additionally defined by separately categorizing different modifiers.

PropBank uses generic labels such as Arg0 and Arg1, as in the following [20]:

President Bush has approved duty-free treatment for imports of certain types of watches.

Relation (REL): approved

Arg0: President Bush

Arg1: duty-free treatment for imports of certain types of watches.

Arguments are numbered Arg0, Arg1, Arg2, as required, depending on the valance of the verb in question [21]. For a particular verb, Arg0 is typically the argument exhibiting the features of a Prototypical Agent, whereas Arg1 is a Prototypical Patient or Theme. No consistent generalizations can be made across verbs for the higher-numbered arguments [18].

The second task of PropBank annotation involves assigning functional tags to all modifiers of the verb, such as manner (MNR), locative (LOC), and temporal (TMP) [19]:

Mr. Bush met him privately, in the White House, on Thursday.

Rel: met

Arg0: Mr. Bush

Arg1: him

ArgM-MNR: privately

ArgM-LOC: in the White House

ArgM-TMP: on Thursday

These functional tags are derived from the labels used in TreeBank II [22] and include TMP (temporal modifier), LOC (location), DIR (direction), and MNR (manner). The TreeBank PRP (purpose) marker has been split into CAU (cause) and a true purpose marker. ADV is retained for sentence-level adverbs (e.g., "unfortunately") and an unavoidable residue of "other" modifiers. The category of DIS was added to mark discourse markers such as "also" and "on the other hand". ArgM is also used to explicitly mark modal verbs and negation particles [21].

Further, each predicate is annotated with its sense ID, alternatively referred to as a roleset or frameset ID [23]. PropBank annotation consists of the selection of a roleset, or a coarse-grained sense of the predicate. The following, for example, is the roleset for the verb fear [20]:

Arg0: entity afraid
Arg1: afraid of what?

PropBank contributes usage-based formulations of verb-argument structures, defining specific rolesets according to the general and idiosyncratic semantic behaviors of the predicates in the data [20].

2.2.1. Frame files

PropBank uses frame files for SRL. Frame files provide guidelines for Propbank annotators and include a list of framesets or coarse-grained senses of the verbs. A frameset represents a set of syntactic frames. A description for each frameset includes the list of verb specific roles and examples of different syntactic realizations of the verb [24].

Frame File for the verb “expect”:

Roles:

Arg0: expecter

Arg1: thing expected

Example: Transitive, active:

Portfolio managers expect further declines in interest rates.

Arg0: Portfolio managers

REL: expect

Arg1: further declines in interest rates

After examining different sentences in a corpus, the question that must be addressed is whether the verb has one or more framesets. The main principle for distinguishing framesets is that two verb meanings are distinguished as different framesets if they have distinct sub-categorization frames. For example, different framesets may have a different number of arguments. Alternatively, the number of arguments could be the same; however, the thematic roles could be critically different, as illustrated by the verb “draw” below:

Frameset draw.01 “art”

Arg0: artist

Arg1: art

Arg2: beneficiary

He (Arg0) was drawing diagrams and sketches (Arg1) for his patron (Arg2)

Frameset draw.02 “pull”

Arg0: puller

Arg1: thing pulled

Arg2: source

The campaign (Arg0) is drawing fire (Arg1) from anti-smoking advocates (Arg2)

Different senses are distinguished as different framesets only if they have distinct syntactic behavior, which correlates with different types of allowable arguments [24].

2.2.2. Korean PropBank

Korean Propbank Annotations are semantic annotations of the Korean English Treebank Annotations and Korean Treebank Version 2.0. Each verb and adjective occurring in the Treebank has been treated as a semantic predicate and the surrounding text has been annotated for arguments and adjuncts of the predicate. The verbs and adjectives have also been tagged with coarse-grained senses. This work was conducted in the Computer and Information Sciences Department at the University of Pennsylvania.

A basic component of the Korean Propbank is the Verb Lexicon. A frame file, consisting of one or more framesets, has been created for each predicate that occurs in the Treebank. These files serve as a reference for both the annotators and users of the data. In total, 2,749 such files have been created.

Figure 1 illustrates a sample Korean frameset. If there are multiple <id>s in the frame file, each <id> conveys a different meaning of the predicate. <edef> indicates the English expression for the predicate and <roleset> is the semantic role. There is more than one frame in a frameset and mapping information and examples are included in each frame. <Mapping> contains the information regarding the semantic role for the case. Text is provided in the example and illustrates the parsing information of the text. Furthermore, <relation> indicates what semantic role is being labeled against what words.

The predicate in Figure 1 is *Mat-da*, which means “be right” in English. This information described in Figure 1 shows that the predicate *Mat-da* has multiple meanings in their corpus. This case shows the second meaning, “be right” between <edef> and </edef>. The <roleset> tag says that the predicate *Mat-da* has only one main argument, which means “entity being right”. In the <mapitem>, the semantic role, when the argument related to the predicate *Mat-da* is subjective (sbj), is “ARG1”.

In the <example> tags, one example sentence and its syntactic parse are shown. Below the parse tree, all arguments related to the predicate are listed with their semantic roles. Mat-da in the sentence has four arguments including the main argument, *geu-geot*. Also the other three modification arguments are listed with their roles, ARGM-DIS, ARGM-CAU, and ARGM-DIS.

Below the arguments, the reflected form of Mat-da, which is used in this sentence, and the auxiliary verbal information are listed.

```
<frameset>
        <id>Mat.02</id>
        <edef>be right</edef>
        <roleset>
            <role argnum="1" argrole="entity being right"/>
        </roleset>
        <frame>
            <mapping>
               <rel>Mat-da</rel>
               <mapitem src="sbj" trg="arg1"/>
            </mapping>
            <example>
               <text> Geu-reo-chi-man in-bal-seo-neun cheol-jo-mang-e wi-jang-hae no-ki-ga a-ju swi-u-ni-kka geu-ge
a-ma ma-jeul geom-ni-da. </text>
               <parse>
(S (ADCP geu-reo-chi-man/ADC)
    (S (S-SBJ (NP-SBJ *pro*)
                 (VP (VP (NP-OBJ in-bal-seon/NNC+ eun/PAU)
                           (VP (NP-ADV cheol-jo-mang/NNC+ e/PAD)
                                (VP (VV wi-jang/NNC+ha/XSV+e/ECS))))
                      no/VX+ki/ENM+ga/PCA))
       (ADJP (ADVP a-ju/ADV)
                (ADJP swip/VJ+ u-ni-kka/ECS)))
    (S (NP-SBJ geu-geot/NPN+i/PCA)
       (VP (VP (ADVP a-ma/ADV)
                  (VP mat/VV+eul/EAN))
             (VX geot/NNX+i/CO+ seum-ni-da/EFN)))
   ./SFN)
               </parse>
               <relation>
                  <Arg n="M">
                     <f>DIS</f>
                     <term> geu-reo-chi-man </term>
                  </Arg>
                  <Arg n="M">
                     <f>CAU</f>
                     <term> in-bal-seo-neun cheol-jo-mang-e wi-jang-hae no-ki-ga a-ju swi-u-ni-kka </term>
                  </Arg>
                  <Arg n="1">
                     <f> i </f>
                     <term> geu-geot </term>
                  </Arg>
                  <Arg n="M">
                     <f>DIS</f>
                     <term> a-ma </term>
                  </Arg>
                  <Rel>
                     <term> ma-jeul </term>
                  </Rel>
                  <Aux>
                     <f> eul </f>
                     <term> geom-ni-da </term>
                  </Aux>
               </relation>
            </example>
        </frame>
      </frameset>
```

Figure 1. Second Frameset of 'Mat-da'

3. CoreNet

To calculate semantic similarity, CoreNet, which is the Korean concept-based lexical semantic network of BOLA (Bank of Language Resources)†, is used. The lexical semantic network is a network of the semantic relations of words and provides useful information in the natural language processing area. Specifically, it acts as an important knowledge base when processing lexical meanings. The name CoreNet is a combination of Corea (Korea) and the words network based on Korean. CoreNet is connected to 2,938 hierarchical concepts and 92,448 lexical items. The CoreNet Korean lexical semantic network systematically and generally provides semantic information with the goal of solving the semantic vagueness that has occurred in computer language processing. This study uses CBL1 (Korean noun) from CoreNet.

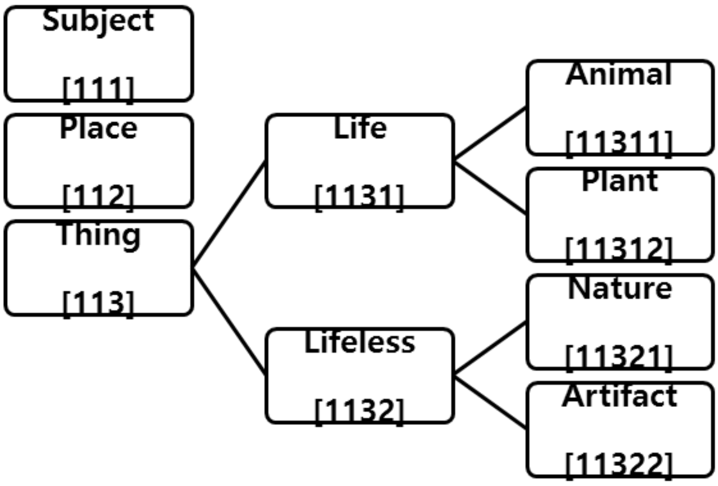

Figure 2. Hierarchy of CBL1

Figure 2 is the hierarchical structure of CBL1. Each number provides sequential information within the concept system, which is the information regarding a higher concept and location. The concept number digits indicate phases and the number whose final digit is deleted belongs to the higher concept. The class decreases as the class numbers move to the first decimal, meaning that two nouns are more similar.

The five-digit number 11311 in Figure 2 represents the Animal class, and indicates that it is a class in the fifth phase of the hierarchical structure. The higher concept is 1131, which represents the Life class, and the next highest class is 113, which represents Things. Subject and Place, which are on the same level as Thing, differ only in the last digit of the concept numbers.

4. Semantic Similarity based on the Weighted Edge Count

The principle of similarity computation is based on the edge-counting method, which is defined as follows. Given ontology Ω formed by a set of nodes and a root node (Root) (Figure 3), Term1 and Term2 represent two ontological elements for which similarity is to be calculated. The principle of similarity computation is based on distance (Depth1 and Depth2) from the Root; Depth separates nodes Term1 and Term2 from the closest common ancestor Common Node [9].

† http://semanticweb.kaist.ac.kr/org/bora/CoreNet_Project/

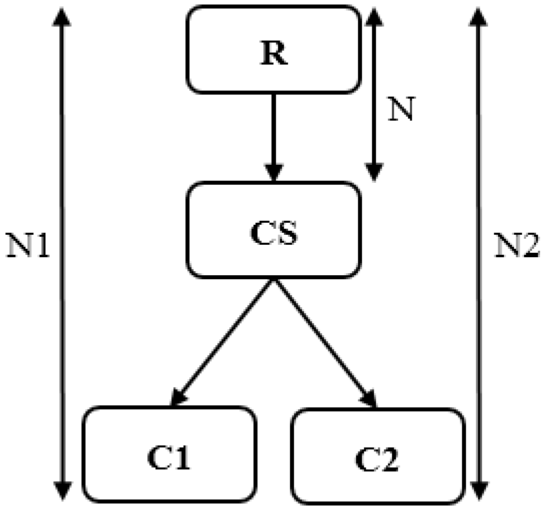

Figure 3. Example of Concept Hierarchy. C1 (Term1) and C2 (Term2) Are Ontological Elements and N1 (Depth1) and N2 (Depth2) Are Distance from Root

The similarity measure [25] is defined by the following expression:

$$Sim_{WP} = 2*N / (N1+N2)$$

The problem resulting from this measure is that the arcs in ontology represent uniform distances (*i.e.*, all the semantic links have the same weight) [26].

In this study, the length of the path between the two noun classes is calculated by setting the weight for each level after categorizing the nouns according to the semantic noun hierarchical structure. The weight of the edge between the highest class and the second highest class and between the second highest class and the third highest class is 1/2 and 1/4, respectively. Hence, the weight is calculated by decreasing the weight of the edge, from the higher class to the lower class, by half.

For example, calculating the distance between the nouns in Nature and Animal in Figure 2, the weight of the edge between Nature and Lifeless is 1/16, between Lifeless and Thing is 1/8, between Thing and Life is 1/8, and between Life and Animal is 1/16, totaling 0.375.

To calculate the similarity between the two items “Machine Civilization” and “Lion,” the result can be obtained by summating the weight of the edge between 11322 and 11311, considering that “Machine Civilization” is a member of “Artifact” and “Lion” is a member of “Animal”. Similarity increases when the summation of the edge weight decreases as the distance between the nouns is shorter.

5. Method

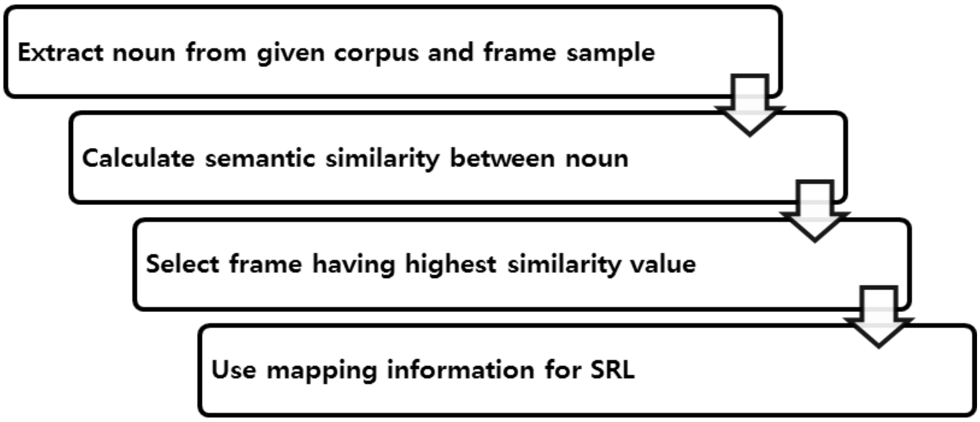

Figure 4 SRL Process

Figure 4 illustrates the process of the SRL method proposed in this study. To conduct SRL using frame files, mapping information must be retrieved by selecting several frames from the frame file; that is, selecting one sense out of several senses of a predicate. Senses of the predicate word are classified according to the semantic and syntactic properties of its arguments. Therefore, to calculate the semantic similarity between the predicate of the given corpus and each frame, we collect the nouns in a sample text of a frame and also collect all of the nouns that appear in a given corpus. The semantic similarities between nouns from the sample texts and a given corpus are measured and the frame with the highest similarity value is selected. For example, provided that the number of the noun corpus is *m* and the number of nouns in the first frame sample is *n*1, the semantic similarity of *m***n*1 must be measured to calculate the mean value. If the number of nouns in the second frame sample is *n*2, the mean value must be calculated by measuring the semantic similarity of *m***n*2. If the mean value of the similarity of the first frame is greater, select the first frame as the proper frame for the predicate; select the second frame as the proper predicate if the mean value of the similarity of the second frame is greater. The mapping information of the selected frame is used for SRL of the given predicate and its arguments. Mapping information is categorized into two groups: one category when the argument to be tagged with a semantic role is subjective, and the other category when the argument to be tagged with a semantic role is objective. A semantic role can also be tagged to the related proper argument.

6. Experiment and Result

When SRL is to be conducted with the following corpus, *Beom-mu-bu gwan-gye-ja-neun "Chin-bu-mo-ga mi-guk si-min-gwon-ja-il ttae-neun ja-nyeo-do i-jung-guk-jeok-ja-ga a-ni-ra mi-guk si-min-in ge* ***mat-da****"myeo, "Ttae-ttae-ro ho-jeo-gi jeong-ni an-dwae i-reon o-ryu-ga bi-jeo-ji-gi-do han-da"ra-go bak-hyeot-da,* against the argument of the predicate Mat-da, nouns of the sample frame text and corpus nouns must be retrieved.

There are ten noun corpora for the above: *Beom-mu-bu_Gwan-gy-eja* "officers of the Ministry of Justice", *Chin-bu-mo* "parents", *Mi-guk_Si-min-gwon-ja* "US citizen", *Ttae* "time", *Ja-nyeo* "children", *I-jung-guk-jeok-ja* "dual citizenship", *Mi-guk_si-min* "US citizen", *Ho-jeok* "family register", *Jeong-ri* "organization", and *O-ryu* "error".

From the first frameset of the frame file, the meaning of Mat-da is "conform" and there is only one frame. Nouns of the frame sample are: *Hu-bang* "back", *Bu-dae* "military", *Gi-gwan* "government office", *Je-dae* "discharge from military", *Jak-jeon* "military operation", *Gye-hoek* "plan", and *Pyeon-seong* "formation". To calculate the semantic similarity between the corpus and the nouns of the first frame example, the distance between nouns is calculated 70 times, and approximately 0.24 is obtained as an average value.

There is one frame in the second frameset (Figure 2), and the sample text nouns are *In-bal-seon* “impression line”, *Cheol-jo-mang* “wire fence”, *Wi-jang* “stomach”, *Geu-geot* “it”, and *Geot* “thing”. The semantic similarity is calculated 50 times. The average value of the distance between nouns is approximately 0.07.

The meaning of *Mat-da* in the third frameset is “hit”, whose frame number is three. The average distance between nouns of the first frame is approximately 0.10, that of the second frame is approximately 0.14, and that of the third frame is approximately 0.10.

The meaning of the fourth frameset is “confront” and it contains only one frame. The average distance between nouns of the frame example and the given corpus is approximately 0.14. Accordingly, the frame in the second frameset, whose average distance between the nouns is selected as the sense of the predicate, and the related mapping information (<mapitem src=“sbj” trg=“arg1”/>) are extracted.

As there is information regarding the tagging of Arg1 on Sbj, which is the subjective argument, the Arg1 semantic role can be tagged to the subjective argument from the argument of the predicate *Mat-da* in the corpus.

In this study, we can assign the semantic roles by applying the frame files to a total of 4,468 arguments. We were able to accurately assign the semantic roles of 4,021 arguments. This tells us that our method shows about 90% accuracy.

However, this method has a high accuracy but also has a problem which cannot be ignored. We have found that this method has a lower recall than expected. The recall ratio was only 29.3%. In Figure 1, *Mat-da* has four arguments, but only one argument, *geu-geot*, could be assigned with its proper semantic role. Despite very high accuracy, this method requires a new complement because of its very low recall.

7. Conclusion

Of the cases that did not match with the manual SRL semantic role, those with no frame, which is equivalent to the verb given to the frame file, were predominant. For the Korean PropBank, a version has been established by the Linguistic Data Consortium of the University of Pennsylvania based on a parallel corpus. However, the ability to apply the current PropBank to semantic analysis is limited. PropBank was established based on a parallel corpus and there are parts that do not match with the processing of Korean because it follows the Penn TreeBank system. Compared to the corpora in English and Chinese, there are limited learning materials for the PropBank and it cannot be used directly as the special characteristics of Korean are not reflected. Currently, information is being collected to redress the deficiency of the Korean Propbank Frame file and the frame file will be adjusted to match with Korean in the future.

Acknowledgments

This paper is a revised and expanded version of a paper entitled “Korean Semantic Role Labeling Using Korean PropBank Frame Files” presented at SIT 2016, Jeju, Korea, December [27].

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and future Planning (2015R1A2A2A01007333), and by Hallym University Research Fund, 2015 (HRF-201512-010).

References

[1] Y. B. Kim, H. Chae, B. Snyder, and Y. S. Kim, "Training a Korean SRL System with Rich Morphological Features", Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL), Baltimore, Maryland, USA, June, **(2014)**, pp. 637-642.
[2] V. Punyakanok, D. Roth and W. T. Yih, "The importance of syntactic parsing and inference in semantic role labeling". Computational Linguistics., vol. 34, no. 2, **(2008)**, pp. 257-287.
[3] K. C. Shet and U. D. Acharya, "A New Similarity Measure for Taxonomy Based on Edge Counting", International Journal of Web & Semantic Technolology (IJWesT)., vol. 3, no. 4, **(2012)**, pp. 23-30.
[4] M. H. Seddiqui and M. Aono, "Metric of Intrinsic Information Content for Measuring Semantic Similarity in an Ontology", Proceedings of the Seventh Asia-Pacific Conference on Conceptual Modelling (APCCM 2010), Brisbane, Australia, January, **(2010)**, pp. 89-96.
[5] Z. Zhou, Y. Wang and J. Gu, "A new model of information content for semantic similarity in WordNet", Proceedings of the Second International Conference on Future Generation Communication and Networking Symposia (FGCNS), Hainan Island, China, December, **(2008)**, pp. 85-89.
[6] G. Pirró, "A semantic similarity metric combining features and intrinsic information content", Data & Knowledge Engineering., vol. 68, no. 11, **(2009)**, pp. 1289-1308.
[7] A. Formica, "Concept similarity in formal concept analysis: An information content approach", Knowledge-Based Systems., vol. 21, no. 1 **(2008)**, pp. 80-87.
[8] T. Pedersen, "Information content measures of semantic similarity perform better without sense-tagged text", Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics, Los Angeles, CA, USA, June, **(2010)**, pp. 329-332.
[9] V. Ganesan, R. Swaminathan, and M. Thenmozhi, "Similarity Measure Based On Edge Counting Using Ontology", International Journal of Engineering Research and Development., vol. 3, no. 3, **(2012)**, pp. 40-44.
[10] J. B. Gao, B. W. Zhang, and X. H. Chen, "A WordNet-based semantic similarity measurement combining edge-counting and information content theory", Engineering Applications of Artificial Intelligence., no. 39, **(2015)**, pp. 80-88.
[11] J. Ge and Y. Qiu, "Concept similarity matching based on semantic distance", Proceedings of the Fourth International Conference on Semantics, Knowledge and Grid, Beijing, China, December, **(2008),** pp. 380-383.
[12] Z. T. T. Myint and K. K. Win, "Conceptual similarity measurement algorithm for domain specific ontology", International Journal of Information Technology, Modeling and Computing (IJITMC)., vol. 2, no. 2, **(2014)**, pp. 29-36.
[13] P. Resnik, "Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language", Journal of Artificial Intelligence Research (JAIR)., no. 11, **(1999)**, pp. 95-130.
[14] K. Hacioglu, "Semantic role labeling using dependency trees," Proceedings of the 20th International Conference on Computational Linguistics, Geneva, Switzerland, August, **(2004)**, pp.1273-1276.
[15] M. Palmer, D. Gildea and N. Xue, "Semantic role Labeling," Synthesis Lectures on Human Language Technologies, Morgan&Claypool Publishers, **(2010)**.
[16] A. Dan and J. Singh, "A Survey on Semantic Role Labeling and Dependency Parsing."
[17] D. Gildea and D. Jurafsky, "Automatic Labeling of Semantic Roles," Computational Linguistics., vol. 28, no. 3, **(2002)**, pp. 245-288.
[18] M. Palmer, D. Gildea, and P. Kingsbury, "The proposition bank: An annotated corpus of semantic roles", Computational Linguistics., vol. 31, no. 1, **(2005)**, pp. 71-106.
[19] C. Bomial, O. Babko-Malaya, J. D. Choi, J. Hwang and M. Palmer, "Propbank annotation guidelines", Available at http://clear.colorado.edu/compsem/documents/propbank_guidelines.pdf, (2010).
[20] C. Bonial, J. Bonn, K. Conger, J. Hwang, and M. Palmer, "PropBank: semantics of new predicate types", Proceedings of the 9th edition of the Language Resources and Evaluation Conference, Reykjavik, Iceland, May, **(2014)**, pp. 3013-3019.
[21] P. Kingsbury, and M. Palmer, "Propbank: the next level of treebank", Proceedings of Treebanks and Lexical Theories, Vaxjo, Sweden, November, **(2003)**.
[22] M. Marcus, G. Kim, M. A. Marcinkiewicz, R. MacIntyre, A. Bies, M. Ferguson, K. Katz and B. Schasberger, "The Penn Treebank: annotating predicate argument structure", Proceedings of the workshop on Human Language Technology, Plainsboro, NJ., USA., March, **(1994)**, pp. 114-119.
[23] J. D. Choi, C. Bonial and M. Palmer, "Propbank Instance Annotation Guidelines Using a Dedicated Editor, Jubilee", Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC), Valletta, Malta, May, **(2010)**, pp. 1871-1875.
[24] O. Babko-Malaya, "Guidelines for Propbank framers", available at http://verbs.colorado.edu/mpalmer_old/old_penn_pages/project_pages/FramingGuidelines.pdf, **(2005)**.
[25] Z. Wu, and M. Palmer, "Verbs semantics and lexical selection", Proceedings of the 32nd annual meeting on Association for Computational Linguistics, Las Cruces, NM., USA., June, **(1994)**, pp. 133-138.

[26] T. Slimani, B. B. Yagahlane, and K. Mellouli, “A New Similarity Measure based on Edge Counting”, Proceedings of the World Academy of Science, engineering and Technology 17. **(2006)**.
[27] M. Seok, H. Song, C. Park, J. Kim and Y. Kim, “Korean Semantic Role Labeling Using Korean PropBank Frame Files,” Advanced Science and Technology Letters., vol. 142 (SIT 2016), **(2016)**, pp. 83-87.

Authors

Miran Seok, **s**he received the B.S. and M.S. Degree from the Department of Convergence Software in Hallym University. She currently works in IT Business Div. of Phill-IT Co. Her research interests focus on virtual input devices.

Hye-Jeong Song, she received the Ph.D. Degree from the Department of Computer Engineering in Hallym University. She is a Research Professor of the Department of Convergence Software, Hallym University. Her research interests focus on biomedical system and bioinformatics.

Chan-Young Park, he received the B.S. and M.S. from Seoul National University and the Ph.D. degree from Korea Advanced Institute of Science and Technology in 1995. From 1991 to 1999 he worked at Samsung Electronics. He is currently a Professor of the Department of Convergence Software, Hallym University. His research interests are in Bio-IT convergence, Intelligent Transportation System, and Sensor Networks.

Jong-Dae Kim, he received the M.S. and the Ph.D. degree from Electrical Engineering from Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1984 and 1990. He worked for Samsung Electronics from 1988 to 2000 as an electrical engineer. He is a Professor of the Department of Convergence Software, Hallym University. His research interests focus on biomedical system and bioinformatics.

Yu-Seop Kim, he received the the Ph.D. degree from Computer Engineering from Seoul National University. He is currently a Professor in the Department of Convergence Softqare at Hallym University, South Korea. His research interests are in the areas of bioinformatics, computational intelligence and natural language processing.

