

Extracting Spatial Entities and Relations in Korean Text

Bogyum Kim and Jae Sung Lee

Dept. of Computer Science

Chungbuk National University

Chungdae-ro 1, Seowon-gu, Cheongju, 28644, Korea

bogyum@cbnu.ac.kr, jasonlee@cbnu.ac.kr

Abstract

A spatial information extraction system retrieves spatial entities and their relationships for geological searches and reasoning. Spatial information systems have been developed mainly for English text, e.g., through the SpaceEval competition. Some of the techniques are useful but not directly applicable to Korean text, because of linguistic differences and the lack of language resources. In this paper, we propose a Korean spatial entity extraction model and a spatial relation extraction model; the spatial entity extraction model uses word vectors to alleviate the over generation and the spatial relation extraction model uses dependency parse labels to find the proper arguments in relations. Experiments with Korean text show that the two models are effective for spatial information extraction.

1 Introduction

A spatial information extraction system retrieves spatially related lexical items and their relationships and then provides the information in a normalized form. This information is used for geological searches and reasoning, and ultimately, for understanding natural language text. For example, from the spatial relations that A is on B and B is on C, a human can simply infer the fact that A is on C. A spatial information extraction system retrieves the relations ‘on (A, B)’ and ‘on (B, C)’ from the text; then, a reasoning program can infer the relation ‘on (A, C)’ from the relations. This enables the system to build a knowledge base with a compact size from text for many intelligent systems such as robot navigation and question-answering systems.

A spatial information extraction task is usually carried out by two sub tasks: spatial entity extraction and spatial relation extraction. Because spatial entity extraction retrieves the entities to be used for spatial relationships, it is different from a place extraction task in named entity recognition systems (Lee et al. 2011), which only retrieves place-related entities. This implies that spatial entity extraction deals with all of the entities involved in spatial relations, such as trajectors, landmarks, and spatial signals. Moreover, spatial signals are usually articles or particles that do not have explicit arguments for spatial relations. Whether some entities or relations are extracted or not depends on the semantic roles in a sentence, which makes the task more complicated and challenging.

Many spatial information extraction systems have been developed for English text, especially those developed through the competition at the SpaceEval conference (Pustejovsky et al. 2015). The application of the techniques directly to Korean text is not simple because of its different linguistic features. The Korean language is a morphologically rich and agglutinative language, which is very different from English (Kim et al. 2016). Moreover, because Korean has a relatively free word order and words are frequently omitted, the order of neighboring words does not always have a significant meaning as in English, where it plays an important role in spatial word classification.

In this paper, we propose two models to extract spatial entities and spatial relations in Korean text. For entity extraction, an ensemble model is used to boost recall, and word vectors are used to tune the results for precision. For relation extraction, a sequence of the dependency labels from the trigger to the argument is used to calculate the argument probability. All of these extraction tasks are based on the ISO-Space mark-up scheme (Pustejovsky et al. 2015). In section 2, related works are briefly reviewed.

The entity extraction method using GloVe word vectors (Pennington et al. 2014) and the relation extraction method using a Bayesian probability follow in sections 3 and 4, respectively. A discussion of the experiments and the conclusions follow in sections 5 and 6, respectively.

2 Related Works

There are generally two approaches to spatial entity extraction, similar to other natural language processing tasks: rule-based and data-driven approaches. The performance of a rule-based approach heavily depends on how much the dictionary covers the open words and how much the rule reflects the linguistic features. This approach usually needs a considerable amount of human labor to encode the rules and fill the dictionary entries manually. Moreover, it is language and domain dependent.

A data-driven approach uses machine learning tools such as CRFs and SVM. A spatial entity extraction task is considered to be a task of sequence labeling, which is solved by using CRFs (Kordjamshidi et al. 2010, Pustejovsky et al. 2015, Roberts and Harabagium 2012, Nichols and Botros 2015) and SVM (Bastianelli et al 2013). Data-driven approaches performed better than rule-based approaches in general and are easily portable to other domains.

The common features for spatial entity extraction based on machine learning are morphemes, named entities, word dependencies, semantic roles, and semantic information such as a WordNet category. Semantic information contributes to the performance. As resources such as WordNet are not easily available to many languages yet, word vectors were used by Bastianelli et al. (2013) and Nichols and Botros (2015), which are generated from a large raw corpus. The word vectors were used to provide semantic information as fine-grained lexical representations and clustered numbers. A spatial entity extraction system for Korean text has been developed by Kim et al. (2015) using a CRFs model, where morphemes, named entities, and parsing results are used as the features. It is based on the SpRL scheme corpus, and preliminary results were provided: with the test using 1,753 annotated sentences from Wikipedia, it was reported that the average F1 score of the entities is 0.610, whereas that of spatial relations is 0.318. As the entities and relations of the annotation scheme in ISO space are different, they are not compared directly to this paper's result.

Dependency parsing results are used for relation extraction. Cross et al. (2011) and Bastianelli et al. (2013) used a parse tree to construct a GRCT (Grammatical Relation Centered Tree) graph for an SVM tree kernel. Jeong et al. (2011) and Kwak et al. (2013) used dependency structures for the relation extraction of Korean sentences; the former used a composite kernel to extract general relations, and the latter built rules to extract spatial relations.

3 Entity Extraction

3.1 Base model

We define a base model called the E1 model, which incorporates the useful features used in prior systems that are applicable to the Korean language. Moreover, we have added more features to improve the performance, such as language-specific features, word phrase spacing, and morpheme-POS (part of speech) tag vectors. For describing the features for the base model, we define the acronyms for the feature description in Table 1. All of the CRFs features for the base model are defined in Table 2 using the acronyms, where a letter means an acronym defined in Table 1, and the attached number is the size of window. For example, MT3 means 'morpheme and POS tag pairs within a 3-morpheme window.' (We use a morpheme window here instead of word window in English text.)

Table 1. Acronyms for the feature elements.

M: morpheme	D: dependency label
T: part of speech tag	H: head's dependency label
B: BI tag of a word phrase spacing	W: main morpheme-POS tag of head
S: sense number of a morpheme	V: cluster number of a morpheme-POS tag vector
N: named entity tag	C: cluster number of the head's morpheme-POS tag vector

Table 2. Features of the E1 model for the CRFs.

All entities		M3, T3, MT3, B3, MS3, MM3, TT3, N3, D1, H1, W1, V3, C1								
M	T	B	S	N	D	H	W	V	C	
...	
충북	NNP	B	00	B-OGG_EDU	B-NP	NP	대학교/NNG	211	181	
대학교	NNG	B	00	I-OGG_EDU	B-NP	NP_AJT	안/NNG	181	298	
안	NNG	B	01	NONE	B-NP_AJT	VP	위치하/VV	298	103	
에	JKB	I	00	NONE	I-NP_AJT	VP	위치하/VV	185	103	
위치하	VV	B	00	NONE	B-VP	VP	있/VX	103	185	
...	
M3:	대학교 안 에				TT3:	NNG-NNG NNG-JKB				
T3:	NNG NNG JKB				W1:	있/VX				
MT3:	대학교-NNG 안-NNG 에-JKB				V3:	181 298 185				

Fig 1. Examples of element features shown in vertical forms and composite features.

3.2 Ensemble model

As the Korean spatial tagged corpus is not large and not well-balanced, machine learning programs such as CRFs are not learned properly. Therefore, we use multiple sub models to overcome the skewness in the data distribution. We assigned the respective features to each entity type, as summarized in Table 3. After testing a candidate with multiple sub models of each entity type, the results are simply accumulated. This is called the E2 model, and this is an interim model for the following final model.

Table 3. Features of each sub model for the CRFs in the E2 model.

Entity type	Feature list
PLACE, PATH	M3, T3, MT3, MM3, TT3, MS3, B3, N3, V3
SPATIAL_ENTITY	M3, T3, MT3, MM3, TT3, MS3, B3
MOTION	M3, T3, MM3, TT3, MS3, B3, V3, D1
MOTION_SIGNAL	M3, T3, MM3, TT3, H1, W1, C1
SPATIAL_SIGNAL	M3, T3, MM3, TT3, N3, H1, W1, C1
MEASURE	M5, T5, MM3, TT3, N3, V3

3.3 Ensemble model using word vectors

As the multiple sub models still produce many false entities, a word vector is used to filter them. The idea is that the common characteristics of entities can be represented in entity tag vectors by summing all of the word vectors learned from the training corpus. The entity tag vectors are used later during testing to check the validity of the candidate entity vectors. Eq. 1 expresses a formula used for the entity tag vector calculation, where the function f converts w_i into a vector representation, and eq. 2 expresses an equation used for the validation method during testing, where θ is the minimum cosine similarity between the entity tag vector (centroid) and a tagged word vector (instance), which is determined during training.

The vectors for the spatial entities are the word vectors of the entities themselves in the training data, as expressed in eq. 3. The function $w2v$ converts an argument word into a vector representation using deep learning programs such as GloVe (Pennington et al. 2014). However, the vectors of the signal entities are the context word vectors of the signal words, as expressed in eq. 4, because the Korean

signals are usually particles, which are too general to be characterized for tag vectors. We propose this model for spatial entity extraction and call it the E3 model.

$$\overrightarrow{STag_j} = \frac{1}{N} \sum_{i=1}^N f(w_i) \quad (1)$$

$$STag(w_i) = \underset{j}{\operatorname{argmax}} (\cos(\overrightarrow{STag_j}, f(w_i)) > \theta_j) \quad (2)$$

$$f(w_i) = w2v(w_i) \quad (3)$$

for $STag(w_i) \in \{\text{PLACE, PATH, SPATIAL_ENTITY, MOTION}\}$ or candidates

$$f(w_i) = \frac{1}{2L} \sum_{l=1}^L (w2v(\text{context}_l(w_i)) + w2v(\text{context}_{-l}(w_i))) \quad (4)$$

for $STag(w_i) \in \{\text{MOTION_SIGNAL, SPATIAL_SIGNAL}\}$ or candidates

4 Relation Extraction

In ISO-Space, a spatial relation consists of two static relations and one dynamic relation. The static relations are the topological relation (QSLink) and orientational link (OLink), which are triggered by SPATIAL_SIGNAL. The extracted relations are represented in a triple format: <trajector, trigger, landmark>. The dynamic relation is the move relation (MoveLink) triggered by a MOTION event. The extracted relation is represented in octuple format: <mover, trigger, source, goal, landmark, mid-point, path, motion signal>. However, it is not easy to extract all the octuplet arguments in most sentences. For a relaxed implementation, the octuple is converted into many triples; then, one or all of the triples are extracted (Nichols and Botros 2015, D’Souza and Ng 2015). We chose to extract one of those triples, <mover, trigger, goal>, because its arguments are filled in most cases, and the triple is also chosen for extraction target in (Nichols and Botros 2015).

4.1 Rule-based model

A rule-based method is straight-forward to implement, if linguistic regularities for spatial relation extraction are easily found. For a performance comparison, we also define a rule-based relation extraction model as a base model and call it the R1 model. Because of the free word order and frequent omission of words in the Korean language, regularities are not easily found. Therefore, the rules do not pose many restrictions, as summarized in Table 4.

Table 4. Relation extraction rules.

Rule 1	Static relations are triggered by SPATIAL_SIGNAL and dynamic relations by MOTION.
Rule 2	SPATIAL_SIGNAL with the type ‘TOPOLOGICAL’ generates QSLink, and that with the type ‘DIRECTIONAL’ generates both OLink and DIR_TOP.
Rule 3	The arguments for the relations are ‘PLACE,’ ‘PATH,’ and ‘SPATIAL_ENTITY.’
Rule 4	All spatial entities for a relation are within the same dependency head (VP, VNP).
Rule 5	When there is more than one argument under a dependency head, the argument closest to the trigger in the dependency relation is classified as a landmark, and the other arguments are classified as trajectors, resulting in multiple relations.
Rule 6	If one and more triggers exist, the arguments cannot cross the other triggers at a sentence position.

4.2 Bayesian model

Dependency parsing is quite effective for free word order languages such as the Korean language and provides useful information for long-distance relations (Lim et al. 2014). We utilize the parsing result to find valid arguments for given triggers such as the SPATIAL_SIGNAL or MOTION tag. In this model, all of the possible argument candidates are searched and verified with a Bayesian probability, which is learned with the training corpus. The argument with the highest probability is chosen, as shown in eq. 5. The probability is calculated as the product of the prior probability of an argument and the conditional probability of the sequence of dependency labels (DPL). The DPL includes the labels from a trigger to the argument, and its conditional probability is approximated in eq. 6.

$$\begin{aligned}
DPL &= (dpl_1, dpl_2, \dots, dpl_n) \\
A &= \{\text{trajector, landmark}\} \\
\text{argmax}_{A_i} P(A_i, DPL) &= \text{argmax}_{A_i} P(DPL|A_i) \cdot P(A_i) \tag{5}
\end{aligned}$$

$$\begin{aligned}
P(DPL|A_i) &= P(dpl_n, dpl_{n-1}, \dots, dpl_1|A_i) \\
&= P(dpl_n|dpl_{n-1}, \dots, dpl_1, A_i) \cdot P(dpl_{n-1}|dpl_{n-2}, \dots, dpl_1, A_i) \cdots P(dpl_1|A_i) \\
&\cong P(dpl_1|A_i) \cdot \prod_{j=2}^n P(dpl_j|dpl_{j-1}) \tag{6}
\end{aligned}$$

5 Experiment

5.1 Experimental setup

As pre-processing steps, a morphological analysis and POS tagging, named entity recognition, and dependency parsing are carried out, and their sources and performance are summarized in Table 5. We used CRFSuite (Okazaki 2007) and the GloVe word vector (Pennington et al. 2014). The word vectors are trained with the morpheme-tagged data in the Sejong corpus (NIKL 2011) to build 300 vector clusters for a feature set.

For the test data, we used the Korean spatial annotation corpus (Kim et al. 2016), which is constructed from 175 documents (1593 sentences) from the Wikitravel web-site¹ following the SpaceEval annotation scheme (Pustejovsky et al. 2015). The testing corpus statistics are listed in Table 6.

The experiment was performed with 5-fold cross validation test. The experiment for entity extraction was directly carried out with the raw corpus data, and relation extraction was performed with the corpus annotated with spatial entities beforehand. (Each experiment corresponds to tasks 1.b and 3.a of SpaceEval)

Table 5. Performance of the pre-processing modules.

Modules	performance	source
Morph. analysis and POS tagging.	99.03% (pre)	(Lee et al. 2016)
Named entity recognition	86.86% (f1)	(Lee et al. 2011)
Dependency parsing	87.63% (LAS)	(Lim et al. 2014)

Table 6. Number of tags in the testing corpus.

Entity						relation		
name	num	ratio	name	num	ratio	name	num	ratio
PLACE	5,636	67.6%	M_SIGNAL	266	3.2%	QSLink	3,548	65.9%
PATH	320	3.8%	S_SIGNAL	1,299	15.6%	OLink	970	18.0%
S_ENTITY	270	3.2%	MEASURE	248	3.0%	MoveLink	868	16.1%
MOTION	294	3.5%						
entity total				8,333	100.0%	relation total	5,386	100.0%

5.2 Results

Table 7 summarized the results of spatial entity extraction. The performance of MEASURE is the second-best because its typical surface form, e.g., the “number + unit” form, is very easily recognized by a program. PLACE is the best performer, which can also be easily found by a named entity recognizer. Moreover, as the ratio of PLACE tag is the largest, 67.6%, in the distribution as presented in Table 6, the prior probability contributes to find more PLACE tags. On the other hand, SPATIAL_ENTITY exhibits the worst performance. We conjecture that the first reason is that the size of the training corpus is too small; the number of Korean spatial entity tags is 270 (3.2%), as presented in Table 6, whereas that of English spatial entity tags is 1670 (23.6%) in the corpus used for SpaceEval 2015 (Pustejovsky et al. 2015). The second reason is that its part of speech tag is a general noun, which is not easy to distinguished from other spatial tags. Moreover, the same word can be either a SPATIAL_ENTITY tag or

¹ <http://www.wikitravel.com/ko/>

none tag depending on the context. In the following example, ‘car’ in the first sentence is tagged SPATIAL_ENTITY, but that in the second sentence is not. This is same for English, but it is more difficult for a free-word-order language to be disambiguated.

철수가 자동차/spatial_entity 에 탔다. (Cheolsu got in a car/spatial_entity.)

철수가 자동차/none 를 샀다. (Cheolsu bought a car/none.)

The E2 model increased the recall but decreased the precision, as indicated in Table 7. However, the E3 model using spatial tag vectors greatly increased the precision by 13.3% point compared to that of the E2 model. Consequently, the F1 measure performance of the E3 model increased by 2.1% point compared to that of the E1 model, which means that use of the tag vector is effective for selecting valid spatial entities.

Table 7 summarizes a comparison of the E3 model with SpRL-CWW (Nichols and Botros 2015), which was the best model at SpaceEval 2015 (Pustejovsky et al. 2015). The overall performance of the E3 model is better than that of SpRL-CWW for all precision, recall, and F1 measure criteria. The performance of SPATIAL_ENTITY and PATH for the E3 model, however, is relatively much lower than that of SpRL-CWW. This means that the ambiguity of general nouns in a semantic role is still problematic and the size of the training corpus is relatively smaller as fore-mentioned.

The performance of relation extraction is summarized in Table 8. All of the values of the precision, recall, and F1 measure for the R2 model are better than those for the R1 model. This implied that the use of the Bayesian probability for selecting arguments is effective. For a general comparison, we have listed the results of the two best approaches in the table, where CWW (Nichols and Botros 2015) is the best machine learning approach, and Pust (Pustejovsky et al. 2015) is the best rule-based approach. Unfortunately, the R2 model exhibits a very low performance compared with both of them.

The spatial relations in Korean text are relatively hard to be retrieved, because the related entities are relatively separated and their appearing order is not consistent. While all related entities are usually located closely to the spatial signal in English text, the entities are sometimes far from the spatial signal in a Korean sentence. Sentence 1 in Fig. 2 shows an example in which the trajector is far from the trigger. In addition, the appearing order is not consistent as shown in Korean sentence 2 in Fig. 2. Both landmark and trajectory appear before the trigger in the first OLINK, whereas landmark appears before and trajectory appears after the trigger in the second OLINK. We used the dependency relations of words to alleviate this problem and thus improved the performance. However, we still need to find more effective methods to overcome Korean linguistic barriers such as the free word order and the lack of language resources; and this problem will be studied in future research.

Table 7. Performance of spatial entity extraction.

Label	Precision				Recall				F1			
	E1	E2	E3	CWW	E1	E2	E3	CWW	E1	E2	E3	CWW
PLACE	0.919	0.917	0.961	0.802	0.928	0.930	0.958	0.777	0.923	0.924	0.960	0.789
PATH	0.848	0.441	0.552	0.815	0.397	0.543	0.539	0.614	0.541	0.487	0.545	0.701
S. ENTITY	0.463	0.210	0.326	0.793	0.213	0.444	0.444	0.653	0.292	0.285	0.376	0.716
MOTION	0.801	0.354	0.544	0.823	0.479	0.713	0.709	0.7	0.600	0.473	0.616	0.756
M. SIGNAL	0.800	0.236	0.556	0.766	0.392	0.698	0.694	0.6	0.536	0.353	0.617	0.673
S. SIGNAL	0.851	0.770	0.892	0.75	0.729	0.794	0.836	0.603	0.786	0.782	0.863	0.668
MEASURE	0.990	0.951	0.951	0.889	0.881	0.906	0.906	0.707	0.936	0.928	0.928	0.788
Overall	0.894	0.728	0.861	0.795	0.849	0.855	0.880	0.674	0.849	0.786	0.870	0.73

E1: base model, E2: ensemble model, E3: proposed ensemble model using word vector, CWW: 5-fold cross validation (Nichols and Botros 2015)

Table 8. Performance of spatial relation extraction.

Relation	Precision				Recall				F1			
	R1	R2	CWW	Pust	R1	R2	CWW	Pust	R1	R2	CWW	Pust
QSLink	0.40	0.49	0.66	-	0.51	0.55	0.54	-	0.45	0.52	0.59	-
OLink	0.12	0.24	0.69	-	0.19	0.48	0.52	-	0.15	0.32	0.59	-
MoveLink	0.18	0.24	0.57	-	0.35	0.65	0.45	-	0.24	0.35	0.5	-
Overall	0.30	0.37	0.64	0.86	0.42	0.54	0.50	0.84	0.35	0.44	0.56	0.85

R1: base model, R2: proposed model using dependency label, CWW: (Nichols and Botros 2015), Pust: Baseline 3.a (Pustejovsky et al. 2015)

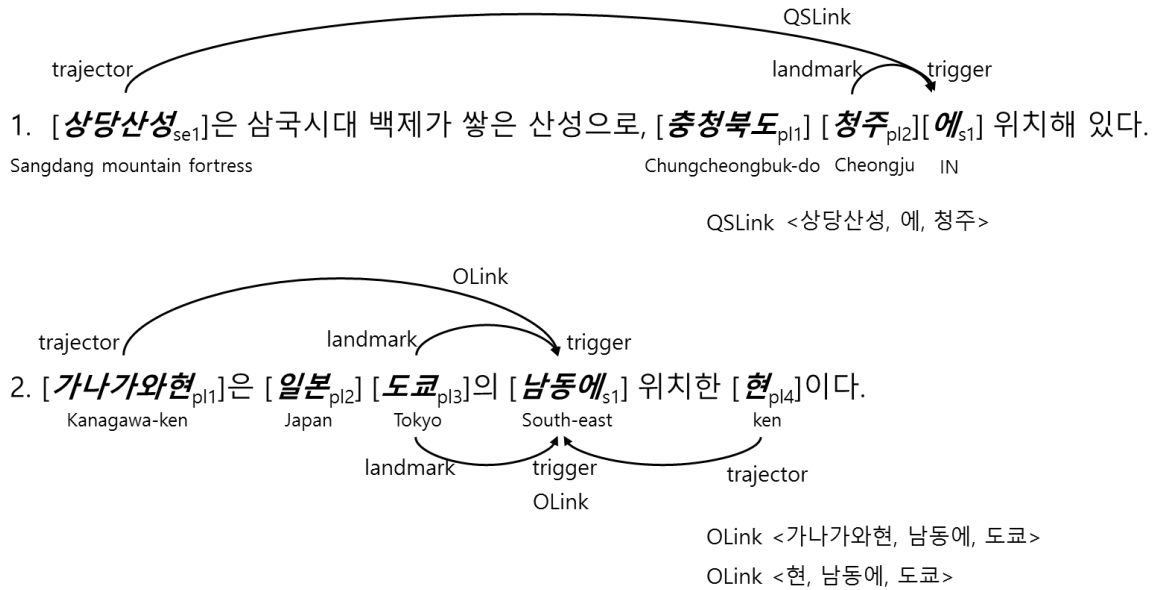


Fig 2. Various types of spatial relation caused by free word order.

6 Conclusion

We have proposed two models for Korean spatial entity extraction and spatial relation extraction. For entity extraction, we utilized the features of prior systems with an adaptation to Korean linguistic features. Moreover, we proposed a new approach to filter false entities using spatial tag vectors, which can be learned automatically from a raw corpus. The experiment showed that the spatial tag vectors are effective for spatial entity extraction and showed better performance than English state-of-the-art performance.

For relation extraction, we proposed a model that uses the dependency label probability to select proper arguments, which is effective and better than a simple rule-based model but much lower than the state-of-the-art performance of an English one. We conjecture that this mainly originates from linguistic differences, especially syntactic structures such as the free word order and word omission, which still require further investigation.

Acknowledgement

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP). (No. R0101-16-0062, Development of Knowledge Evolutionary WiseQA Platform Technology for Human Knowledge Augmented Services)

Reference

- Bogyum Kim, Yongmin Park, and Jae Sung Lee. 2015. Automatic space information extraction from Korean text. *Journal of Information*, 18(7):2953-2962.
- Bogyum Kim, Myung Yun Kang, and Jae Sung Lee. 2016. Issues in spatial information annotation in Korean texts. *The First International Workshop on Spatial/Temporal Information Extraction from Unstructured Texts*, pages 458-461.
- Changki Lee, Pum-Mo Ryu, and HyunKi Kim. 2011. Named entity recognition using a modified Pegasos algorithm. In *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, pages 2337-2340. ACM.
- Chang-Hoo Jeong, Sung-Pil Choi, Yun-Soo Choi, Sa-Kwang Song, and Hong-Woo Chun. 2011. Relation extraction based on composite kernel combining pattern similarity of predicate-argument structure. *Journal of Internet Computing and Services*, 12(5):73-85.
- Chung-Hee Lee, Joon-Ho Lim, Soojong Lim, and HyunKi Kim. 2016. Syllable-based Korean POS tagging based on combining a pre-analyzed dictionary with machine learning. *Journal of Korean Institute of Information Scientists and Engineers*, 43(3):362-369.
- Danilo Croce, Alessandro Moschitti, and Roberto Basili. 2011. Structured lexical similarity via convolution kernels on dependency trees. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1034-1046, Association for Computational Linguistics.
- Emanuele Bastianelli, Danilo Croce, and Roberto Basili. 2013. UNITO-HMM-TK: Structured kernel-based learning for spatial role labelling. In *Second Joint Conference on Lexical and Computational Semantics, SemEval 2013*, pages 573-579.
- Eric Nichols and Fadi Botros. 2015. SpRL-CWW: Spatial relation classification with independent multi-class models. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, pages 895-901.
- Haritz Salaberri, Olatz Arregi, and Beñat Zepirain. 2015. IXAGroupEHUSpaceEval: (X-Space) A WordNet-based approach towards the automatic recognition of spatial information following the ISO-Space annotation scheme. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, pages 856-891.
- James Pustejovsky, Parisa Kordjamshidi, Marie-Francine Moens, Aaron Levine, Seth Dworkman, and Zachary Yocum. 2015. SemEval-2015 task 8: SpaceEval. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, pages 884-894.
- Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global vectors for word representation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1532-1543.
- Jennifer D'Souza and Vincent Ng. 2015. UTD: Ensemble-based spatial relation extraction. In *Proceedings of the 9th International Workshop on Semantic Evaluation*, pages 862-869.
- Joon-Ho Lim, Yeo-Chan Yoon, Yongjin Bae, Su-Jong Im, Hyunki Kim, and Kyu-Chul Lee. 2014. Korean dependency parsing model based on transition system using head final constraint. In *Proceedings of the 27th Annual Conference on Human & Cognitive Language Technology*, pages 81-86.
- Kirk Roberts and Sanda M. Harabagium. 2012. UTD-SpRL: A joint approach to spatial role labeling. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics Volume 2: Proceedings of the 6th International Workshop on Semantic Evaluation*, pages 419-424. Association for Computational Linguistics.
- Naoaki Okazaki. 2007. CRFsuite: A first implementation of conditional random fields (CRFs). <http://www.chokkan.org/software/crfsuite/>.
- NIKL (National Institute of Korean Language). 2011. 21st century Sejong project final result, revised edition.
- Parisa Kordjamshidi, Martijn Van Otterlo, and Marie-Francine Moens. 2010. Spatial role labeling: Task definition and annotation scheme. In *Proceedings of the 7th Conference on International Language Resources and Evaluation*, pages 413-420, European Language Resources Association.
- Sujeong Kwak, Bogyum Kim, and Jae Sung Lee. 2013. Triplet extraction using Korean dependency parsing result. In *Proceedings of the 25th Annual Conference on Human & Cognitive Language Technology*, pages 86-89.